AT&T Technical Journal
Vol. 64, No. 8, October 1985
Printed in U.S.A.

Nonlinear Input-Output Maps and Approximate Representations*

By I. W. SANDBERG†

(Manuscript received May 20, 1985)

An approximation theorem is given for causal time-invariant nonlinear maps that take one set of functions defined on $[0, \infty)$ into another. The theorem is used to show that, under some typically very reasonable conditions, an input-output map can be approximated arbitrarily well in a meaningful sense by a finite Volterra series, even though it may not have a Volterra series expansion. The set of inputs on which the approximation holds need not be compact, and the inputs need not be continuous.

I. INTRODUCTION

In this paper an approximation theorem is given for causal time-invariant nonlinear maps that take one set of functions defined on $[0, \infty)$ into another. The theorem is used to show that, under some typically very reasonable conditions, an input-output map can be approximated arbitrarily well in a meaningful sense by a finite Volterra series, even though it may not have a Volterra series expansion. The set of inputs on which the approximation holds need not be compact. A more detailed introduction follows.

1.1 Background

Researchers have long been interested in a variety of questions concerning the mathematical representation of systems that need not

* This paper was presented at the Midwest Symposium on Circuits and Systems, Louisville, Kentucky, August 19–20, 1985.
† AT&T Bell Laboratories.

be linear. In recent years much has been learned about the existence, determination, and properties of power-series-like expansions for expressing a system's outputs in terms of its inputs (see, for instance, Refs. 1 through 7). An important example of the form of such an expansion is

$$w(t) = \sum_{q=1}^{\infty} \int_0^t \cdots \int_0^t k_q(\tau_1, \cdots, \tau_q)u(t - \tau_1) \cdots u(t - \tau_q)d\tau_1 \cdots d\tau_q, \qquad t \geq 0, \quad (1)$$

in which w is the output, u is the input, the kernels k_q are functions determined by the system, and u is drawn from a set of bounded real-valued inputs such that the right side of (1) converges uniformly in t for every input. While certain more general versions of (1), where the k_q are *symbolic* functions and vector-valued inputs are taken into account, are frequently needed, in all cases one has

$$w = \sum_{q=1}^{\infty} K_q(u), \qquad u \in U, \quad (2)$$

in which U is a set of inputs and each K_q is a homogeneous map of degree q. Under weak assumptions these K_q are *uniquely* determined, and as such are very special associates of the system represented by (2).

The right side of (1) is an example of what is often called a Volterra series. Actually, Volterra considered not (1) but related expansions in which the integration limits are constants and the $u(t - \tau_j)$ are replaced with $u(\tau_j)$. These expansions were used by Volterra as a model of a nonlinear functional in his path-breaking studies of operations on functionals. A comprehensive account of this work is given in Ref. 8, where attention is directed to a representation result (Ref. 8, page 20) due to Fréchet that concerns, in particular, the uniform approximation of continuous functionals on compact sets, using a finite number of terms in Volterra's series. Volterra also mentions the analogy between this aspect of Fréchet's result and the Weierstrass approximation theorem for continuous real functions on a real compact interval.

While Fréchet's Weierstrass-like result is certainly interesting and important, it does have significant limitations with regard to the representation of input-output maps: (1) it concerns *approximations* rather than expansions in the usual sense, (2) these approximations are on *compact* sets, (3) it directly concerns functionals rather than mappings from one function space to another, etc. These limitations, as well as those of the more general representation result of Fréchet described in Ref. 8, p. 20, do not appear to have been always appreciated by early writers concerned with applications of Volterra series, who sometimes cited Ref. 8 as though it contained a justification for the use of relations of the form (1).

There can be very basic differences between an arbitrarily good approximation and an expansion, and, more to the point, between knowing that one, as opposed to the other, exists. For example, an approximation known to exist may not have properties that facilitate its determination. Nevertheless, existence results concerning approximations can sometimes be useful, especially when an expansion does not exist. Thus, it is clearly of interest to consider the extent to which input-output maps of systems can be approximated in some meaningful sense by a finite sum,

$$\sum_{q=1}^{Q} \int_0^t \cdots \int_0^t k_q(\tau_1, \cdots, \tau_q) u(t - \tau_1) \cdots u(t - \tau_q) d\tau_1 \cdots d\tau_q, \qquad t \geqslant 0, \quad (3)$$

of terms of the form that appear on the right side of (1), or by a finite sum of suitably more general terms if u is vector valued. Of course, approximations involving larger classes of finite sums of iterated integrals can be of interest too.

Related questions have in fact been considered for many years,[9–15] and as one might expect, the main mathematical tool that has been used is the Stone-Weierstrass theorem. In this earlier work the input signals considered are assumed to belong to a Hilbert space (e.g., an L_2 space) and/or to be defined on only a *finite* interval, and the inputs are taken to belong to a *compact* set. In contrast, in Ref. 16 an approximation result is given for input-output maps that act between certain subsets of the Banach space $C(\mathbb{R})$ of bounded, continuous, real-valued functions defined on the doubly infinite interval $(-\infty, \infty)$, with the usual norm. The maps considered there are assumed to be time invariant and to have a "fading-memory" property that enables one to prove that a certain set of functions defined on $(-\infty, 0]$ is compact. The extent to which the results in Ref. 16 bear on the main problem considered in this paper, where the input and output signals are defined on $[0, \infty)$, is not discussed in Ref. 16. Although the results in this paper are considerably different from those in Ref. 16, there are some similarities: an approximately-finite memory hypothesis (related to hypotheses in Ref. 17, Section 2.2) plays a central role, and we too depend on a form of the Stone-Weierstrass theorem. On the other hand, the compact sets with which we deal are always sets of functions defined on a *finite* interval $[0, \omega]$.

1.2 Outline of this paper's results

In Section II, attention is focussed on a class of causal time-invariant maps G that take S into S_0, where S and S_0 are sets of signals (i.e., sets of functions) defined on $[0, \infty)$, and the elements of S_0 are real valued. The maps G are assumed to possess a factorization FH, where H takes S into S_1, and F maps S_1 into S_0, where S_1 is a third set of signals on $[0, \infty)$. Certain hypotheses on S, S_0, S_1, and the factors F and H are introduced in Theorem 1 of Section 2.2. Under those conditions, the theorem shows that given any $\epsilon > 0$, there are a constant $\Delta \geqslant 0$, and a map P having an important special form (that involves a real polynomial p in several variables together with a certain "fundamental set" of maps) such that the approximation

$$|G(u)(t) - (PH)(u_{\max[0,t-\Delta]})(t)| < \epsilon, \qquad t \geqslant 0$$

holds for all $u \in S$, where u_ω for arbitrary nonnegative ω is defined by $u_\omega(t) = 0$ if $0 \leqslant t < \omega$, and $u_\omega(t) = u(t)$ for $t \geqslant \omega$. One of the main hypotheses used is that the memory of G is "approximately finite," to the extent that for any $\delta_0 > 0$ there is a $\delta > 0$ for which

$$|Gu(t) - Gu_{\max[0,t-\delta]}(t)| < \delta_0$$

for $t \geqslant 0$ and $u \in S$. This hypothesis can be shown to be satisfied in many cases of interest. More will be said about this later.

The hypotheses of Theorem 1 are of an abstract nature and so is its conclusion. The theorem is used as a "tool theorem" in the proof of Theorem 2 in Section 2.4 which addresses a case that is of direct interest in applications. In Theorem 2, the memory of G is assumed to be approximately finite, S is taken to be a set of uniformly bounded vector-valued functions on $[0, \infty)$, S_1 is a similar set, H is a convolution, and F is causal, time invariant, and continuous in a certain typically reasonable sense (see the theorem for additional details). Assume now for the sake of simplifying the discussion that the elements of S are *scalar* valued. According to the theorem, under the conditions stated there, Gu can be uniformly approximated arbitrarily well on S by a finite sum of the form (3). Again, the reader is referred to the theorem for the details. Related results concerning discrete-time cases and composites of maps that have approximately-finite memory are given in Sections 2.5 and 2.3, respectively.

The case considered in Theorem 2 arises often. This is discussed in Appendix B, where an important class of input-output maps is addressed, and where a technique for showing that the memory of a nonlinear map is approximately finite is illustrated.

II. INPUT-OUTPUT MAPS AND APPROXIMATIONS

2.1 Preliminaries

Throughout Section II, V is a linear space, Ω denotes the interval $[0, \infty)$, t and ω are elements of Ω, S and S_0 are two sets of functions on Ω, the elements of S and S_0 take values in V and $(-\infty, \infty)$, respectively, and G is a map from S to S_0.

We use S_1 to denote a third set of functions on Ω; these take values in a normed linear space V_1. It is assumed that

$$G = FH,$$

where H maps S into S_1 and F takes S_1 into S_0.

The set S, which is our set of inputs, is assumed to have the following properties:

(i) $u \in S \Rightarrow (u)_\omega \in S$ for each ω, where $(u)_\omega(t) = u(t)$ for $t \leq \omega$, and $(u)_\omega(t) = 0$ (here the zero element of V) otherwise.

(ii) $u \in S \Rightarrow (T_\omega u) \in S$ for $\omega \neq 0$, in which T_ω is defined by $(T_\omega u)(t) = 0$ $(0 \leqslant t < \omega)$ and $(T_\omega u)(t) = u(t - \omega)$ for $t \geqslant \omega$.

(iii) $u \in S \Rightarrow u_\omega \in S$ for $\omega \neq 0$, where u_ω is given by $u_\omega(t) = 0$ for $0 \leqslant t < \omega$ and $u_\omega(t) = u(t)$ for $t \geqslant \omega$. [Note the distinction between u_ω and $(u)_\omega$ of Property (i).]

(iv) $\omega \neq 0$ and $u \in S$ with $u(t) = 0$ $(0 \leqslant t < \omega) \Rightarrow v \in S$, where $v(t) = u(t + \omega)$, $t \geqslant 0$.

Notice that (i)–(iv) simply require that S be closed under certain elementary operations. With regard to S_1 we assume that

(v) Properties (ii) and (iii) hold with S replaced with S_1.

We use the standard definitions of causality and time invariance. That is, a map M from S to S_0, from S to S_1, or from S_1 to S_0 is *causal* if u_1 and u_2 in the domain of M with $u_1(t) = u_2(t)(0 \leqslant t \leqslant \omega)$ always implies that $(Mu_1)(t) = (Mu_2)(t)$ for $0 \leqslant t \leqslant \omega$; M is *time invariant* if $\omega \neq 0$ and u in the domain of $M \Rightarrow (MT_\omega u)(t) = 0$ for $0 \leqslant t < \omega$ and $(MT_\omega u)(t) = (Mu)(t - \omega)$, $t \geqslant \omega$. Also, by $M \in \mathscr{A}(S)$ or $M \in \mathscr{A}(S_1)$ we mean that the domain of M is S or S_1, respectively, and that M has "approximately-finite memory" in the sense that for each constant $\delta_0 > 0$ there is a positive $\delta \in \Omega$ such that

$$|(Mu)(t) - (Mu_{\max[0,t-\delta]})(t)| < \delta_0, \qquad t \geqslant 0$$

for all u in the domain of M. Here $|\cdot|$ denotes simply the absolute value if the range of M is S_0; it denotes the norm in V_1 otherwise.

The set of functions x defined on $[0, \omega]$ with values in V_1 such that $x(t) = y(t)$ $(0 \leqslant t \leqslant \omega)$ for some $y \in S_1$ is denoted by $(S_1)_\omega$ for each ω. We use $H(S)_\omega$ to stand for the set of functions x in $(S_1)_\omega$ such that $x(t) = (Hu)(t)$ $(0 \leqslant t \leqslant \omega)$ for some $u \in S$. It is assumed in Theorem 1 (below) that the $H(S)_\omega$ have the following property:

(vi) There is a family of metric spaces $\{(X_\omega, \rho_\omega):\omega > 0\}$ such that for each $\omega > 0$ we have $H(S)_\omega \subset X_\omega \subset (S_1)_\omega$ and X_ω is compact (i.e., compact in itself) with respect to the metric ρ_ω.

For $\omega \neq 0$ and each causal map M from S_1 to S_0, M_ω denotes the functional on $(S_1)_\omega$ defined by

$$M_\omega x = (My)(\omega), \qquad x \in (S_1)_\omega, \tag{4}$$

where $y \in S_1$ satisfies $x(t) = y(t)$, $0 \leq t \leq \omega$. The following is assumed with regard to F in Theorem 1 (below):

(vii) F is causal and for each $\omega \neq 0$, F_ω is *continuous* on X_ω with respect to ρ_ω [where X_ω and ρ_ω are described in (vi)].

Finally, by a *fundamental set* $\{F_\alpha : \alpha \in \Lambda\}$ of maps from S_1 to S_0 relative to (vi), we mean that the F_α are causal and time invariant, and that for $\omega \neq 0$ the corresponding family $\{F_{\alpha\omega} : \alpha \in \Lambda\}$ of functionals on $(S_1)_\omega$ is continuous on X_ω with respect to ρ_ω, and separates the points of X_ω. (By "separates the points of X_ω" is meant Ref. 18, p. 41 that for each pair of distinct elements x_1 and x_2 of X_ω there is an $\alpha \in \Lambda$ such that $F_{\alpha\omega}x_1 \neq F_{\alpha\omega}x_2$.)

2.2 The main approximation result

In this section we prove the following:

Theorem 1: Let (i)–(vii) be met, with F and H causal and time invariant, and with $G \in \mathscr{A}(S)$. Suppose that there is a fundamental set $\{F_\alpha : \alpha \in \Lambda\}$ of maps from S_1 to S_0 relative to (vi). Then for each $\epsilon > 0$ there are a $\Delta \in \Omega$, a positive integer k, elements $F_{\alpha_1}, \cdots, F_{\alpha_k}$ of $\{F_\alpha : \alpha \in \Lambda\}$, and a real polynomial p in k real variables with $p(0, \cdots, 0) = 0$ such that

$$|G(u)(t) - (PH)(u_{\max[0,t-\Delta]})(t)| < \epsilon, \qquad t \geq 0 \tag{5}$$

for every $u \in S$, where P is the map from S_1 into S_0 given by $(Py)(t) = p[F_{\alpha_1}(y)(t), \cdots, F_{\alpha_k}(y)(t)]$, $t \geq 0$ for $y \in S_1$. In addition, the map $Q: S \to S_0$ defined by $(PH)(u_{\max[0,t-\Delta]})(t) = (Qu)(t)$ for $u \in S$ and $t \geq 0$ is causal and time invariant.

2.2.1 Proof of Theorem 1

Proof: Given ϵ, choose a positive $\Delta \in \Omega$ so that

$$|(Gu)(t) - G(u_{\max[0,t-\Delta]})(t)| < \epsilon/2, \qquad t \geq 0 \tag{6}$$

for $u \in S$. Observe that S contains an element θ such that $\theta(t) = 0$ for $t \geq 0$. By the time-invariance of H, $H(S)$ also contains such an element, and therefore there is $e \in X_\Delta$ such that $e(t) = 0$ for $t \in [0, \Delta]$. By the causality and time invariance of F, we have $F_\Delta e = 0$. Using a version of the Stone-Weierstrass theorem (see Ref. 18, p. 46), Condition (vii), and the hypothesis that there is a fundamental set $\{F_\alpha : \alpha \in \Lambda\}$ of maps from S_1 to S_0 relative to (vi),* there are a positive integer k, a polynomial p as described, and elements $F_{\alpha_1}, \cdots, F_{\alpha_k}$ of the fundamental set such that

$$|F_\Delta x - P_\Delta x| < \epsilon/2, \qquad x \in X_\Delta,$$

* It was necessary to establish that $F_\Delta e = 0$ because $F_{\alpha\Delta} e = 0$ for all $\alpha \in \Lambda$ (see Ref. 18, Theorem 5).

where the functional P_Δ is the associate [see (4)] of P described in the statement of the theorem. Thus, using $H(S)_\Delta \subset X_\Delta$, we have

$$|(FHu)(\Delta) - (PHu)(\Delta)| < \epsilon/2, \qquad u \in S. \tag{7}$$

Now let any $u \in S$ and $t \in \Omega$ be given. Suppose first that $t > \Delta$. Let v be defined by $v(\tau) = u_{(t-\Delta)}[\tau + (t - \Delta)]$, $\tau \geqslant 0$ [notation of Condition (iii)]; here we have used Conditions (iii) and (iv). By the time invariance of $G = FH$ and of PH in (7), one has $G[u_{(t-\Delta)}](t) = G(v)(\Delta)$ and $(PH)[u_{(t-\Delta)}](t) = (PH)(v)(\Delta)$. Thus, by (7),

$$|G[u_{(t-\Delta)}](t) - (PH)[u_{(t-\Delta)}](t)| < \epsilon/2.$$

Using this and (6),

$$|G(u)(t) - (PH)[u_{\max(0,t-\Delta]}](t)| \leqslant |G(u)(t) - G(u_{\max[0,t-\Delta]})(t)|$$
$$+ |G(u_{\max[0,t-\Delta]})(t) - (PH)(u_{\max[0,t-\Delta]})(t)| < \epsilon/2 + \epsilon/2 = \epsilon.$$

Suppose now that $t < \Delta$. Then by Conditions (i) and (ii) and the causality and time invariance of G, we see that $G(u)(t) = G[(u)_t](t) = G[T_{(\Delta-t)}(u)_t](\Delta)$ [notation of Conditions (i) and (ii)], and similarly $(PH)(u)(t) = (PH)[T_{(\Delta-t)}(u)_t](\Delta)$. Thus, using (7),

$$|(Gu)(t) - (PH)(u)(t)| < \epsilon/2. \tag{8}$$

Since (8) holds also for $t = \Delta$, and obviously $u = u_{\max[0,t-\Delta]}$ when $t \leq \Delta$, we have (5). At this point it suffices to prove the following:

Lemma: Let K be a causal time-invariant map of S into S_0, and with $\Delta \in \Omega$, let M be the map from S to S_0 given by

$$(Mu)(t) = K(u_{\max[0,t-\Delta]})(t), \qquad t \geqslant 0$$

for $u \in S$. Then M is causal and time invariant.

Proof: Let u_a and u_b in S satisfy $u_a(\tau) = u_b(\tau)$ for $0 \leqslant \tau \leqslant \omega$, and let $t \in [0, \omega]$. Since $(u_{a\max[0,t-\Delta]})_t = (u_{b\max[0,t-\Delta]})_t$ and K is causal, it is clear that $(Mu_a)(t) = K[(u_{a\max[0,t-\Delta]})_t](t) = K[(u_{b\max[0,t-\Delta]})_t(t) = (Mu_b)(t)$, showing that M is causal.

Now let $u \in S$ and let ω in Ω be nonzero. For $t < \omega$, $(MT_\omega u)(t) = K[\{T_\omega u\}_{\max[0,t-\Delta]}](t) = 0$, because K is time invariant and $\{T_\omega u\}_{\max[0,t-\Delta]}(\tau) = 0$ for $\tau < \omega$. Suppose that $t \geqslant \omega$. Since (as can easily be verified) $\{T_\omega u\}_{\max[0,t-\Delta]} = T_\omega u_{\max[0,t-\omega-\Delta]}$, one has $(MT_\omega u)(t) = K(T_\omega u_{\max[0,t-\omega-\Delta]})(t) = K(u_{\max[0,t-\omega-\Delta]})(t-\omega) = (Mu)(t-\omega)$, by the time invariance of K. Thus M is time invariant. This completes the proof of the lemma and of the theorem.

2.3 Comments

All of the material in Sections 2.1 and 2.2 remains valid if Ω is replaced throughout with $\{0, 1, \cdots\}$ (with the understanding that then $[0, \omega]$ means $\{0, \cdots, \omega\}$).

The following result concerning the hypothesis that $G \in \mathscr{A}(S)$ is frequently useful. By "$F \in \text{Lip}(H(S))$" below is meant that there is a positive constant c such that

$$|(Fy_1)(t) - (Fy_2)(t)| \leq c \sup_{\tau\in[0,t]} |y_1(\tau) - y_2(\tau)|$$

for $t \geq 0$ and y_1 and y_2 in $H(S)$, where $|\cdot|$ on the right side denotes the norm in V_1.

Proposition: Let (i)–(v) be met,* with $H \in \mathscr{A}(S)$, $F \in \mathscr{A}(S_1)$, and $F \in \text{Lip}[H(S)]$. Then $G \in \mathscr{A}(S)$.

The proposition is proved in Appendix A. For an application, see Appendix B.

2.4 Approximations and finite Volterra series

In the following theorem, n and p are arbitrary positive integers and $L_\infty(n)$ and $L_\infty(p)$, respectively, denote the normed linear spaces of real n-vector valued and real p-vector valued Lebesgue measurable functions u defined on Ω such that $\|u\| \triangleq \sup_t |u(t)| < \infty$, in which $|u(t)| = \max_j |u_j(t)|$ and $u_j(t)$ stands for the jth component of $u(t)$. By a "vector," we mean a column vector. Also, for any positive integer q and any q n-vectors $a_1, \cdots, a_q$, we use $\chi[a_1, \cdots, a_q]$ to denote the vector of order n^q whose elements are the n^q distinct products $(a_1)_{\lambda_1} \cdots (a_q)_{\lambda_q}$, corresponding to distinct sequences $\lambda_1, \cdots, \lambda_q$ with each λ_j drawn from $\{1, \cdots, n\}$, arranged in an arbitrary predetermined order that depends only on q and n. Of course, $\chi[a_1, \cdots, a_q]$ is simply the product $a_1 \cdots a_q$ if $n = 1$. Finally, we use $K(q, s)$ (q, s positive; q an integer) to denote the set of all functions k from $[0, \infty)^q$ to the $1 \times n^q$ matrices such that

$$k_j(\tau_1, \cdots, \tau_q) = \sum_{r=1}^{R_j} \prod_{i=1}^{q} \phi_{jir}(\tau_i), \tag{9}$$

for all $\tau_1, \cdots, \tau_q$ and all $j \in \{1, \cdots, n^q\}$, where $R_j < \infty$ and the ϕ_{jir} are real valued and continuous on $[0, s]$ and vanish on (s, ∞). [Notice that $K(q, s)$ is simply a set of row-matrix-valued functions whose elements have a certain nice finite sum of products representation.]

Theorem 2: Let $G \in \mathscr{A}(S)$, with $S = \{u \in L_\infty(n): \|u\| \leq \beta\}$, where β is a positive constant. Assume that H is defined by

$$(Hu)(t) = \int_0^t h(t - \tau)u(\tau)d\tau, \qquad t \geq 0$$

for $u \in S$, where h is a real $p \times n$ matrix-valued function on Ω such that each h_{ij} is (Lebesgue) integrable on Ω. Take S_1 to be $\{y \in L_\infty(p): \|y\|$

---

* For the sake of ease of exposition, (i)–(v) are assumed to be satisfied. However, only (iii) and (iii) with S replaced with S_1 are used.

$\leqslant \beta_1\}$, *in which* β_1 *is any number that satisfies* $\sup_t |(Hu)(t)| \leqslant \beta_1$ *for* $u \in S$.* *Assume also that F is causal and time invariant on* S_1, *and that F satisfies the continuity condition that given a continuous y in* S_1, *and numbers* $t \in (0, \infty)$ *and* $\delta_1 > 0$, *there is a* $\delta_2 > 0$ *such that*

$$|(Fy)(t) - (Fz)(t)| < \delta_1$$

whenever $z \in S_1$, *z is continuous, and* $\max_{\tau \in [0,t]} |y(\tau) - z(\tau)| < \delta_2$. *Under these conditions, given any* $\epsilon > 0$, *there is a positive integer Q, an* $s \in (0, \infty)$, *and elements* k_q *of* $K(q, s)$ $(1 \leqslant q \leqslant Q)$ *such that*

$$|(Gu)(t) - (Vu)(t)| < \epsilon, \qquad t \geqslant 0 \tag{10}$$

for all $u \in S$, *where*

$$(Vu)(t) = \sum_{q=1}^{Q} \int_0^t \cdots \int_0^t k_q(\tau_1, \cdots, \tau_q) \chi[u(t - \tau_1), \cdots, u(t - \tau_q)] d\tau_1 \cdots d\tau_q. \tag{11}$$

2.4.1 Proof of Theorem 2

Proof: We use Theorem 1. Conditions (i)–(v) are met with F and H causal and time invariant, and with $G \in \mathscr{A}(S)$.

Consider Condition (vi). Let ω be any positive number. For $x \in H(S)_\omega$,

$$x(t) = \int_0^t h(t - \tau) u(\tau) d\tau, \qquad t \in [0, \infty)$$

for some $u \in S$. In particular,

$$\sup_{t \in [0,\omega]} |x(t)| \leqslant \beta_1, \tag{12}$$

and there is a function λ from $(-\infty, \infty)$ to $[0, \infty)$, which depends only on h, such that $\lambda(\alpha) \to 0$ as $\alpha \to 0$, and

$$|x(t + \alpha) - x(t)| \leqslant \beta\lambda(\alpha) \tag{13}$$

for t and $(t + \alpha)$ in $[0, \omega]$. The existence of such a λ follows directly from the result (see Ref. 19, p. 14) that

$$\int_{-\infty}^{\infty} |r(t + \alpha) - r(t)| dt \to 0 \quad \text{as} \quad \alpha \to 0$$

when r is integrable on $(-\infty, \infty)$. Now let $\{X_\omega, \rho_\omega\}$ be the metric space of all functions x from $[0, \omega]$ to the real p-vectors such that (12) as well as (13) are satisfied, and

* Our conditions here on S and S_1 are clearly consistent with (i)–(v) of Section 2.1.

$$\rho_\omega(x_1, x_2) = \max_{t\in[0,\omega]} |x_1(t) - x_2(t)|.$$

This space is easily seen to be closed. For $p = 1$ its elements are uniformly bounded and equicontinuous. Thus, the space is compact for $p = 1$. Using this fact and, for example, the proposition that compactness in a metric space is equivalent to sequential compactness, it follows that the space is in fact compact for any positive integer p. Clearly, $H(S)_\omega \subset X_\omega \subset (S_1)_\omega$ which shows that (vi) holds.

By the continuity condition on F in Theorem 2, (vii) is satisfied.

Now let $\{F_\alpha : \alpha \in \Lambda\}$ be the set of all maps M defined on S_1 having the representation

$$(My)(t) = \int_0^t m(t - \tau)y(\tau)d\tau, \qquad t \geqslant 0 \tag{14}$$

where m is a real $(1 \times p)$ matrix-valued function on $[0, \infty)$ such that for any component m_j of m there is a real $\sigma > 0$ for which m_j is continuous on $[0, \sigma]$ and vanishes on (σ, ∞). Let $\omega > 0$, and observe that the $F_{\alpha\omega}$ are continuous on X_ω with respect to ρ_ω.

To see that they separate points of X_ω, let x_1 and x_2 be distinct points of X_ω. Let $i \in \{1, \cdots, p\}$ be such that $x_{1i}(t) - x_{2i}(t) \neq 0$ on some subinterval of $[0, \omega]$. Let $\alpha \in \Lambda$ be such that $(F_\omega y)(t)$ is given by the right side of (14) with $m_i(t) = [x_{1i}(\omega - t) - x_{2i}(\omega - t)]$ for $t \in [0, \omega]$ and $m_i(t) = 0$ otherwise, and with m_j vanishing on $[0, \infty)$ for $j \neq i$. Then

$$F_{\alpha\omega}x_1 - F_{\alpha\omega}x_2 = \int_0^\omega [x_{1i}(\tau) - x_{2i}(\tau)]^2 d\tau > 0.$$

Thus $\{F_\alpha : \alpha \in \Lambda\}$ is a fundamental set in the sense of Theorem 1, and by Theorem 1 given $\epsilon > 0$ there are Δ, k, $F_{\alpha_1}, \cdots, F_{\alpha_k}$, p, and P as described there such that (10) holds with $(Vu)(t) = (PH)(u_{\max[0,t-\Delta]})(t)$.

For $t \geqslant 0$, we have

$$(PH)(u_{\max[0,t-\Delta]})(t) = p[(F_{\alpha_1}H)u_{\max[0,t-\Delta]}(t), \cdots, (F_{\alpha_k}H)u_{\max[0,t-\Delta]}(t)]. \tag{15}$$

It is not difficult to verify that for any $j \in \{1, \cdots, k\}$ the operator $(F_{\alpha_j}H)$ is equivalent to a convolution C whose $1 \times n$ matrix-valued kernel c has elements that are continuous and integrable on Ω.

Also, one finds that

$$\int_0^t c(t - \tau)u_{\max[0,t-\Delta]}(\tau)d\tau = \int_0^t b(t - \tau)u(\tau)d\tau, \qquad t \geqslant 0,$$

where $b(t) = c(t)$ for $t \in [0, \Delta]$ and $b(t)$ equals the $1 \times n$ zero matrix otherwise. This together with (15), and just the observation that products of integrals can be written as iterated integrals, shows that V is as described in Theorem 2.

2.5 Comments

Cases in which the conditions of Theorem 2 are met arise often in applications. This is illustrated in Appendix B where the theorem is used to show that an important large class of input-output maps have finite Volterra series approximations.

In the proof of Theorem 2, the F_α are taken to be linear operators. It is clear that related additional approximation theorems can be obtained by allowing the F_α to be nonlinear.

Equation (15) in the proof of Theorem 2 shows that G in the theorem can be approximated arbitrarily well by a linear dynamic subsystem followed by a memoryless nonlinear subsystem with "polynomial nonlinearities." The existence of approximate system representations involving linear subsystems with an additional (and constant) input and only nonlinearities that take *absolute values* can be proved using Theorem 3 of Ref. 18.

The proof of Theorem 2 can easily be modified to establish a corresponding result for the discrete-time case in which Ω is replaced with $\Omega_d \triangleq \{0, 1, \cdots\}$. In fact, for that case the proof simplifies in an important conceptual way because then for any positive integer ω, $(S_1)_\omega$ is compact with respect to the usual discrete-time analog of ρ_ω in the proof of the theorem. In particular, in the discrete-time case we can set $n = p$, set $S = S_1$, and take H to be the identity map from S onto itself. This leads to the following theorem in which $\mathscr{L}_\infty(n)$ is $L_\infty(n)$ with Ω replaced with Ω_d, and $k(q)$ stands for the collection of all functions k from Ω_d^q to the $1 \times n^q$ matrices such that (11) holds for all $\tau_1, \cdots, \tau_q$ and all j where $R_j < \infty$ and the $\phi_{jir}(\tau_i)$ are real and are nonzero for at most a finite number of values of τ_i.

Theorem 3: Let $U = \{u \in \mathscr{L}_\infty(n) : \|u\| \leq \beta\}$ in which β is a positive number, and let K be a map from U to the real-valued functions defined on Ω_d such that K is causal, time invariant and an element of $\mathscr{A}(U)$ in the sense of Section 2.1 with Ω replaced with Ω_d. Let K satisfy the continuity condition that given $y \in U$ and numbers $t \in \{1, 2, \cdots\}$ and $\delta_1 > 0$, there is a $\delta_2 > 0$ such that $|(Ky)(t) - (Kz)(t)| < \delta_1$ whenever $z \in U$ and $\max_{\tau \in \{0, \cdots, t\}} |y(\tau) - z(\tau)| < \delta_2$. Then, given any $\epsilon > 0$, there is a positive integer Q, and elements k_q of $k(q)$ $(1 \leq q \leq Q)$ such that

$$|(Ku)(t) - (Vu)(t)| < \epsilon, \qquad t \in \Omega_d$$

for all $u \in U$, where

$$(Vu)(t) = \sum_{q=1}^{Q} \sum_{\tau_1=0}^{t} \cdots \sum_{\tau_q=0}^{t} k_q(\tau_1, \cdots, \tau_q)\chi[u(t - \tau_1), \cdots, u(t - \tau_q)].$$

REFERENCES

1. I. W. Sandberg, "Expansions for Nonlinear Systems," B.S.T.J., *61*, No. 2 (February 1982), pp. 159–99.
2. I. W. Sandberg, "On Volterra Expansions for Time-varying Nonlinear Systems," IEEE Trans. Circuits Syst., *CAS-30* (February 1983), pp. 61–7.
3. I. W. Sandberg, "The Mathematical Foundations of Associated Expansions for Mildly Nonlinear Systems," IEEE Trans. Circuits Syst., *CAS-30* (July 1983), pp. 441–55.
4. M. Fliess, M. Lamnabhi, and F. Lamnabhi-Lagarrique, "An Algebraic Approach to Nonlinear Functional Expansions," IEEE Trans. Circuits Syst. *CAS-30* (August 1983), pp. 554–70.
5. R. J. P. de Figueiredo, "A Generalized Fock Space Framework for Nonlinear System and Signal Analysis," IEEE Trans. Circuits Syst., *CAS-30* (September 1983), pp. 637–47.
6. S. Boyd, "Volterra Series: Engineering Fundamentals," dissertation, University of California, Berkeley, March 1985.
7. I. W. Sandberg, "Criteria for the Global Existence of Functional Expansions for Input/Output Maps," AT&T Tech. J., *64*, No. 7 (September 1985), pp. 1639–58.
8. V. Volterra, *Theory of Functionals and of Integral and Integro-Differential Equations*, New York: Dover, 1959.
9. M. B. Brilliant, "Theory of the Analysis of Nonlinear Systems," MIT RLE Report, No. 345, 1958.
10. D. George, "Continuous Nonlinear Systems," MIT RLE Report, No. 355, 1959.
11. W. Porter and T. Clark, "Causality Structure and the Weierstrass Theorem," Journal of Math. Analy. Applications, *52* (November 1975), pp. 351–63.
12. W. Root, "On the Modeling of Systems for Identification, Part I: ϵ-Representations of Classes of Systems," SIAM J. Control, *13* (July 1975) pp. 927–44.
13. P. Gallman and K. S. Narendra, "Representations of Nonlinear Systems via the Stone-Weierstrass Theorem," Automatica, *12* (November 1976) pp. 619–22.
14. W. Porter, "Approximation by Bernstein Systems," Math. Syst. Theory, *11* (November 1978), pp. 259–74.
15. W. J. Rugh, *Nonlinear System Theory: The Volterra/Wiener Approach*, Baltimore: Johns Hopkins Univ. Press, 1981, pp. 34–7.
16. S. Boyd and L. O. Chua, "Fading Memory and the Problem of Approximating Nonlinear Operators with Volterra Series," Univ. of Calif. Memorandum No. UCB/ERL M84/96, November 29, 1984.
17. I. W. Sandberg, "Existence and Evaluation of Almost Periodic Steady-State Responses of Mildly Nonlinear Systems," IEEE Trans. Circuits Syst., *CAS-31* (August 1984), pp. 689–701.
18. M. H. Stone, "A Generalized Weierstrass Approximation Theorem," *Studies in Modern Analysis, Vol. 1*, R. C. Buck, ed. Englewood Cliffs: Prentice Hall, 1962.
19. N. Wiener, *The Fourier Integral and Certain of its Applications*, New York: Dover, 1933.
20. I. W. Sandberg, "Some Results on the Theory of Physical Systems Governed by Nonlinear Functional Equations," B.S.T.J., *44*, No. 5 (May–June, 1965), pp. 871–98.

APPENDIX A

Proof of the Proposition

Let any $\delta > 0$, $u \in S$, and $t \geqslant 0$ be given. Choose real $\delta_1 > 0$ and $\delta_2 > 0$ such that

$$2\delta_1 + 2c\delta_2 < \delta,$$

and let Δ_1 and Δ_2 be elements of Ω for which

$$|(Fy)(\tau) - Fy_{\max[0,\tau-\Delta_1]}(\tau)| < \delta_1, \qquad \tau \geqslant 0 \tag{16}$$

for $y \in H(S)$, and

$$|(Hv)(\tau) - Hv_{\max[0,\tau-\Delta_2]}(\tau)| < \delta_2, \qquad \tau \geqslant 0 \tag{17}$$

for $v \in S$. Choose $\Delta \in \Omega$ so that $\Delta > \Delta_1 + \Delta_2$, and consider

$$\phi \triangleq |(FHu)(t) - FHu_{\max[0,t-\Delta]}(t)|.$$

If $t \in [0, \Delta]$, $u = u_{\max[0,t-\Delta]}$ and $\phi = 0$. Now let $t > \Delta$. Clearly,

$$\phi = |(FHu)(t) - (FH\hat{u})(t)|,$$

where $\hat{u} = u_{(t-\Delta)}$. Using (16), we have

$$|(FHu)(t) - F(Hu)_{(t-\Delta_1)}(t)| < \delta_1,$$

and

$$|(FH\hat{u})(t) - F(H\hat{u})_{(t-\Delta_1)}(t)| < \delta_1,$$

and one finds that

$$\begin{aligned}\phi \leqslant |(FHu)(t) - F(Hu)_{(t-\Delta_1)}(t) + F(Hu)_{(t-\Delta_1)}(t) \\ - (FH\hat{u})(t) + F(H\hat{u})_{(t-\Delta_1)}(t) - F(H\hat{u})_{(t-\Delta_1)}(t)| \\ \leqslant 2\delta_1 + c \sup_{t\in[(t-\Delta_1),t]} |(Hu)(\tau) - (H\hat{u})(\tau)|.\end{aligned}$$

By (17), $|(Hu)(\tau) - Hu_{(\tau-\Delta_2)}(\tau)| < \delta_2$ and $|(H\hat{u})(\tau) - H\hat{u}_{(\tau-\Delta_2)}(\tau)| < \delta_2$ for $\tau > \Delta_2$. Note that $\tau \geqslant (t - \Delta_1)$ and $t > \Delta \Rightarrow \tau > \Delta_2$; and that for $\tau \geqslant (t - \Delta_1)$, $Hu_{(\tau-\Delta_2)}(\tau) = H\hat{u}_{(\tau-\Delta_2)}(\tau)$. Thus,

$$\sup_{\tau\in[(t-\Delta_1),t]} |(Hu)(\tau) - (H\hat{u})(\tau)| < 2\delta_2,$$

which shows that $\phi \leqslant 2\delta_1 + 2c\delta_2$. Since this implies that $\phi < \delta$, the proposition is proved.

APPENDIX B

An Example of an Application of Theorem 2

In this Appendix we consider systems governed by the model

$$y = Nx \tag{18}$$

$$x = Av + Cy \tag{19}$$

$$w = Dv + By, \tag{20}$$

in which v is the input, w is the output, A, B, C, and D are linear operators, N is nonlinear, and x and y can be viewed as the input and output, respectively, of the nonlinear portion of the system. Models of this kind have been used in Ref. 2 and in other papers. Here we suppose that v, w, x, and y belong to $L_\infty(n)$, $L_\infty(1)$, $L_\infty(p)$, and $L_\infty(p)$, respectively, that N is memoryless and defined by $(Nx)(t) = \eta[x(t)]$ where η is a map from $\mathbb{R}^p$ to $\mathbb{R}^p$ which takes the zero element of $\mathbb{R}^p$ into itself, that η satisfies a global Lipschitz condition $|\eta(x_a) - \eta(x_b)|$

$\leqslant \gamma | x_a - x_b |$ where $| \cdot |$ is as in the definition of the norm in $L_\infty(p)$, and that A, B, C, and D, respectively, are causal time-invariant bounded linear maps from $L_\infty(n)$ to $L_\infty(p)$, $L_\infty(p)$ to $L_\infty(1)$,, $L_\infty(p)$ to $L_\infty(p)$, and $L_\infty(n)$ to $L_\infty(1)$. In particular, we assume that C has the convolution representation

$$(Cy)(t) = \int_0^t c(t - \tau) y(\tau) d\tau, \qquad t \geqslant 0$$

for $y \in L_\infty(p)$, where $c(\cdot)$ is $p \times p$ and has integrable elements [that is, has elements that are integrable on $[0, \infty)$]. The equations of a very large class of systems with a single output can be put in this form with A, B, and D convolutions whose matrix-valued kernels are either integrable, or integrable with the exception of an impulse at the origin (see Ref. 2, Appendices I and II).

B.1 Further assumptions, and approximations

Assume in the remainder of this Appendix that $(I - CN)$ is an invertible map of $L_\infty(p)$ onto $L_\infty(p)$, where I is the identity operator on $L_\infty(p)$, and that $(I - CN)^{-1}$ is causal, time invariant, and globally Lipschitz. Conditions under which these assumptions are met can be obtained from standard existence theory and results in the area of stability theory [see, for example, Ref. 20, Theorem 3 and Corollary 3(*a*)]. It follows that $w = Dv + BN(I - CN)^{-1} Av$ for all $v \in L_\infty(n)$.

With r an arbitrary positive constant, let us now restrict our inputs v to the ball $\Lambda = \{v \in L_\infty(n): \|v\| \leqslant r\}$. Let $\Lambda_1 = \{u \in L_\infty(p): \|u\| \leqslant r\|A\|\}$. In addition to the assumptions introduced above, suppose that A is a convolution with an integrable kernel, and that $(I - CN)^{-1}$, which takes Λ_1 into $L_\infty(p)$, belongs to $\mathscr{A}(\Lambda_1)$ in the sense of Section 2.1. Using the proposition in Section 2.3, and by considering one component of $(I - CN)^{-1} A$ at a time, we see that $(I - CN)^{-1} A \in \mathscr{A}(\Lambda)$, since, as can easily be verified, $A \in \mathscr{A}(\Lambda)$. Similarly, $N(I - CN)^{-1} A: \Lambda \to L_\infty(p)$ and finally $BN(I - CN)^{-1} A$ both belong to $\mathscr{A}(\Lambda)$. Thus, by Theorem 2 [with $H = A$ and $F = BN(I - CN)^{-1}$], and in the sense of Theorem 2, $BN(I - CN)^{-1} A$ can be approximated arbitrarily well on Λ by a finite Volterra series.

Before proceeding to the important matter of how one might show that $(I - CN)^{-1} \in \mathscr{A}(\Lambda_1)$ under some reasonable conditions, suppose that the assumptions described above are met, with the exception that A is *not* a convolution. Assume instead that $(Av)(t) = av(t)$, where a is a $p \times n$ matrix of constants. (This case arises naturally in the study of feedback systems.) Using the identity $(I - CN)^{-1} = (I - CN)^{-1}CN + I$, one has $w = Dv + BNAv + BN(I - CN)^{-1}CNAv$. The term $BNAv$ has a simple representation as is; and if, for example, B is a convolution with an integrable kernel and $n = p = 1$, then, by the Weierstrass approximation theorem for real-valued continuous functions on a compact real interval, it is clear that it can be approximated arbitrarily well on Λ by a finite series having the form

$$\sum_{q=1}^{Q} \int_0^t b_q(t - \tau)v(\tau)^q d\tau, \qquad t \geqslant 0. \tag{21}$$

Consider now the more interesting term $BN(I - CN)^{-1}CNAv$. By Theorem 2 (this time with $H = C$) we see that it can be approximated arbitrarily well on Λ by a finite series of the form (11) with each $u(t - \tau_j)$ replaced with $\eta[av(t - \tau_j)]$. In particular, using the fact that the k_q in (11) satisfy

$$\int_{[0,\infty)^q} |k_{qj}(\tau_1, \cdots, \tau_q)| d(\tau_1, \cdots, \tau_q) < \infty$$

for each j, and the Weierstrass approximation theorem for real-valued continuous functions of several real variables, it follows that the term can be approximated arbitrarily well throughout Λ by a finite Volterra-like series in the sense of the sets of iterated integrals $K(m)$ in Ref. 2. [These Volterra-like series, which are frequently needed in *exact* expansion representations, can be viewed as Volterra series with symbolic kernels that include certain delta functions. A simple example of a Volterra-like series is (21).]

B.2 $(I - CN)^{-1}$ and the memory condition

The hypothesis that $(I - CN)^{-1} \in \mathscr{A}(\Lambda_1)$ plays a key role in the discussion above. We begin our comments concerning this hypothesis with the observation that with arbitrary $t \geqslant 0$, $\Delta \geqslant 0$, and $u \in \Lambda_1$, one has $[(I - CN)^{-1}u](t) - [(I - CN)^{-1}u_{\max[0,t-\Delta]}](t)$ equal to $x(t) - \tilde{x}(t)$, where x and $\tilde{x}$ are elements of $L_\infty(p)$ such that

$$x(\alpha) + \int_0^\alpha c(\alpha - \tau)\eta[x(\tau)]d\tau = u(\alpha) \tag{22}$$

$$\tilde{x}(\alpha) + \int_0^\alpha c(\alpha - \tau)\eta[\tilde{x}(\tau)]d\tau = u_{\max[0,t-\Delta]}(\alpha) \tag{23}$$

for $\alpha \geqslant 0$.

With σ any positive constant,

$$y(\alpha) + \int_0^\alpha \hat{c}(\alpha - \tau)\hat{\eta}[y(\tau), \tau]d\tau = z(\alpha) \tag{24}$$

$$\tilde{y}(\alpha) + \int_0^\alpha \hat{c}(\alpha - \tau)\hat{\eta}[\hat{y}(\tau), \tau]d\tau = \hat{z}(\alpha) \tag{25}$$

for $\alpha \in [0, \infty)$, where $y(\alpha) = x(\alpha)e^{\sigma\alpha}$, $\tilde{y}(\alpha) = \tilde{x}(\alpha)e^{\sigma\alpha}$, $\hat{c}(\alpha) = c(\alpha)e^{\sigma\alpha}$, $\hat{\eta}[y(\tau), \tau] = e^{\sigma\tau}\eta[e^{-\sigma\tau}y(\tau)]$, $z(\alpha) = u(\alpha)e^{\sigma\alpha}$, and $\hat{z}(\alpha) = e^{\sigma\alpha}u_{\max[0,t-\Delta]}(\alpha)$. Let Σ denote the set of positive σ such that the elements of $\hat{c}$ are square integrable, and suppose that Σ is not empty. Let us now make the key assumption, which we shall refer to as A.0, that from (24) and (25) it can be concluded that for some $\sigma \in \Sigma$ there is a constant λ which depends only on c, η, and σ such that

$$\|y - \tilde{y}\|_2 \leqslant \lambda \|z - \tilde{z}\|_2,$$

where

$$\|y - \tilde{y}\|_2^2 = \int_0^\infty [y(t) - \tilde{y}(t)]^{Tr}[y(t) - \tilde{y}(t)]dt,$$

"Tr" denotes the transpose, and similarly for $\|z - \tilde{z}\|_2$. Much is known about conditions under which A.0 is met [see Ref. 20, Corollary 1(a) and Theorem 6], and it is known that A.0 is met in certain specific important cases.

For $t \leq \Delta$, obviously $x(t) - \tilde{x}(t) = 0$. Now let $t > \Delta$.

Notice that $\|y - \tilde{y}\|_2 \leqslant \xi e^{\sigma(t-\Delta)}$ where $\xi = \lambda p^{1/2} r \|A\|$. Using (24) and (25), and the Schwarz inequality,

$$|y(t) - \tilde{y}(t)| \leqslant \max_i \sum_{j=1}^{p} \left(\int_0^t |\hat{c}(\tau)_{ij}|^2 d\tau\right)^{1/2} \cdot \left(\int_0^t |\hat{\eta}_j[\tilde{y}(\tau), \tau] - \hat{\eta}_j[y(\tau), \tau]|^2 d\tau\right)^{1/2}.$$

Since

$$\int_0^t |\hat{\eta}_j[\tilde{y}(\tau), \tau] - \hat{\eta}_j[y(\tau), \tau]|^2 d\tau \leq \int_0^t |\hat{\eta}[\tilde{y}(\tau), \tau] - \hat{\eta}[y(\tau), \tau]|^2 d\tau \leqslant \gamma^2 \int_0^t |\tilde{y}(\tau) - y(\tau)|^2 d\tau$$
$$\leq \gamma^2 \int_0^t [\tilde{y}(\tau) - y(\tau)]^{Tr}[\tilde{y}(\tau) - y(\tau)]d\tau \leqslant \gamma^2 \|\tilde{y} - y\|_2^2,$$

we find that

$$|y(t) - \tilde{y}(t)| \leqslant \xi_1 \gamma \xi e^{\sigma(t-\Delta)},$$

where

$$\xi_1 = \max_i \sum_{j=1}^{p} \left(\int_0^t |\hat{c}(\tau)_{ij}|^2 d\tau \right)^{1/2}.$$

Thus,

$$|x(t) - \tilde{x}(t)| \leq \xi_1 \gamma \xi e^{-\sigma\Delta}.$$

This shows that $(I - CN)^{-1} \in \mathscr{A}(\Lambda_1)$ under the conditions described, and therefore that the input-output maps of a very large class of systems have finite Volterra series approximations in the strong sense of this Appendix. [For example, using material in Ref. 20 (see the comment at the bottom of p. 875 there), it is not difficult to show that this class includes a large family of electrical networks consisting of sources, passive elements, and monotone nonlinear resistors.]

AUTHOR

Irwin W. Sandberg, B.E.E., 1955, M.E.E., 1956, and D.E.E., 1958, Polytechnic Institute of Brooklyn; AT&T Bell Laboratories, 1958—. Mr. Sandberg has been concerned with analysis of radar systems for military defense, synthesis and analysis of active and time-varying networks, with several fundamental studies of properties of nonlinear systems, and with some problems in communication theory and numerical analysis. His more recent interests have included compartmental models, the theory of digital filtering, global implicit-function theorems, and functional expansions for nonlinear systems. IEEE Centennial Medalist, Former Vice Chairman IEEE Group on Circuit Theory, and Former Guest Editor IEEE Transactions on Circuit Theory Special Issue on Active and Digital Networks. Fellow and member, IEEE; member, American Association for the Advancement of Science, Eta Kappa Nu, Sigma Xi, Tau Beta Pi, National Academy of Engineering.