THE BELL SYSTEM TECHNICAL JOURNAL
Vol. 61, No. 2, February 1982
Printed in U.S.A.

Expansions for Nonlinear Systems*

By I. W. SANDBERG

(Manuscript received April 22, 1981)

In this paper we study operator-type models of dynamic nonlinear physical systems, such as communication channels and control systems. Attention is focused on the problem of determining conditions under which there exists a power-series-like expansion, or a polynomial-type approximation, for a system's outputs in terms of its inputs. Related problems concerning properties of the expansions are also considered and nonlocal, as well as local, results are given. In particular, we show for the first time the existence of a locally convergent Volterra-series representation for the input-output relation of an important large class of nonlinear systems containing an arbitrary finite number of nonlinear elements.

I. INTRODUCTION

In this paper we study operator-type models of dynamic nonlinear physical systems, such as communication channels and control systems. Attention is focused on the problem of determining conditions under which there exists a power-series-like expansion, or a polynomial-type approximation, for a system's outputs in terms of its inputs. Related problems concerning properties of the expansions are also considered and nonlocal, as well as local, results are presented. In particular, we show for the first time the existence of a locally convergent Volterra-series representation for the input-output relation of an important large class of nonlinear systems containing an arbitrary finite number of nonlinear elements.

With regard to background material, functional power series of the form

* The material given in this paper was described in the writer's "Conference Course" at the 1981 European Conference on Circuit Theory and Design, The Hague, August 1981.

$$k_0 + \sum_{m=1}^{\infty} \int_a^b \cdots \int_a^b k_m(t, \tau_1, \cdots, \tau_m) u(\tau_1) \cdots u(\tau_m) d\tau_1 \cdots d\tau_m, \quad (0)$$

in which k_0 is a constant, t is a parameter, and u and the k_m for $m \geq 1$ are continuous functions, were considered in 1887 by Vito Volterra[1,2] in connection with his studies of functions of functions. (These studies provided much of the initial motivation to develop the field now known as functional analysis.) About twenty years later, Fréchet[3] proved that a continuous real functional (i.e., a continuous real scalar-valued map) defined on a compact set of real continuous functions on $[a, b]$ could be approximated by a sum of a finite number of terms in Volterra's series (0), but with (in analogy with the well-known Weierstrass approximation theorem) the number of terms, as well as the k_m, dependent on the degree of approximation.

It was Norbert Wiener[4] who first used a Volterra-series representation in the analysis of a nonlinear system.* The form of Volterra's expansion provided also the basis for Wiener's later work (see, for example, Refs. 5 and 6) on nonlinear analysis and synthesis. His studies, which were concerned mainly with the modeling of systems when only input-output data (rather than the system's equations) are available, stimulated considerable interest concerning Volterra and other† functional expansions for nonlinear systems. It was appreciated from the outset that such expansions, when they exist, could provide important insight of a qualitative nature concerning the input-output behavior of a system, and that they could be useful in connection with, for example, the estimation and/or equalization of distortion caused by nonlinearities.

There is a fairly large literature related directly or indirectly to the material of the present paper (see, for example, Refs. 1 through 29 and the references cited there). In most cases the functional expressions considered are Volterra series (or truncated Volterra series). With regard to systems for which the governing equations are known, with relatively few exceptions, questions concerning the existence of an expansion, its convergence, and/or the nature of the approximation provided by a truncated series are either not addressed or are left unanswered. (See, for instance, the remarks in Ref. 6, p. 137, on the lack of understanding concerning convergence.)

On the other hand, some material regarding the range of validity and specific properties of functional expansions has appeared, both for systems governed by ordinary differential equations defined on a finite time interval $[0, \tau]$ (Refs. 15, 16, 20, and 23 are representative refer-

* In Ref. 4 Wiener considers the problem of evaluating the output moments of a specific type of detector circuit driven by a random input.

† See, for instance, Ref. 7.

ences), and for polynomic systems, which are modeled by operator generalizations of ordinary functions of polynomial form.[26–29] The work on differential equations is concerned mainly with the particular case of bilinear systems and with "linear-analytic" systems.[16,23] The main result obtained asserts that under certain conditions (see, for example, Ref. 23) there does exist a locally convergent Volterra series for the solution (but with the size of the region of convergence dependent on τ).

The studies of polynomic systems, which are operator theoretic in nature and which draw on the theory of multilinear forms, are more closely related to the results reported in this paper. The most pertinent earlier proposition[29] is one to the effect that if a certain contraction mapping condition is met, then it is possible to construct in a particular way a local inverse of a certain generalized power series. While we do not use previous results in the polynomic systems area, there are some points of contact with the earlier material, and this is discussed at appropriate places in Section II.

We now briefly outline the remainder of the paper. Section II begins with some mathematical preliminaries. In Section 2.1, we introduce the general setting of concern throughout the rest of Section II. This involves two maps f and g related by $f[g(u)] = u$ for u in a certain set that can be thought of as a set of system inputs such that each u contained produces an output $g(u)$. The remaining portion of Section II describes, proves, and discusses results concerning expansions and approximations of $g(u)$. The material in Section II is somewhat abstract. Examples which illustrate how the material can be used to obtain more specific results of general interest are given in Section III.

II. APPROXIMATIONS AND EXPANSIONS

Throughout the paper, $\mathscr{B}$ and $\mathscr{B}_0$ denote two Banach spaces, each with real or complex scalars, and X denotes a nonempty open subset of $\mathscr{B}_0$. We use the symbol $\|\cdot\|$ for the norm associated with $\mathscr{B}$, as well as for the norm associated with $\mathscr{B}_0$, and θ is used to denote the zero element of $\mathscr{B}$ and of $\mathscr{B}_0$.

It will become clear shortly that a central role in our development is played by first and higher order Fréchet derivatives (see, for instance, Ref. 30). Before proceeding, we recall a few pertinent facts and definitions.

Let F map X into $\mathscr{B}$, and let x_0 be a point in X. If there is a bounded linear map L_{x_0} from $\mathscr{B}_0$ to $\mathscr{B}$ such that $\|F(x_0 + h) - F(x_0) - L_{x_0}h\| = o(\|h\|)$ as $\|h\| \to 0$, then F is said to be *Fréchet differentiable* at x_0 with Fréchet derivative L_{x_0}, which we denote by $dF(x_0)$. If F is Fréchet differentiable at every point in X, then we say that F is differentiable on X. Similarly, if F is Fréchet differentiable on X and $dF(\cdot)$ is continuous, then F is said to be *continuously differentiable* on X. Higher order Fréchet derivatives of F are defined in the usual inductive manner.

Note that the second-order Fréchet derivative $d^2F(x_0)$, when it exists, is a bounded linear map from $\mathscr{B}_0$ into the space $L(\mathscr{B}_0, \mathscr{B})$ of bounded linear maps from $\mathscr{B}_0$ into $\mathscr{B}$. For h_1 and h_2 in $\mathscr{B}_0$, $d^2F(x_0)h_1h_2$, by which we mean $[d^2F(x_0)h_1]h_2$, is an element of $\mathscr{B}$, and, therefore, $d^2F(x_0)$ can be regarded as a bilinear map from $\mathscr{B}_0 \times \mathscr{B}_0$ into $\mathscr{B}$, i.e., it can be *identified* in the obvious way with such a map. This bilinear map satisfies the symmetry condition that $[d^2F(x_0)h_1]h_2 = [d^2F(x_0)h_2]h_1$. In general,[30] the mth order Fréchet derivative $d^mF(x_0)$ for $m > 1$ is a bounded linear map from $\mathscr{B}_0$ into a Banach space of bounded linear maps with $\| d^mF(x_0) \|$ defined in the usual way in terms of induced norms, and $d^mF(x_0)$ can be regarded alternatively as a symmetric m-linear map from $\mathscr{B}_0^m$ into $\mathscr{B}$. Moreover, for $1 \leq l < m$ and $h_1, h_2, \cdots, h_l$ elements of $\mathscr{B}_0$, $d^mF(x_0)h_1h_2 \cdots h_l$ is a bounded linear map from $\mathscr{B}_0$ into a Banach space of bounded linear operators.

A result that we shall use is the following essentially standard inverse function proposition (Ref. 30, p. 273).*

Lemma 1: Let $F{:}X \to \mathscr{B}$ be continuously Fréchet differentiable on X, and let $x_0 \in X$. If $dF(x_0)$ is an invertible map of $\mathscr{B}_0$ onto $\mathscr{B}$, there is an open neighborhood $V \subset X$ of x_0 such that F restricted to V is a homeomorphism of V onto an open neighborhood of $F(x_0)$ in $\mathscr{B}$. In addition, if F is r times continuously differentiable on V, the inverse mapping G of $F(V)$ onto V is r times continuously differentiable on $F(V)$.

Also of importance in our work are derivatives of Banach-space-valued maps defined on an open subset S of the real or complex numbers. If F maps S into $\mathscr{B}$ and $s_0 \in S$, then F is said to have a *derivative $dF(s_0)/ds$* at s_0 if $dF(s_0)/ds$ is an element of $\mathscr{B}$ and we have

$$\lim_{|\gamma| \to 0} \| \gamma^{-1}[F(s_0 + \gamma) - F(s_0)] - dF(s_0)/ds \| = 0.$$

Again, higher order derivatives are defined in the usual inductive way. Here derivatives are elements of $\mathscr{B}$.

2.1 *f* and *g*

Throughout the paper, $f{:}X \to \mathscr{B}$ denotes a map with the property that there is a nonempty open convex subset U of $\mathscr{B}$ such that for each $u \in U$, there is in X a unique x_u such that $f(x_u) = u$. (Recall that X is a nonempty open subset of $\mathscr{B}_0$).

---

* With regard to a difference in a hypothesis of Lemma 1 and the cited proposition in Ref. 30, we note that if A is a bounded linear invertible map of $\mathscr{B}_0$ onto $\mathscr{B}$ with inverse A^{-1}, then, by a result of Banach, this inverse is a *bounded* linear map of $\mathscr{B}$ onto $\mathscr{B}_0$.

We shall be concerned primarily with the map $g: U \to X$ defined by $f[g(u)] = u$ for every $u \in U$.

2.2 A representation theorem for g(•)

We shall refer to the following two hypotheses:

H.1: For some positive integer p, the mth order F-derivative (i.e., Fréchet derivative) $d^m f$ exists and is continuous on X for $m = 1, 2, \cdots, (p+1)$.

H.2: $[df(x)]^{-1}$ exists [i.e., $df(x)$ is an invertible map of $\mathscr{B}_0$ onto $\mathscr{B}$] for $x \in X$.

Theorem 1: Suppose that H.1 and H.2 are met, and let u and u_0 be points in U. For $m = 1, 2, \cdots, p+1$ and each $\beta \in [0, 1]$, let $g_m(u_0, u-u_0, \beta)$ be defined as follows:

1.(a) $g_1(u_0, u-u_0, \beta) = df\{g[u_0 + \beta(u-u_0)]\}^{-1}(u-u_0)$

1.(b) $g_m(u_0, u-u_0, \beta) =$

$$-df\{g[u_0+\beta(u-u_0)]\}^{-1} \sum_{l=2}^{m} (l!)^{-1} \sum_{\substack{k_1+k_2+\cdots+k_l=m \\ k_j>0}}$$

$$\cdot\, d^l f\{g[u_0+\beta(u-u_0)]\} g_{k_1}(u_0, u-u_0, \beta) g_{k_2}(u_0, u-u_0, \beta) \cdots g_{k_l}(u_0, u-u_0, \beta)^* \quad (1)$$

for $m = 2, \cdots, p+1$.

Then $g_{p+1}(u_0, u-u_0, \beta)$ depends continuously on β for $\beta \in [0, 1]$, and we have

$$g(u) = g(u_0) + g_1(u_0, u-u_0) + \cdots + g_p(u_0, u-u_0) + (p+1)\int_0^1 (1-\beta)^p g_{p+1}(u_0, u-u_0, \beta)\, d\beta,$$

in which $g_m(u_0, u-u_0) = g_m(u_0, u-u_0, 0)$ for $m = 1, \cdots, p$.

Moreover,

2.(a) *there are positive constants ρ and σ, which do not depend on u, such that*

$$\left\| g(u) - g(u_0) - \sum_{m=1}^{p} g_m(u_0, u-u_0) \right\| \le \rho \|u-u_0\|^{p+1} \quad \text{for} \quad \|u-u_0\| \le \sigma,$$

2.(b) *there are positive constants $\rho_1, \rho_2, \cdots, \rho_p$, which do not depend on u, such that $\|g_m(u_0, u-u_0)\| \le \rho_m \|u-u_0\|^m$ for $m = 1, 2, \cdots, p$, and*

---

* In (1), $\sum_{\substack{k_1+k_2+\cdots+k_l=m \\ k_j>0}}$ denotes a sum over all positive $k_1, \cdots, k_l$ that add to m.

2.(c) *for* $m = 1, 2, \cdots, p$ *we have* $g_m(u_0, u_a - u_0) = r^m g_m(u_0, u_b - u_0)$ *for* u_a *and* u_b *in* U *such that* $(u_a - u_0) = r(u_b - u_0)$ *for some real number* r.

Proof: By H.1, H.2, and Lemma 1, the mth order F-derivative $d^m g(w)$ exists and is continuous for $m = 1, 2, \cdots, (p+1)$ and $w \in U$.

By a version of Taylor's theorem for Fréchet differentiable maps (Ref. 30, pp. 190–1),

$$g(u) = g(u_0) + dg(u_0)(u - u_0) + \frac{1}{2!} d^2 g(u_0)(u - u_0)^2 + \cdots + \frac{1}{p!} d^p g(u_0)(u - u_0)^p + \int_0^1 \frac{(1-\beta)^p}{p!} d^{(p+1)} g[u_0 + \beta(u - u_0)](u - u_0)^{p+1} d\beta, \quad (2)$$

where $(u - u_0)^m$, for $m = 1, 2, \cdots, p+1$, denotes $[(u - u_0), (u - u_0), \cdots, (u - u_0)]$ with m terms.

Since $d^m g(w)$ exists for $m = 1, 2, \cdots, p+1$ for w belonging to the convex set U, $d^m g[u_0 + \beta(u - u_0)]/d\beta^m$ exists and is equal to $d^m g[u_0 + \beta(u - u_0)](u - u_0)^m$ for $\beta \in [0, 1]$ and $m = 1, 2, \cdots, p+1$ [31, p. 198].*

Let $q(\beta)$ denote $g[u_0 + \beta(u - u_0)]$ for all real β such that $u_0 + \beta(u - u_0) \in U$, and let $q^{(m)}(\beta)$ stand for the mth derivative of $q(\beta)$ with respect to β at the arbitrary point $\beta \in [0, 1]$.

We have $f[q(\beta)] = u_0 + \beta(u - u_0)$ when $u_0 + \beta(u - u_0) \in U$. By a version of the chain rule (Ref. 31, p. 173) for the derivative of a composition function,

$$df[q(\beta)]q^{(1)}(\beta) = u - u_0, \qquad \beta \in [0, 1] \quad (3)$$

since $df[q(\beta)]$ and $q^{(1)}(\beta)$ exist for $\beta \in [0, 1]$. Thus,

$$dg[u_0 + \beta(u - u_0)](u - u_0) = q^{(1)}(\beta) = df[g(u_0 + \beta(u - u_0))]^{-1}(u - u_0), \qquad \beta \in [0, 1] \quad (4)$$

Now let $2 \le m \le (p+1)$, and let $f_0: \mathscr{B}_0 \to \mathscr{B}$ and $q_0: (-\infty, \infty) \to \mathscr{B}_0$ be defined by

$$f_0(y) = \sum_{l=1}^{m} (l!)^{-1} d^l f[q(\beta)][y - q(\beta)]^l \quad (5)$$

---

* More specifically, since with $h = (u - u_0)$, $g[u_0 + (\beta + \sigma)h] = g(u_0 + \beta h) + dg(u_0 + \beta h)\sigma h + o(\|\sigma h\|)$ as $\sigma \to 0$ (for $\beta \in [0, 1]$), it is clear that $dg(u_0 + \beta h)/d\beta$ exists and is equal to $dg(u_0 + \beta h)h$ for $\beta \in [0, 1]$. Similarly, using $d^{m-1}g[u_0 + (\beta + \sigma)h] = d^{m-1}g(u_0 + \beta h) + d^m g(u_0 + \beta h)\sigma h + o(\|\sigma h\|)$ as $\sigma \to 0$ for $2 \le m \le (p+1)$, we see that $d^m g(u_0 + \beta h)/d\beta^m = d^m g(u_0 + \beta h)h^m$ for $1 \le m \le (p+1)$.

$$q_0(r) = q(\beta) + \sum_{l=1}^{m} (l!)^{-1}(r-\beta)^l q^{(l)}(\beta) \tag{6}$$

for $y \in \mathscr{B}_0$ and $r \in (-\infty, \infty)$, where $\beta \in [0, 1]$. We will use the following generalization of the classical rule for differentiating a product of two differentiable functions.

Proposition 1: With S a Banach space and A an open interval in $(-\infty, \infty)$ containing a point r_0, let $L(r)$ denote a bounded linear map from $\mathscr{B}_0$ into S for each $r \in A$, and let $e(\cdot)$ be a map from A to $\mathscr{B}_0$. If $dL(r_0)/dr$ and $de(r_0)/dr$ exist, then $d[L(r)e(r)]/dr$ exists at $r = r_0$, and $d[L(r)e(r)]/dr = L(r)de(r)/dr + [dL(r)/dr]e(r)$ at $r = r_0$.

A proof of Proposition 1 is given in Appendix A. Using Proposition 1 and the observation that $d^m f[q(\beta)]/d\beta^m = \theta$ for $\beta \in [0, 1]$ when $2 \leq m \leq (p + 1)$, we show in Appendix B that $d^m\{f_0[q_0(r)]\}/dr^m|_{r=\beta} = \theta$ for $\beta \in [0, 1]$.

Since we have

$$f_0[q_0(r)] = \sum_{l=1}^{m} (l!)^{-1} d^l f[q(\beta)] \left(\sum_{k=1}^{m} (k!)^{-1}(r-\beta)^k q^{(k)}(\beta) \right)^l$$

for every r, and each $d^l f[q(\beta)]$ can be regarded as a l-linear operator on $\mathscr{B}_0^l$, it follows that for $\beta \in [0, 1]$:

$$\begin{aligned} \theta &= d^m f_0[q_0(r)]/dr^m|_{r=\beta} \\ &= m! \sum_{l=1}^{m} (l!)^{-1} \sum_{\substack{k_1+k_2+\cdots+k_l=m \\ k_j>0}} (k_1!k_2! \cdots k_l!)^{-1} \\ &\qquad \cdot d^l f[q(\beta)] q^{(k_1)}(\beta) q^{(k_2)}(\beta) \cdots q^{(k_l)}(\beta). \end{aligned} \tag{7}$$

Referring now to the $g_m(u_0, u - u_0, \beta)$ in the statement of the theorem, notice that $g_1(u_0, u - u_0, \beta) = q^{(1)}(\beta)$ for $\beta \in [0, 1]$, and that, by (7), $g_m(u_0, u - u_0, \beta) = (m!)^{-1} q^{(m)}(\beta)$ for $2 \leq m \leq (p + 1)$ and $\beta \in [0, 1]$. Therefore, using (2) and $q^{(m)}(\beta) = d^m g[u_0 + \beta (u - u_0)] \cdot (u - u_0)^m$ for $\beta \in [0, 1]$ and each m, we obtain the formula for $g(u)$ given in the theorem, in which $g_{p+1}(u_0, u - u_0, \beta)$, which is equal to

$$[(p+1)!]^{-1} d^{(p+1)} g[u_0 + \beta(u - u_0)](u - u_0)^{(p+1)},$$

depends continuously on β for $0 \leq \beta \leq 1$.

Since

$$g_m(u_0, u - u_0) = (m!)^{-1} d^m g(u_0)(u - u_0)^m$$

for $m = 1, 2, \cdots, p$ in which $d^m g(u_0)$ is an mth order Fréchet derivative,* it follows at once that properties 2.(b) and 2.(c) hold.

* In particular, $d^m g(u_0)$ and $d^m g(u_0) v_1 \cdots v_{m-l}$ for $1 \leq l \leq m - 1$ and $v_j \in \mathscr{B}$ for $1 \leq j \leq m - l$ are bounded linear maps on $\mathscr{B}$.

Finally, and referring to (2), since

$$\|d^{(p+1)}g(u_0)(u-u_0)^{(p+1)}\| \le \|d^{(p+1)}g(u_0)\|\cdot\|u-u_0\|^{(p+1)},$$

property 2.(a) is a consequence of the result (Ref. 30, pp. 190–1) that for every $\sigma_0 > 0$ there is a $\sigma > 0$ such that

$$\Bigg\| g(u) - dg(u_0)(u-u_0) - \frac{1}{2!}\,d^2g(u_0)(u-u_0)^2 - \cdots - \frac{1}{(p+1)!}\,d^{(p+1)}g(u_0)(u-u_0)^{(p+1)}\Bigg\| \le \sigma_0\|u-u_0\|^{(p+1)}$$

for $\|u-u_0\| \le \sigma$.

2.3 Corollary to Theorem 1

Corollary 1: Suppose that H.1 is met, and that $u_0 \in U$ is such that $df[g(u_0)]$ is invertible (i.e., is an invertible map of $\mathscr{B}_0$ onto $\mathscr{B}$). Then there are positive constants ρ and σ such that for $u \in U$ with $\|u-u_0\| \le \sigma$,

$$\|g(u) - g(u_0) - g_1(u_0, u-u_0) - \cdots - g_p(u_0, u-u_0)\| \le \rho\|u-u_0\|^{(p+1)},$$

in which $g_1(u_0, u-u_0), \cdots, g_p(u_0, u-u_0)$ are defined in Theorem 1. *In addition* 2.(b) *and* 2.(c) *of Theorem* 1 *hold.*

Proof: Since $df(\cdot)$ is continuous on X, and, by a theorem of Banach,[32] $df[g(u_0)]^{-1}$ is bounded, by a standard type of argument (see, for example, Ref. 30, pp. 154–5) $df(\cdot)^{-1}$ exists and is continuous* in some open neighborhood Γ of $g(u_0)$ in X. Also, since $df[g(u_0)]^{-1}$ exists, Γ contains an open neighborhood $N_{g(u_0)}$ and U contains an open neighborhood N_{u_0} of u_0 such that f restricted to $N_{g(u_0)}$ is a homeomorphism of $N_{g(u_0)}$ onto N_{u_0} (see Lemma 1). Let Ξ be a nonempty open convex subset of N_{u_0}. At this point the corollary follows from Theorem 1 with $X = N_{g(u_0)}$ and $U = \Xi$.

2.4 Comments

For a fixed u_0, the expansion given in Theorem 1 has the properties that the homogeneity condition 2.(c) is met and the remainder, the integral, is bounded above by $\rho\|u-u_0\|^{(p+1)}$ for $\|u-u_0\| \le \sigma$ for some positive constants ρ and σ. A proof given in Ref. 33, p. 174 can easily be modified to show that the expansion is *unique* in the sense that there is no other similar [i.e., $g(u_0)$ plus p terms plus remainder] expansion of $g(u)$ valid for all $u \in U$ with these homogeneity and

---

* The continuity is not used in the present proof. It is used in Section 2.7 and in Appendix C, where reference is made to this proof.

remainder properties. Of course, a corresponding uniqueness proposition holds in the case of the truncated expansion in Corollary 1.

In some cases, $d^l f[g(u_0)]$ of Theorem 1 is the zero map whenever l is even and $2 \le l \le p$. Then $g_m(u_0, u - u_0) = \theta$ for m even with $2 \le m \le p$. This follows from a simple inductive argument using 1.(b) and the observation that $k_1 + k_2 + \cdots + k_l$ is an odd number if l is odd and each k_j is positive and odd.

Referring to Corollary 1, we can establish the existence of an expansion in a more general setting. Specifically, let $\mathcal{B}_1$ be a third Banach space and let $h(\cdot, \cdot)$ be a $(p + 1)$-times Fréchet continuously differentiable map of $S_0 \times S$ into $\mathcal{B}_1$, where S_0 and S are nonempty open subsets of $\mathcal{B}_0$ and $\mathcal{B}$, respectively. Let x_0 and u_0 be elements of S_0 and S, respectively, such that $h(x_0, u_0) = \theta$, in which here θ is used to denote the zero element of $\mathcal{B}_1$. Finally, assume that $D_1 h(x_0, u_0)$ the Fréchet partial derivative of $h(x, u)$ with respect to x, at the point (x_0, u_0), is an invertible map of $\mathcal{B}_0$ onto $\mathcal{B}_1$.

By the implicit function theorem in Ref. 30, p. 270 and a related proposition in Ref. 30, Result (10.2.3), it follows that there is an open convex neighborhood N of u_0 in S, and a $(p + 1)$-times Fréchet continuously differentiable map w of N into S_0 such that $w(u_0) = x_0$ and $h[w(u), u] = \theta$ for $u \in N$.* Therefore (2), with g replaced with w, is a representation about u_0 of w valid for $u \in N$ (Ref. 30, pp. 190–1). It can be shown that the terms in the representation can be determined by successively differentiating $h\{w[u_0 + \beta(u - u_0)], u_0 + \beta(u - u_0)\}$ with respect to β and setting the result equal to θ. [Recall that $D_1 h(x_0, u_0)$ is assumed to be invertible, and see the proof of Theorem 1.]†

The proof of Theorem 1 shows that the recursive relation (1) arises in a natural way. Such formulas concerning the inversion of ordinary power series and/or the derivatives of composite ordinary functions are probably well known in some circles. In Ref. 29, similar relations are given in an abstract setting for the different problem of constructing a local inverse of a mapping that has a power series expansion.

The expansion of $g(u)$ in Theorem 1, and its associated truncation in Corollary 1, each contains a constant term $g(u_0)$, a term $g_1(u_0, u - u_0)$ that can be written as $L_{u_0}(u - u_0)$ in which L_{u_0} is a bounded linear map, and a sum $R_{u_0}(u - u_0)$ of higher order terms such

---

* Also, N can be chosen so that w is the *only* continuous map of N into S_0 such that $w(u_0) = x_0$ and $h[w(u), u] = \theta$ for $u \in N$.

† Similar remarks apply also in the case of Theorem 4, below. By applying either Theorem 2 or Theorem 5, below, to the map $H: S_0 \times S \to B_1 \times B$ defined by $H(x, v) = [h(x, v), v]$ for $(x, v) \in S_0 \times S$, the writer has obtained an *explicit* expansion, involving partial derivatives of $h(\cdot, \cdot)$, for the solution x of $h(x, u) = w$ in terms of u and w, under certain reasonable assumptions concerning $h(\cdot, \cdot)$ and the sets from which u and w are drawn. The details will be given in another paper.

that $R_{u_0}(h) = o(\|h\|)$ as $\|h\| \to 0$. The following result shows that the hypothesis of Theorem 1 that H.2 is met, and of Corollary 1 that $df[g(u_0)]$ is invertible, are not merely ones of convenience which allow an explicit expression to be given for the terms.

Proposition 2: Let u_0 be an element of U such that $df[g(u_0)]$ exists [respectively, such that f is continuously F-differentiable on a neighborhood of $g(u_0)$]. Suppose that there is a constant $\sigma > 0$ such that

$$g(u) = g(u_0) + L(u - u_0) + R(u - u_0) \tag{8}$$

for $u \in U$ with $\|u - u_0\| < \sigma$, in which L is a bounded linear map from $\mathscr{B}$ into $\mathscr{B}_0$, and $R(\cdot)$ is a map of $\mathscr{B}$ into $\mathscr{B}_0$ with the property that $R(h) = o(\|h\|)$ as $\|h\| \to 0$. Then $df[g(u_0)]^{-1}$ exists [respectively, $df(x)^{-1}$ exists and is continuous in x for x in some neighborhood of $g(u_0)$].

The proposition is proved in Appendix C. It shows, for example, that H.2 is a consequence of the hypotheses that $g(U)$ is an open set, $X = g(U)$, f is differentiable on X, and (8) holds for each $u_0 \in U$. In this connection, notice that because U is open, $g(U)$ is open under merely the condition that f is *continuous* on X.

2.5 Results for complex Banach spaces

Theorem 2: Suppose that $\mathscr{B}$ and $\mathscr{B}_0$ are over the complex field, that the F-derivative $d^m f(x)$ exists at each $x \in X$ for all m, and that H.2 is met. Let $u_0 \in U$, and let ρ be a positive constant with the property that $u_0 + v \in U$ for $\|v\| < \rho$. Then for $u \in U$ such that $\|u - u_0\| < \rho$, we have

$$g(u) = g(u_0) + \sum_{m=1}^{\infty} g_m(u_0, u - u_0), \tag{9}$$

in which

$$g_1(u_0, u - u_0) = df[g(u_0)]^{-1}(u - u_0), \tag{10}$$

and

$$g_m(u_0, u - u_0) = -df[g(u_0)]^{-1} \sum_{l=2}^{m} (l!)^{-1} \sum_{\substack{k_1+k_2+\cdots+k_l=m \\ k_j>0}} d^l f[g(u_0)] g_{k_1}$$

$$\cdot (u_0, u - u_0) g_{k_2}(u_0, u - u_0) \cdots g_{k_l}(u_0, u - u_0), \qquad m \geq 2. \tag{11}$$

Proof: Here $d^m g$ exists on U for each m (see the proof of Theorem 1). In particular, dg exists throughout U, and therefore we see that

$$\lim_{z \to 0} z^{-1}[g(x + zh) - g(x)]$$

exists in $\mathscr{B}_0$ (and equals $dg(x)h$) for each $x \in U$ and $h \in \mathscr{B}$, where z is a *complex* scalar variable. Thus, by Theorem 3.16.2 of Ref. 34, p. 111 and the development preceding it,

$$d^m g[u_0 + z(u - u_0)]/dz^m|_{z=0}$$

exists for each $m \geq 1$, and we have

$$g(u) = g(u_0) + \sum_{m=1}^{\infty} (m!)^{-1} d^m g[u_0 + z(u - u_0)]/dz^m|_{z=0}$$

for $\|u - u_0\| < \rho$.

Notice that

$$d^m g[u_0 + z(u - u_0)]/dz^m|_{z=0} = d^m g[u_0 + r(u - u_0)]/dr^m|_{r=0}$$

in which r is a real variable. Since (see the proof of Theorem 1)

$$d^m g[u_0 + r(u - u_0)]/dr^m|_{r=0} = m! g_m(u_0, u - u_0),$$

the proof is complete.

Theorem 3: Let the hypotheses of Theorem 2 hold. Then for each u_0 there is a $\sigma > 0$ such that the series on the right side of (9) *converges* uniformly *for* $\|u - u_0\| < \sigma$.

Proof: The proof of Theorem 2 shows, using the openness of U, that for each u_0 there is a $\rho > 0$ such that the series converges for $\|u - u_0\| < \rho$. Since dg exists on U, g is continuous on U. The map g is, therefore, *locally bounded* on U in the sense of Ref. 34, Definition 3.17.1, and thus the proof of Theorem 3.17.1 of Ref. 34, p. 112, shows that, given u_0, there is a $\sigma > 0$ such that the convergence is uniform for $\|u - u_0\| < \sigma$.

The following result is obtained from Theorems 2 and 3 in the same way that Corollary 1 is proved.

Theorem 4: Assume that $\mathscr{B}$ and $\mathscr{B}_0$ are over the complex field, and that $d^m f$ exists on X for each m. Let $u_0 \in U$, and suppose that $df[g(u_0)]$ is an invertible map of $\mathscr{B}_0$ onto $\mathscr{B}$. Then there is a $\sigma > 0$ such that the expansion

$$g(u) = g(u_0) + \sum_{m=1}^{\infty} g_m(u_0, u - u_0)$$

is valid and uniformly convergent for $u \in U$ with $\|u - u_0\| < \sigma$, where $g_1(u_0, u - u_0), g_2(u_0, u - u_0), \cdots$ are defined by (10) *and* (11).

2.6 Comments

Under the conditions of Theorem 2 (respectively, Theorem 4) the infinite sum $R(u - u_0)$of the terms $g_2(u_0, u - u_0)$, $g_3(u_0, u - u_0)$, $\cdots$ has the property that $R(h) = o(\|h\|)$ as $\|h\| \to 0$. (This follows from the fact that $dg(u_0) = df[g(u_0)]^{-1}$.) Therefore, Proposition 2, as well as remarks similar to those of Section 2.4, apply here too with regard to the necessity of the hypothesis that H.2 is met (respectively, $df[g(u_0)]$ is invertible).

Following is an interesting corollary of Theorems 2 and 3.

Corollary 2: Suppose that $\mathscr{B}$ *and* $\mathscr{B}_0$ *are complex Banach spaces, and that* f *is a* C^∞*-diffeomorphism of* X *onto* U *(i.e., that* f *is a homeomorphism of* X *onto* U *such that* f *and its inverse* g *have F-derivatives of all orders on* X *and* U*, respectively). Let* $u_0 \in U$*, and let* ρ *be a positive constant with the property that* $u_0 + v \in U$ *for* $\|v\| < \rho$*. Then the series representation* (9), *in which the* $g_m(u_0, u - u_0)$ *are given by* (10) *and* (11), *is valid for* $\|u - u_0\| < \rho$*, and there is a* $\sigma > 0$ *such that the series on the right side of* (9) *converges uniformly for* $\|u - u_0\| < \sigma$*.*

Proof: Since $(df)^{-1}$ exists on X under the conditions of Corollary 2*, Corollary 2 follows from Theorems 2 and 3.

2.7 Discussion

Uniqueness propositions similar to the one described in Section 2.4 apply in the cases of Theorems 2 and 4, as well as Corollary 2. Consider, for example, Theorem 2 and assume that its hypotheses are met. From (10) and (11) (or from the proofs of Theorems 1 and 2), we see that the expansion on the right side of (9) has the homogeneity property that for each m, $g_m(u_0, u_a - u_0) = r^m g_m(u_0, u_b - u_0)$ for $u_a, u_b \in U$ such that $\|u_a - u_0\| < \rho$, $\|u_b - u_0\| < \rho$, and $(u_a - u_0) = r(u_b - u_0)$ for some real r. Suppose that $g(u_0) + \sum_{m=1}^{\infty} h_m(u_0, u - u_0)$ is also an expansion of $g(u)$ about u_0 valid for $\|u - u_0\| < \rho$, and that it has the corresponding homeogeneity property. Assuming, for the purpose of induction that $h_m(u_0, u - u_0) = g_m(u_0, u - u_0)$ for $\|u - u_0\| < \rho$ and $1 \le m \le n$ for some nonnegative integer n, we see† that for any fixed u such that $\|u - u_0\| < \rho$,

$$h_{n+1}(u_0, u - u_0) - g_{n+1}(u_0, u - u_0) = \sum_{m=n+2}^{\infty} [g_m(u_0, u - u_0) - h_m(u_0, u - u_0)] r^{(m-n-1)} \quad (12)$$

for $0 < |r| < 1$. Since

$$\sup_m \| g_m(u_0, u - u_0) - h_m(u_0, u - u_0) \|$$

* We have $f[g(u)] = u$ and $g[f(x)] = x$ for each $u \in U$ and each $x \in X$. Thus, by a version of the chain rule for differentiating a composite function (Ref. 31, pp. 171–2), $df[g(u)]dg(u) = I$ and $dg[f(x)]df(x) = I_0$ for each u and x, where I and I_0 are the identity maps on $\mathscr{B}$ and $\mathscr{B}_0$, respectively. This shows that $df(x)$ has both a right inverse and a left inverse for each $x \in X$, and, therefore, that $(df)^{-1}$ exists on X.

† This type of observation is used in Ref. 33, p. 174, to prove the uniqueness result given there.

is finite, the right side of (12) approaches zero as $r \to 0$. Thus $h_{n+1} \cdot (u_0, u - u_0) = g_{n+1}(u_0, u - u_0)$, and, therefore, $h_m(u_0, u - u_0) = g_m(u_0, u - u_0)$ for all m and all u such that $\|u - u_0\| < \rho$.

The proof of Theorem 2 is based in part on basically well-known results concerning Banach space valued functions of a complex variable. Such results are used also in, for example, Refs. 20 and 24 for related but different purposes.

A comparison of Theorems 1 and 4 leads us to ask whether Theorem 1 can be used to prove a result along the lines of Theorem 4 for cases in which $\mathscr{B}$ and $\mathscr{B}_0$ are not necessarily over the complex field, but the $\|d^m f(x)\|$ are sufficiently small in some not too restrictive sense for x near $g(u_0)$. In this connection, we have the following.

Theorem 5: Let $d^m f$ exist on X for each m, and let u_0 be a point in U. Suppose that there are positive constants δ and γ, and a neighborhood N_0 of $g(u_0)$ in X, such that $\|d^m f(x)\| \le m!\delta\gamma^m$ for $x \in N_0$ and every $m \ge 2$. Assume that $df[g(u_0)]$ is an invertible map from $\mathscr{B}_0$ onto $\mathscr{B}$. Then the conclusion of Theorem 4 holds.

Proof: By the proof of Corollary 1, it suffices to show that there is a $\sigma > 0$ such that

$$(p + 1) \int_0^1 (1 - \beta)^p g_{p+1}(u_0, u - u_0, \beta) d\beta,$$

the remainder in the expansion for $g(u)$ of Theorem 1, approaches θ as $p \to \infty$ uniformly for $\|u - u_0\| < \sigma$. Since $\int_0^1 (1 - \beta)^p d\beta = (p + 1)^{-1}$, it is enough to prove that there are constants $\sigma > 0$, $d > 0$, and $c > 0$ such that for all p, all $\beta \in [0, 1]$, and all $u \in U$ with $\|u - u_0\| < \sigma$, we have $\|g_p(u_0, u - u_0, \beta)\| \le ce^{-dp}$. That we do as follows.

By the continuity of g at u_0, and the continuity of $(df)^{-1}$ at $g(u_0)$ (see the first part of the proof of Corollary 1), choose $\sigma > 0$ so that U contains the open ball of radius σ centered at u_0, and $\|df\{g[u_0 + \beta(u - u)]\}^{-1}\| \le c_1$ and $\|d^l f\{g[u_0 + \beta(u - u_0)]\}\| \le l!\delta\gamma^l$ for some constant c_1 whenever $\beta \in [0, 1]$ and $\|u - u_0\| < \sigma$.

Choose any positive number d. For each m, let h_m denote $\sup\{\|e^{dm} g_m(u_0, u - u_0, \beta)\| : \beta \in [0, 1], \|u - u_0\| < \sigma\}$. From Parts 1.(a) and 1.(b) of Theorem 1, and our hypotheses, the h_m are finite, and we have

$$\sum_{m=1}^{p} h_m \le e^d c_1 \sigma + c_1 \delta \sum_{m=2}^{p} \sum_{l=2}^{m} \sum_{\substack{k_1+k_2+\cdots+k_l=m \\ k_j>0}} \gamma^l h_{k_1} h_{k_2} \cdots h_{k_l}$$

for each $p > 1$. Since

$$\left(\sum_{m=1}^{p-1} h_m \right)^l = \sum h_{k_1} h_{k_2} \cdots h_{k_l}$$

over $(k_1, k_2, \cdots, k_l) \in \{1, 2, \cdots, (p-1)\}^l$, it easily follows that

$$\sum_{m=2}^{p} \sum_{l=2}^{m} \sum_{\substack{k_1+k_2+\cdots+k_l=m \\ k_j>0}} \gamma^l h_{k_1} h_{k_2} \cdots h_{k_l} \le \sum_{l=2}^{p} \gamma^l \left(\sum_{m=1}^{p-1} h_m \right)^l .$$

Thus, with

$$s_p = \sum_{m=1}^{p} h_m,$$

we have

$$s_p \le e^d c_1 \sigma + c_1 \delta \sum_{l=2}^{p} (\gamma s_{p-1})^l$$

for $p > 1$.

Now let $c > 0$ be chosen such that

$$c_1 \delta \sum_{l=2}^{\infty} (\gamma c)^l \le \frac{1}{2} c,$$

and, if necessary, reduce σ so that $e^d c_1 \sigma \le \frac{1}{2}c$. Since $s_1 \le \frac{1}{2}c$, and $s_{(p-1)} \le c$ implies that $s_p \le c$ for $p > 1$, it is clear that $s_p \le c$ (and hence $h_p \le c$) for all p, which completes the proof.

2.8 Comments

Since the hypotheses of Theorem 5 can be shown to ensure the local existence of a Fréchet "power series" expansion (see Ref. 30, p. 190) of f about $g(u_0)$, another way to prove Theorem 5 is to use the result stated in Ref. 29. The observation concerning the representation of $(s_{p-1})^l$ as a sum of products used in our proof to obtain the inequality involving s_p and s_{p-1} (but not the exponential weighting approach) is used also in Ref. 27 for a case that corresponds here to the one in which only a finite number of the $\| d^l f \|$ do not vanish.

Theorem 1 can also be used to prove *nonlocal* convergence results when $\mathscr{B}$ and $\mathscr{B}_0$ are not necessarily complex spaces. For example, let ρ be the radius of any finite open ball contained in U and centered at u_0. Suppose that for every $x \in X$ the following is true: $d^m f(x)$ exists for each m, $df(x)^{-1}$ exists, and $\| df(x)^{-1} \| \le \rho_0$ for some constant ρ_0. Then, using Theorem 1, it can be shown that if the $\| d^m f(\cdot) \|$ satisfy certain smallness conditions on X for $m \ge 2$, the expansion described in the conclusion of Theorem 2 converges uniformly to $g(u)$ for $\| u - u_0 \| < \rho$. The "smallness conditions" are met if, for example,

$$\sup_{x \in X} \| d^m f(x) \| = 0 \quad \text{for} \quad m \ge M \quad \text{for some} \quad M \ge 2,$$

and each nonzero

$$\sup_{x \in X} \| d^m f(x) \|$$

with $m > 1$ is sufficiently small. The details and a proof will be given in a later paper.

2.9 Properties 1 and 2

We conclude this section with a proof of a proposition used in Section III, where attention is directed to cases in which the elements of $\mathscr{B}$ and $\mathscr{B}_0$ are functions of a time variable t. The proposition is used to show that under certain very reasonable conditions, causality and time invariance (or periodicity of variation)* are properties which, when possessed by g, are inherited by the terms $g_1(u_0, u - u_0), \cdots,$ $g_p(u_0, u - u_0)$ in Theorem 1, and by the terms $g_1(u_0, u - u_0)$, $g_2(u_0, u - u_0), \cdots$ in Theorem 2. We first introduce some preliminaries.

Let Ω denote a nonempty set of real numbers. For each $\omega \in \Omega$, let T_ω and $T_{0\omega}$ denote linear transformations of $\mathscr{B}$ and $\mathscr{B}_0$, respectively, such that $\| T_{0\omega} w \| \le \| w \|$ for $\omega \in \Omega$ and $w \in \mathscr{B}_0$. Let S be a subset of $\mathscr{B}$ such that $T_\omega S \subseteq S$ for $\omega \in \Omega$. Let J denote an open interval in the set R^1 of real numbers.

We say that a map $F: J \times S \to \mathscr{B}_0$ has *Property 1* on S at a point $r \in J$ if

$$T_{0\omega} F(r, v) = T_{0\omega} F(r, T_\omega v)$$

for all $v \in S$ and $\omega \in \Omega$. Finally, we say that $F: J \times S \to \mathscr{B}_0$ has *Property 2* on S at a point r in J if

$$T_{0\omega} F(r, v) = F(r, T_\omega v)$$

for $v \in S$ and $\omega \in \Omega$.

Proposition 3: Suppose that $F: J \times S \to \mathscr{B}_0$ has Property 1 (respectively, Property 2) on S for each $r \in J$, and that for an arbitrary $v \in S$ the derivative $dF(r, v)/dr$ exists at each $r \in J$. Then the map $H: J \times S \to \mathscr{B}_0$, defined by $H(r, v) = dF(r, v)/dr$ for each r and each v, has Property 1 (respectively, Property 2) on S for each $r \in J$.

Proof: Assume initially that F has Property 1. Let arbitrary $r \in J$ and $v \in S$ be given and let β be a real variable. Using

$$\lim_{\beta \to 0} \| \beta^{-1} [F(r + \beta, v) - F(r, v)] - H(r, v) \| = 0,$$

we have

$$\lim_{\beta \to 0} \| \beta^{-1} [T_{0\omega} F(r + \beta, v) - T_{0\omega} F(r, v)] - T_{0\omega} H(r, v) \| = 0 \quad (13)$$

* See Section 3.1 for the pertinent definitions.

for any $\omega \in \Omega$. Since (13) holds also with v replaced with $T_\omega v$, and using the hypotheses that F has Property 1 on S at r and at $(r + \beta)$ for sufficiently small β, we find that

$$\lim_{\beta \to 0} \| \Delta(\beta) + T_{0\omega} H(r, v) - T_{0\omega} H(r, T_\omega v) \| = 0,$$

in which $\Delta(\beta) = \beta^{-1}[T_{0\omega} F(r + \beta, v) - T_{0\omega} F(r, v)] - T_{0\omega} H(r, v)$. By (13), $\| \Delta(\beta) \| \to 0$ as $\beta \to 0$. Therefore, we have $T_{0\omega} H(r, v) = T_{0\omega} H(r, T_\omega v)$ for arbitrary $\omega \in \Omega$, as claimed. The Property 2 part of the proposition can be proved in essentially the same way.

III. APPLICATIONS AND EXAMPLES

Throughout this section, we consider cases where each element of $\mathscr{B}$, and also of $\mathscr{B}_0$, is a function of a time variable t. Specifically, we now assume that each element of $\mathscr{B}$ is a map from a set T of real numbers into a linear space V with zero element θ_V, and, similarly, that the elements of $\mathscr{B}_0$ are maps from T into a linear space V_0 with zero element θ_V.

We shall be concerned mainly with the cases where either $\mathscr{B} = \mathscr{B}_0 = L_\infty(\mathrm{R})$ or $\mathscr{B} = \mathscr{B}_0 = L_\infty(\mathrm{C})$, in which by $L_\infty(\mathrm{R})$ [respectively, $L_\infty(\mathrm{C})$] we mean the real (respectively, complex) Banach space of (Lebesgue) measurable* real (respectively complex) column n-vector valued functions v defined on the interval $[0, \infty)$ such that the jth component v_j of v satisfies

$$\sup_{t \geq 0} | v_j(t) | < \infty \quad \text{for} \quad j = 1, 2, \cdots, n,$$

and where the norm $\| \cdot \|$ on $L_\infty(\mathrm{R})$ or $L_\infty(\mathrm{C})$ is given by

$$\| v \| = \max_j \sup_t | v_j(t) |.$$

(As usual, n denotes an arbitrary positive integer.) If, for example, $\mathscr{B} = \mathscr{B}_0 = L_\infty(\mathrm{R})$, then $T = [0, \infty)$ and we can take V and V_0 to be R^n.

3.1 Causality and time-invariance

Referring to Proposition 3 and the associated definitions, let $\Omega = T$, and initially let T_ω (respectively, $T_{0\omega}$) be the "time-truncation" operator defined on $\mathscr{B}$ (respectively, $\mathscr{B}_0$) by $(T_\omega v)(t) = v(t)$ for $t \leq \omega$, and $(T_\omega v)(t) = \theta_V$ for $t > \omega$ (respectively, by $(T_{0\omega} v)(t) = v(t)$ for $t \leq \omega$, and $(T_{0\omega} v)(t) = \theta_{V_0}$ for $t > \omega$) for each v and each ω. Assume that for $\omega \in T$, T_ω and $T_{0\omega}$ map $\mathscr{B}$ and $\mathscr{B}_0$ into themselves, and that $\| T_{0\omega} v \| \leq \| v \|$

* See, for example, Ref. 35.

for $\omega \in T$ and all v. [Notice that these assumptions are satisfied if, for example, $\mathscr{B} = \mathscr{B}_0 = L_\infty(R)$ or $\mathscr{B} = \mathscr{B}_0 = L_\infty(C)$.]*

Suppose that the hypotheses of Theorem 1 are met, that U is an open ball centered at $u_0 = \theta$, that $T_\omega U \subseteq U$ for $\omega \in T$ [which is clearly satisfied if $\mathscr{B} = L_\infty(R)$ or $L_\infty(C)$] and initially assume that g is causal on U in the sense that $T_{0\omega} g(v) = T_{0\omega} g(T_\omega v)$ for any $v \in U$ and $\omega \in T$.

Let $G_m(u)$ denote $g_m(\theta, u)$ of Theorem 1 for $m = 1, 2, \cdots, p$ and $u \in U$. We observe that $g(\beta u)$ is an element of $\mathscr{B}_0$ for each $\beta \in (-1, 1)$ and each $u \in U$. By the proof of Theorem 1, $d^m g(\cdot)$ exists on U for $m = 1, 2, \cdots, p$, from which it follows that $d^m g(\beta u)/d\beta^m$ exists for $\beta \in (-1, 1)$, $u \in U$, and $m = 1, 2, \cdots, p$ [31, p. 198]. By the proof of Theorem 1, $m! G_m(u) = d^m g(\beta u)/d\beta^m$ at $\beta = 0$ for each m and u. Therefore, by the Property 1 part of Proposition 3 [with $S = U$ and $J = (-1, 1)$] and an obvious inductive argument, it follows that each $G_m: U \to \mathscr{B}_0$ is causal in the same sense that g is causal.†

Now suppose that T is one of the four sets $[0, \infty)$, $(-\infty, \infty)$, $\{0, 1, 2, \cdots\}$, or $\{0, \pm 1, \pm 2, \cdots\}$. Again take $\Omega = T$. Let T_ω (respectively, $T_{0\omega}$) denote the "time delay operator" defined by $(T_\omega v)(t) = \theta_V$ for $t < \omega$ and $(T_\omega v)(t) = v(t - \omega)$ for $t \geq \omega$ when either $T = [0, \infty)$ or $T = \{0, 1, 2, \cdots\}$, and by $(T_\omega v)(t) = v(t - \omega)$ when $T = (-\infty, \infty)$ or $T = \{0, \pm 1, \pm 2, \cdots\}$ (respectively, $(T_{0\omega} v)(t) = \theta_{V_0}$ for $t < \omega$ and $(T_{0\omega} v)(t) = v(t - \omega)$ for $t \geq \omega$ if T is either $[0, \infty)$ or $\{0, 1, 2, \cdots\}$, and $(T_{0\omega} v)(t) = v(t - \omega)$ when T is either $(-\infty, \infty)$ or $\{0, \pm 1, \pm 2, \cdots\}$). Assume here, as above, that T_ω and $T_{0\omega}$ map $\mathscr{B}$ and $\mathscr{B}_0$, respectively, into themselves, that $\| T_{0\omega} v \| \leq \| v \|$ for each v and ω, that the hypotheses of Theorem 1 are met, that U is an open ball centered at θ, and that $u_0 = \theta$. Consider the case in which g is causal on U, and g maps the zero element of $\mathscr{B}$ into the zero element of $\mathscr{B}_0$. Assume that g is *time invariant* on U in the sense that $T_{0\omega} g(u) = g(T_\omega u)$ for $u \in U$ and $\omega \in T$. Let G_m be as defined in the preceding paragraph. By the Property 2 part of Proposition 3, and the observations concerning $d^m g(\beta u)/d\beta^m$ in the preceding paragraph, we see that each G_m ($m = 1, 2, \cdots, p$) is time invariant on U.

The material just described can be modified to address the case in which g is periodically varying with a given period τ. Specifically, suppose that T is $[0, \infty)$, $(-\infty, \infty)$, $\{0, 1, 2, \cdots\}$, or $\{0, \pm 1, \pm 2, \cdots\}$, and that τ is a positive element of T. Let T_ω and $T_{0\omega}$ be as defined in the preceding paragraph, but with Ω taken to be the single-element set $\{\tau\}$ rather than T. Then, in the setting described in the preceding

* The assumptions are not met for $\mathscr{B}$ the set of bounded *continuous* functions from $[0, \infty)$ to R^1.

† Our definition is consistent with the one introduced in Ref. 36, p. 888, concerning causality for operators between abstract spaces. Also, a related result is given in Ref. 37, p. 40, for polynomial operators.

paragraph, $T_{0\omega}g(u) = g(T_\omega u)$ for $u \in U$ and $\omega \in \Omega$ means that g is *periodically varying with period* τ on U, and we see that if g has this property, it is inherited by the G_m.

In the case of Theorem 2, each $d^m g(\cdot)$ exists on U (because H.2 holds), and $g_m(\theta, u) = (m!)^{-1} d^m g(\beta u)/d\beta^m$ at $\beta = 0$ for each $m = 1, 2, \cdots$ and $\|u\| < \rho$. Therefore, results essentially the same as those developed in the preceding four paragraphs hold also with regard to the terms in (9).*

3.2 An application of theorem 1

Our first example, as well as the example in Section 3.3, concerns a nonlinear integral equation that plays an important role in the theory of feedback systems. To introduce the equation, we need the following definitions.

Let α_0 and α_1 be positive numbers with $\alpha_0 \leq \alpha_1$, and let $\psi_1, \psi_2, \cdots, \psi_n$ be a collection of $(p+1)$-times continuously differentiable functions from R^1 onto R^1 such that $\psi_i(0) = 0$ and $\alpha_0 \leq d\psi_i(\lambda)/d\lambda \leq \alpha_1$ for all i and all λ. Below, for convenience, we shall use $\psi_i^{(m)}$ to denote the mth derivative of ψ_i. Let ψ denote the map from R^n into R^n defined by $[\psi(s)]_i = \psi_i(s_i)$ for $i = 1, 2, \cdots, n$ and all $s \in R^n$, in which $[\psi(s)]_i$ and s_i are the ith components of $\psi(s)$ and s, respectively.

Let k denote an $n \times n$ matrix-valued function defined on $[0, \infty)$ such that each k_{ij} is measurable, bounded, and satisfies

$$\int_0^\infty |k_{ij}(\tau)| d\tau < \infty.$$

In this and the following section, each k_{ij} is assumed to be real valued.

Consider the equation

$$x(t) + \int_0^t k(t-\tau)\psi[x(\tau)]d\tau = u(t), \qquad t \geq 0, \tag{14}$$

as well as the related equation

$$y(t) + \int_0^t k(t-\tau)D(\tau)y(\tau)d\tau = v(t), \qquad t \geq 0, \tag{15}$$

in which u and v are elements of $L_\infty(R)$, and D is a real $n \times n$ diagonal-matrix-valued function defined on $[0, \infty)$ such that each D_{ii} is measurable on $[0, \infty)$ and satisfies $\alpha_0 \leq D_{ii}(\tau) \leq \alpha_1$ for each τ. Since ψ satisfies

* Specifically, the development remains valid if U is taken to be $\{u \in \mathscr{B}: \|u\| < \rho\}$, and if "Theorem 1" and "$m = 1, 2, \cdots, p$", respectively, are replaced with "Theorem 2" and "1, 2, $\cdots$."

a global Lipschitz condition on R^n, and k is as indicated, a standard successive-approximations approach (see, in particular, Ref. 38, Section 1.13) can be used to show that for each u and v in $L_\infty(R)$, (14) and (15) have solutions x and y, respectively, in the space E of functions w from $[0, \infty)$ into R^n such that

$$\int_0^\lambda |w_i(\tau)|^2 d\tau < \infty$$

for each $\lambda \in (0, \infty)$ and each i, and that x and y are unique in E.* To fix ideas, assume (this is often very easy to justify) that the only solutions of (14) of interest to us are those that are contained in E.

Let A.1 denote the hypothesis that for each u and v in $L_\infty(R)$, $L_\infty(R)$ contains any solution x of (14) in E, as well as any solution y of (15) in E. For explicit conditions on k, α_0, and α_1 under which A.1 holds, see Ref. 36, Theorem 3.

Assume initially that A.1 is met, and notice that then for each pair of elements $u, v \in L_\infty(R)$, the space $L_\infty(R)$ contains exactly one element x such that (14) is satisfied and exactly one element y such that (15) is met. In particular, we see that we can take f in Section 2.1 to be the map of $L_\infty(R)$ into itself defined by

$$f(w)(t) = w(t) + \int_0^t k(t-\tau)\psi[w(\tau)]d\tau, \qquad t \geq 0$$

for each $w \in L_\infty(R)$, and can take $\mathscr{B}$, $\mathscr{B}_0$, X, and U to be $L_\infty(R)$. In this example, g is defined on all of $\mathscr{B}$.

To discuss the example in more detail, it is convenient to let K and Ψ denote the maps of $L_\infty(R)$ into $L_\infty(R)$ defined by

$$(Kw)(t) = \int_0^t k(t-\tau)w(\tau)d\tau, \qquad t \geq 0$$

$$(\Psi w)(t) = \psi[w(t)], \qquad t \geq 0$$

for each $w \in L_\infty(R)$. Thus, here $f = I + K\Psi$ in which I is the identity map on $L_\infty(R)$. Consider Ψ.

Proposition 4: *$d\Psi$ exists on $L_\infty(R)$, and for w and h in $L_\infty(R)$, we have $[d\Psi(w)h](t) = D_0(t)h(t)$ for $t \geq 0$ in which $D_0(t)$ is the diagonal matrix* $\text{diag}[\psi_1^{(1)}[w_1(t)], \psi_2^{(1)}[w_2(t)], \cdots, \psi_n^{(1)}[w_n(t)]\}$.

---

* The integral on the left side of (14) can easily be shown to be an element of R^n for each t whenever $x \in E$. Since the value of the integral is unchanged if x is replaced by any element of E that agrees with x almost everywhere, (14) has a solution if there is an element of E that satisfies the equation almost everywhere, and, moreover, any solution $x \in E$ is unique and not merely *essentially* unique. Similar remarks apply in the case of (15).

Proof: For w and h in $L_\infty(\mathrm{R})$,

$$\Psi(w+h)(t) - \Psi(w)(t)$$

$$= D_0(t)h(t) + \int_0^1 [D_{0\beta}(t) - D_0(t)]d\beta \cdot h(t), \qquad t \geq 0, \quad (16)$$

in which $D_{0\beta}(t)$ is the diagonal matrix of order n whose ith diagonal element is $\psi_i^{(1)}[\beta[w_i(t) + h_i(t)] + (1-\beta)w_i(t)]$. Since each $\psi_i^{(1)}$ is uniformly continuous on compact subsets of R^1, we see that for any fixed w, the difference $(D_{0\beta i} - D_{0i})$ of ith diagonal elements satisfies

$$\sup_{t \geq 0} |D_{0\beta i}(t) - D_{0i}(t)| \to 0 \quad \text{as} \quad \|h\| \to 0.$$

Thus, with $\delta \in L_\infty(\mathrm{R})$ defined by $\delta(t) = \int_0^1 [D_{0\beta}(t) - D_0(t)]d\beta \cdot h(t)$ for $t \geq 0$, we see that $\|\delta\| = o(\|h\|)$, which proves the proposition.

Proposition 4, together with the discussion above concerning (14) and (15), show that df exists, that (using the linearity of K) $df = I + Kd\Psi(w)$ for each w, and that for any w the map $df(w)$ is an invertible map of $L_\infty(\mathrm{R})$ onto itself.

To make further progress, we need the following result which is more general than Proposition 4.

Proposition 5: For each $m = 1, 2, \cdots, (p+1)$, $d^m\Psi$ exists and is continuous on $L_\infty(\mathrm{R})$, and, with $h_1, h_2, \cdots, h_m$ any m elements of $L_\infty(\mathrm{R})$, we have

$$[d^m\Psi(w)h_1 \cdots h_m(t)]_i = \psi_i^{(m)}[w_i(t)] \prod_{j=1}^{m} h_{ji}(t), \qquad t \geq 0$$

for each $i = 1, 2, \cdots, n$ and any $w \in L_\infty(\mathrm{R})$.

Proof: Let w be given. By Proposition 4 and, (with regard to continuity) the observation that $\|d\Psi(w+v)h - d\Psi(w)h\| \leq \max_i \sup_{t\geq 0} |\psi_i^{(1)}[w_i(t) + v_i(t)] - \psi_i^{(1)}[w_i(t)]|$ for $\|h\| = 1$, the assertion for $m = 1$ is true. Suppose that the assertion is true for m such that $1 \leq m \leq l$ with $l < (p+1)$.

Let $\tilde{Q}_l(w)$ denote the continuous multilinear mapping of $L_\infty(\mathrm{R})^{(l+1)}$ into $L_\infty(\mathrm{R})$ defined by

$$[\tilde{Q}_l(w)(p_1, p_2, \cdots, p_{l+1})(t)]_i = \psi_i^{(l+1)}[w_i(t)] \prod_{j=1}^{l+1} p_{ji}(t), \qquad t \geq 0$$

for each i and for $p_j \in L_\infty(\mathrm{R})$ for all j. We shall use $Q_l(w)$ to denote the usual associate (Ref. 31, p. 318) of $\tilde{Q}_l$ that belongs to $L(L_\infty(\mathrm{R}), L(L_\infty(\mathrm{R}), \cdots, L(L_\infty(\mathrm{R}), L_\infty(\mathrm{R}))\cdots))$ with $(l+1)$ L's, in which $L(A_1, A_2)$ stands for the set of continuous linear operators from the Banach space A_1 into the Banach space A_2.*

* For example, if $l = 2$, $L(L_\infty(\mathrm{R}), L(L_\infty(\mathrm{R}), \cdots, L(L_\infty(\mathrm{R}), L_\infty(\mathrm{R}))\cdots)) = L(L_\infty(\mathrm{R}), L(L_\infty(\mathrm{R}), L(L_\infty(\mathrm{R}), L_\infty(\mathrm{R}))))$.

By our induction hypothesis, and with h, as well as h_j for $j = 1, 2, \cdots, l$, elements of $L_\infty(\mathrm{R})$,

$$\|d^l\Psi(w+h)h_1 \cdots h_l - d^l\Psi(w)h_1 \cdots h_l - Q_l(w)hh_1 \cdots h_l\|$$
$$= \max_i \sup_{t\geq 0} \left| \psi_i^{(l)}[w_i(t)+h_i(t)] \prod_{j=1}^{l} h_{ji}(t) - \psi_i^{(l)}[w_i(t)] \prod_{j=1}^{l} h_{ji}(t) - \psi_i^{(l+1)}[w_i(t)]h_i(t) \prod_{j=1}^{l} h_{ji}(t) \right| \leq c(h) \prod_{j=1}^{l} \|h_j\|,$$

where

$$c(h) = \max_i \sup_{t\geq 0} |\psi_i^{(l)}[w_i(t)+h_i(t)] - \psi_i^{(l)}[w_i(t)] - \psi_i^{(l+1)}[w_i(t)]h_i(t)|.$$

Thus, $\|d^l\Psi(w+h) - d^l\Psi(w) - Q_l(w)h\| \leq c(h)$. By the uniform continuity on compact sets of the $\psi_i^{(l+1)}$ (see the proof of Proposition 4), we have $c(h) = o(\|h\|)$ as $\|h\| \to 0$, which shows that $d^{(l+1)}\Psi(w)$ exists and that $d^{(l+1)}\Psi(w) = Q_l(w)$. Since $\|Q_l(w+h)p_1 \cdots p_{l+1} - Q_l(w)p_1 \cdots p_{l+1}\| \leq$

$$\max_i \sup_{t\geq 0} |\psi_i^{(l+1)}[w_i(t)+h_i(t)] - \psi_i^{(l+1)}[w_i(t)]| \cdot \prod_{j=1}^{l+1} \|p_j\|,$$

and the $\psi_i^{(l+1)}$ are continuous, we see that

$$\|Q_l(w+h) - Q_l(w)\| \to 0 \quad \text{as} \quad \|h\| \to 0,$$

showing that $d^{(l+1)}\Psi(w)$ depends continuously on w. This completes the proof.

Returning now to our example, by Proposition 5 and the linearity and boundedness of K, we see that $d^m f(w)$ exists and is continuous for $w \in L_\infty(\mathrm{R})$ and $m = 1, 2, \cdots, (p+1)$. (Of course, $d^m f = Kd^m\Psi$ for $1 < m \leq (p+1)$.) Therefore, the hypotheses of Theorem 1 are satisfied. Now choose $u_0 = \theta$ in Theorem 1 and notice that $g(\theta) = \theta$.

Referring to the standard successive approximations technique (Ref. 38, Section 1.13) that can be used to construct a unique solution x in E of (14) for each $u \in L_\infty(\mathrm{R})$, by the rule by which the iterates are generated it follows that g is causal and time invariant on $L_\infty(\mathrm{R})$ in the sense of Section 3.1. Therefore, the material in Section 3.1 shows that $g_m(\theta, \cdot)$ of Theorem 1 is both causal and time invariant on $L_\infty(\mathrm{R})$ for each $m = 1, 2, \cdots, p$.*

The terms $g_1(\theta, u), g_2(\theta, u), \cdots, g_p(\theta, u)$ in the expansion in Theorem 1 can be determined using 1.(a) and 1.(b).† For example, with H

* The same conclusion can be reached by considering the specific form of the $g_m(\theta, u)$, and using the fact that the operator H introduced below can be shown to be causal and time invariant. (In this connection, see Lemma 2.)

† The recursive process is straightforward in principle, but the complexity mounts rapidly with increasing order.

denoting $[I + Kd\Psi(\theta)]^{-1}$, we have $g_1(\theta, u) = Hu$, and, using $g_1(\theta, u) = Hu$, we find that when $p \geq 2$,

$$g_2(\theta, u) = -\tfrac{1}{2}HKd^2\Psi(\theta)(Hu)^2 \tag{17}$$

since $k_1 + k_2 = 2$ with k_1 and k_2 positive integers requires that $k_1 = k_2 = 1$, and

$$g_3(\theta, u) = -\tfrac{1}{6}HKd^3\Psi(\theta)(Hu)^3 + \tfrac{1}{2}HKd^2\Psi(\theta)(Hu)[HKd^2\Psi(\theta)(Hu)^2]. \tag{18}$$

[The derivation of (18) uses $g_1(\theta, u) = Hu$, (17), and the observation that $k_1 + k_2 = 3$ is met if either $k_1 = 1$ and $k_2 = 2$ or $k_1 = 2$ and $k_2 = 1$.]

Proposition 5 provides an important interpretation of the terms in the expressions for the $g_m(\theta, u)$. For example, by Proposition 5, we see that $d^3\Psi(\theta)(Hu)^3$, which appears in the first term on the right side of (18), is the element s of $L_\infty(\mathrm{R})$ given by $s_i(t) = \psi_i^{(3)}(0)[(Hu)(t)_i]^3$ for $t \geq 0$ and each i. Similarly, the ith component of $d^2\Psi(\theta)(Hu)\cdot[HKd^2\psi(\theta)(Hu)^2]$ in (18) has values $\psi_i^{(2)}(0)[(Hu)(t)_i]q_i(t)$, where $q = HKd^2\Psi(\theta)(Hu)^2$. Of course, q also has an immediate interpretation.

3.3 An application of corollary 1

Corollary 1 is in many respects a local version of Theorem 1. Here we give an example of an application of the corollary. As in Section 3.2, let $\mathscr{B} = \mathscr{B}_0 = L_\infty(\mathrm{R})$, and let K, Ψ, I, and u_0 be as defined there, but here it is not required that A.1 be met.

Let $F: L_\infty(\mathrm{R}) \to L_\infty(\mathrm{R})$ be given by $F = I + K\Psi$. As in Section 3.2, d^mF exists and is continuous on $L_\infty(\mathrm{R})$ for $m = 1, 2, \cdots, (p + 1)$. Of course, $dF(\theta) = I + Kd\Psi(\theta)$.

Let z be a complex scalar variable, and let $\tilde{K}$, the Laplace transform of k, be given by

$$\tilde{K}(z) = \int_0^\infty k(t)e^{-zt}dt, \qquad \mathrm{Re}(z) \geq 0.$$

Assume that

$$\det[1_n + \tilde{K}(z)\mathrm{diag}\{\psi_1^{(1)}(0), \cdots, \psi_n^{(1)}(0)\}] \neq 0$$

for $\mathrm{Re}(z) \geq 0$, in which 1_n is the identity matrix of order n. As a consequence, it can be shown (see Lemma 2 in Section 3.5) that $dF(\theta): L_\infty(\mathrm{R}) \to L_\infty(\mathrm{R})$ is invertible. Thus, by Lemma 1, there are open subsets S_1 and S_2 of $L_\infty(\mathrm{R})$, each containing θ, such that F restricted to S_1 is a homeomorphism of S_1 onto S_2.

Therefore, the hypotheses of Corollary 1 are met if X is chosen to be S_1, f is taken to be the restriction of F to X, and U is any open convex set contained in S_2 and containing θ. This establishes the existence of, and shows how to obtain, a series approximation with error $o(\|u\|^p)$ as $\|u\| \to 0$ of the locally unique solution x of $Fx = u$, for u of sufficiently small norm.

3.4 Physical systems, and an application of theorem 4

In the following, $H(\mathrm{C})$ denotes the linear space of complex column n-vector-valued functions h defined on $[0, \infty)$ such that, with T_ω the "time truncation" operator of Section 3.1, we have $T_\omega h \in L_\infty(\mathrm{C})$ for $\omega \in [0, \infty)$ (i.e., such that any truncation of h is bounded and measurable). Clearly, unlike $L_\infty(\mathrm{C})$, $H(\mathrm{C})$ can contain unbounded functions.

Consider a physical system with an input v drawn from $L_\infty(\mathrm{C})$ and an output w contained in $H(\mathrm{C})$. Let the system be composed of linear elements, as well as nonlinear elements. Suppose that the nonlinear elements can be viewed as collectively introducing a constraint that can be written as $y = Nx$, in which N is a map from one subset of $H(\mathrm{C})$ into another, and where x and y, respectively, are the $H(\mathrm{C})$ input and output of the nonlinear part of the system.

With regard to the remainder of the system, which is linear, assume that there are linear maps A, B, C, and D of $H(\mathrm{C})$ into itself such that

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

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

Concerning the degree of generality of the model, and the assumption that the values of v, w, x, and y have the *same dimension* n, notice that we have not ruled out the possibility that some components of v, x, and/or y have no effect on the system, and, similarly, that certain of the components of w can be ignored. Nonzero initial conditions, if any, are assumed to be able to be taken into account either in N or as inputs to the system. A signal-flow-graph representation of the relations under consideration is given in Fig. 1.*

In this section, we use Theorem 4 to obtain a result concerning the response w of the system to inputs v having sufficiently small norm. To state the result, we introduce the following hypotheses and definition.

B.1: The restrictions of A, B, C, and D to $L_\infty(\mathrm{C})$ are bounded linear maps of $L_\infty(\mathrm{C})$ into itself.

B.2: There are open neighborhoods S_1 and S_2 of θ in $L_\infty(\mathrm{C})$ such that N restricted to S_1 is an invertible map of S_1 onto S_2. The map N also satisfies $N(\theta) = \theta$.

* This type of representation of a system has been used in different but related settings in Refs. 39, 40, and 41. The maps A, B, C, and D exist for a very large class of systems, but it is not difficult to give examples in which one or more map does not exist (see Ref. 40, pp. 244–5, for a very simple linear example along these lines).

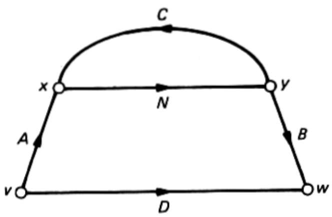

Fig. 1—Signal flow graph.

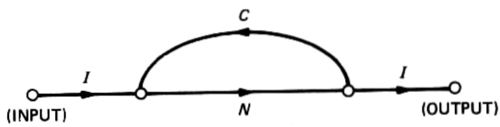

Fig. 2—Feedback part of the flow graph of Fig. 1.

Definition: When B.2 holds, the inverse of the restriction of N to S_1 is denoted by Φ. (Of course, Φ maps S_2 onto S_1, and we have $\Phi(\theta) = \theta$.)
B.3: There is an open neighborhood S_0 of θ in $L_\infty(\mathrm{C})$ such that $d^m\Phi$ exists on S_0 for $m = 1, 2, \cdots$.
B.4: $[d\Phi(\theta) - C_\infty]$ is an invertible map of $L_\infty(\mathrm{C})$ onto $L_\infty(\mathrm{C})$, where C_∞ is the restriction of C to $L_\infty(\mathrm{C})$.
Theorem 6: When B.1, B.2, B.3, and B.4 are met, there is a positive number δ and a neighborhood S of θ in $L_\infty(\mathrm{C})$ with the following properties:

(*i*) *For each $v \in L_\infty(\mathrm{C})$ with $\|v\| < \delta$, there exist unique y, x, and w of S, S_1, and $L_\infty(\mathrm{C})$, respectively, such that* (19), (20), *and $y = Nx$ hold.*

(*ii*) *The function w described in* (*i*) *is given by*

$$w = Dv + \sum_{m=1}^{\infty} Bg_m(Av) \tag{21}$$

for $\|v\| < \delta$, in which $g_1(Av) = [d\Phi(\theta) - C_\infty]^{-1}Av$, and

$$g_m(Av) = -[d\Phi(\theta) - C_\infty]^{-1}$$

$$\cdot \sum_{l=2}^{m} (l!)^{-1} \sum_{\substack{k_1+k_2+\cdots+k_l=m \\ k_j>0}} d^l\Phi(\theta) g_{k_1}(Av) g_{k_2}(Av) \cdots g_{k_l}(Av)$$

for $m \geq 2$, and the series on the right side of (21) *converges uniformly for $\|v\| < \delta$.*

Proof: Using B.4, there are open neighborhoods S and S_3 of θ in $L_\infty(\mathrm{C})$, with $S \subseteq S_2$, such that for each $p \in S_3$ there is in S a unique y with the property that $\Phi(y) - C_\infty y = p$ (see Lemma 1). By the boundedness of the restriction of A to $L_\infty(\mathrm{C})$, $\delta_1 > 0$ can be chosen so that $Av \in S_3$ when $v \in L_\infty(\mathrm{C})$ and $\|v\| < \delta_1$, and thus so that for each such v there is a unique $y \in S$ such that $\Phi(y) - C_\infty y = Av$. For each such v and its associate y, let $x = \Phi(y)$ and $w = Dv + By$. Observe that for $\|v\| < \delta_1$ the corresponding triple (y, x, w) has the properties indicated in (*i*).

Now assume, as we may without loss of generality, that S_3 in the preceding paragraph is convex, and that S_0 of B.3 satisfies $S \subseteq S_0$. Choose $\mathscr{B}$, $\mathscr{B}_0$, X, U, and f, respectively, of Section 2.1 to be $L_\infty(\mathrm{C})$, $L_\infty(\mathrm{C})$, S, S_3, and the map defined by $f(s) = \Phi(s) - C_\infty s$ for $s \in S$. We have $df = d\Phi - C_\infty$ on X, $d^m f = d^m\Phi$ for $m = 2, 3, \cdots$ on X, and, by B.4, $df(\theta)$ is an invertible map. Thus, by Theorem 4, there is a $\sigma > 0$ such that S_3 contains an open ball centered at θ of radius σ, and the solution $s \in S$ of $\Phi(s) - C_\infty s = p$ for $p \in L_\infty(\mathrm{C})$ with $\|p\| < \sigma$ is given by the uniformly convergent series

$$\sum_{m=1}^{\infty} g_m(p),$$

in which $g_1(p) = [d\Phi(\theta) - C_\infty]^{-1}p$, and where

$$g_m(p) = -[d\Phi(\theta) - C_\infty]^{-1} \sum_{l=2}^{m} (l!)^{-1} \sum_{\substack{k_1+k_2+\cdots+k_l=m \\ k_j>0}} d^l\Phi(\theta) g_{k_1}(p) \cdots g_{k_l}(p)$$

for $m \geq 2$. Therefore, (ii) of Theorem 6 holds if for some $\delta \in (0, \delta_1)$ we have $\|Av\| < \sigma$ whenever $\|v\| < \delta$, and this condition is met because by B.1 the restriction of A to $L_\infty(\mathrm{C})$ is bounded. [Notice that here, and with regard to the *uniform* convergence of the series, we also use the boundedness of the restriction of B to $L_\infty(\mathrm{C})$.]

3.5 Volterra-series representations

In this, our final section, we first consider the single-input case in which v in Theorem 6 satisfies $[v(t)]_i = 0$ for $t \geq 0$ and $i > 1$. We give a theorem which provides explicit conditions on A, B, C, D, and N under which the hypotheses of Theorem 6 are met and the series for w can be interpreted as a *Volterra series* in v_1 the function on $[0, \infty)$ whose values are $[v(t)]_1$. Towards the end of the section, an n-input extension of our result is given.

We will use the following definitions and hypotheses.

Definition: For any positive integers l and q, let $S_q^{(l)}$ denote the set of all functions h from the l-dimensional interval $[0, \infty)^l$ into the set of complex $n \times q$ matrices such that each h_{ij} is measurable and bounded on $[0, \infty)^l$, and satisfies

$$\int_{[0,\infty)^l} |h_{ij}(\tau_1, \tau_2, \cdots, \tau_l)| \, d(\tau_1, \tau_2, \cdots, \tau_l) < \infty. \tag{22}$$

Definition: If r and s are two complex column n-vectors, then rs denotes the column n-vector defined by $(rs)_i = r_i s_i$ for $i = 1, 2, \cdots, n$.

Definition: ${}^{(1)}L_\infty(\mathrm{C})$ denotes $L_\infty(\mathrm{C})$ with $n = 1$.

C.1: Referring to A, B, C, and D of (19) and (20), there are elements a, b, c, and d of $S_n^{(1)}$ such that for each $p \in L_\infty(\mathrm{C})$,

$$(Ap)(t) = \int_0^t a(t-\tau)p(\tau)d\tau$$

$$(Bp)(t) = \int_0^t b(t-\tau)p(\tau)d\tau$$

$$(Cp)(t) = \int_0^t c(t-\tau)p(\tau)d\tau$$

$$(Dp)(t) = \int_0^t d(t-\tau)p(\tau)d\tau$$

for $t \geq 0$.

C.2: For some $\gamma > 0$, N of Section 3.4 is defined on $\Gamma = \{s \in L_\infty(\mathrm{C}): \|s\| < \gamma\}$ by

$$[(Ns)(t)]_j = n_j\{[s(t)]_j\}, \qquad t \geq 0 \tag{23}$$

for $j = 1, 2, \cdots, n$, in which each n_j maps complex numbers z with $|z| < \gamma$ into complex numbers such that $n_j(0) = 0$ and such that $dn_j(z)/dz$ exists for $|z| < \gamma$ and is nonzero at $z = 0$. (Clearly, when C.2 is met, N restricted to Γ can be represented by n single-input single-output memoryless nonlinear operators.)

C.3: $\det\{1_n - \mathrm{diag}[dn_1(0)/dz, \cdots, dn_n(0)/dz] \int_0^\infty c(\tau)e^{-z\tau}d\tau\} \neq 0$ for $\mathrm{Re}(z) \geq 0$.*

A result (see Lemma 3, below) concerning elements of $S_1^{(l)}$ that we shall use is the following:

Proposition 6: If $k_l \in S_1^{(l)}$ for some l, then the iterated integral

$$\int_0^t \cdots \int_0^t k_l(t-\tau_1, t-\tau_2, \cdots, t-\tau_l)\mu(\tau_1)\mu(\tau_2) \cdots \mu(\tau_l)d\tau_1 d\tau_2 \cdots d\tau_l$$

exists for $t \geq 0$ and $\mu \in {}^{(1)}L_\infty(\mathrm{C})$, and $V_{k_l}(\mu)$ defined on $[0, \infty)$ by

$$V_{k_l}(\mu)(t) = \int_0^t \cdots \int_0^t k_l(t-\tau_1, t-\tau_2, \cdots, t-\tau_l)\mu(\tau_1)\mu(\tau_2) \cdots \mu(\tau_l)d\tau_1 d\tau_2 \cdots d\tau_l$$

for an arbitrary $\mu \in {}^{(1)}L_\infty(\mathrm{C})$, is an element of $L_\infty(\mathrm{C})$.

Theorem 7: Suppose that C.1, C.2, and C.3 are met. Then

(i) The hypotheses of Theorem 6 are satisfied.

(ii) For each $l = 1, 2, \cdots$ there is a $k_l \in S_1^{(l)}$ such that, under the condition that $[v(t)]_i = 0$ for $t \geq 0$ and $i > 1$, we have

* As in Section 3.3, 1_n denotes the identity matrix of order n.

$$w = \sum_{l=1}^{\infty} V_{k_l}(v_1) \quad for \quad \|v\| < \delta,$$

with the series uniformly convergent for $\|v\| < \delta$, *where* v, w, *and* δ *are described in Theorem* 6, *and* $V_{k_l}(\cdot)$ *is as indicated in Proposition* 6.

(*iii*) *Each* k_l *can be taken to be continuous* on* $[0, \infty)^l$ *when* a *and* d *are continuous on* $[0, \infty)$.

Proof of Theorem 7: Hypothesis B.1 is clearly met. Also, with N as described in C.2, an easy modification of Proposition 4 shows† that $dN{:}L_\infty(C) \to L_\infty(C)$ exists on Γ, and that for $p \in \Gamma$ and $h \in L_\infty(C)$,

$$[dN(p)h](t) = D(t)h(t), \qquad t \geq 0, \tag{24}$$

where $D(t) = \text{diag}\{dn_1[p_1(t)]/dz, \cdots, dn_n[p_n(t)]/dz\}$.

Similarly, an easy modification of Proposition 5 establishes that $d^mN(p)$ exists for $p \in \Gamma$ and all m. (Observe that, because z is complex, the existence of $dn_j(z)/dz$ for each j and $|z| < \gamma$ means that the derivatives of each n_j of *all orders* exist for $|z| < \gamma$.) Since each $dn_j(z)/dz$ is not zero at $z = 0$, and hence each is nonzero throughout a neighborhood of $z = 0$, it is clear that the $|dn_j(z)/dz|$ are bounded away from zero on some neighborhood of $z = 0$. It follows from (24) that $dN(p)$ is invertible for p in a neighborhood of θ in $L_\infty(C)$. Thus, there are open neighborhoods S_1 and S_2 of θ such that N restricted to S_1, which we denote by N_0, is an invertible map of S_1 onto S_2 and $d^m(N_0^{-1})$ exists throughout S_2 for each $m = 1, 2, \cdots$ (see Lemma 1). Therefore, B.2 and B.3 are satisfied.

Let $\Psi{:}S_2 \to S_1$ denote N_0^{-1}, and notice that Ψ satisfies

$$[(\Psi s)(t)]_j = n_j^{-1}\{[s(t)]_j\}, \qquad t \geq 0$$

for each j and all s in some neighborhood of θ, where n_j^{-1}, defined in a neighborhood of the origin of the complex plane, is a local inverse of n_j. Since each $dn_j(z)/dz$ exists throughout a neighborhood of the origin, and does not vanish at the origin, we see that for each j and m, $d^mn_j^{-1}(z)/dz^m$ exists in a neighborhood of the origin. Therefore, by a direct modification of Proposition 5,

$$[d^m\Psi(\theta)h_1 \cdots h_m(t)]_i = d^mn_j^{-1}(0)/dz^m \prod_{j=1}^{m} h_{ji}(t) \tag{25}$$

for $t \geq 0$, each m and i, and any $h_1, h_2, \cdots, h_m$ in $L_\infty(C)$.

We shall use (25) subsequently. At the moment, concerning Ψ,

---

* Of course, by k_l "continuous" we mean that each component of k_l is continuous.
† Notice that if $dn_j[z(p_j(t) + h_j(t)) + (1 - z)p_j(t)]/dz$ exists at a point $(\alpha, 0)$, then $dn_j[\beta(p_j(t) + h_j(t)) + (1 - \beta)p_j(t)]$, with β a real variable, exists at $\beta = \alpha$ and the values of the two derivatives are the same.

note merely that $[d\Psi(\theta)]^{-1} = dN(\theta)$. Since $[d\Psi(\theta) - C_\infty] = d\Psi(\theta)\{I - [d\Psi(\theta)]^{-1}C_\infty\}$, where I is the identity transformation in $L_\infty(\mathrm{C})$ and C_∞ is defined in B.4, by C.3 and Lemma 2 below, we see that the hypotheses of Theorem 6 are satisfied.

In Lemma 2 we refer to the following.

D.1: $\lambda \in S_n^{(1)}$, Λ denotes the map of $L_\infty(\mathrm{C})$ into itself defined by

$$(\Lambda p)(t) = \int_0^t \lambda(t-\tau)p(\tau)d\tau, \qquad t \geq 0$$

for $p \in L_\infty(\mathrm{C})$, and, with z a scalar complex variable, $\tilde{\Lambda}(z)$ denotes

$$\int_0^\infty \lambda(t)e^{-zt}dt, \qquad \mathrm{Re}(z) \geq 0.$$

Lemma 2: Let D.1 hold, and suppose that $\det\{1_n - \tilde{\Lambda}(z)\} \neq 0$ *for* $\mathrm{Re}(z) \geq 0$. *Then*

(*i*) $(I - \Lambda)$ *is an invertible map of* $L_\infty(\mathrm{C})$ *onto itself,*

(*ii*) *there is a* $\kappa \in S_n^{(1)}$ *such that*

$$(I-\Lambda)^{-1}p(t) = p(t) + \int_0^t \kappa(t-\tau)p(\tau)d\tau, \qquad t \geq 0$$

for $p \in L_\infty(\mathrm{C})$, *and*

(*iii*) *if* λ *is continuous for* $t \geq 0$, *then* κ *can be taken to be continuous on* $[0, \infty)$.*

Lemma 2 is proved in Appendix D.

We also need the following two lemmas which are proved in Appendices E and F.

Lemma 3: Suppose that $h \in S_1^{(l)}$ *for some* $l \geq 1$, *that* $s \in S_n^{(1)}$, *and that* u *is a bounded measurable function from* $[0, \infty)^l$ *into the complex numbers. Then*

(*i*) *With* $\tilde{h}$ *defined on* $(-\infty, \infty)^l$ *by* $\tilde{h} = h$ *on* $[0, \infty)^l$ *and* $\tilde{h} = 0_{n1}$ (*the zero* $n \times 1$ *matrix*) *otherwise, the function* k, *defined by*

$$k(\alpha_1, \alpha_2, \cdots, \alpha_l) = \int_0^\infty s(\tau)\tilde{h}(\alpha_1 - \tau, \cdots, \alpha_l - \tau)d\tau$$

for $(\alpha_1, \alpha_2, \cdots, \alpha_l) \in [0, \infty)^l$, *belongs to* $S_1^{(l)}$.

(*ii*) *If* h *is continuous on* $[0, \infty)^l$, *then so is* k.

(*iii*) *The iterated integrals*

$$\int_0^t \cdots \int_0^t h(t-\tau_1, \cdots, t-\tau_l)u(\tau_1, \cdots, \tau_l)d\tau_1 \cdots d\tau_l$$

* Part (*iii*) is not used in the proof of Theorem 7, and is included because it is useful for other purposes.

and

$$\int_0^t \cdots \int_0^t k(t-\tau_1, \cdots, t-\tau_l)u(\tau_1, \cdots, \tau_l)d\tau_1 \cdots d\tau_l$$

exist, and are invariant with respect to interchanges of orders of integration, for $t \geq 0$; and p defined by

$$p(t) = \int_0^t \cdots \int_0^t h(t-\tau_1, \cdots, t-\tau_l)u(\tau_1, \cdots, \tau_l)d\tau_1 \cdots d\tau_l, \; t \geq 0$$

is an element of $L_\infty(C)$.

(*iv*) *We have*

$$\int_0^t s(t-\tau) \int_0^\tau \cdots \int_0^\tau h(\tau-\tau_1, \cdots, \tau-\tau_l)u(\tau_1, \cdots, \tau_l)d\tau_1 \cdots d\tau_l d\tau$$

$$= \int_0^t \cdots \int_0^t k(t-\tau_1, \cdots, t-\tau_l)u(\tau_1, \cdots, \tau_l)d\tau_1 \cdots d\tau_l, \qquad t \geq 0.$$

Lemma 4: If $h \in S_1^{(p)}$ and $k \in S_1^{(q)}$, then the function s, defined on $[0, \infty)^{p+q}$ by*

$$s(\tau_1, \cdots, \tau_{p+q}) = h(\tau_1, \cdots, \tau_p)k(\tau_{p+1}, \cdots, \tau_{p+q})$$

for $(\tau_1, \cdots, \tau_{p+q}) \in [0, \infty)^{p+q}$, belongs to $S_1^{(p+q)}$.

We now return to the proof of Theorem 7.

With v, w, and δ as in Part (*ii*) of Theorem 6, we have

$$w = Dv + \sum_{m=1}^{\infty} Bg_m(Av), \qquad \|v\| < \delta$$

in which the series converges uniformly, $g_1(Av) = [d\Psi(\theta) - C_\infty]^{-1}Av$, and

$$g_m(Av) = -[d\Psi(\theta) - C_\infty]^{-1} \cdot \sum_{l=2}^{m} (l!)^{-1} \sum_{\substack{k_1+k_2+\cdots+k_l=m \\ k_j>0}} d^l\Psi(\theta) g_{k_1}(Av) \cdots g_{k_l}(Av), \qquad m \geq 2.$$

Assume now that $[v(t)]_i = 0$ for $t \geq 0$ and $i > 1$. By Lemma 3, it suffices to show that for each m there is a $q_m \in S_1^{(m)}$ such that

$$g_m(Av)(t) = \int_0^t \cdots \int_0^t q_m(t-\tau_1, \cdots, t-\tau_m)v_1(\tau_1) \cdots v_1(\tau_m)d\tau_1 \cdots d\tau_m \quad (26)$$

* See the second definition at the beginning of Section 3.5.

for $t \geq 0$ and $\|v\| < \delta$, and that q_m is continuous on $[0, \infty)^m$ when a is continuous on $[0, \infty)$. Thus, suppose for the purpose of induction that (26) holds with q_m as indicated for $m = 1, 2, \cdots, p$ for some $p \geq 1$. Observe that by Lemma 3 and Part (*ii*) of Lemma 2 the induction hypothesis is met for $p = 1$.

We have

$$g_{p+1}(Av) = -[d\Psi(\theta) - C_\infty]^{-1} \cdot \sum_{l=2}^{(p+1)} (l!)^{-1} \sum_{\substack{k_1+k_2+\cdots+k_l=p+1 \\ k_j>0}} d^l\Psi(\theta)g_{k_1}(Av) \cdots g_{k_l}(Av). \quad (27)$$

Now let $l \geq 2$ be fixed, and let $k_1, \cdots, k_l$ be any l positive integers such that $k_1 + k_2 + \cdots + k_l = p + 1$. By Lemma 4, (25), and Lemma 3, r defined by

$$r(\tau_1, \cdots, \tau_{p+1}) = q_{k_1}(\tau_1, \cdots, \tau_{k_1})q_{k_2}(\tau_{k_1+1}, \cdots, \tau_{k_1+k_2}) \cdots q_{k_l}(\tau_{k_1+\cdots+k_{l-1}+1}, \cdots, \tau_{p+1})$$

for $(\tau_1, \cdots, \tau_{p+1}) \in [0, \infty)^{p+1}$ belongs to $S_1^{(p+1)}$, and

$$d^l\Psi(\theta)g_{k_1}(Av) \cdots g_{k_l}(Av)(t) = D\int_0^t \cdots \int_0^t r(t-\tau_1, \cdots, t-\tau_{p+1})v_1(\tau_1) \cdots v_1(\tau_{p+1})d\tau_1 \cdots d\tau_{p+1}$$

for $t \geq 0$, in which D is the diagonal matrix of order n whose jth diagonal element is $d^l n_j^{-1}(0)/dz^l$. Observe that r is continuous on $[0, \infty)^{p+1}$ when each q_{k_j} is continuous on $[0, \infty)^{k_j}$.

Therefore, using (27), and by Lemmas 2 and 3, there is a $q_{p+1} \in S_1^{(p+1)}$, which is continuous on $[0, \infty)^{p+1}$ when a is continuous on $[0, \infty)$, such that (26) holds with m replaced with $(p + 1)$. This completes the proof of the theorem.

Comments: By Lemma 2 and Proposition 7 (in Appendix G), an interpretation of C.3 is simply that the "feedback part" of the graph of Fig. 1, shown in Fig. 2, is bounded-input bounded-output stable* when C.1 and C.2 are met, N is replaced with its linearization $dN(\theta)$ extended in the natural way to all of $H(\mathrm{C})$, and $C{:}H(\mathrm{C}) \rightarrow H(\mathrm{C})$ is defined by $(Cp)(t) = \int_0^t c(t-\tau)p(\tau)d\tau$, $t \geq 0$ for each $p \in H(\mathrm{C})$.

The k_l in Theorem 7 can be taken to be real valued (i.e., to have

* By this we mean that for each $L_\infty(\mathrm{C})$ input to the graph of Fig. 2, there is in $H(\mathrm{C})$ a unique output, and the output belongs to $L_\infty(\mathrm{C})$. In this case, existence and uniqueness of an output in $H(\mathrm{C})$ follows from the hypotheses concerning c via the usual successive approximations approach (Ref. 38, Section 1.13). In Fig. 2, I denotes the identity transformation in $H(\mathrm{C})$.

zero imaginary part) if a, b, c, and d are real valued, and $d^m n_j^{-1}(0)/dz^m$ is real for all m and j. (It is not difficult to show that the $d^m n_j^{-1}(0)/dz^m$ are real when $d^m n_j(0)/dz^m$ is real for every m and j.) In particular, we see that Theorem 7 establishes the existence of a Volterra-series expansion for the important corresponding case in which v_1, w, x, and y in Fig. 1 are restricted to be real valued, C.1 is met, and N (which then would be a map between real-valued function spaces) can be analytically extended so that C.2 and C.3 are satisfied.* † In this connection, Theorem 5 can be used to prove results along the same lines as Theorems 6 and 7, but with $L_\infty(\mathrm{C})$ replaced throughout with $L_\infty(\mathrm{R})$. Similarly, Corollary 1 can be used to obtain corresponding pth order *approximation* results under weaker differentiability hypotheses concerning N.

Theorem 6 provides explicit expressions for the $g_m(Av)$. For example,

$$g_1(Av) = [d\Psi(\theta) - C_\infty]^{-1}Av$$

$$g_2(Av) = -\tfrac{1}{2}[d\Psi(\theta) - C_\infty]^{-1}d^2\Psi(\theta)\{[d\Psi(\theta) - C_\infty]^{-1}Av\}^2$$

$$g_3(Av) = -[d\Psi(\theta) - C_\infty]^{-1}d^2\Psi(\theta)g_1(Av)g_2(Av)$$
$$- \tfrac{1}{6}[d\Psi(\theta) - C_\infty]^{-1}d^3\Psi(\theta)[g_1(Av)]^3,$$

and so on. Therefore, assuming merely that we can write down the representation of $[d\Psi(\theta) - C_\infty]$ along the lines of Part (*ii*) of Lemma 2, notice that it is not difficult in principle to give an explicit expression for any of the k_l of Theorem 7.

Theorem 7 (and Proposition 6) can be extended to cover the case in which the restriction that $[v(t)]_i = 0$ for $t \geq 0$ and $i > 1$ is not met.‡ Specifically, using the results and techniques described in connection with the proof of Theorem 7, it is not difficult to prove the following extension in which for each l, $\chi[v(\tau_1), \cdots, v(\tau_l)]$ denotes the column vector of order n^l whose elements are the n^l distinct products $v_{\omega_1}(\tau_1)v_{\omega_2}(\tau_2) \cdots v_{\omega_l}(\tau_l)$, corresponding to distinct sequences $\omega_1, \omega_2, \cdots, \omega_l$ with each ω_j drawn from $\{1, 2, \cdots, n\}$, arranged in an arbitrary predetermined order.

---

* Observe that this extendability condition is often met. (In particular, polynomial nonlinearities frequently arise in locally-valid models, and polynomials in z are entire functions.)

† Theorem 7 bears on problems concerning the transmission of digital signals over analog communication channels. Discussions with this writer's colleague, J. Salz, concerning such problems provided part of the motivation to formulate the system model in Section 3.4 and to develop a theorem along the lines of Theorem 7.

‡ The case of more than one input is of importance, for example, in connection with studies of the effects of initial conditions. Also, straightforward modifications suffice to establish corresponding results for the *time-discrete* case in which t takes values in $\{0, 1, 2, \cdots\}$.

Theorem 8: Under the hypotheses of Theorem 7, *for every* $l = 1, 2, \cdots$ *there is a* $k_l \in S_{nl}^{(l)}$, *with* k_l *continuous on* $[0, \infty)^l$ *when a and d are continuous on* $[0, \infty)$, *such that*

(*i*) *The iterated integral*

$$\int_0^t \cdots \int_0^t k_l(t-\tau_1, \cdots, t-\tau_l)\chi[v(\tau_1), \cdots, v(\tau_l)]d\tau_1 \cdots d\tau_l$$

exists for $t \geq 0$ *and* $v \in L_\infty(C)$, *and* $V_{k_l}(v)$ *defined on* $[0, \infty)$ *by*

$$V_{k_l}(v)(t) = \int_0^t \cdots \int_0^t k_l(t-\tau_1, \cdots, t-\tau_l)\chi[v(\tau_1), \cdots, v(\tau_l)]d\tau_1 \cdots d\tau_l$$

for any $v \in L_\infty(C)$ *is an element of* $L_\infty(C)$.

(*ii*) *With* $V_{k_l}(\cdot)$ *as indicated in* (*i*) *above, and with* v, w, *and* δ *as described in Theorem* 6, *the expansion*

$$w = \sum_{l=1}^{\infty} V_{k_l}(v)$$

converges uniformly for $\|v\| < \delta$.

It is not difficult to verify that Theorems 7 and 8 remain valid if C.1 is modified to allow a constant multiple (or, more generally, a constant $n \times n$-matrix multiple) of the identity operator on $H(C)$ to be added to B.

We do not give here the details of other extensions of Theorem 7,* but it is worthwhile to appreciate that in some extensions in which C.1 is weakened, series representations can arise in which, unless generalized-function kernels are admitted, the terms do not have the form normally associated with a Volterra series. Consider, for example, that if $n = 1$, if A, as well as B, is the identity transformation on $H(C)$, if C and D have the representations on $L_\infty(C)$ given in C.1, and if C.2 and C.3 are satisfied with $d^2 n_1^{-1}(0)/dz^2 = 2$, then the hypotheses of Theorem 6 are met and the second term in the sum in (21) is a function whose values are

* We leave for another paper results concerning cases in which N can be of a more general form, and the restrictions to $L_\infty(C)$ of A, B, C, and D are not necessarily time invariant, and generalized functions may be involved in their representations. Also left for a later paper are applications to differential equations. Assuming merely that A, B, C, and D are defined on *all* of $H(C)$ as convolution operators with kernels a, b, c, and d, it is a simple corollary of Theorem 8 that there exists a Volterra-series expansion also for the case in which the conditions of Theorem 8 are met, with the exception that the $dn_j(z)/dz$ are not necessarily *nonzero* at $z = 0$. More specifically, N can be replaced with N plus a multiple βI of the identity operator I on $H(C)$, with $|\beta|$ sufficiently small that the four linear parts of the system can be modified accordingly and remain $S_n^{(1)}$ convolutions.

$$-\rho^3 v(t)^2 - 2\rho^2 v(t) \int_0^t k(t-\tau)v(\tau)d\tau + \rho^2 \int_0^t k(t-\tau)v(\tau)^2 d\tau$$

$$+ 2\rho \int_0^t k(t-\tau)v(\tau) \int_0^\tau k(\tau-\tau_1)v(\tau_1)d\tau_1 d\tau$$

$$+ \int_0^t \int_0^t k_a(t-\tau_1, t-\tau_2)v(\tau_1)v(\tau_2)d\tau_1 d\tau_2,$$

in which ρ is a nonzero constant, $k \in S_1^{(1)}$ and $k_a \in S_1^{(2)}$.

APPENDIX A

Proof of Proposition 1

Let σ be real, nonzero, and such that $(r_0 + \sigma) \in A$. Then, using the linearity of $L(r_0 + \sigma)$,

$$\sigma^{-1}[L(r_0+\sigma)e(r_0+\sigma) - L(r_0)e(r_0)]$$

$$= L(r_0+\sigma)\sigma^{-1}[e(r_0+\sigma) - e(r_0)] + \sigma^{-1}[L(r_0+\sigma) - L(r_0)]e(r_0)$$

$$= L(r_0)de(r_0)/dr + [dL(r_0)/dr]e(r_0) + \Delta_1(\sigma) + \Delta_2(\sigma) + \Delta_3(\sigma) + \Delta_4(\sigma),$$

where

$$\Delta_1(\sigma) = L(r_0)\{\sigma^{-1}[e(r_0+\sigma) - e(r_0)] - de(r_0)/dr]\}$$

$$\Delta_2(\sigma) = [L(r_0+\sigma) - L(r_0)]\{\sigma^{-1}[e(r_0+\sigma) - e(r_0)] - de(r_0)/dr\}$$

$$\Delta_3(\sigma) = [L(r_0+\sigma) - L(r_0)]de(r_0)/dr$$

$$\Delta_4(\sigma) = \{\sigma^{-1}[L(r_0+\sigma) - L(r_0)] - dL(r_0)/dr\}e(r_0).$$

Since $L(r_0)$ is a *bounded* operator and $\|\sigma^{-1}[e(r_0+\sigma) - e(r_0)] - de(r_0)/dr\| \to 0$ as $\sigma \to 0$, $\Delta_1(\sigma) \to 0$ in S as $\sigma \to 0$. It is clear that $\Delta_2(\sigma)$, $\Delta_3(\sigma)$, and $\Delta_4(\sigma)$ approach zero in S as $\sigma \to 0$. This completes the proof.

APPENDIX B

Part of the Proof of Theorem 1

Let k be any integer such that $1 \le k \le p$, and let $v_1(\cdot)$, $v_2(\cdot)$, $\cdots$, $v_k(\cdot)$ denote k maps from an open subset of $(-\infty, \infty)$ containing $[0, 1]$ into $\mathscr{B}_0$ such that each $v_j(\cdot)$ is differentiable on $[0, 1]$. With l an integer such that $0 \le l \le k-1$, consider $d^k f[q(\beta)]v_1(\beta) \cdots v_{k-l-1}(\beta)v_{k-l}(\beta)$. Since $d^k f[q(\beta)]$ for $0 \le \beta \le 1$ is a Fréchet derivative, $d^k f[q(\beta)]v_1(\beta) \cdots v_{k-l-1}(\beta)$ is a bounded linear map from $\mathscr{B}_0$ into a Banach space S for each $\beta \in [0, 1]$. By the version of the chain rule in Ref. 31, p. 173, $d\{d^k f[q(\beta)]\}/d\beta$ exists for $\beta \in [0, 1]$, and

$$d\{d^k f[q(\beta)]\}/d\beta = d^{k+1} f[q(\beta)]q^{(1)}(\beta), \qquad \beta \in [0, 1]. \tag{28}$$

By Proposition 1, (28), and an obvious inductive argument, $d\{d^k f[q(\beta)]v_1(\beta) \cdots v_{k-l-1}(\beta)\}/d\beta$ exists for $\beta \in [0, 1]$. Thus, by Proposition 1, $d\{d^k f[q(\beta)]v_1(\beta) \cdots v_{k-l}(\beta)\}/d\beta$ exists and satisfies

$$\begin{aligned} d\{d^k f[q(\beta)]v_1(\beta) \cdots v_{k-l}(\beta)\}/d\beta \\ = d^k f[q(\beta)]v_1(\beta) \cdots v_{k-l-1}(\beta)[dv_{k-l}(\beta)/d\beta] \\ + d\{d^k f[q(\beta)]v_1(\beta) \cdots v_{k-l-1}(\beta)\}/d\beta \; v_{k-l}(\beta) \end{aligned} \tag{29}$$

for $\beta \in [0, 1]$.

By relations (28) and (29), and the fact that $df[q(\beta)]/d\beta = df[q(\beta)]q^{(1)}(\beta)$ for $0 \le \beta \le 1$, we see that an expression $E_m(\beta)$ can be given for $d^m f[q(\beta)]/d\beta^m$ for $\beta \in [0, 1]$ which depends only on $d^l f[q(\beta)]$ and $q^{(l)}(\beta)$ for $l = 1, 2, \cdots, m$. For example, $E_2(\beta) = df[q(\beta)]q^{(2)}(\beta) + d^2 f[q(\beta)]q^{(1)}(\beta)q^{(1)}(\beta)$. Since $df[q(\beta)]/d\beta = (u - u_0)$ for $\beta \in [0, 1]$, we see that $E_m(\beta) = \theta$ for $\beta \in [0, 1]$.

Now consider (5) and (6). Since $d^l f_0[q_0(r)]$ and $d^l q_0(r)/dr^l$ exist at $r = \beta$, with $d^l f_0[q_0(\beta)] = d^l f[q(\beta)]$ and $d^l q_0(r)/dr^l = q^{(l)}(\beta)$ for $r = \beta$ and $l = 1, 2, \cdots, m$, by Proposition 1 and observations similar to those of the preceding three paragraphs, we see that, as claimed in Section 2.2, $d^m f_0[q(r)]/dr^m|_{r=\beta}$ exists and that it equals $E_m(\beta)$.

APPENDIX C

Proof of Proposition 2

Under the conditions indicated, $dg(u_0)$ exists (and equals L). Thus, using $f[g(u)] = u$ for $u \in U$, and with I the identity operator on $\mathscr{B}$, we have $df[g(u_0)]dg(u_0) = I$. This shows that $df[g(u_0)]$ has a *right* inverse.

On the other hand, for $u \in U$ with $\|u - u_0\| < \sigma$, $g(u) - g(u_0) = Lf[g(u)] - Lf[g(u_0)] + R\{f[g(u)] - f[g(u_0)]\}$ and, thus,

$$\begin{aligned} \{I_0 - Ldf[g(u_0)]\}[g(u) - g(u_0)] = R\{df[g(u_0)][g(u) - g(u_0)] \\ + R_1[g(u) - g(u_0)]\} + R_2[g(u) - g(u_0)] \end{aligned} \tag{30}$$

in which I_0 is the identity operator on $\mathscr{B}_0$, $R_1(h) = o(\|h\|)$ as $\|h\| \to 0$, and, using the boundedness of L, $R_2(h) = o(\|h\|)$ as $\|h\| \to 0$.

Now let $v \in \mathscr{B}$ be arbitrary. By the continuity of f at $g(u_0)$, and the openness of X and U, choose $\beta > 0$ so that $g(u_0) + \alpha v \in X$, $\|f[g(u_0) + \alpha v] - f[g(u_0)]\| < \sigma$, and $f[g(u_0) + \alpha v] \in U$ for $|\alpha| < \beta$. Notice that for each α with $|\alpha| < \beta$, and with $u = f[g(u_0) + \alpha v]$, we have $\|u - u_0\| < \sigma$, as well as $f[g(u)] = f[g(u_0) + \alpha v]$, and hence* $g(u) = g(u_0) + \alpha v$.

Therefore, by (30),

$$\{I_0 - Ldf[g(u_0)]\}v = \alpha^{-1}R\{df[g(u_0)]\alpha v + R_1(\alpha v)\} + \alpha^{-1}R_1(\alpha v) \quad (31)$$

for $0 < |\alpha| < \beta$. Since the right side of (31) approaches θ as $\alpha \to 0$, and since v is arbitrary, it follows that $df[g(u_0)]$ has a *left* inverse. Since $df[g(u_0)]$ has both a left inverse and a right inverse, $df[g(u_0)]^{-1}$ exists. Finally, a standard type of argument shows that $(df)^{-1}$ exists and is continuous throughout some neighborhood of $g(u_0)$ when f is continuously differentiable on a neighborhood of $g(u_0)$ (see the proof of Corollary 1 and the references given there).

APPENDIX D

Proof of Lemma 2

In the following, we use L_1 to denote the set of complex-valued functions summable over $[0, \infty)$.

For $\mathrm{Re}(z) \geq 0$, we have $[1_n - \tilde{\Lambda}(z)]^{-1} = M(z)\{\det[1_n - \tilde{\Lambda}(z)]\}^{-1}$, in which M is the matrix of transposed cofactors of $[1_n - \tilde{\Lambda}(z)]$ if $n > 1$ and $M = 1$ if $n = 1$. Since $\lambda_{ij} \in L_1$ for each i and j, and the convolution of any two bounded L_1 functions belongs to L_1 and is bounded, it follows that there is a $q \in L_1$, and an $r \in S_n^{(1)}$, such that

$$\det[1_n - \tilde{\Lambda}(z)] = 1 - \int_0^\infty q(t)e^{-zt}dt$$

$$M(z)\tilde{\Lambda}(z) = \int_0^\infty r(t)e^{-zt}dt$$

for $\mathrm{Re}(z) \geq 0$.

Since $\det[1_n - \tilde{\Lambda}(z)] \neq 0$ for $\mathrm{Re}(z) \geq 0$, it follows (see Ref. 42, pp. 60–63) that there is an element s of L_1 such that

$$\{\det[1_n - \tilde{\Lambda}(z)]\}^{-1} = 1 + \int_0^\infty s(\tau)e^{-z\tau}d\tau, \qquad \mathrm{Re}(z) \geq 0.^\dagger$$

Thus, κ defined by

$$\kappa(t) = r(t) + \int_0^t r(t-\tau)s(\tau)d\tau, \qquad t \geq 0$$

---

* Here we use the hypothesis that for each $u \in U$, there is exactly one $x \in X$ such that $f(x) = u$.

† Notice that when $\tilde{\Lambda}(z)$ is *rational* in z, it is a simple matter to show the existence of such an s.

belongs to $S_n^{(1)}$, and $H(z)$ given by

$$H(z) = \int_0^\infty \kappa(\tau)e^{-z\tau}d\tau, \qquad \text{Re}(z) \geq 0$$

satisfies $H(z) = [1_n - \tilde{\Lambda}(z)]^{-1}\tilde{\Lambda}(z)$ for all $\text{Re}(z) \geq 0$.

Since $[1_n - \tilde{\Lambda}(z)]^{-1}\tilde{\Lambda}(z) = \tilde{\Lambda}(z) + \tilde{\Lambda}(z)[1_n - \tilde{\Lambda}(z)]^{-1}\tilde{\Lambda}(z)$ for $\text{Re}(z) \geq 0$, we have

$$\kappa(t) = \lambda(t) + \int_0^t \lambda(\tau)\kappa(t-\tau)d\tau \tag{32}$$

for almost every $t \geq 0$.*

For an arbitrary $p \in L_\infty(\text{C})$, let $q \in L_\infty(\text{C})$ be defined by

$$q(t) = p(t) + \int_0^t \kappa(t-\tau)p(\tau)d\tau, \qquad t \geq 0. \tag{33}$$

We have for $t \geq 0$,

$$q(t) - (\Lambda q)(t) = p(t) + \int_0^t \kappa(t-\tau)p(\tau)d\tau - \int_0^t \lambda(t-\tau)p(\tau)d\tau - \int_0^t \lambda(t-\tau_1)\int_0^{\tau_1} \kappa(\tau_1-\tau_2)p(\tau_2)d\tau_2 d\tau_1.$$

Since

$$\int_0^t \lambda(t-\tau_1)\int_0^{\tau_1} \kappa(\tau_1-\tau_2)p(\tau_2)d\tau_2 d\tau_1$$

can be expressed as

$$\int_0^t \int_0^{t-\beta} \lambda(\alpha)\kappa(t-\beta-\alpha)p(\beta)d\alpha d\beta$$

for $t \geq 0$ [the justification of the interchange of order of integration being provided by Theorems of Fubini and Tonelli (Ref. 43, pp. 137–45)], we have, using (32), $(I - \Lambda)q = p$. Thus, $(I - \Lambda)$ maps $L_\infty(\text{C})$ onto itself. Similarly, (32) holds with λ and κ interchanged in the integral, and (32) so modified can be used to show (see the proof of Theorem I of Ref. 39) that whenever there is a solution $q \in L_\infty(\text{C})$ of $(I - \Lambda)q = p$ with $p \in L_\infty(\text{C})$, then (33) holds. This establishes Parts (*i*) and (*ii*) of the lemma.

Suppose now that λ is continuous on $[0, \infty)$. Since $\kappa \in S_n^{(1)}$, by Part

* Equation (32) is a matrix-valued-function version of the usual equation for the resolvent kernel.

(*ii*) of Lemma 3, the integral in (32) depends continuously on t for $t \geq 0$. Thus, by (32), the values of κ agree almost everywhere on $[0, \infty)$ with those of a continuous function. This completes the proof of the lemma.

APPENDIX E

Proof of Lemma 3

Consider k. That each k_{i1} is measurable on $[0, \infty)^l$ and satisfies

$$\int_{[0,\infty)^l} |k_{i1}(\tau_1, \cdots, \tau_l)|\, d(\tau_1, \cdots, \tau_l) < \infty$$

follows from a direct application of Theorems of Fubini and Tonelli (Ref. 43, pp. 137–45). (See the proof in Ref. 44, pp. 99–100, for the $l = 1$ case.) Since every s_{ij} is summable over $[0, \infty)$, and h is bounded, we see that k is bounded. Thus, (*i*) holds.

Suppose now that h is continuous on $[0, \infty)^l$. Let $\alpha_m \geq 0$ be given for $m = 1, 2, \cdots, l$, let δ_m for $m = 1, 2, \cdots, l$ be real variables such that each $(\alpha_m + \delta_m)$ is nonnegative, and let Δ be defined by

$$\Delta(\alpha + \delta) = \int_0^\infty s_{ij}(\tau)\tilde{h}_{j1}(\alpha_1 + \delta_1 - \tau, \cdots, \alpha_l + \delta_l - \tau)\, d\tau$$

for any i and j. Let $\gamma > 0$ be given. With b_1 and b_2 such that $|\tilde{h}_{j1}(\tau_1, \cdots, \tau_l)| \leq b_1$ and $|s_{ij}(\tau)| \leq b_2$ for $(\tau_1, \cdots, \tau_l) \in [0, \infty)^l$ and $\tau \in [0, \infty)$, choose $\tau_0 \in (0, \infty)$ so that

$$2b_1 \int_{\tau_0}^\infty |s_{ij}(\tau)|\, d\tau \leq \frac{1}{2}\gamma,$$

and observe that

$$|\Delta(\alpha + \delta) - \Delta(\alpha)| \leq \frac{1}{2}\gamma + b_2 \int_0^{\tau_0} |\tilde{h}_{j1}(\alpha_1 + \delta_1 - \tau, \cdots, \alpha_l + \delta_l - \tau) - \tilde{h}_{j1}(\alpha_1 - \tau, \cdots, \alpha_l - \tau)|\, d\tau. \quad (34)$$

Since γ is arbitrary, and, by the boundedness and uniform continuity of h on compact subsets of $[0, \infty)^l$, the value of the integral in (34) can be made arbitrarily small by choosing

$$\sum_{m=1}^{l} |\delta_m|$$

to be sufficiently small, we see that k is continuous on $[0, \infty)^l$, which proves (*ii*).

Since $h(t-\tau_1, \cdots, t-\tau_l)u(\tau_1, \cdots, \tau_l)$ and $k(t-\tau_1, \cdots, t-\tau_l)(\tau_1, \cdots, \tau_l)$ are bounded and measurable on $(\tau_1, \cdots, \tau_l) \in [0, t]^l$, the multiple integrals

$$\int_{[0,t]^l} h(t-\tau_1, \cdots, t-\tau_l)u(\tau_1, \cdots, \tau_l)d(\tau_1, \cdots, \tau_l)$$

and

$$\int_{[0,t]^l} k(t-\tau_1, \cdots, t-\tau_l)u(\tau_1, \cdots, \tau_l)d(\tau_1, \cdots, \tau_l)$$

exist. Therefore, two repeated [i.e., two $(l-1)$-fold] applications of Fubini's theorem (Ref. 43, p. 137) show that the iterated integrals in (*iii*) exist, that each equals the corresponding multiple integral, and that each is invariant under changes in the order of integration. Notice that the existence of

$$\int_0^t \left[\int_0^t \cdots \int_0^t h(t-\tau_1, \cdots, t-\tau_l)u(\tau_1, \cdots, \tau_l)d\tau_1 \cdots d\tau_l\right]d\tau \qquad (35)$$

for any $t > 0$ can be established in essentially the same way.

Now let p be defined on $[0, \infty)$ by

$$p(t) = \int_0^t \cdots \int_0^t h(t-\tau_1, \cdots, t-\tau_l) \cdot u(\tau_1, \cdots, \tau_l)d\tau_1 \cdots d\tau_l, \qquad t \geq 0. \qquad (36)$$

Since $h \in S_1^{(l)}$, and u is bounded on $[0, \infty)^l$, it is clear from the relationship between the iterated integral in (36) and the corresponding multiple integral that p is bounded on $[0, \infty)$. That p is *measurable* on $[0, \infty)$, is a consequence of the existence of (35) for all $t > 0$.

Similarly, again using Fubini's theorem and the fact that a bounded measurable function defined on a set E of finite measure is summable over E, we have, for any $t \geq 0$,

$$\int_0^t s(t-\tau) \int_0^\tau \cdots \int_0^\tau h(t-\tau_1, \cdots, t-\tau_l)u(\tau_1, \cdots, \tau_l)d\tau_1 \cdots d\tau_l d\tau$$

$$= \int_0^t s(\tau) \int_0^{t-\tau} \cdots \int_0^{t-\tau} h(t-\tau-\tau_1, \cdots, t-\tau-\tau_l) \cdot u(\tau_1, \cdots, \tau_l)d\tau_1 \cdots d\tau_l d\tau$$

$$= \int_0^t s(\tau) \int_{[0,t]^l} \tilde{h}(t-\tau-\tau_1, \cdots, t-\tau-\tau_l)$$

$$\cdot u(\tau_1, \cdots, \tau_l) d\tau_1 \cdots d\tau_l d\tau$$

$$= \int_{[0,t]^l} \int_0^t s(\tau) \tilde{h}(t - \tau - \tau_1, \cdots, t - \tau - \tau_l) d\tau$$

$$\cdot u(\tau_1, \cdots, \tau_l) d(\tau_1, \cdots, \tau_l)$$

$$= \int_{[0,t]^l} k(t - \tau_1, \cdots, t - \tau_l) u(\tau_1, \cdots, \tau_l) d(\tau_1, \cdots, \tau_l)$$

$$= \int_0^t \cdots \int_0^t k(t - \tau_1, \cdots, t - \tau_l) u(\tau_1, \cdots, \tau_l) d\tau_1 \cdots d\tau_l,$$

showing that (iv) is met. This completes the proof of the lemma.

APPENDIX F

Proof of Lemma 4

By Fubini's theorem (Ref. 43, p. 137), and the proposition that bounded measurable functions on a set E of finite measure are summable on E,

$$\int_{[0,T]^{(p+q)}} |h_{i1}(\tau_1, \cdots, \tau_p) k_{i1}(\tau_{p+1}, \cdots, \tau_{p+q})| d(\tau_1, \cdots, \tau_{p+q})$$

$$\leq \int_{[0,\infty)^p} |h_{i1}(\tau_1, \cdots, \tau_p)| d(\tau_1, \cdots, \tau_p)$$

$$\cdot \int_{[0,\infty)^q} |k_{i1}(\tau_{p+1}, \cdots, \tau_{p+q})| d(\tau_{p+1}, \cdots, \tau_{p+q})$$

for each i and any finite $T > 0$, from which it is clear that the lemma holds.

APPENDIX G

On the Necessity of the Condition That Det$[1_n - \tilde{\Lambda}(z)] \neq 0$ for Re$(z) \geq 0$

Proposition 7: Let D.1 (which appears just before Lemma 2) hold. If for each $q \in L_\infty(\mathrm{C})$ there is a $p \in L_\infty(\mathrm{C})$ such that $(I - \Lambda)p = q$, then $\det[1_n - \tilde{\Lambda}(z)] \neq 0$ for $\mathrm{Re}(z) \geq 0$.

Proof:

Since $\lambda \in S_n^{(1)}$, by a standard successive approximations approach (Ref. 38, Section 1.13), it can be shown that there is an $n \times n$ matrix-valued function κ defined on $[0, \infty)$ such that each κ_{ij} is square summable on any finite interval $[0, \beta]$, and

$$\kappa(t) = \lambda(t) + \int_0^t \lambda(\tau)\kappa(t - \tau) d\tau, \qquad t \geq 0 \tag{37}$$

(i.e., and such that (32) is met for $t \geq 0$). From (37) and the Schwarz inequality, the κ_{ij} are bounded on finite intervals. Using Fubini's theorem (see the proof of Theorem I of Ref. 39), if $(I - \Lambda)p = q$ with q and p in $L_\infty(\mathrm{C})$, then

$$p(t) = q(t) + \int_0^t \kappa(t-\tau)q(\tau)d\tau, \qquad t \geq 0.$$

Thus, by the hypothesis of the proposition, each κ_{ij} is summable on $[0, \infty)$. In particular, the Laplace transform $\tilde{H}(z)$ of κ, given by

$$\tilde{H}(z) = \int_0^\infty \kappa(t)e^{-zt}dt$$

exists for all $\mathrm{Re}(z) \geq 0$.

We have $(I - \Lambda)(I + H)q = q$ for every $q \in L_\infty(\mathrm{C})$, in which H is the convolution transformation defined in $L_\infty(\mathrm{C})$ in the usual way in terms of κ. Now let q be given by $q(t) = e^{-t}c_i$ for $t \geq 0$, in which c_i is the column n-vector whose ith component is unity and all other components are zero. Upon taking the Laplace transform of both sides of $(I - \Lambda)(I + H)q = q$, we find that $[1_n - \tilde{\Lambda}(z)][1_n + \tilde{H}(z)]c_i = c_i$ for $\mathrm{Re}(z) \geq 0$ and each i. Therefore, $[1_n - \tilde{\Lambda}(z)][1_n + \tilde{H}(z)] = 1_n$ for $\mathrm{Re}(z) \geq 0$, which shows that $\det[1_n - \tilde{\Lambda}(z)] \neq 0$ for $\mathrm{Re}(z) \geq 0$.

REFERENCES

1. V. Volterra, "Sopra le Funzioni che Dipendono da Altre Funzioni, Nota 1," Rend. Lincei Ser. 4, 3 (1887), pp. 97–105.
2. V. Volterra, *The Theory of Functionals and of Integral and Integro-differential Equations*, New York: Dover, 1959.
3. M. Fréchet, "Sur les Fonctionnelles Continues," Ann. de L'Ecole Normale sup., 3rd Series, *27* (1910).
4. N. Wiener, "Response of a Non-linear Device to Noise," Massachusetts Institute of Technology Radiation Laboratory Report 129 (April 1942).
5. N. Wiener, *Nonlinear Problems in Random Theory*, Cambridge, Mass.: M.I.T. Press, 1958.
6. M. Schetzen, *The Volterra and Wiener Theories of Nonlinear Systems*, New York: John Wiley, 1980.
7. L. A. Zadeh, "On the Representation of Nonlinear Operators," IRE Wescon Convention Record, Part 2 (1957), pp. 105–13.
8. E. Bedrosian and S. O. Rice, "The Output Properties of Volterra Systems Driven by Harmonic and Gaussian Inputs," Proc. IEEE, *59* (December 1971), pp. 1688–707.
9. M. B. Brilliant, "Theory of the Analysis of Nonlinear Systems," Massachusetts Institute of Technology Research Laboratory of Electronics, Report No. 345 (1958).
10. P. d'Allessandro, A. Isidori, and A. Ruberti, "Realizations and Structure Theory of Bilinear Dynamical Systems," SIAM J. Control, *12*, No. 3 (August 1974), pp. 517–35.
11. S. B. Karmaker, "Approximate Analysis of Non-Linear Systems by Laplace Transform," J. Sound and Vibration, *69*, No. 4 (April 1980), pp. 597–602.
12. J. Katzenelson and L. A. Gould, "The Design of Nonlinear Filters and Control Systems, Part II," Information and Control, *7* (June 1964), pp. 117–45.
13. E. Weber, "Complex Convolution Applied to Non-linear Problems," Proc. Symp. Non-linear Circuit Analysis, Polytechnic Inst. of Brooklyn, *VI* (1956), pp. 409–27.
14. B. J. Leon and D. J. Schaefer, "Volterra Series and Picard Iteration for Nonlinear

Circuits and Systems," IEEE Trans. Circuits and Systems, *25* No. 9 (September 1978), pp. 789–93.
15. C. Bruni, G. DiPillo, and G. Koch, "On the Mathematical Models of Bilinear Systems," Ricerche Di Automatica, *2*(*i*), (1971).
16. R. W. Brockett, "Volterra Series and Geometric Control Theory," Automatica *12* (1976), pp. 167–76. See also R. W. Brockett and E. G. Gilbert, "An Addendum to Volterra Series and Geometric Control Theory," Automatica, *12* (1976), p. 635.
17. R. B. Parente, "Nonlinear Differential Equations and Analytic Systems Theory," SIAM J. Appl. Math., *18* (January 1970), pp. 41–66.
18. H. L. Van Trees, "Functional Techniques for the Analysis of the Nonlinear Behavior of Phase-locked Loops," Proc. IEEE, *52* (August 1964), pp. 894–911.
19. S. Narayanan, "Application of Volterra Series to Intermodulation Distortion Analysis of a Transistor Feedback Amplifier," IEEE Trans. Circuit Theory, *17* (November 1970), pp. 518–27.
20. E. G. Gilbert, "Functional Expansions for the Response of Nonlinear Differential Systems," IEEE Trans. Automatic Control, *22* (December 1977), pp. 909–21.
21. L. O. Chua and C. Y. Ng, "Frequency Domain Analysis of Nonlinear Systems: Formulation of Transfer Functions," Elect. Circuits and Systems, *3* No. 6 (November 1979), pp. 257–69.
22. S. Narayanan, "Transistor Distortion Analysis using Volterra Series Representations," B.S.T.J., *46* No. 5 (May–June 1967), pp. 991–1023.
23. C. Lesiak and A. J. Krener, "The Existence and Uniqueness of Volterra Series for Nonlinear Systems," IEEE Trans. Automatic Control, *23* No. 6 (December 1978), pp. 1090–5.
24. A. V. Balakrishnan, *Applied Functional Analysis*, New York: Springer Verlag, 1976, pp. 158–61.
25. M. Fliess, "A Note on Volterra Series for Nonlinear Differential Systems," IEEE Trans. Automatic Control, *25* (February 1980), pp. 116–7.
26. W. A. Porter, "An Overview of Polynomic Systems Theory," Proc. IEEE, *64* (January 1976), pp. 18–23.
27. A. Halme, J. Orava, and H. Blomberg, "Polynomial Operators in Nonlinear System Theory," Int. J. Syst. Sci., *2* No. 1 (July 1971), pp. 25–47.
28. P. A. Prenter, "On Polynomial Operators and Equations," *Nonlinear Functional Analysis and Applications*, L. B. Rall, ed., New York: Academic Press (1971).
29. J. Orava and A. Halme, "Inversion of Generalized Power Series Representation," J. of Mathematical Analysis and Applications, *45* No. 1 (January 1974), pp. 136–41.
30. J. Dieudonné, *Foundations of Modern Analysis*, New York: Academic Press, 1969.
31. T. M. Flett, *Differential Analysis*, London: Cambridge University Press, 1980.
32. I. J. Maddox, *Elements of Functional Analysis*, London: Cambridge University Press, 1977, pp. 118–9.
33. L. M. Graves, "Reimann Integration and Taylor's Theorem in General Analysis," Trans. Amer. Mathematical Soc., *29* (January 1927), pp. 163–77.
34. E. Hille and R. S. Phillips, *Functional Analysis and Semi-groups*, Providence: Amer. Mathematical Soc., 1957.
35. F. Riesz and B. Sz.-Nagy, *Functional Analysis*, New York: Frederick Ungar Publishing Co., 1955.
36. I. W. Sandberg, "Some Results on the Theory of Physical Systems Governed by Nonlinear Functional Equations," B.S.T.J., *44* No. 5 (May 1965), pp. 871–98.
37. A. Halme, "Polynomial Operators for Nonlinear Systems Analysis," Acta Polytech. Scand. Ma, *24* (1972), pp. 1–63. (Dissertation).
38. F. G. Tricomi, *Integral Equations*, New York: Interscience, 1957.
39. I. W. Sandberg, "Signal Distortion in Nonlinear Feedback Systems," B.S.T.J., *42* No. 6 (November 1963), pp. 2533–49.
40. I. W. Sandberg, "On the Theory of Linear Multi-loop Feedback Systems," *Feedback Systems*, J. B. Cruz, Jr., ed., New York: McGraw-Hill, 1972.
41. J. G. Truxal, *Automatic Feedback Control System Synthesis*, New York: McGraw-Hill, 1955, p. 114.
42. R. E. Paley and N. Wiener, *Fourier Transforms in the Complex Domain*, Providence: Amer. Mathematical Soc., 1934.
43. S. McShane, *Integration*, Princeton: Princeton Univ. Press, 1944.
44. S. Bochner and K. Chandrasekharan, *Fourier Transforms*, Princeton: Princeton Univ. Press, 1949.