Functional Expansions for the Response of Nonlinear Differential Systems

ELMER G. GILBERT

Abstract—**This paper is concerned with representing the response of nonlinear differential systems by functional expansions. An abstract theory of variational expansions, similar to that of L. M. Graves (1927), is developed. It leads directly to concrete expressions (multilinear integral operators) for the functionals of the expansions and sets conditions on the differential systems which insure that the expansions give reasonable approximations of the response. Similarly, it is shown that the theory of analytic functions in Banach spaces leads directly to conditions which imply uniform convergence of functional series. The main results on differential systems are summarized in a set of theorems, some of which overlap and extend the recent results of Brockett on Volterra series representations for the response of linear analytic differential systems. Other theorems apply to more general nonlinear differential systems. They provide a rigorous foundation for a large body of previous research on Volterra series expansions. The multilinear integral operators are obtained from systems of differential equations which characterize exactly the variations. These equations are of much lower order than those obtained by the technique of Carleman. A nonlinear feedback system serves as an example of an application of the theory.**

I. Introduction

THE USE of functional expansions to represent the response of dynamic systems is a well-established concept, dating back to 1942 when N. Wiener characterized the response of a nonlinear device by a Volterra series. Since then functional expansions (usually Volterra series) have played an important role in the modeling of nonlinear systems, both when the underlying system equations are known and when the system is characterized only by the availability of input–output data. This paper is concerned with the former situation. The dynamic system is described by a system of nonlinear differential equations and the objective is to obtain a local approximation of the system output by a functional expansion operating on the input. Usually, although not always, the expansion is a truncated power series.

There is a sizeable literature of prior research in this direction. In the 1950's and 1960's, Volterra series were derived and exploited in a variety of situations. References [3], [4], [10], [11], [18], [21]–[23], and [26] give a good, although by no means complete, perspective of this work and include additional references to the literature. The main concern was with relatively simple, stationary differential systems. Multidimensional Laplace (or Fourier) transforms appeared frequently in both system characterization and response evaluation. For the most part, little attention was given to the validity of the Volterra series other than to assume without justification its uniform convergence. In this respect the paper by Bruni, DiPillo, and Koch [7] was a marked advance. For a general bilinear differential system it showed that the input–output map was a uniformly convergent Volterra series. By a technique due to Carleman, Krener [17] has shown that a rather general class of nonlinear differential systems can be approximated by bilinear differential systems. This provides a path for extending, in a variety of ways, the results in [7] to general nonlinear systems. An example is the recent paper [5] by Brockett on "linear analytic" systems.

A different approach to a general theory is pursued here and in [12]. The differential system is viewed abstractly as a mapping P from a Banach space of inputs into a Banach space of outputs. Then P is expanded in an appropriate abstract series and the series is interpreted concretely to obtain the desired functional expansion. Balakrishnan [2, sect. 3.6] takes a similar point of view in considering a bilinear differential system, although the details of his analysis are quite different from those which follow.

The most obvious candidate for the abstract series is a Frechet power series. The precise details may be found in [9] and, stated briefly, are as follows. Let $P^{(i)}(v_0)[w_1][w_2]\cdots[w_i]$ denote the ith Frechet differential of P at v_0 with increments $w_1,\cdots,w_i$. If P is k times continuously differentiable in a neighborhood of v_0,

$$P(v_0+v)=P(v_0)+\sum_{i=1}^{k}\frac{1}{i!}P^{(i)}(v_0)[v][v]\cdots[v]+R_k(v) \quad (1.1)$$

where $\|R_k(v)\|<\epsilon\|v\|^k$ if $\|v\|<\delta(\epsilon)$. While (1.1) is appealing and establishes connections with current research in polynomic systems theory (see [24] for a review), it is not obvious that the differentials exist and are continuous or that they can be determined easily from the description of the differential system. At least part of the reason for this difficulty is that $P^{(i)}(v_0)[w_1][w_2]\cdots[w_i]$ contains much more information than is needed, since in (1.1), $w_1,\cdots,w_i$ are all set equal to v. The "essential information" is contained in the ith variation of P at v_0, which is defined by

Manuscript received August 25, 1976; revised August 17, 1977. Paper recommended by G. L. Blankenship, Past Chairman of the Stability, Nonlinear, and Distributed Systems Committee. This work was supported in part by the Air Force Office of Scientific Research, Air Force Systems Command, USAF, under Grant AFOSR-77-3158.

The author is with the Department of Aerospace Engineering and the Program in Computer, Information, and Control Engineering, University of Michigan, Ann Arbor, MI 48104.

$$\delta^i P_{v_0}(v) = \left(\frac{d}{d\alpha}\right)^i P(v_0 + \alpha v)\bigg|_{\alpha=0}. \tag{1.2}$$

In particular,

$$P(v_0 + v) = P(v_0) + \sum_{i=1}^{k} \frac{1}{i!}\delta^i P_{v_0}(v) + R_k(v). \tag{1.3}$$

Because $\delta^i P_{v_0}(v)$ is obtained by examining P on a one-dimensional subspace, it can be determined for differential systems by examining the solution of a differential equation with a parameter. As will be seen, this leads immediately to concrete characterizations of $\delta^i P_{v_0}(v)$ and to conditions on the differential systems which assure that $\|R_k(v)\|$ is bounded in a reasonable fashion. The expansion (1.3) was first considered by Graves [14] in 1927 and seems to have been neglected in recent years, except as it pertains to the theory of analytic functions in complex Banach spaces [15].

The plan and content of the paper are now summarized. In Section II the theory of (1.3) is developed for fixed k. It is pointed out that (1.3) is a sum of homogeneous functions and not necessarily powers of v. Conditions are given which guarantee that $\|R_k(v)\| < \epsilon\|v\|^k$ when $\|v\| < \delta(\epsilon)$ or $\|R_k(v)\| < \mu\|v\|^{k+1}$ when $\|v\| < \rho$. Complete proofs are given because they are simple and informative and take little more space than proofs based on the results of Graves [14]. The theory of analytic functions in complex Banach spaces is reviewed in Section III. The results are taken directly from Hille and Phillips [15] and lead to simple conditions which imply uniform convergence of (1.3) as $k \to \infty$. In Section IV the methods of Sections II and III are applied to the differential system:

$$\dot{x}(t) = f(x(t),t) + v(t)g(x(t),t), \qquad x(0) = \xi \tag{1.4}$$

$$y(t) = h(x(t),t). \tag{1.5}$$

Here, $x(t) \in \mathfrak{R}^n$, $v(t) \in \mathfrak{R}$, $y(t) \in \mathfrak{R}$, and $t \in [0,T]$. The terms in (1.3) are given exactly by the solution of differential equations whose order is much lower than those arising from the technique of Carleman. These differential equations lead directly to the characterization of (1.3) as a truncated Volterra series. Precise conditions on f,g,h which assure the validity of (1.3) are given in Theorems 4.1–4.4. Brockett's result [5], [6] on the uniform convergence of the Volterra series is included. Similar results are obtained in Section V for the more general differential system:

$$\dot{x}(t) = f(x(t),u(t),t), \qquad x(0) = \xi \tag{1.6}$$

$$y(t) = h(x(t),u(t),t) \tag{1.7}$$

where $x(t) \in \mathfrak{R}^n$, $u(t) \in \mathfrak{R}^m$, $y(t) \in \mathfrak{R}^l$, and $t \in [0,T]$. Here v is the pair $(u(\cdot),\xi)$, so that the functional expansion for $y(t)$ includes the effect of changes in the initial condition. The derivation of the variational equations is usually much simplified when f and h have a specific form. This is illustrated in Section VI by the analysis of a nonlinear control system. A discussion of the results and other applications of the approach is contained in Section VII.

The reader who is interested mainly in the application of the basic theory of (1.4)–(1.7) may jump to Section IV without loss of continuity. Sections II and III are of general interest because they constitute a methodology by which other types of systems may be analyzed.

II. Approximation by a Sum of Homogeneous Functions

In what follows it is assumed that $\mathfrak{R}$ denotes the real numbers, $\mathcal{V}$ and $\mathcal{W}$ are real Banach spaces, $\mathfrak{N}$ is an open set in $\mathcal{V}$, and P is a function from $\mathfrak{N}$ into $\mathcal{W}$.

The objective is to obtain a representation for P of the form

$$P(v_0 + v) = P(v_0) + \sum_{i=1}^{k} Q_i(v) + R_k(v), \qquad v \in N(\rho) \tag{2.1}$$

where $\rho > 0$,

$$N(\rho) = \{v : \|v\| < \rho,\ v \in \mathcal{V}\}, \tag{2.2}$$

and for $i = 1,\cdots,k$, $Q_i : \mathcal{V} \to \mathcal{W}$ is homogeneous of degree i, that is, for all $v \in \mathcal{V}$ and $\alpha \in \mathfrak{R}$, $Q_i(\alpha v) = \alpha^i Q_i(v)$. Generally, $R_k(v)$ is to be small in some reasonable way, e.g., $\|R_k(v)\| < \mu\|v\|^{k+1}$.

Before proceeding, it should be emphasized that (2.1) is not a natural generalization of Taylor's formula. This is because $P(v_0) + \sum_{i=1}^{k} Q_i(v)$ is not necessarily a polynomial in v.

Definition 2.1: The function $Q_i : \mathcal{V} \to \mathcal{W}$ is an i-power if there exist functions $c_j : \mathcal{V}^2 \to \mathcal{W}$, $j = 0,\cdots,i$, such that for all $v, \bar{v} \in \mathcal{V}$ and $\alpha, \beta \in \mathfrak{R}$,

$$Q_i(\alpha v + \beta \bar{v}) = \sum_{j=0}^{i} c_j(v,\bar{v})\alpha^{i-j}\beta^j. \tag{2.3}$$

The function $Q_0 + \sum_{i=1}^{k} Q_i(v)$, $Q_0 \in \mathcal{W}$, is a polynomial of degree k if for $i = 1,\cdots,k$, Q_i is an i-power.

Example 2.2: Let $P(v) = r^2 f(\theta)$ where $\mathcal{V} = \mathfrak{R}^2$, $\mathcal{W} = \mathfrak{R}$, and r and θ are polar coordinates in the plane. Clearly, P is homogeneous of degree 2. However, it is a 2-power if and only if it is a quadratic form, i.e., there are real numbers a_1, a_2, a_3 such that $f(\theta) = a_1\cos^2\theta + a_2\cos\theta\sin\theta + a_3\sin^2\theta$. Thus, for $k \geq 2$ and $v_0 = 0$, $P = r^2\cos^4\theta$ can be represented by (2.1), but the sum is not a polynomial.

It is necessary to consider derivatives of continuous functions $f : \mathfrak{R} \to \mathcal{W}$. The notation $(d/d\alpha)f(\alpha) = g(\alpha)$ means $g(\alpha) \in \mathcal{W}$ satisfies

$$\lim_{\beta \to 0} \|g(\alpha) - \beta^{-1}(f(\alpha+\beta) - f(\alpha))\| = 0. \tag{2.4}$$

If f is defined on a closed interval of $\mathfrak{R}$ and α is an end point, the appropriate one-sided limit is used. Higher order derivatives are formed in the obvious way. For example, $(d/d\alpha)^2 f(\alpha)=h(\alpha)$ where $h(\alpha)=(d/d\alpha)g(\alpha)$. Only the simplest results from the theory of integration are needed. If $f:\mathfrak{R}\to\mathcal{W}$ has a continuous derivative g on $[0,\alpha]$, then the Riemann integral [14], [15] of g exists and $\int_0^\alpha g(\sigma)\,d\sigma=f(\alpha)-f(0)$. Moreover, if $\gamma:\mathfrak{R}\to\mathfrak{R}$ is integrable and $\|g(\alpha)\|\leqslant\gamma(\alpha)$, then $\|f(\alpha)-f(0)\|\leqslant\int_0^\alpha\gamma(\sigma)\,d\sigma$.

Definition 2.3: For $\bar v\in\mathfrak{N}$ and $v\in\mathcal{V}$ adopt the notation

$$P^0_{\bar v;v}(\alpha)=P(\bar v+\alpha v)$$

$$P^i_{\bar v;v}(\alpha)=\left(\frac{d}{d\alpha}\right)^i P^0_{\bar v;v}(\alpha),\qquad i=1,\cdots,k. \tag{2.5}$$

For all $v\in\mathcal{V}$ assume that there is an open interval, $(-\rho(\bar v,v),\rho(\bar v,v))\subset\mathfrak{R}$, such that $P^i_{\bar v;v}$ is continuous on the interval for $i=0,\cdots,k$. Then P is said to have a smooth kth variation at $\bar v$ and for $i=1,\cdots,k$, $\delta^iP_{\bar v}:\mathcal{V}\to\mathcal{W}$, which is given by

$$\delta^iP_{\bar v}(v)=\left(\frac{d}{d\alpha}\right)^i P(\bar v+\alpha v)\bigg|_{\alpha=0}=P^i_{\bar v;v}(0), \tag{2.6}$$

is called the ith variation of P at $\bar v$.

Some observations are in order. Simple examples where $\mathcal{V}=\mathfrak{R}^2$ and $\mathcal{W}=\mathfrak{R}$ show that the continuity of $P^k_{\bar v,v}(\alpha)$ in α does not imply the continuity of $\delta^kP_{\bar v}(v)$ in either $\bar v$ or v. The terminology "smooth" kth variation is an attempt to distinguish this difference. Since $(d/d\alpha)^iP(\bar v+\alpha\beta v)=\beta^i(d/d(\alpha\beta))^iP(\bar v+\alpha\beta v)$, it follows that

$$\delta^iP_{\bar v}(\beta v)=\beta^i\delta^iP_{\bar v}(v),\qquad v\in\mathcal{V},\beta\in\mathfrak{R}. \tag{2.7}$$

Thus, $\delta^iP_{v_0}$ is homogeneous of degree i and (1.3) meets the requirements of (2.1). Finally, when P has a smooth kth variation at $\bar v+\alpha v$, the identity $(d/d\alpha)^iP(\bar v+\alpha v)=(d/d\beta)^iP(\bar v+\alpha v+\beta v)|_{\beta=0}$ implies

$$\delta^iP_{\bar v+\alpha v}(v)=P^i_{\bar v;v}(\alpha),\qquad v\in\mathcal{V},0<i\leqslant k. \tag{2.8}$$

The first approximation theorem can now be stated.

Theorem 2.4: Suppose that there exists a $\rho_0>0$ such that $\{v_0\}+N(\rho_0)\subset\mathfrak{N}$ and P has a smooth kth variation at $\bar v$ for all $\bar v\in\{v_0\}+N(\rho_0)$. For all $\epsilon>0$ assume that there exists a $\delta(\epsilon)$, $0<\delta(\epsilon)<\rho_0$ such that the following condition is satisfied:

$$\|\delta^kP_{v_0+\alpha v}(v)-\delta^kP_{v_0}(v)\|<\epsilon\|v\|^k,$$
$$\text{for all } v\in N(\delta(\epsilon)),\ \alpha\in[0,1]. \tag{2.9}$$

Then for any $\epsilon>0$, $R_k(v)$ in (1.3) satisfies

$$\|R_k(v)\|<\frac{\epsilon}{k!}\|v\|^k,\qquad \text{for all } v\in N(\delta(\epsilon)). \tag{2.10}$$

Proof: Choose $\epsilon>0$. Then

$$r(v,\alpha)=P(v_0+\alpha v)-P(v_0)-\sum_{i=1}^k\frac{1}{i!}\delta^iP_{v_0}(\alpha v)$$
$$=P(v_0+\alpha v)-P(v_0)-\sum_{i=1}^k\frac{1}{i!}P^i_{v_0;v}(0)\alpha^i \tag{2.11}$$

is defined for all $v\in N(\delta(\epsilon))$ and $\alpha\in[0,1]$, and $r(v,1)=R_k(v)$. Moreover, for all $v\in N(\delta(\epsilon))$, $\alpha\in[0,1]$, $i=1,\cdots,k$, the function $P^i_{v_0+\alpha v;v}(\beta)$ is continuous for $|\beta|$ sufficiently small. Since $P^i_{v_0;v}(\alpha+\beta)=P^i_{v_0+\alpha v;v}(\beta)$, this implies that the functions

$$r_1(v,\alpha)=\left(\frac{d}{d\alpha}\right)r(v,\alpha)$$
$$=P^1_{v_0;v}(\alpha)-\sum_{i=1}^k\frac{1}{(i-1)!}P^i_{v_0;v}(0)\alpha^{i-1}$$
$$r_2(v,\alpha)=\left(\frac{d}{d\alpha}\right)^2r(v,\alpha)$$
$$=P^2_{v_0;v}(\alpha)-\sum_{i=2}^k\frac{1}{(i-2)!}P^i_{v_0;v}(0)\alpha^{i-2}$$
$$\vdots\qquad\vdots\qquad\vdots$$
$$r_k(v,\alpha)=\left(\frac{d}{d\alpha}\right)^kr(v,\alpha)=P^k_{v_0;v}(\alpha)-P^k_{v_0;v}(0) \tag{2.12}$$

are defined and continuous in α for $v\in N(\delta(\epsilon))$ and $\alpha\in[0,1]$. By (2.8) and (2.9), $\|r_k(v,\alpha)\|<\epsilon\|v\|^k$. By the right side of (2.12), $r_{k-1}(v,0)=0$. Thus, integration of $r_k(v,\alpha)$ with respect to α gives

$$\|r_{k-1}(v,\alpha)\|=\|r_{k-1}(v,\alpha)-r_{k-1}(v,0)\|$$
$$\leqslant\int_0^\alpha\|r_k(v,\sigma)\|\,d\sigma\leqslant\alpha\epsilon\|v\|^k. \tag{2.13}$$

Repeated integration (noting that $r_i(v,0)=0$, $i=1,\cdots,k-2$, and $r(v,0)=0$) gives $\|r(v,\alpha)\|<(\alpha^k/k!)\epsilon\|v\|^k$. Setting $\alpha=1$ completes the proof.

Theorem 2.5: Suppose that there exists a $\rho_0>0$ such that $\{v_0\}+N(\rho_0)\subset\mathfrak{N}$ and P has a smooth $(k+1)$th variation at $\bar v$ for all $\bar v\in\{v_0\}+N(\rho_0)$. Assume that there exists an $M>0$ and a ρ, $0<\rho\leqslant\rho_0$ such that the following condition is satisfied:

$$\|\delta^{k+1}P_{v_0+\alpha v}(v)\|<M\|v\|^{k+1},$$
$$\text{for all } v\in N(\rho),\ \alpha\in[0,1]. \tag{2.14}$$

Then $R_k(v)$ in (1.3) satisfies

$$\|R_k(v)\|<\frac{M}{(k+1)!}\|v\|^{k+1},\qquad\text{for all } v\in N(\rho). \tag{2.15}$$

Proof: Proceed as in the previous proof with $\delta(\epsilon)$ replaced by ρ and add to (2.12) the function

$$r_{k+1}(v,\alpha)=\left(\frac{d}{d\alpha}\right)^{k+1}r(v,\alpha)=P^{k+1}_{v_0;v}(\alpha) \tag{2.16}$$

which is continuous in α for $\alpha\in[0,1]$. Using $r_k(v,0)=0$ and the bound (2.14) gives, by integration, $\|r_k(v,\alpha)\|<M\alpha\|v\|^{k+1}$. Repeated integration, as before, completes the proof.

These theorems are in the spirit of the results obtained by Graves [14] in 1927. In his Theorem 5, $\mathcal{V}$ is a linear metric space, $\mathcal{W}$ is a complete linear metric space, and $R_k(v)$ is an integral. Notice that the completeness of $\mathcal{V}$ has not been used in the above proof. The normed spaces have the advantage that they permit the main conditions, (2.9) and (2.14), and the error bounds, (2.10) and (2.15), to be expressed in particularly useful ways.

Remark 2.6: Consider the conditions

$$\|\delta^k P_{\bar{v}}(v) - \delta^k P_{v_0}(v)\| < \epsilon \|v\|^k, \quad \text{for all } v \in \mathcal{V},\ \bar{v} \in \{v_0\} + N(\delta(\epsilon)) \tag{2.17}$$

and

$$\|\delta^{k+1} P_{\bar{v}}(v)\| < M\|v\|^{k+1}, \quad \text{for all } v \in \mathcal{V},\ \bar{v} \in \{v_0\} + N(\rho). \tag{2.18}$$

Because of the homogeneity of $\delta^k P_{\bar{v}}$ and $\delta^{k+1} P_{\bar{v}}$, these conditions imply, respectively, conditions (2.9) and (2.14). In many applications of Theorems 2.4 and 2.5, these conditions are easier to work with and can be verified. Example 2.2 shows that (2.9) and (2.14) can be satisfied when (2.17) and (2.18) cannot.

It is of interest to know whether or not the expansions presented above are in some sense unique. A result of Graves [14], modified (proof omitted) to suit the present situation, answers the question affirmatively.

Theorem 2.7: Let $Q_i : \mathcal{V} \to \mathcal{W}$, $i = 1, \cdots, k$ be homogeneous of degree i. Suppose P satisfies the hypotheses of Theorem 2.4. If for any $\epsilon > 0$ there exists a $\bar{\delta}(\epsilon)$ such that

$$\left\| P(v_0 + v) - P(v_0) - \sum_{i=1}^{k} Q_i(v) \right\| < \frac{\epsilon}{k!} \|v\|^k, \quad \text{for all } v \in N(\bar{\delta}(\epsilon)), \tag{2.19}$$

then

$$Q_i(v) = \frac{1}{i!} \delta^i P_{v_0}(v), \qquad i = 1, \cdots, k. \tag{2.20}$$

Suppose P satisfies the hypothesis of Theorem 2.5. If there exist $\bar{M} > 0$ and $\bar{\rho}$, $0 < \bar{\rho} \leq \rho_0$, such that

$$\left\| P(v_0 + v) - P(v_0) - \sum_{i=1}^{k} Q_i(v) \right\| < \frac{\bar{M}}{(k+1)!} \|v\|^{k+1}, \quad \text{for all } v \in N(\bar{\rho}), \tag{2.21}$$

then (2.20) holds.

In specific applications of the preceding theory it may be important to determine that $P(v_0) + \sum_{i=1}^{k} (1/i!) \delta^i P_{v_0}(v)$ is a polynomial of degree k. Often, as is the case in Sections IV, V, and VI, this can be done by simply inspecting the concrete forms of $\delta^i P_{v_0}(v)$ for $i = 1, \cdots, k$.

III. Power Series

It is possible to extend the above ideas to infinite series. The most natural framework for doing this is the theory of analytic functions in Banach spaces. This theory is well established and a suitable account appears in Hille and Phillips [15]. In this section several key results from [15] are stated in the notation of Section II. Although [15] encompasses a much richer theory, what is used here seems sufficient for many applications. It leads to easily verified conditions which ensure the uniform convergence of (1.3) for $k \to \infty$. As a bonus, these conditions imply that (1.3) is a power series, i.e., for all $i \geq 1$, $\delta^i P_{v_0}$ is an i-power.

In this section it is understood that $\mathcal{R}$ is replaced by the complex field $\mathcal{C}$, and that $\mathcal{V}$ and $\mathcal{W}$ are complex Banach spaces. The notation $(d/d\alpha) f(\alpha) = g(\alpha)$ means (2.4) holds for all complex $\beta \to 0$. Otherwise, the notation is the same as in the previous section.

Definition 3.1 (Definition 3.17.2 of [15]): Let $\mathfrak{N} \subset \mathcal{V}$ be an open set. The function $P : \mathfrak{N} \to \mathcal{W}$ is analytic in $\mathfrak{N}$ if: 1) P has a first variation at all $\bar{v} \in \mathfrak{N}$, i.e.,

$$\left(\frac{d}{d\alpha}\right) P(\bar{v} + \alpha v)\Big|_{\alpha=0} = \delta^1 P_{\bar{v}}(v) \tag{3.1}$$

exists for all $\bar{v} \in \mathfrak{N}$ and $v \in \mathcal{V}$; and 2) P is locally bounded, i.e., for all $\bar{v} \in \mathfrak{N}$ there is a $\rho(\bar{v}) > 0$ and a finite $M(\bar{v}) > 0$ such that $\{\bar{v}\} + N(\rho(\bar{v})) \subset \mathfrak{N}$ and

$$\|P(v)\| \leq M(\bar{v}), \qquad \text{for all } v \in \{\bar{v}\} + N(\rho(\bar{v})). \tag{3.2}$$

Theorem 3.2 (Theorems 3.17.1 and 26.3.5 of [15]): Assume that P is analytic in the open set $\mathfrak{N} \subset \mathcal{V}$. Then $\delta^k P_{\bar{v}} : \mathcal{V} \to \mathcal{W}$ exists for all positive integers k and $\bar{v} \in \mathfrak{N}$. Moreover, $\delta^k P_{\bar{v}}$ is a k-power.

Theorem 3.3 (Theorem 3.17.1 of [15]): Let P satisfy the hypothesis of Theorem 3.2. Then for $v_0 \in \mathfrak{N}$ there exists a $\rho > 0$, dependent on v_0, such that

$$P(v_0 + v) = P(v_0) + \sum_{i=1}^{\infty} \frac{1}{i!} \delta^i P_{v_0}(v) \text{ uniformly in } N(\rho). \tag{3.3}$$

That is, given any $\epsilon > 0$, there exists a positive integer $k(\epsilon)$ such that $R_k(v)$ in (1.3) satisfies

$$\|R_k(v)\| < \epsilon \qquad \text{for all } k > k(\epsilon),\ v \in N(\rho). \tag{3.4}$$

IV. The Differential System (1.4)–(1.5)

Functional expansions for both $x(t)$ and $y(t)$ will be obtained. To distinguish between the two corresponding mappings of v, the notations

$$x(t) = P(v)(t) \tag{4.1}$$

$$y(t) = p(v)(t) \tag{4.2}$$

are adopted. Before considering the expansions for P and p, some additional notations and assumptions are needed.

Let $\mathcal{C}([0,T], \mathcal{R}^n)$ be the (Banach) space of continuous functions from $[0,T]$ into $\mathcal{R}^n$ with norm $\|e\| =$

$\sup_{[0,T]}|e(t)|$ where $|e(t)|$ is the sup norm of $e(t)$ over the components of $e(t)$. It is assumed in (1.4) that $v \in \mathcal{V} = \mathcal{C}([0,T],\mathfrak{R})$. The notation $f,g,h \in C_x^{(k)}$ is used when $f,g:\mathfrak{R}^n \times [0,T] \to \mathfrak{R}^n$ and $h:\mathfrak{R}^n \times [0,T] \to \mathfrak{R}$ are continuous and have continuous partial derivatives of order k with respect to the components of x in $\mathfrak{R}^n \times [0,T]$. It is assumed that $f,g,h \in C_x^{(k)}$ with $k \geq 1$. Let $\mathcal{Y} = \mathcal{C}([0,T],\mathfrak{R})$ and $\mathcal{X} = \mathcal{C}([0,T],\mathfrak{R}^n)$. For P, $\mathcal{X}$ plays the role of $\mathcal{W}$ in Sections II and III; for p, $\mathcal{Y}$ plays the role of $\mathcal{W}$.

In general, P and p are not defined on $\mathcal{V}$ because the solutions of (1.4) may have finite escape time. This question is treated in the Appendix, Theorem A.1 and Remark A.2, and it is shown that if (1.4) has a solution on $[0,T]$ for $v=v_0 \in \mathcal{V}$, then there is a neighborhood of v_0, $\mathfrak{N}$, such that $P:\mathfrak{N} \to \mathcal{X}$ and $p:\mathfrak{N} \to \mathcal{Y}$. Thus, for such v_0 it is possible to proceed according to the plan of Section II. To clarify the presentation, the characterizations for $\delta^i P_{\bar{v}}$ and $\delta^i p_{\bar{v}}$ are derived initially without full justification. Then, the conditions required in Theorems 2.4, 2.5, and 3.3 are verified rigorously. The results concerning the functional expansions for P and p are summarized in a set of theorems.

Derivatives of f,g,h with respect to x appear in the development and it is necessary to have a compact way of denoting them. If $f \in C_x^{(k)}$, then f has a kth Frechet differential with respect to x [9]. This differential has k increments, $w_1,\cdots,w_k \in \mathfrak{R}^n$, and is written $f^{(k)}(x,t)[w_1]\cdots[w_k]$. This is a compact notation because each of the n components of $f^{(k)}$ is a linear combination of products of k numbers, where in each product the numbers are taken from (one from each) the components of $w_1,\cdots,w_k$. The differential is symmetric (the ordering of $w_1,\cdots,w_k$ is immaterial) and k-linear (linear in each w_i taken separately) [9]. For brevity, the designation (x,t), which indicates where the differential is evaluated, will be omitted if it does not cause confusion. When $f^{(k)}$ has repeated arguments, a power-like notation is used, e.g., $f^{(4)}[w_1][w_3][w_2][w_1] = f^{(4)}[w_1]^2[w_2][w_3]$. Because of the symmetry, this is only a slight abuse of notation. The same notation and remarks apply to g and h.

Let $\bar{v} \in \mathfrak{N}$ and $v \in \mathcal{V}$. To obtain the characterization of $\delta^i P_{\bar{v}}(v)$, define

$$z(t,\alpha) = P(\bar{v}+\alpha v)(t), \qquad z_i(t,\alpha) = \left(\frac{d}{d\alpha}\right)^i P(\bar{v}+\alpha v)(t). \tag{4.3}$$

It is clear that $z(t,\alpha)$ is given by the solution of

$$\dot{z} = f(z,t) + \bar{v}(t)g(z,t) + \alpha v(t)g(z,t), z(0,\alpha) = \xi. \tag{4.4}$$

Differentiation of this equation with respect to α yields differential equations for the $z_i(t,\alpha)$. For example, if $\bar{f}(z,t) = f(z,t) + \bar{v}(t)g(z,t)$,

$$\begin{aligned} \dot{z}_1 &= \bar{f}^{(1)}[z_1] + \alpha v g^{(1)}[z_1] + vg, z_1(0,\alpha) = 0 \\ \dot{z}_2 &= \bar{f}^{(1)}[z_2] + \alpha v g^{(1)}[z_2] + \bar{f}^{(2)}[z_1]^2 + \alpha v g^{(2)}[z_1]^2 \\ &\quad + 2vg^{(1)}[z_1], z_2(0,\alpha) = 0 \end{aligned} \tag{4.5}$$

where, for simplicity, arguments have been omitted, and it is understood that g and the derivatives of f and g are evaluated at $(z(t,\alpha),t)$ where $z(t,\alpha)$ is the solution of (4.4). Define

$$\bar{x}(t) = P(\bar{v})(t), \qquad x_i(t) = \delta^i P_{\bar{v}}(v)(t). \tag{4.6}$$

Differential equations for the variations of P are obtained from $x_i(t) = z_i(t,0)$. For example, (4.4) and (4.5) yield

$$\begin{aligned} \dot{\bar{x}} &= f(\bar{x},t) + \bar{v}(t)g(\bar{x},t), \qquad \bar{x}(0) = \xi \\ \dot{x}_1 &= A(\bar{x},t)x_1 + v(t)B(\bar{x},t), \qquad x_1(0) = 0 \\ \dot{x}_2 &= A(\bar{x},t)x_2 + A^2_{1,1}(\bar{x},t)[x_1]^2 \\ &\quad + v(t)\{B^2_1(\bar{x},t)[x_1]\}, \qquad x_2(0) = 0 \end{aligned} \tag{4.7}$$

where $A(\bar{x},t)x_1 = f^{(1)}(\bar{x},t)[x_1] + \bar{v}(t)g^{(1)}(\bar{x},t)[x_1]$, $B(\bar{x},t) = g(\bar{x},t)$, $A^2_{1,1}(\bar{x},t)[x_1]^2 = f^{(2)}(\bar{x},t)[x_1]^2 + \bar{v}(t)g^{(2)}(\bar{x},t)[x_1]^2$, and $B^2_1(\bar{x},t)[x_1] = 2g^{(1)}(\bar{x},t)[x_1]$. Continuing in a similar fashion and omitting the arguments $\bar{x},t$ gives

$$\begin{aligned} \dot{x}_3 &= Ax_3 + A^3_{1,2}[x_1][x_2] + A^3_{1,1,1}[x_1]^3 \\ &\quad + v\{B^3_2[x_2] + B^3_{1,1}[x_1]^2\}, \qquad x_3(0) = 0 \\ \dot{x}_4 &= Ax_4 + A^4_{2,2}[x_2]^2 + A^4_{1,3}[x_1][x_3] \\ &\quad + A^4_{1,1,2}[x_1]^2[x_2] + A^4_{1,1,1,1}[x_1]^4 \\ &\quad + v\{B^4_3[x_3] + B^4_{1,2}[x_1][x_2] \\ &\quad + B^4_{1,1,1}[x_1]^3\}, \qquad x_4(0) = 0 \\ &\vdots \\ \dot{x}_k &= Ax_k + F^k(x_1,\cdots,x_{k-1}) \\ &\quad + vG^k(x_1,\cdots,x_{k-1}), \qquad x_k(0) = 0. \end{aligned} \tag{4.8}$$

A few comments concerning these equations may be useful. The equation for x_1 is the usual linearization of (1.4) about the reference pair $\bar{x}(t)$, $\bar{v}(t)$. For x_k, $k>1$, the equations have (the same) linear dynamics, but with forcing terms F^k and vG^k. F^k is a sum of j-linear, symmetric functions of the type $A^k_{i_1,i_2,\cdots,i_j}[x_{i_1}]\cdots[x_{i_j}]$. The indices $i_1,\cdots,i_j$ satisfy $0<i_1 \leq i_2 \cdots \leq i_j$ and $i_1+i_2+\cdots+i_j = k$. The last result follows because x_k must be a homogeneous function of degree k in u [see (2.7)]. Similar remarks apply to the term G^k and the functions $B^k_{i_1,i_2,\cdots,i_j}[x_{i_1}]\cdots[x_{i_j}]$ except $i_1+i_2+\cdots+i_j = k-1$. It is clear that the complexity of F^k and G^k mounts rapidly with k.

By noting that

$$p(\bar{v}+\alpha v)(t) = h(z(t,\alpha),t) \tag{4.9}$$

and proceeding similarly, equations for the variations

$$y_i(t) = \delta^i p_{\bar{v}}(v)(t) \tag{4.10}$$

can be derived. For $i=1,2$,

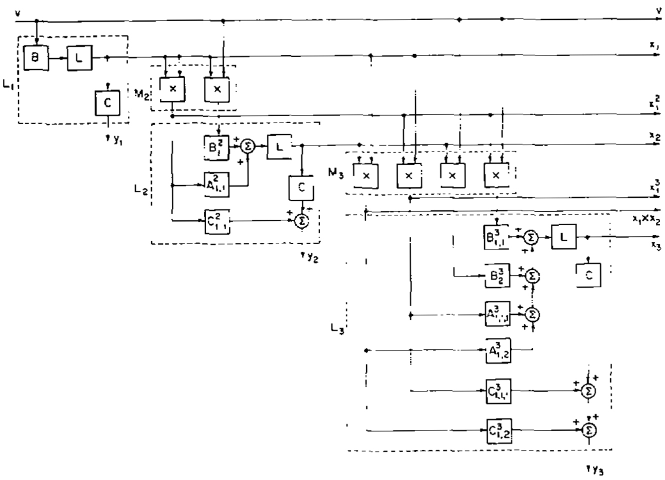

Fig. 1. Structural representation of the variational equations for the system (1.4)–(1.5).

$$
\begin{aligned}
y_1 &= C(\bar{x},t)x_1 \\
y_2 &= C(\bar{x},t)x_2 + C^2_{1,1}(\bar{x},t)[x_1]^2
\end{aligned} \tag{4.11}
$$

where $C(\bar{x},t)x_1 = h^{(1)}(\bar{x},t)[x_1]$ and $C^2_{1,1}(\bar{x},t)[x_1]^2 = h^{(2)}(\bar{x},t)[x_1]^2$. For $i=3,4,\cdots,k$,

$$
\begin{aligned}
y_3 &= Cx_3 + C^3_{1,2}[x_1][x_2] + C^3_{1,1,1}[x_1]^3 \\
y_4 &= Cx_4 + C^4_{2,2}[x_2]^2 + C^4_{1,3}[x_1][x_3] \\
&\quad + C^4_{1,1,2}[x_1]^2[x_2] + C^4_{1,1,1,1}[x_1]^4 \\
&\ \vdots \\
y_k &= Cx_k + H^k(x_1,\cdots,x_{k-1})
\end{aligned} \tag{4.12}
$$

where H^k is a sum of j-linear, symmetric functions of the type $C^k_{i_1,i_2,\cdots,i_j}[x_{i_1}]\cdots[x_{i_j}]$ where the indices $i_1,\cdots,i_j$ satisfy $0<i_1\leqslant i_2\cdots\leqslant i_j$ and $i_1+i_2+\cdots+i_j=k$.

Equations (4.6)–(4.8) and (4.10)–(4.12) characterize the variation of P and p. As indicated in Fig. 1, these equations have a special structure involving a patterned interconnection of vector multipliers M_i and linear dynamic systems L_i. In this representation L is the linear map of elements $e\in\mathfrak{X}$ into elements $w\in\mathfrak{X}$ which is defined by

$$\dot{w} = A(\bar{x},t)w + e(t), \qquad w(0)=0. \tag{4.13}$$

The notation $w_1\times w_2$ indicates a vector whose components are all possible products of components of w_1 times components of w_2 with a systematic scheme of ordering. Thus, if w_1 has dimension n_1 and w_2 has dimension n_2, $w_1\times w_2$ has dimension n_1n_2. Similarly, $w_1\times w_2\times w_3=(w_1\times w_2)\times w_3$, $w^2=w\times w$, and so on. Using this notation a symmetric, k-linear form $D[w_1][w_2]\cdots[w_k]$ can be written $Dw_1\times w_2\times\cdots\times w_k$ where D is interpreted as a linear mapping. Specifically, if $D[w_1]\cdots[w_k]$ is from $\mathfrak{R}^{kn}$ into $\mathfrak{R}^m$, D is an $m\times n^k$ matrix. Because of the symmetry, the matrix D has the property that $Dw_1\times w_2\times\cdots\times w_k = Dw_{i_1}\times w_{i_2}\times\cdots\times w_{i_k}$ where $i_1,\cdots,i_k$ is any permutation of the integers $1,2,\cdots,k$. In Fig. 1 this notation helps to sort out those "parts" of (4.7)–(4.8) and (4.11)–(4.12) which are linear (the L_i) and those which are nonlinear (the M_i).

Because of the special structure of (4.7)–(4.8) and (4.11)–(4.12), it is easy to integrate successively the equations to obtain integral formulas for $\delta^iP_{\bar{v}}$ and $\delta^iP_{\bar{v}}(v)$. To illustrate, consider the equations (4.7). Let $\bar{\Phi}(t)$ be the fundamental matrix defined by $\dot{\bar{\Phi}}=A(\bar{x}(t),t)\bar{\Phi}$, $\bar{\Phi}(0)=I$. Then the variation of the parameters formula gives

$$\delta^1P_{\bar{v}}(v)(t) = x_1(t) = \int_0^T W^1_{\bar{v}}(t,\sigma)v(\sigma)\,d\sigma \tag{4.14}$$

where

$$
\begin{aligned}
W^1_{\bar{v}}(t,\sigma) &= \bar{\Phi}(t)\bar{\Phi}^{-1}(\sigma)B(\bar{x}(\sigma),\sigma), \qquad 0\leqslant\sigma\leqslant t\leqslant T \\
&= 0, \qquad 0\leqslant t<\sigma\leqslant T.
\end{aligned} \tag{4.15}
$$

Using this result to express $x_1(t)$ in the last equation of (4.7) yields

$$\delta^2P_{\bar{v}}(v)(t) = x_2(t) = \int_0^T\int_0^T W^2_{\bar{v}}(t,\sigma_1,\sigma_2)v(\sigma_1)v(\sigma_2)\,d\sigma_1\,d\sigma_2 \tag{4.16}$$

where

$$
\begin{aligned}
W^2_{\bar{v}}(t,\sigma_1,\sigma_2) &= \int_0^t \bar{\Phi}(t)\bar{\Phi}^{-1}(\sigma) \\
&\qquad A^2_{1,1}(\bar{x}(\sigma),\sigma)\left[W^1_{\bar{v}}(\sigma,\sigma_1)\right]\left[W^1_{\bar{v}}(\sigma,\sigma_2)\right]d\sigma \\
&\qquad + \bar{\Phi}(t)\bar{\Phi}^{-1}(\sigma_1) \\
&\qquad B^2_1(\bar{x}(\sigma_1),\sigma_1)\left[W^1_{\bar{v}}(\sigma_1,\sigma_2)\right], \\
&\qquad\qquad 0\leqslant\sigma_1,\sigma_2\leqslant t\leqslant T \\
&= 0, \quad 0\leqslant t<\sigma_1\leqslant T \text{ or } 0\leqslant t<\sigma_2\leqslant T.
\end{aligned} \tag{4.17}
$$

Although the formulas become even more complex, it is clear that the process can be continued for $i>2$. Thus, there are functions $W^i_{\bar{v}}(t,\sigma_1,\cdots,\sigma_i)$ such that

$$
\begin{aligned}
x_i(t) = \delta^iP_{\bar{v}}(v)(t) &= \int_0^T\cdots\int_0^T W^i_{\bar{v}}(t,\sigma_1,\cdots,\sigma_i) \\
&\qquad \cdot v(\sigma_1)\cdots v(\sigma_i)\,d\sigma_1\cdots d\sigma_i.
\end{aligned} \tag{4.18}
$$

Similarly,

$$
\begin{aligned}
y_i(t) = \delta^ip_{\bar{v}}(v)(t) &= \int_0^T\cdots\int_0^T w^i_{\bar{v}}(t,\sigma_1,\cdots,\sigma_i) \\
&\qquad \cdot v(\sigma_1)\cdots v(\sigma_i)\,d\sigma_1\cdots d\sigma_i
\end{aligned} \tag{4.19}
$$

where

$$w^1_{\bar{v}}(t,\sigma) = C(\bar{x}(t),t)W^1_{\bar{v}}(t,\sigma) \tag{4.20}$$

and

$$
\begin{aligned}
w^2_{\bar{v}}(t,\sigma_1,\sigma_2) &= C(\bar{x}(t),t)W^2_{\bar{v}}(t,\sigma_1,\sigma_2) \\
&\quad + C^2_{1,1}(\bar{x}(t),t)\left[W^1_{\bar{v}}(t,\sigma_1)\right]\left[W^1_{\bar{v}}(t,\sigma_2)\right].
\end{aligned} \tag{4.21}
$$

Substituting these characterizations for $\delta^i P_{v_0}(v)$ and $\delta^i p_{v_0}$ into (1.3) gives concrete expansions for $x(t)$ and $y(t)$. Let $x_0(t)$ and $y_0(t)$ be given, respectively, by (1.4) and (1.5) with $v(t)=v_0(t)$. Then setting $\bar{v}(t)=v_0(t)$ in (4.7), (4.8), (4.11), and (4.12) yields

$$x(t)=x_0(t)+\sum_{i=1}^{k}\frac{1}{i!}x_i(t)+R_k(v)(t) \tag{4.22}$$

$$y(t)=y_0(t)+\sum_{i=1}^{k}\frac{1}{i!}y_i(t)+r_k(v)(t). \tag{4.23}$$

If the integral formulas (4.18) and (4.19) are used (with $\bar{v}=v_0$) in these equations, they become truncated Volterra series for $x(t)$ and $y(t)$.

It remains to be shown that the preceding steps are justified and that the remainder terms in (4.22) and (4.23) are bounded in a suitable fashion. This will be done using standard tools from the theory of differential equations and Theorems 2.4, 2.5, and 3.3. The results are contained in the following theorems.

Theorem 4.1: Let $f,g,h\in C_x^{(k)}$ and suppose (1.4) has a solution on $[0,T]$ for $v=v_0\in\mathcal{V}$. Then there exists an open set $\mathfrak{N}, v_0\in\mathfrak{N}\subset\mathcal{V}$, such that P and p are defined in $\mathfrak{N}$ and have smooth kth variations in $\mathfrak{N}$. For $i=1,\cdots,k$ and $\bar{v}\in\mathfrak{N}$, $\delta^i P_{\bar{v}}(v)$ and $\delta^i p_{\bar{v}}(v)$ are given by (4.6)–(4.8) and (4.10)–(4.12) or (4.18) and (4.19). The kernel functions $W_{\bar{v}}^i(t,\sigma_1,\cdots,\sigma_i)$ and $w_{\bar{v}}^i(t,\sigma_1,\cdots,\sigma_i)$ are calculated from (4.6)–(4.8) and (4.10)–(4.12) by the process described above and are continuous for all $t,\sigma_1,\cdots,\sigma_i\in[0,T]$ except for $t=\sigma_i$ or $\sigma_i=\sigma_j$, $i,j=1,\cdots,k$ where jump discontinuities may appear.

Proof: By Theorem A.1 and Remark A.2 of the Appendix, there exists a $\rho>0$ such that P is defined in $\mathfrak{N}=\{v_0\}+N(\rho)$. Since for all $\bar{v}\in\mathfrak{N}$ and $v\in\mathcal{V}$, $\bar{v}+\alpha v\in\mathfrak{N}$ for $|\alpha|$ sufficiently small, (4.4) has a solution for $|\alpha|$ sufficiently small. Moreover, if (4.4) is written as

$$\dot{z}=F(z,\alpha,t), \qquad z(0)=\xi, \tag{4.24}$$

it is clear that F is k times continuously differentiable in z and α. From this it is known (see [19, ch. II, sect. 4]) that $z(t,\alpha)$ is k times continuously differentiable in α for α in a neighborhood of $\alpha=0$. This proves that P has a smooth kth variation at $\bar{v}$ for all $\bar{v}\in\mathfrak{N}$. Because of the differentiability of z, f, and g, the steps leading to (4.7) and (4.8) are valid. Thus, these differential equations define $\sigma^i P_{\bar{v}}(v)=x_i$ for $i=1,\cdots,k$. From (4.9), $h\in C_x^{(k)}$, and the differentiability of $z(t,\alpha)$ with respect to α, it follows that p has a smooth variation in $\mathfrak{N}$ and (4.10)–(4.12) are valid. Because $\bar{x}(t)$ is continuous, all the terms in (4.7), (4.8), (4.11), and (4.12) are continuous in t. This justifies the use of the variation of parameters formula and shows that $W_{\bar{v}}^i$ and $w_{\bar{v}}^i$ are continuous on the indicated subset of $[0,T]^{i+1}$.

Theorem 4.2: Let $f,g,h\in C_x^{(k+1)}$ and suppose (1.4) has a solution on $[0,T]$ for $v=v_0\in\mathcal{V}$. Then there exists a $\rho>0$ and $\mu>0$ such that

$$\|R_k(v)\|,\ \|r_k(v)\|<\mu\|v\|^{k+1} \tag{4.25}$$

for all $v\in\mathcal{V}$ such that $\|v\|<\rho$.

Proof: First, $R_k(v)$ is considered. Because of Theorem 4.1 and Remark 2.6, it is clear that the hypotheses of Theorem 2.5 are satisfied if (2.18) is satisfied. From Remark A.3, $\bar{x}=P(\bar{v})$ satisfies $\|\bar{x}\|<K_0$ for $\bar{v}\in\{v_0\}+N(\rho)=\mathfrak{N}$. This means that there are constants $\overline{M}_1$ and $\overline{M}_2$ such that $|A(\bar{x},t)x_1|\leqslant\overline{M}_1|x_1|$ and $|B(\bar{x},t)|\leqslant\overline{M}_2$ for all $t\in[0,T]$, $\bar{v}\in\mathfrak{N}$. By the Gronwall inequality [19] it is known that if w satisfies (4.13), then $\|w\|\leqslant(T\exp\overline{M}_1T)\|e\|$. Combining the results of the preceding two sentences and applying them to the differential equation for x_1 shows that there exists an M_1 such that $\|x_1\|<M_1\|v\|$ for all $\bar{v}\in\mathfrak{N}$ and $v\in\mathcal{V}$. Using this result and similar reasoning, it can be seen from the differential equation for x_2 that there exists an M_2 such that $\|x_2\|<M_2\|v\|^2$. The process can be repeated until it is shown that there exists an M such that $\|x_{k+1}\|<M\|v\|^{k+1}$ for all $\bar{v}\in\mathfrak{N}$ and $v\in\mathcal{V}$, which implies that (2.18) is satisfied. Thus, the bound (2.15) follows from Theorem 2.5. In a similar way, (4.11) and (4.12) with $\|x_i\|<M_i\|v\|^i$ show that there exist m_i such that $\|y_i\|<m_i\|v\|^i$. Using the same argument and defining $\mu(k+1)!=\max\{M_{k+1},m_{k+1}\}$ completes the proof.

Theorem 4.3: Let $f,g,h\in C_x^{(k)}$ and suppose that (1.4) has a solution on $[0,T]$ for $v=v_0\in\mathcal{V}$. Then for any $\epsilon>0$, there exists a $\delta(\epsilon)$ such that

$$\|R_k(v)\|,\ \|r_k(v)\|<\epsilon\|v\|^k \tag{4.26}$$

for all $v\in\mathcal{V}$ such that $\|v\|<\delta(\epsilon)$.

Proof: Because of Theorem 4.1 and Remark 2.6, it is sufficient to show that P and p satisfy (2.17). Since the notation becomes very burdensome, this will be done only for P and $k=2$. From this it should be clear how the proof proceeds for p and $k>2$. Define $\hat{z}_i(t)=\delta^i P_{\bar{v}}(v)(t)-\delta^i P_{v_0}(v)(t)$. Then it is clear from (4.7) that

$$\dot{\hat{z}}_1=A(x_0,t)\hat{z}_1+\big(A(\bar{x},t)-A(x_0,t)\big)x_1 + v(t)\big(B(\bar{x},t)-B(x_0,t)\big), \qquad \hat{z}_1(0)=0. \tag{4.27}$$

By Theorem A.1 it follows that $\|\bar{x}-x_0\|<K\|\tilde{v}\|$, $\tilde{v}=\bar{v}-v_0$. Using the continuity of $f^{(1)}(x,t)$ and $g^{(1)}(x,t)$, it is clear from this that there exists a $\delta_1(\epsilon)$ such that $|(A(\bar{x},t)-A(x_0,t))x_1|<\epsilon|x_1|$ and $|v(t)(B(\bar{x},t)-B(x_0,t))|<\epsilon|v(t)|$ if $\|\tilde{v}\|<\bar{\delta}(\epsilon)$. Applying the Gronwall bound of the previous proof and recalling that $\|x_1\|<M_1\|v\|$, it follows that there exists a $\delta_1(\epsilon)$ such that $\|\hat{z}_1\|<\epsilon\|v\|$ if $\|\hat{v}\|<\delta_1(\epsilon)$. Now consider the proof of (2.17). From (4.7) it follows that

$$\dot{\hat{z}}_2=A(x_0,t)\hat{z}_2+\big(A(\bar{x},t)-A(x_0,t)\big)x_2+A_{1,1}^2(\bar{x},t)[x_1]^2 - A_{1,1}^2(x_0,t)[x_1-\hat{z}_1]^2+v(t)B_1^2(\bar{x},t)[x_1] - v(t)B_1^2(x_0,t)[x_1-\hat{z}_1], \qquad \hat{z}_2(0)=0. \tag{4.28}$$

Using the representation for multilinear forms which was introduced in the discussion of Fig. 1, this becomes

$$\begin{aligned}\dot{\hat{z}}_2 = {} & A(x_0,t)\hat{z}_2 + (A(\bar{x},t) - A(x_0,t))x_2 + (A^2_{1,1}(\bar{x},t) \\ & - A^2_{1,1}(x_0,t))x_1^2 + 2A^2_{1,1}(x_0,t)x_1 \times \hat{z}_1 \\ & - A^2_{1,1}(x_0,t)\hat{z}_1^2 + v(t)(B^2_1(\bar{x},t) \\ & - B^2_1(x_0,t))x_1 + v(t)B^2_1(x_0,t)\hat{z}_1, \quad \hat{z}_2(0)=0. \quad (4.29)\end{aligned}$$

Because of the continuity of $f^{(2)}(x,t)$ and $g^{(2)}(x,t)$ and the definitions of $A^2_{1,1}(x,t)$ and $B^2_1(x,t)$, $\|x_2\| < M_2\|v\|^2$, $\|x_1\| < M_1\|v\|$, and $\|\hat{z}_1\| \le \epsilon\|v\|$ for $\|\tilde{v}\| < \delta_1(\epsilon)$, it can be argued that there exists a $\delta_2(\epsilon)$ such that $\|\tilde{v}\| < \delta_2(\epsilon)$ implies that the sum of the last four terms on the right side of (4.29) are bounded in norm by $\epsilon\|v\|^2$. This and the Gronwall bound establish the existence of $\delta_2(\epsilon)$ such that $\|\hat{z}_2\| < \epsilon\|v\|^2$ if $\|\tilde{v}\| = \|\bar{v} - v_0\| < \delta_2(\epsilon)$. The definition of $\hat{z}_2$ shows that (2.17) is satisfied for $k=2$.

Theorem 4.4: Assume that f,g,h are analytic functions x, i.e., $f,g,h \in C_x^{(1)}$ when the complex field $\mathcal{C}$ replaces $\mathcal{R}$. Suppose that (1.4) has a solution on $[0,T]$ for $v = v_0 \in \mathcal{V}$. Then the variations $\delta^i P_{v_0}(v)$ and $\delta^i p_{v_0}(v)$ are defined by (4.6)–(4.8) and (4.10)–(4.12) or (4.18) and (4.19) for all $i > 0$ and (4.22), (4.23) converge uniformly in a neighborhood of v_0 as $k \to \infty$. Specifically, there exists a $\rho > 0$ such that the following statement is true: given any $\epsilon > 0$ there exists a positive integer $k(\epsilon)$ such that

$$\|R_k(v)\|, \; \|r_k(v)\| < \epsilon \qquad (4.30)$$

for all $k \ge k(\epsilon)$ and $v \in \mathcal{V}$ which satisfy $\|v\| < \rho$.

Proof: In all of what follows $\mathcal{R}$ has been replaced by $\mathcal{C}$. As indicated in Remark A.4, this does not change the results of the Appendix. Thus, as before, (4.24) defines $z(t,\alpha) = P(\bar{v} + \alpha v)(t)$ for all $\bar{v} \in \{v_0\} + N(\rho) = \mathcal{N}$, $v \in \mathcal{V}$, and $\alpha \in \mathcal{C}$, $|\alpha|$ sufficiently small. Moreover, since $F(z,\alpha,t)$ is analytic in (z,α) for each $t \in [0,T]$, it is known [19, ch. II, sect. 5] that the solution is analytic in α. Thus condition 1) of Definition 3.1 is satisfied. Condition 2) follows because Remark A.3 implies that P is bounded on $\mathcal{N}$. Hence, P is analytic in $\mathcal{N}$. Because of the analyticity of h and $z(t,\alpha)$, (4.9) shows that $p(\bar{v} + \alpha v)(t)$ is analytic in α. The bound on P implies a bound on p and, hence, p is analytic in $\mathcal{N}$. The results of the theorem follow immediately from Theorems 3.2 and 3.3.

V. The Differential System (1.6)–(1.7)

In this section functional expansions for the system (1.6)–(1.7) are examined. Although there is an appreciable increase in the complexity of the notation, the schema of the previous section applies with only minor modifications. Because of this, the emphasis is on results, and many details concerning formulas and proofs are omitted.

As before, the notations (4.1) and (4.2) are adopted, where now $x(t)$ is the solution of (1.6), $y(t)$ is given by (1.7), and v is the pair $(u(\cdot),\xi)$. It is assumed that $v \in \mathcal{V} = \mathcal{C}([0,T],\mathcal{R}^m) \times \mathcal{R}^n$ and $\|v\| = \|u\| + |\xi|$. The notation $f,h \in C^{(k)}_{(x,u)}$ is used when $f: \mathcal{R}^n \times \mathcal{R}^m \times [0,T] \to \mathcal{R}^n$ and $h: \mathcal{R}^n \times \mathcal{R}^m \times [0,T] \to \mathcal{R}^l$ are continuous and have continuous partial derivatives of order k with respect to the components of x and u in $\mathcal{R}^n \times \mathcal{R}^m \times [0,T]$. It is assumed that $f,h \in C^{(k)}_{(x,u)}$ with $k \ge 1$. If (1.6) has a solution on $[0,T]$ for $v_0 \in \mathcal{V}$, then Theorem A.1 of the Appendix implies the existence of an open set $\mathcal{N} \subset \mathcal{V}$ such that (1.6) has a solution for all $v \in \mathcal{V}$. Thus, the maps $P: \mathcal{N} \to \mathcal{X}$ and $p: \mathcal{N} \to \mathcal{Y}$, where $\mathcal{X} = \mathcal{C}([0,T],\mathcal{R}^n)$ and $\mathcal{Y} = \mathcal{C}([0,T],\mathcal{R}^l)$, are defined. The objective is to expand P and p in expressions of the form (4.22) and (4.23).

Mimicking the development of the previous section, $z(t,\alpha)$ and $z_i(t,\alpha)$ are defined by (4.3). When $\bar{v} \in \mathcal{N}$, $v \in \mathcal{V}$ and $|\alpha|$ is sufficiently small, $z(t,\alpha)$ is defined and is the solution of

$$\dot{z} = f(z, \bar{u}(t) + \alpha u(t), t), \qquad z(0,\alpha) = \bar{\xi} + \alpha\xi. \qquad (5.1)$$

If $f \in C^{(k)}_{(x,u)}$, $z(t,\alpha)$ is k times continuously differentiable with respect to α for α in a neighborhood of $\alpha = 0$ [19, ch. II, sect. 4]. Thus, it is permissible to differentiate (5.1) k times with respect to α. Let $f^{(i)}(x,u,t)[(z_1,w_1)] \cdots [(z_i,w_i)]$ denote the ith Frechet differential of f with respect to (x,u) with increments $(z_1,w_1), \cdots, (z_i,w_i) \in \mathcal{R}^n \times \mathcal{R}^m$. Then noting that $(d/d\alpha)^i (z, \bar{u} + \alpha u) = (z_i, 0)$ for $i > 1$, it is clear that

$$\begin{aligned}
\dot{z}_1 &= f^{(1)}[(z_1,u)], \qquad z_1(0,\alpha) = \xi \\
\dot{z}_2 &= f^{(1)}[(z_2,0)] + f^{(2)}[(z_1,u)]^2, \qquad z_2(0,\alpha) = 0 \\
\dot{z}_3 &= f^{(1)}[(z_3,0)] + f^{(2)}[(z_1,u)][(z_2,0)] \\
&\quad + f^{(2)}[(z_2,0)][(x_1,u)] \\
&\quad + f^{(2)}[(z_1,u)][(z_2,0)] \\
&\quad + f^{(3)}[(z_1,u)]^3, \qquad z_3(0,\alpha) = 0 \\
&\;\vdots \qquad \vdots
\end{aligned} \qquad (5.2)$$

where it is understood that the differentials are evaluated at $(z(t,\alpha), \bar{u}(t) + \alpha u(t), t)$. Finally, from (4.6), $x_i(t) = z_i(t,0)$, and the symmetry of the differentials, it follows that the (smooth, up to order k) variations of P at $\bar{v}$ are given by

$$\begin{aligned}
\dot{\bar{x}} &= f(\bar{x}, \bar{u}(t), t), \qquad \bar{x}(0) = \bar{\xi} \\
\dot{x}_1 &= f^{(1)}(\bar{x}, \bar{u}(t), t)[(x_1, u(t))], \qquad x_1(0) = \xi \\
\dot{x}_2 &= f^{(1)}(\bar{x}, \bar{u}(t), t)[(x_2, 0)] \\
&\quad + f^{(2)}(\bar{x}, \bar{u}(t), t)[(x_1, u(t))]^2, \qquad x_2(0) = 0, \\
\dot{x}_3 &= f^{(1)}(\bar{x}, \bar{u}(t), t)[(x_3, 0)] \\
&\quad + 3f^{(2)}(\bar{x}, \bar{u}(t), t)[(x_2, 0)][(x_1, u(t))] \\
&\quad + f^{(3)}(\bar{x}, \bar{u}(t), t)[(x_1, u(t))]^3, \qquad x_3(0) = 0 \\
&\;\vdots \qquad \vdots
\end{aligned} \qquad (5.3)$$

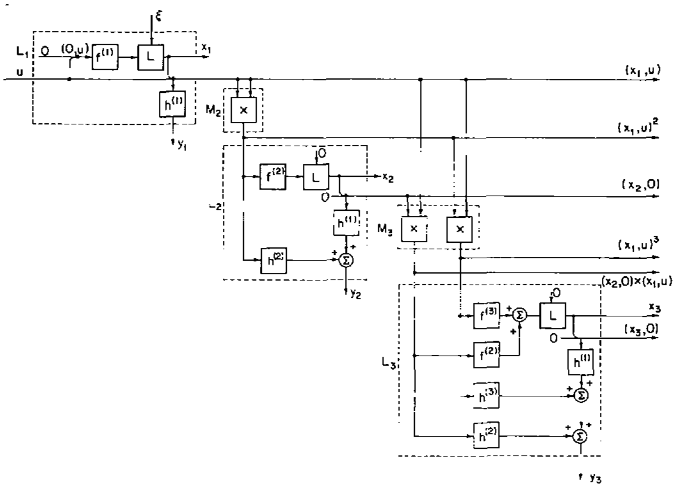

Fig. 2. Structural representation of the variational equations for the system (1.6)–(1.7).

Observing that $p(\bar{v}+\alpha v)(t) = h(z(t,\alpha), \bar{u}(t)+\alpha u(t), t)$, using the notation (4.10), and assuming that $h \in C^{(k)}_{(x,u)}$ shows that the (smooth, up to order k) variations of p are given by

$$y_1 = h^{(1)}(\bar{x},\bar{u}(t),t)\big[(x_1,u(t))\big]$$

$$y_2 = h^{(1)}(\bar{x},\bar{u}(t),t)\big[(x_2,0)\big] + h^{(2)}(\bar{x},\bar{u}(t),t)\big[(x_1,u(t))\big]^2$$

$$y_3 = h^{(1)}(\bar{x},\bar{u}(t),t)\big[(x_3,0)\big] + 3h^{(2)}(\bar{x},\bar{u}(t),t)\big[(x_2,0)\big]\big[(x_1,u(t))\big] + h^{(3)}(\bar{x},\bar{u}(t),t)\big[(x_1,u(t))\big]^3$$

$$\vdots \qquad \vdots \tag{5.4}$$

The variational equations have a special structure which is indicated in Fig. 2. As in Fig. 1, the multilinear forms are interpreted as linear mappings of "product vectors." The figure shows that the system (5.3)–(5.4) is a interconnection of vector multipliers (M_i) and linear dynamic systems (L_i). The operator L, which maps elements $(e,\eta) \in \mathfrak{X} \times \mathfrak{R}^n$ into elements $w \in \mathfrak{X}$, is defined by

$$\dot{w} = f^{(1)}(\bar{x}(t),\bar{u}(t),t)\big[(w,0)\big] + e(t), \qquad w(0) = \eta. \tag{5.5}$$

To obtain explicit formulas for the variations, it is helpful to write (5.3)–(5.4) in greater detail. For example, the equations for x_i, y_i, $i = 1,2$, are written

$$\dot{x}_1 = Ax_1 + Bu, \qquad x_1(0) = \xi$$

$$\dot{x}_2 = Ax_2 + A^2_{1,1}[x_1]^2 + B^2_{1,1}[u]^2 + B^2_{1;1}[u][x_1], \qquad x_2(0) = 0$$

$$y_1 = Cx_1 + Du$$

$$y_2 = Cx_2 + C^2_{1,1}[x_1]^2 + D^2_{1,1}[u]^2 + D^2_{1;1}[u][x_1] \tag{5.6}$$

where A, B, C, and D are t-dependent matrices and the remaining terms are t-dependent bilinear (not necessarily symmetric) functions. For $i > 2$, similar expressions can be written, although the number of multilinear terms which appear grows rapidly with i. When the variation of parameters formula is applied to (5.6), it is seen that the variations can be written as follows:

$$\delta^1 P_{\bar{v}}(v)(t) = x_1(t) = \int_0^T W^1_1(t,\sigma_1)\big[u(\sigma_1)\big]\,d\sigma_1 + \Phi^1_0(t)[\xi] \tag{5.7}$$

$$\delta^2 P_{\bar{v}}(v)(t) = x_2(t) = \int_0^T\!\!\int_0^T W^2_{1,2}(t,\sigma_1,\sigma_2)\big[u(\sigma_1)\big]\big[u(\sigma_2)\big]\,d\sigma_1\,d\sigma_2 + \int_0^T W^2_{1,1}(t,\sigma_1)\big[u(\sigma_1)\big]^2 d\sigma_1 + \int_0^T \Phi^2_{1;0}(t,\sigma_1)\big[u(\sigma_1)\big][\xi]\,d\sigma_1 + \Phi^2_{0,0}(t)[\xi]^2 \tag{5.8}$$

$$\delta^1 p_{\bar{v}}(v)(t) = y_1(t) = \int_0^T w^1_1(t,\sigma_1)\big[u(\sigma_1)\big]\,d\sigma_1 + w^1_0(t)\big[u(t)\big] + \phi^1_0(t)[\xi] \tag{5.9}$$

$$\delta^2 p_{\bar{v}}(v)(t) = y_2(t) = \int_0^T\!\!\int_0^T w^2_{1,2}(t,\sigma_1,\sigma_2)\big[u(\sigma_1)\big]\big[u(\sigma_2)\big]\,d\sigma_1\,d\sigma_2 + \int_0^T w^2_{1,1}(t,\sigma_1)\big[u(\sigma_1)\big]^2 d\sigma_1 + \int_0^T w^2_{0,1}(t,\sigma_1)\big[u(t)\big]\big[u(\sigma_1)\big]\,d\sigma_1 + w^2_{0,0}(t)\big[u(t)\big]^2 + \int_0^T \phi^2_{1;0}(t,\sigma_1)\big[u(\sigma_1)\big][\xi]\,d\sigma_1 + \phi^2_{0,0}(t)[\xi]^2 + \phi^2_{0;0}(t)\big[u(t)\big][\xi]. \tag{5.10}$$

The functions labeled W, w, Φ, ϕ utilize the notation which has been established for linear and bilinear forms and are determined from the data in (5.6). For example,

$$\Phi^2_{1;0}(t,\sigma_1)[u][\xi] = \bar{\Phi}(t)\bar{\Phi}^{-1}(\sigma_1)\, B^2_{1;1}(\bar{x}(\sigma_1),\bar{u}(\sigma_1),\sigma_1)[u]\big[\Phi^1_0(\sigma_1)[\xi]\big] + 2\int_0^{t} \bar{\Phi}(t)\bar{\Phi}^{-1}(\sigma)\, A^2_{1,1}(\bar{x}(\sigma),\bar{u}(\sigma),\sigma)\big[W^1_1(\sigma,\sigma_1)[u]\big]\big[\Phi^1_0(\sigma)[\xi]\big]\,d\sigma, \qquad 0 \leqslant \sigma_1 \leqslant t \leqslant T$$

$$= 0, \qquad 0 \leqslant t < \sigma_1 \leqslant T \tag{5.11}$$

where $\bar{\Phi}(t)$ is the fundamental matrix corresponding to A. The multiplier 2 arises from the fact that $A^2_{1,1}$ is symmetric; the required change in order of integration is permissible because the bilinearity of $A_{1,1}^2$ leads to the identity

$$\int_0^t \bar{\Phi}^{-1}(\sigma)A_{1,1}^2(\bar{x}(\sigma),\bar{u}(\sigma),\sigma)\left[\int_0^T W_1^1(\sigma,\sigma_1)[u(\sigma_1)]\,d\sigma_1\right][\Phi_0^1(\sigma)[\xi]]\,d\sigma$$
$$=\int_0^t\int_0^T \bar{\Phi}^{-1}(\sigma)A_{1,1}^2(\bar{x}(\sigma),\bar{u}(\sigma),\sigma)[W_1^1(\sigma,\sigma_1)[u(\sigma_1)]][\Phi_0^1(\sigma)[\xi]]\,d\sigma_1\,d\sigma. \quad (5.12)$$

It is easy to see that $W_{1,2}^2$ and $w_{1,2}^2$ are zero for $0\leqslant t<\sigma_1\leqslant T$ or $0\leqslant t<\sigma_2\leqslant T$, and are continuous in t,σ_1,σ_2 except at $t=\sigma_1,\sigma_2$ or $\sigma_1=\sigma_2$; W_1^1, $W_{1,1}^2$, $\Phi_{1;0}^2$, w_1^1, $w_{1,1}^2$, $w_{0,1}^2$, and $\phi_{1;0}^2$ are zero for $0\leqslant t<\sigma_1\leqslant T$, and continuous in t,σ_1 for $0\leqslant\sigma_1\leqslant t\leqslant T$; Φ_0^1, $\Phi_{0,0}^2$, w_0^1, ϕ_0^1, $w_{0,0}^2$, $\phi_{0,0}^2$, and $\phi_{0;0}^2$ are continuous in t for $0\leqslant t\leqslant T$. Similar results hold for $2<i\leqslant k$, although the complexity of the formulas is formidable.

Even when $\xi=0$, the above formulas do not lead to a truncated Volterra series. This is because of the terms corresponding to $W_{1,1}^2$, $w_{1,1}^2$, $w_{0,1}^2$, $w_{0,0}(t)$. If "impulsive" kernels or Stieltjes integrals are allowed, the terms may take on the appearance of terms in a Volterra series. For example, if $\xi=0$, $\delta^2 p_{\bar{v}}(v)(t)$ can be written

$$\delta^2 p_{\bar{v}}(v)(t)=\int_0^T\int_0^T w^2(t,\sigma_1,\sigma_2)[u(\sigma_1)][u(\sigma_2)]\,d\sigma_1\,d\sigma_2 \quad (5.13)$$

where

$$w^2(t,\sigma_1,\sigma_2)[u_1][u_2]=w_{1,2}^2(t,\sigma_1,\sigma_2)[u_1][u_2]+\delta(\sigma_1-\sigma_2)w_{1,1}^2(t,\sigma_1)[u_1][u_2]+\delta(t-\sigma_1)\delta(t-\sigma_2)w_{0,0}^2(t)[u_1][u_2] \quad (5.14)$$

and $\delta(t)$ is the Dirac "function."

The preceding characterizations of P and p can be substituted into (4.22), (4.23) to obtain functional expansions for P and p. The results are made precise in the following theorems. Proofs are omitted, since they are similar to those of the previous section and involve rather lengthy notations.

Theorem 5.1: Let $f,h\in C_{(x,u)}^{(k)}$ and suppose that (1.6) has a solution on $[0,T]$ for $v=(u,\xi)=v_0\in\mathcal{V}$. Then there exists an open set $\mathfrak{N}$, $v_0\in\mathfrak{N}\subset\mathcal{V}$, such that P and p are defined in $\mathfrak{N}$ and have smooth kth variations in $\mathfrak{N}$. For $i=1,\cdots,k$ and $\bar{v}\in\mathfrak{N}$, $\delta^i P_{\bar{v}}(v)$ and $\delta^i p_{\bar{v}}(v)$ are characterized by the system of equations (5.3)–(5.4) and integral formulas of the type (5.7)–(5.10).

Theorem 5.2: Let $f,h\in C_{(x,u)}^{(k+1)}$ and suppose that (1.6) has a solution on $[0,T]$ for $v=v_0\in\mathcal{V}$. Then there exists a $\rho>0$ and $\mu>0$ such that (4.25) is satisfied for all $v\in\mathcal{V}$ such that $\|v\|<\rho$.

Theorem 5.3: Let $f,h\in C_{(x,u)}^{(k)}$ and suppose that (1.6) has a solution on $[0,T]$ for $v=v_0\in\mathcal{V}$. Then for any $\epsilon>0$, there exists a $\delta(\epsilon)$ such that (4.26) is satisfied for all $v\in\mathcal{V}$ such that $\|v\|<\delta(\epsilon)$.

Theorem 5.4: Assume that f,h are analytic functions of x and u, i.e., $f,h\in C_{(x,u)}^{(1)}$ when the complex field $\mathcal{C}$ replaces $\mathfrak{R}$. Suppose that (1.6) has a solution on $[0,T]$ for $v=v_0\in\mathcal{V}$. Then $\delta^i P_{v_0}(v)$ and $\delta^i p_{v_0}(v)$ are defined for all $i>0$, and (4.22), (4.23) converge uniformly in a neighborhood of v_0 as $k\to\infty$.

VI. An Example

General characterizations of the variations, such as (4.7), (4.8), (4.11), (4.12) or (5.3), (5.4), are unnecessarily complex for many applications of the preceding theory. Frequently, it is simpler to derive the variational equations from scratch, using (1.2) as the basic mathematical tool.

To illustrate this point, consider the nonlinear feedback system shown in Fig. 3(a). The linear dynamic system $\tilde{L}$ is defined by the equations

$$\dot{x}(t)=\tilde{A}(t)x(t)+\tilde{b}(t)\tilde{u}(t), \qquad x(0)=\xi$$
$$y(t)=\tilde{c}(t)x(t) \quad (6.1)$$

where $x(t)\in\mathfrak{R}^n$, $\tilde{u}(t)\in\mathfrak{R}$, $y(t)\in\mathfrak{R}$, and the matrices $\tilde{A},\tilde{b}$, $\tilde{c}$ are continuous in $[0,T]$; the nonlinearity $\psi:\mathfrak{R}\to\mathfrak{R}$ is k times continuously differentiable; the input to the feedback system is $u\in\mathfrak{U}=\mathcal{C}([0,T],\mathfrak{R})$. It is desired to obtain a functional expansion for the map $p(v)(t)=y(t)$ where $v=(u,\xi)\in\mathfrak{U}\times\mathfrak{R}^n$. Clearly,

$$\dot{x}(t)=\tilde{A}(t)x(t)+\tilde{b}(t)\psi(u(t)-\tilde{c}(t)x(t)), \qquad x(0)=\xi$$
$$y(t)=\tilde{c}(t)x(t) \quad (6.2)$$

so the system is of the form (1.6)–(1.7) where $f,h\in C_{(x,u)}^{(k)}$. Since Theorems 5.1–5.3 (Theorem 5.4 when ψ is analytic) are relevant and establish the validity of the functional expansion, it remains only to determine the characterization of the variations.

The approach of Sections IV and V is repeated, but the special structure of (6.2) is exploited. Assume that (6.2) has a solution on $[0,T]$ for $\xi=\xi_0$ and $u=u_0$. Then, for $|\alpha|$ sufficiently small, $z(t,\alpha)$ and $w(t,\alpha)$ are defined by

$$\dot{z}=\tilde{A}(t)z+\tilde{b}(t)\psi(u_0(t)+\alpha u(t)-\tilde{c}(t)z),\; z(0,\alpha)=\xi_0+\alpha\xi$$
$$w=\tilde{c}(t)z. \quad (6.3)$$

Moreover, for $i=1,\cdots,k$, the derivatives $(d/d\alpha)^i z(t,\alpha)=z_i(t,\alpha)$ and $(d/d\alpha)^i w(t,\alpha)=w_i(t,\alpha)$ exist, are continuous, and are defined by differential equations obtained by differentiating (6.3) with respect to α. For example, if $k\geqslant 2$,

$$\dot{z}_1=\tilde{A}(t)z_1+\tilde{b}(t)\psi^{(1)}(u_0(t)+\alpha u(t)-w(t,\alpha))(u(t)-w_1)$$
$$w_1=\tilde{c}(t)z_1, \qquad z_1(0,\alpha)=\xi$$
$$\dot{z}_2=\tilde{A}(t)z_2+\tilde{b}(t)\big(\psi^{(1)}(u_0(t)+\alpha u(t)-w(t,\alpha))(-w_2)+\psi^{(2)}(u_0(t)+\alpha u(t)-w(t,\alpha))(u(t)-w_1)^2\big)$$
$$w_2=\tilde{c}(t)z_2, \qquad z_2(0,\alpha)=0 \quad (6.4)$$

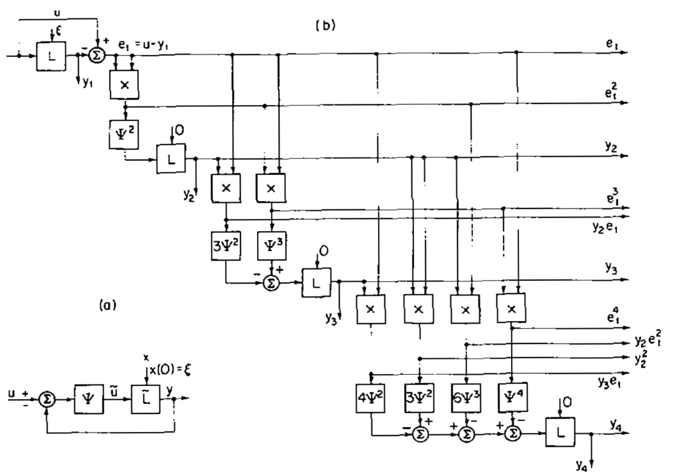

Fig. 3. (a) The feedback system. (b) Structural representation for the output variations with $u_0 + u \to u$ and $\xi_0 + \xi \to \xi$.

where $\psi^{(i)}$ is the ith derivative of ψ. Setting $x_i(t) = z_i(t,0)$ and $y_i(t) = w_i(t,0)$ gives $\delta^i p_{v_0}(v)(t) = y_i(t)$ for $i = 1, \cdots, k$. For instance, if $k \geq 4$ it follows directly that

$$\dot{x}_1 = A(t)x_1 + \tilde{b}(t)\psi_0^{(1)}(t)u(t),\ y_1 = \tilde{c}(t)x_1,\ x_1(0) = \xi$$
$$\dot{x}_2 = A(t)x_2 + \tilde{b}(t)\psi_0^{(2)}(t)(u(t) - y_1)^2,$$
$$y_2 = \tilde{c}(t)x_2,\ x_2(0) = 0$$
$$\dot{x}_3 = A(t)x_3 + \tilde{b}(t)\big(3\psi_0^{(2)}(t)(-y_2)(u(t) - y_1)$$
$$+ \psi_0^{(3)}(t)(u(t) - y_1)^3\big),\ y_3 = \tilde{c}(t)x_3,\ x_3(0) = 0$$
$$\dot{x}_4 = A(t)x_4 + \tilde{b}(t)\big(4\psi_0^{(2)}(t)(-y_3)(u(t) - z_1) + 3\psi_0^{(2)}(t)(y_2)^2$$
$$+ 6\psi_0^{(3)}(t)(-y_2)(u(t) - y_1)^2 + \psi_0^{(4)}(t)(u(t) - y_1)^4\big)$$
$$y_4 = \tilde{c}(t)x_4,\ x_4(0) = 0 \tag{6.5}$$

where

$$\psi_0^{(i)}(t) = \psi^{(i)}(u_0(t) - y_0(t))$$
$$A(t) = \tilde{A}(t) - \tilde{b}(t)\psi_0^{(1)}(t)\tilde{c}(t) \tag{6.6}$$

and $y_0(t)$ is given by

$$\dot{x}_0 = \tilde{A}(t)x_0 + \tilde{b}(t)\psi(u_0(t) - \tilde{c}(t)x_0),$$
$$y_0 = \tilde{c}(t)x_0,\ x_0(0) = \xi_0. \tag{6.7}$$

Fig. 3(b) shows the structure of the variational equations (6.5). For this representation, it is assumed that $\psi_0^{(1)}(t) \neq 0$ for $t \in [0,T]$ and $\Psi^i(t) = \psi_0^{(i)}(t)(\psi_0^{(1)}(t))^{-1}$. The linear dynamic system L describes the linearized closed-loop system, i.e., the first line of (6.5).

Because the dynamical behavior of the system (6.5) is expressed by repeated applications of L, simplifications occur in the calculation of the integral characterizations of $\delta^i p_{v_0}(v)(t)$. In particular, for $\psi_0^{(1)}(t) \neq 0$, $t \in [0,T]$, the variations can all be expressed in terms of $w_1^1(t,\sigma_1)$, the weighting function for L, and $\phi_0^1(t)$ [see (5.9)]. For example, if $\xi = 0$,

$$\delta^2 p_{v_0}(v)(t) = y_2(t) = \int_0^T w_1^1(t,\sigma)\Psi^{(2)}(\sigma)(u(\sigma) - y_1(\sigma))^2 d\sigma. \tag{6.8}$$

Using

$$\delta^1 p_{v_0}(v)(t) = y_1(t) = \int_0^T w_1^1(t,\sigma_1)u(\sigma_1)\,d\sigma \tag{6.9}$$

in this expression gives

$$\delta^2 p_{v_0}(v)(t) = \int_0^T\int_0^T w_{1,2}^2(t,\sigma_1,\sigma_2)u(\sigma_1)u(\sigma_2)\,d\sigma_1\,d\sigma_2 + \int_0^T w_{1,1}^2(t,\sigma_1)(u(\sigma_1))^2 d\sigma_1 \tag{6.10}$$

where

$$w_{1,2}^2(t,\sigma_1,\sigma_2) = -2w_1^1(t,\sigma_1)w_1^1(\sigma_1,\sigma_2)\Psi^2(\sigma_1) + \int_0^T w_1^1(t,\sigma)w_1^1(\sigma,\sigma_1)w_1^1(\sigma,\sigma_2)\Psi^2(\sigma)\,d\sigma$$
$$w_{1,1}^2(t,\sigma_1) = w_1^1(t,\sigma_1)\Psi^2(\sigma_1). \tag{6.11}$$

In the above formulas the bracket notation of (5.9) and (5.10) has been suppressed because $u(t)$ is scalar valued. Similar calculations can be made for $i > 2$.

The preceding results are easily interpreted because, contrary to Figs. 1 and 2, all the operations indicated in Fig. 3 are on scalar variables. It is seen that the effects of nonlinear distortion (measured by $y_2, y_3, y_4, \cdots$) can be kept small by keeping $|e_1|$ small. This can be achieved by high open-loop gain, which tends to make the error (e_1) in the linearized model (L) small. If L has much "resonant peaking," the distortion "near" the resonant frequency will be large. This follows for two reasons: 1) in this frequency range $|e_1|$ is large, and 2) the operator L, which amplifies the multiplier outputs, tends to make $|y_2|$, $|y_3|, |y_4|, \cdots$ large. More specific analyses of distortion effects in feedback systems are described in [21], [22], and [26].

VII. Conclusion

Sections II and III present a general methodology for obtaining functional expansions which represent the response of nonlinear differential systems. The variational approach provides a convenient recipe for deriving the concrete form of the terms in the expansions, and precise results concerning the approximation errors (Theorems 2.4 and 2.5) or convergence of the expansions (Theorem 3.3) are available. These features are illustrated by the developments in Sections IV, V, and VI. In addition, the variational formula (1.2) provides a simple justification for a variety of techniques [10], [18], [25], and [26] which have been proposed for the derivation of Volterra series. The methodology extends to other types of dynamic systems as well. For instance, discrete-time systems, similar in form to (1.4)–(1.5) or (1.6)–(1.7), can be treated in an analogous fashion [13] and the results of Section VI (e.g., Fig. 3) can be generalized to the case where $\tilde{L}$ is characterized abstractly as a linear, causal operator.

Theorems 4.1–4.4 and 5.1–5.4 give precise conditions on general classes of differential systems which assure the validity of functional expansions. Not surprisingly, these conditions are reminiscent of those which occur in the usual theory of power series. For the system (1.4)–(1.5), Theorem 4.4 gives an alternative path to Brockett's Theorem 1 [5], [6]. Theorems 4.2 and 4.3 show that many of the results in [5] are valid when f,g,h are finitely differentiable instead of analytic in x. Theorems 5.1–5.4 extend these results to the much more general system (1.6) and (1.7). For the special case of stationary systems (t does not appear in f and h) where $x_0(t),u_0(t)$ are constant, Theorems 5.1–5.4 justify developments which have appeared in many papers and reports [4], [10], [11], [18], [21], [22], [23], [25], [26]. However, in opposition to a frequently expressed belief, the functional expansions are not necessarily Volterra series (see the remarks preceding Theorem 5.1). In [7] the response of bilinear differential systems to initial conditions is characterized; Section V shows that similar characterizations can be obtained for very general nonlinear differential systems. It is worth noting that Theorems like 4.1–4.4 and 5.1–5.4 can be proved with little modification when the t dependence is more general. For example, in Section V, f, h and u need only be measurable in t if simple integrability conditions are introduced. For $k=1$, Theorems 4.2, 4.3 and 5.2, 5.3 justify rigorously the validity of linear models for nonlinear systems. Such justification is usually neglected; [8], which applies to $t\in[0,+\infty)$ instead of $t\in[0,T]$, is an exception.

The form of the variational equations has a number of important implications with respect to the general theory of nonlinear systems, which will be only suggested in what follows. The variations are given *exactly* by the solution of differential equations of *relatively* low order (compare with [5]). This is a potential advantage in developing efficient numerical techniques for the evaluation of the kernel functions which appear in the functional expansions. It also shows that if a Volterra series is realizable in the form (1.4)–(1.5), then each term in the series is individually realizable. This generalizes a conclusion contained in Brockett's Theorem 4 [5]. Alternative approaches to many other results in [5] follow from the formulas of Section IV. For example, necessary and sufficient conditions that (1.4)–(1.5) have a finite Volterra series can be given in terms of the A's, B's and C's which appear in (4.7), (4.8), (4.11), and (4.12). If it is desired to characterize the variations (4.6) and (4.10) as the solutions of bilinear differential equations, the trick used in the Carleman bilinearization [5], [17] can be applied to (4.7), (4.8), (4.11), and (4.12). To illustrate, the term $x_1^2=z_{1,1}$, which appears in Fig. 1, can be obtained by solving

$$\dot{z}_{1,1}=\dot{x}_1\times x_1+x_1\times\dot{x}_1=(Ax_1+Bu)\times x_1 + x_1\times(Ax_1+Bu)=\mathcal{A}_{1,1}z_{1,1}+\mathcal{B}_{1,1}(x_1\times v),$$

which is bilinear in $x_1,z_{1,1}$, and v. This approach differs from [5] in that the bilinear equations give the variations exactly (in [5] there is a remainder term resulting from the truncation of the infinite order Carleman system). Finally, Figs. 1 and 2 can be interpreted as general structural results concerning the realization of nonlinear, causal operators as differential systems. For instance, if an operator is a 2-power of the class (1.6)–(1.7), it can be realized by the first two "layers" of Fig. 1. This realization has similarities with the realization of bilinear operators discussed in [1], [16], and [20]. In fact, the principal results of these papers can be obtained very simply using the tools of Section V [13].

Appendix

Theorem A.1: Assume that $f:\mathfrak{R}^n\times\mathfrak{R}^m\times[0,T]\to\mathfrak{R}^n$ is continuous and has continuous first partial derivatives with respect to x and u in $\mathfrak{R}^n\times\mathfrak{R}^m\times[0,T]$. Let $\mathfrak{U}=\mathcal{C}([0,T],\mathfrak{R}^m)$ and $\mathfrak{X}=\mathcal{C}([0,T],\mathfrak{R}^n)$. Suppose that (1.6) has a solution x_0 for $u=u_0\in\mathfrak{U}$ and $\xi=\xi_0\in\mathfrak{R}^n$. Then there exist constants $\rho>0$ and $K>0$ such that for all $u\in\mathfrak{U}$ and $\xi\in\mathfrak{R}^n$ satisfying $\|u-u_0\|+|\xi-\xi_0|<\rho$, 1) the system (1.6) has a solution $x\in\mathfrak{X}$, and 2) $\|x-x_0\|<K(\|u-u_0\|+|\xi-\xi_0|)$.

Proof: Omitted. The general idea is to write an integral equation for $\hat{x}=x-x_0$ and show that for $\|u-u_0\|$ and $|\xi-\xi_0|$ sufficiently small it corresponds to $\hat{x}=T(\hat{x})$ where T is a contraction.

Remark A.2: By an obvious change in notation, Theorem A.1 applies to (1.4).

Remark A.3: For $\|u-u_0\|+|\xi-\xi_0|<\rho$, it follows that $\|x\|\leqslant K_0=K\rho+\|x_0\|$.

Remark A.4: The theorem is valid if $\mathfrak{R}$ is replaced by $\mathcal{C}$.

Acknowledgment

This research was carried out while the author was on leave in the Department of Electrical Engineering, The Johns Hopkins University, Baltimore, MD. Conversations with W. J. Rugh were most helpful.

References

[1] M. A. Arbib, "A characterization of multilinear systems," *IEEE Trans. Automat. Contr.*, vol. AC-14, pp. 699–702, 1969.
[2] A. V. Balakrishnan, *Applied Functional Analysis*. New York: Springer Verlag, 1976.
[3] E. Bedrosian and S. O. Rice, "The output properties of Volterra systems driven by harmonic and Gaussian inputs," *Proc. IEEE*, vol. 59, pp. 1688–1707, 1971.
[4] M. B. Brilliant, "Theory of the analysis of nonlinear systems," M.I.T. Res. Lab. Electron., Cambridge, MA, Tech. Rep. 345, Mar. 1958.
[5] R. W. Brockett, "Volterra series and geometric control theory," *Automatica*, vol. 12, pp. 167–176, 1976.
[6] R. W. Brockett and E. G. Gilbert, "An addendum to Volterra series and geometric control theory," *Automatica*, vol. 12, p. 635, 1976.

[7] C. Bruni, G. DiPillo, and G. Koch, "On the mathematical models of bilinear systems," *Ricerche di Automatica*, vol. 2, pp. 11–26, 1971.
[8] C. A. Desoer and K. K. Wong, "Small signal behavior of nonlinear lumped networks," *Proc. IEEE*, vol. 56, pp. 14–22, 1968.
[9] J. Dieudonné, *Foundations of Modern Analysis*. New York: Academic, 1969.
[10] R. H. Flake, "Volterra series representation of time-varying nonlinear systems," in *Proc. 2nd Int. IFAC Congr.*, 1963, pp. 91–99.
[11] D. A. George, "Continuous nonlinear systems," M.I.T. Res. Lab. Electron., Cambridge, MA, Tech. Rep. 355, July 1959.
[12] E. G. Gilbert, "Volterra series and the response of nonlinear differential systems: A new approach," in *Proc. of the 1976 Conf. on Information Sciences and Systems*, The Johns Hopkins Univ., Baltimore, MD, pp. 394–400.
[13] ——, "Bilinear and 2-power input-output maps: Finite dimensional realizations and the role of functional series," *IEEE Trans. Automat. Contr.*, to be published, Feb. 1978.
[14] L. M. Graves, "Riemann integration and Taylor's theorem in general analysis," *Trans. Amer. Math. Soc.*, vol. 29, pp. 163–177, 1927.
[15] E. Hille and R. S. Phillips, *Functional Analysis and Semi-Groups*. Providence, RI: Amer. Math. Soc., 1957.
[16] R. E. Kalman, "Pattern recognition properties of multilinear machines," presented at the IFAC Symp. on Technical and Biological Problems of Control, Yerevan, Armenian SSR, 1968.
[17] A. J. Krener, "Linearization and bilinearization of control systems," in *Proc. 1974 Allerton Conf. on Circuit and System Theory*, 1974, pp. 834–843.
[18] Y. H. Ku and A. A. Wolf, "Volterra-Wiener functionals for the analysis of nonlinear systems," *J. Franklin Inst.*, vol. 281, pp. 9–26, 1966.
[19] S. Lefschetz, *Differential Equations: Geometric Theory*. New York: Interscience, 1957.
[20] G. Marchesini and G. Picci, "Some results on the abstract realization theory of multilinear systems," in *Theory and Application of Variable Structure Systems*, R. R. Mohler and A. Ruberti, Ed. New York: Academic, 1972, pp. 109–135.
[21] S. Narayanan, "Transistor distortion analysis using Volterra series representation," *Bell Syst. Tech. J.*, vol. 46, pp. 991–1023, 1967.
[22] S. Narayanan, "Application of Volterra series to intermodulation distortion analysis of a transistor feedback amplifier," *IEEE Trans. Circuit Theory*, vol. CT-17, pp. 518–527, 1970.
[23] R. B. Parente, "Nonlinear differential equations and analytic systems theory," *SIAM J. Appl. Math.*, vol. 18, pp. 41–66, 1970.
[24] W. A. Porter, "An overview of polynomic systems theory," *Proc. IEEE*, vol. 64, pp. 18–23, 1976.
[25] K. V. Rao and R. R. Mohler, "On the synthesis of Volterra kernels of bilinear systems," *Automat. Contr. Theory and Appl.*, vol. 3, pp. 44–46, 1975.
[26] H. L. Van Trees, "Functional techniques for the analysis of the nonlinear behavior of phase-locked loops," *Proc. IEEE*, vol. 52, pp. 894–911.

Elmer G. Gilbert was born in Joliet, IL. He received the Ph.D. degree from the University of Michigan, Ann Arbor, in 1957.

Since 1957 he has been on the faculty of the College of Engineering, the University of Michigan, and is now a Professor in the Department of Aerospace Engineering. Most of his teaching and research activities are in the Computer, Information, and Control Engineering Program at Michigan. His most recent leave, 1974–1976, was as a Visiting Professor in the Department of Electrical Engineering, The Johns Hopkins University, Baltimore, MD. His current research interests are focused in two main areas: optimal periodic control and the theory of nonlinear dynamical systems.

Dr. Gilbert has served on a number of national committees and until this year was a member of the Editorial Board for the *SIAM Journal on Control and Optimization*.

Convergence of Recursive Adaptive and Identification Procedures Via Weak Convergence Theory

HAROLD J. KUSHNER, FELLOW, IEEE

Abstract—**Results and concepts in the theory of weak convergence of a sequence of probability measures are applied to convergence problems for a variety of recursive adaptive (stochastic approximation-like) methods. Similar techniques have had wide applicability in areas of operations research and in some other areas in stochastic control. It is quite likely that they will play a much more important role in control theory than they do at present, since they allow relatively simple and natural proofs for many types of convergence and approximation problems. Part of the aim of the paper is tutorial: to introduce the ideas and to show how they might be applied. Also, many of the results are new, and they can all be generalized in many directions.**

Manuscript received November 22, 1976; revised April 25, 1977. Paper recommended by Y. Bar-Shalom, Chairman of the Stochastic Control Committee. This work was supported in part by the Air Force Office of Scientific Research under Grant AF-AFOSR-76-3063, by the National Science Foundation under Grant Eng-73-03846-A01, and by the Office of Naval Research under Grant NONR N000 14-76-C-0279.
The author is with the Division of Applied Mathematics and Engineering, Brown University, Providence, RI 02912.

I. Introduction

The aims of this paper are twofold. The first aim is tutorial. The technique of and the results in the theory of *weak convergence of a sequence of probability measures* have found many useful applications in many areas of operations research and statistics [1], [2]. Their role in control theory has been relatively limited, being confined mainly to the work in [3], [4] which deals with control problems on diffusion models. Yet, its intrinsic power, as well as the nature of the past successes, suggest that its role in control theory should be deeper than it is at present. The techniques are particularly valuable when convergence or approximation ideas are being dealt with.

In order to illustrate the possibilities, the ideas of weak convergence theory will be applied (the second goal of the