Identification and Control of a Nonlinear Discrete-Time System Based on its Linearization: A Unified Framework

Lingji Chen, *Member, IEEE,* and Kumpati S. Narendra, *Life Fellow, IEEE*

***Abstract*—This paper presents a unified theoretical framework for the identification and control of a nonlinear discrete-time dynamical system, in which the nonlinear system is represented explicitly as a sum of its linearized component and the residual nonlinear component referred to as a "higher order function." This representation substantially simplifies the procedure of applying the implicit function theorem to derive local properties of the nonlinear system, and reveals the role played by the linearized system in a more transparent form. Under the assumption that the linearized system is controllable and observable, it is shown that: 1) the nonlinear system is also controllable and observable in a local domain; 2) a feedback law exists to stabilize the nonlinear system locally; and 3) the nonlinear system can exactly track a constant or a periodic sequence locally, if its linearized system can do so. With some additional assumptions, the nonlinear system is shown to have a well-defined relative degree (delay) and zero-dynamics. If the zero-dynamics of the linearized system is asymptotically stable, so is that of the nonlinear one, and in such a case, a control law exists for the nonlinear system to asymptotically track an arbitrary reference signal exactly, in a neighborhood of the equilibrium state. The tracking can be achieved by using the state vector for feedback, or by using only the input and the output, in which case the nonlinear autoregressive moving-average (NARMA) model is established and utilized. These results are important for understanding the use of neural networks as identifiers and controllers for general nonlinear discrete-time dynamical systems.**

***Index Terms*—Control, discrete-time, identification, linearization, neural networks, nonlinear autoregressive moving-average (NARMA) model, nonlinear dynamical systems, normal form, tracking.**

I. Introduction

IN 1990, it was first proposed [1] that neural networks should be used as components in dynamical systems, and that identification and control of such systems using neural networks should be undertaken within a unified framework of systems theory. Since that time, a great deal of progress has been made both in the theory and the practice of neural network based control (e.g., [2]–[7]). In a number of papers, system theoretic issues such as stabilization [8], representation and identification of single-input–single-output (SISO) [9] and multivariable [10] nonlinear systems, existence of controllers for exact and asymptotic tracking [11], and disturbance rejection [12] have been addressed. In all the above cases, the results are local in nature, and are valid in a domain around the equilibrium state. More specifically, let a nonlinear system Σ be described by

$$\begin{aligned} \Sigma : x(k+1) &= F(x(k), u(k)) \\ y(k) &= H(x(k)) \end{aligned} \tag{1}$$

where the input $u(k)$, the state $x(k)$, and the output $y(k)$ at time k belong respectively to $\mathbb{R}^r$, $\mathbb{R}^n$, and $\mathbb{R}^m$, the functions $F(\cdot)$ and $H(\cdot)$ are smooth, and the origin is an equilibrium state. The local results for Σ are based on properties of its linearized system Σ_L described by

$$\begin{aligned} \Sigma_L : x(k+1) &= Ax(k) + Bu(k) \\ y(k) &= Cx(k) \end{aligned} \tag{2}$$

where

$$\begin{aligned} A &\triangleq \frac{\partial F}{\partial x}\Big|_{(x=0,u=0)} \\ B &\triangleq \frac{\partial F}{\partial u}\Big|_{(x=0,u=0)} \\ C &\triangleq \frac{\partial H}{\partial x}\Big|_{x=0} . \end{aligned}$$

The results are derived, for the most part, by using the Implicit Function Theorem: First the nonlinear mapping of interest is derived (which usually involves compositing nonlinear mappings $F(\cdot)$ and $H(\cdot)$), and then the Jacobian of this mapping is obtained and identified with its counterpart in the linearized system. Since Σ_L is assumed to be controllable and observable, the Jacobian can be shown to be nonsingular and, therefore, the implicit function theorem can be applied to derive the desired local results.

While the procedure is conceptually straightforward, determining the Jacobian is known to be an involved process, as is evident from [11]. In this paper an attempt is made to simplify the above procedure and to display in a more revealing fashion the role played by the linearized system Σ_L. Toward this end, we rewrite the system equations for Σ in the following equivalent form

$$\begin{aligned} \Sigma : x(k+1) &= Ax(k) + Bu(k) + f(x(k), u(k)) \\ y(k) &= Cx(k) + h(x(k)) \end{aligned} \tag{3}$$

where the nonlinear functions $f(\cdot)$ and $h(\cdot)$ are naturally defined by the differences, and will be called "higher order functions."

Manuscript received June 5, 2002; revised August 11, 2003. This work was supported by the NSF by Grant ECS-0113239.

L. Chen is with Scientific Systems Company, Inc., Woburn, MA 01801 USA.
K. S. Narendra is with Yale University, New Haven, CT 06510 USA.

Digital Object Identifier 10.1109/TNN.2004.826206

A block-diagram representation of a SISO system is shown in Fig. 1.

There are several reasons for using the representation given in (3) when considering system theoretic properties and problems of identification, stabilization, regulation, and tracking.

i) From a pedagogic point of view, many of the results that are currently available can be derived in a unified fashion by using such a representation, so that the underlying theoretical concepts become transparent.
ii) Since the best developed part of control theory deals with linear systems [of the form Σ_L described in (2)], the methods proposed in i) may suggest how the same concepts can also be applied to problems in the nonlinear domain.
iii) Once the basic concepts become clear, it is possible to suggest new methods for solving difficult identification and control problems. For an example, see [13].
iv) After years of computer simulation studies in which neural networks have been used to identify and control dynamical systems, there is general agreement at present in the community that linear identifiers and controllers should be first designed before nonlinear identification and control is attempted. The representation of the system Σ as given in (3) lends itself readily to such control.

This paper is organized as follows. Section II presents mathematical preliminaries and problem statement. Section III examines system theoretic properties such as controllability and observability. The problem of tracking a constant, a periodic sequence, and an arbitrary reference signal are considered in Section IV when the state vector is accessible. The case when only the input and the output are available is considered in Section V. Conclusions are drawn in Section VI.

II. Mathematical Preliminaries and Statement of the Problem

A. Higher Order Functions and Two Theorems

Since we will explicitly use addition, multiplication, and composition of "higher order terms" as functions of their arguments, we now formally define "higher order functions" as follows.

Definition 1: A continuously differentiable function $F(x): \mathbb{R}^m \to \mathbb{R}^n$ is called a "Higher Order Function" if: a) $F(0) = 0$, and b) $\partial F/\partial x|_{(x=0)} = 0$. We denote the class of higher order functions by $\mathcal{H}$.

Thus, any function can be equivalently written as a sum of a linear function and a higher order function. When the system in (1) is written in the form of (3), the functions $f(\cdot)$ and $h(\cdot)$ are higher order functions.

The following properties of higher order functions $\mathcal{H}$ are found to be useful and can be verified in a straightforward manner:

i) If A is a constant matrix and $F(\cdot) \in \mathcal{H}$, then $AF(\cdot) \in \mathcal{H}$.
ii) If $F_1, F_2 \in \mathcal{H}$, then $F_1F_2 \in \mathcal{H}$ and $F_1 + F_2 \in \mathcal{H}$.
iii) If $F_1(\cdot) \in \mathcal{H}$ and $F_2(0) = 0$ and is continuously differentiable, then the composition $F_1(F_2(\cdot)) \in \mathcal{H}$.

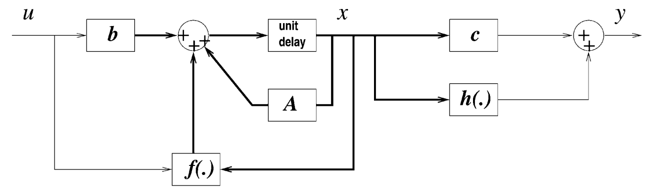

Fig. 1. A SISO system represented using higher order function.

Two theorems that play an important role in nonlinear identification and control are the inverse function theorem and the implicit function theorem [14]. We now present their specialized forms for cases where some of the relevant functions belong to $\mathcal{H}$.

Theorem 1 (Inverse Function Theorem): In some neighborhood $U_1 \subset \mathbb{R}^n$ of the origin, let

$$Ax + f(x) = y, \quad x \in U_1, \ y \in \mathbb{R}^n$$

where A is a nonsingular matrix and $f(\cdot) \in \mathcal{H}$. Then there exists an open set $U \subset U_1$ containing the origin such that the set V, defined by $V \triangleq AU + f(U)$, is open, and for all $x \in U$

$$x = A^{-1}y + g(y), \quad y \in V, \ g(\cdot) \in \mathcal{H}.$$

Theorem 2 (Implicit Function Theorem): Let U_1 be an open subset of $\mathbb{R}^{n+k}$ containing the origin. Let an element of U_1 be denoted by (x, y) with $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^k$. Let $F(x, y) = Ax + By + f(x, y)$ be a function from U_1 to $\mathbb{R}^n$, where A is a nonsingular matrix and $f(\cdot) \in \mathcal{H}$. Then there exists an open set $U \subset \mathbb{R}^k$ containing the origin such that

$$x = -A^{-1}By + g(y), \quad y \in U, \ g(\cdot) \in \mathcal{H}$$

satisfies the equation $F(x, y) = 0$.

By the above two theorems, when the underlying nonlinear functions are expressed as the sum of linear and higher order functions, the application of the inverse function theorem and the implicit function theorem becomes a matter of simple manipulation of equations. No differentiation or function evaluation is needed, because the Jacobian matrix is explicitly given. Note also that the vector y can be a concatenation of several vectors. This leads to the following corollary:

Corollary 1: If $Ax + By + f(x, y) = z$, $f(\cdot) \in \mathcal{H}$, and A is nonsingular, then locally $x = A^{-1}(z - By) + \bar{g}(y, z)$, $\bar{g}(\cdot) \in \mathcal{H}$.

Sometimes it is helpful to inspect a system in its block diagram form. For linear systems, the inverse operations of "sum" and "gain" can be obtained in a straightforward manner. It is no longer the case for systems with nonlinear operations. However, with the help of the above theorems, some inverse operations (in a neighborhood of the equilibrium point) can be obtained by inspection, and the relationship between the signals x and y can be established as in Fig. 2.

B. Statement of the Problem

Let a single-input–single-output (SISO) nonlinear systems be described by (3). The order of the system and the nonlinear functions are all assumed to be known. We are interested in the following problems related to Σ:

i) system representation;
ii) stabilization using state feedback;

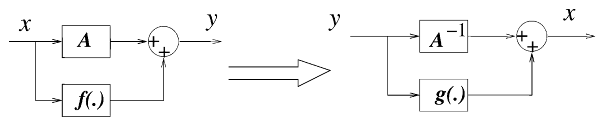

Fig. 2. Inverse function theorem expressed in block diagram form.

iii) existence of a controller for regulation;
iv) existence of a controller for tracking.

We will first study the case when the state vector x is accessible, and then study the case when only the input and the output are accessible.

Prior to addressing the problems stated above, we consider system theoretic properties of Σ that are important for all the discussions.

III. System Theoretic Properties

A. Controllability

Qualitatively speaking, controllability is the ability to transfer the state of a system from one point to another in the state space. For the nonlinear system Σ represented by (3), the following definition of local controllability is relevant.

Definition 2 (Local Controllability): A system is locally controllable if there exists a neighborhood Ω_c of the origin such that, for any $p,\ q\in\Omega_c$ there is a sequence of inputs $U\triangleq[u(0),u(1),\ldots,u(m)]$ that transfers the system from p to q.

It is, in general, quite difficult to demonstrate the controllability of a complex nonlinear system [15]. However, if the linearized system Σ_L is controllable, sufficient conditions can be stated for Σ, as shown below in Theorem 3.

Theorem 3: If the system Σ_L is controllable, then the system Σ is locally controllable.

Proof: We will show the existence of a control sequence U that transfers the SISO system from $x(0)$ to $x(n)$ in n steps. The properties of higher order functions discussed in Section II will be repeatedly used in the following derivations.

$$\begin{aligned}
x(1) &= Ax(0)+bu(0)+f(x(0),u(0))\\
x(2) &= Ax(1)+bu(1)+f(x(1),u(1))\\
&= A^2x(0)+Abu(0)+Af(x(0),u(0))+bu(1)\\
&\quad +f(Ax(0)+bu(0)+f(x(0),u(0)),u(1))\\
&\triangleq A^2x(0)+Abu(0)+bu(1)+f_2(x(0),u(0),u(1))\\
&\ \ \vdots\\
x(n) &= Ax(n-1)+bu(n-1)+f(x(n-1),u(n-1))\\
&= A^nx(0)+[\,A^{n-1}b\quad\cdots\quad Ab\quad b\,]\begin{bmatrix}u(0)\\ \vdots\\ u(n-2)\\ u(n-1)\end{bmatrix}\\
&\quad +f_n(x(0),u(0),\ldots,u(n-2),u(n-1))\\
&\triangleq A^nx(0)+W_cU+f_n(x(0),U)
\end{aligned}\tag{4}$$

where

$$\begin{aligned}
W_c &\triangleq [\,A^{n-1}b\quad\cdots\quad Ab\quad b\,]\\
U &\triangleq [\,u(0)\quad\cdots\quad u(n-2)\quad u(n-1)\,]^T
\end{aligned}$$

and $f_2(\cdot),\ldots,f_n(\cdot)$ are all properly defined higher order functions. Since (A,b) is controllable, its controllability matrix W_c is nonsingular. By Corollary 1 in Section II, in a neighborhood Ω_c of the origin in the space $\mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^n$, the desired control sequence is given by

$$U=W_c^{-1}\left(x(n)-A^nx(0)\right)+g(x(0),x(n)),\quad g(\cdot)\in\mathcal{H}.\tag{5}$$

■

The above result is well known. However, it has been derived in a much more straightforward manner in the proof, without cumbersome function composition and Jacobian calculation. The same is also true for many other results presented in the rest of the paper.

B. Stability and Stabilizability

The linear unforced system

$$x(k+1)=Ax(k)\tag{6}$$

is asymptotically stable if (and only if) the eigenvalues of the matrix A lie within the unit circle of the complex plane, or equivalently, if (and only if) the Lyapunov equation

$$A^TPA-P=-Q,\quad Q=Q^T>0$$

has a positive definite solution $P=P^T>0$. In such a case

$$V(x)=x^TPx$$

is a Lyapunov function for (6). The same Lyapunov function can be used to prove the (local) asymptotic stability of the nonlinear system

$$x(k+1)=Ax(k)+\eta(x(k)),\quad \eta(\cdot)\in\mathcal{H}$$

since

$$\begin{aligned}
V(x(k+1))-V(x(k)) &= -x^TQx+2x^TA^TP\eta+\eta^TP\eta\\
&<0
\end{aligned}\tag{7}$$

in a neighborhood of the origin[1].

A controllable linear system can be stabilized by a feedback control law of the form

$$u(k)=\Gamma x(k)\tag{8}$$

so that the closed-loop system

$$x(k+1)=(A+b\Gamma)x(k)$$

is asymptotically stable. It follows that the same feedback law stabilizes Σ also, since the closed-loop system

$$x(k+1)=(A+b\Gamma)x(k)+f(x(k),\Gamma x(k)),\quad f(\cdot)\in\mathcal{H}$$

is also asymptotically stable in a neighborhood of the origin, as discussed earlier.

For discrete-time linear systems, the feedback gain Γ can be chosen so that $A_s\triangleq A+b\Gamma$ becomes nil-potent, i.e., $A_s^n=0$. In such a case, any state can be brought back to the origin in no more than n steps by a state feedback law ("deadbeat control"). This result can be extended to the nonlinear system Σ.

[1]By definition of higher order functions, for any $\epsilon>0$, there exists a neighborhood of the origin in which η is bounded by ϵx.

Theorem 4 (Finite Stabilizability): If the system Σ_L is controllable, then the system Σ can be stabilized in at most n steps by a state feedback law.

Proof: Let $x(n)$ be zero in (5):

$$U = -W_c^{-1}A^n x(0) + g(x(0), 0).$$

This is the (unique) control sequence of length n that brings any state $x(0)$ to the origin. The first element of this sequence is

$$U(0) \triangleq \Gamma x(0) + \gamma(x(0)), \quad \gamma(\cdot) \in \mathcal{H}.$$

We claim that the state feedback law

$$u(k) = \Gamma x(k) + \gamma(x(k))$$

stabilizes the system in at most n steps, as a consequence of the following two facts:

i) the solution (5) is unique in the neighborhood Ω_c of the origin;
ii) the origin is an equilibrium state.

For a complete proof, see Appendix. ■

C. Observability

Definition 3 (Local Observability): If $x(0) \in \Omega$ can be re-constructed from a sequence of observations $y(0), y(1), \ldots, y(l)$ and $u(0), u(1), \ldots, u(l-1)$, then the system is locally observable.

Theorem 5: If the system Σ_L is observable, then the system Σ is locally observable.

Proof: The procedure follows along the same line as in the case of controllability in Section III-A. Define the vectors

$$Y_{(0,n-1)} \triangleq [y(0), y(1), \ldots, y(n-1)]^T$$
$$U_{(0,n-2)} \triangleq [u(0), u(1), \ldots, u(n-2)]^T$$

and the matrix

$$P = \begin{bmatrix} 0 & 0 & \cdots & 0 \\ cb & 0 & \cdots & 0 \\ cAb & cb & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ cA^{n-2}b & \cdots & cAb & cb \end{bmatrix}$$

and it can be shown that

$$Y_{(0,n-1)} = W_o x(0) + PU_{(0,n-2)} + h(x(0), U_{(0,n-2)})$$
$$h(\cdot) \in \mathcal{H}$$

where

$$W_o \triangleq \begin{bmatrix} c \\ cA \\ \cdots \\ cA^{n-1} \end{bmatrix}$$

is the observability matrix W_o of Σ_L, which is assumed to be nonsingular. Therefore

$$x(0) = W_o^{-1}(Y_{(0,n-1)} - PU_{(0,n-2)}) + \eta(Y_{(0,n-1)}, U_{(0,n-2)}) \quad (9)$$

where $\eta(\cdot) \in \mathcal{H}$. ■

In summary, the results stated in this section merely make precise, in specific contexts, the well-known linearization principle, that design based on linearizations works locally for nonlinear systems. The controllability (stabilizability) and observability of Σ_L assure the local controllability (stabilizability) and observability of Σ.

IV. Regulation and Tracking; State Vector Accessible

As stated earlier, our interest is in examining critically the role played by Σ_L in many of the identification and control problems commonly addressed in the context of the nonlinear system Σ. We first consider the problem of set-point regulation or regulation around an equilibrium state that is different from the origin.

A. Set-Point Regulation

The output $y(k)$ of a stabilizable linear system Σ_L can be regulated around a constant value r asymptotically, using static state feedback, if $c(I-A)^{-1}b \neq 0$, or equivalently, if the transfer function of Σ_L, $H(z) \triangleq c(zI-A)^{-1}b$, does not have a zero at $z = 1$.

We now consider the set-point regulation of the nonlinear system Σ described by (3). Since Σ_L is stabilizable, so is Σ. Let

$$u = \Gamma x + v$$

be the input to Σ, where $\bar{A} \triangleq A + b\Gamma$ is a stable matrix (and hence $(I-\bar{A})^{-1}$ is well-defined), and v is a constant. The closed loop system has the form (for notational convenience, we drop the time k and denote $k+1$ by a superscript $+$)

$$x^+ = \bar{A}x + bv + f(x, \Gamma x + v)$$
$$y = cx + h(x).$$

We will first show that with a constant input v, the state x, and hence the output y, will tend to constant values, and then the "nonlinear gain" from v to y can be determined. For a discrete-time system, consider the equality

$$x^* = \bar{A}x^* + bv + f(x^*, \Gamma x^* + v).$$

Let

$$x^* = (I-\bar{A})^{-1}bv + \eta(v)$$

be an explicit representation of x^* in terms of v by using the implicit function theorem in Section II. With the coordinate transformation

$$z = x - x^*$$

x^* is transformed into the origin of the z-coordinates. The representation of the system now becomes (with definitions $f_x \triangleq \partial f(x,u)/\partial x$ and $f_u \triangleq \partial f(x,u)/\partial u$)

$$\begin{aligned} z^+ &= \bar{A}z + f(z + x^*, \Gamma z + \Gamma x^* + v) - f(x^*, \Gamma x^* + v) \\ &= \left(\bar{A} + f_x(x^*, \Gamma x^* + v) + f_u(x^*, \Gamma x^* + v)\Gamma\right) z + o(z^2) \\ &\triangleq \bar{A}(v)z + o(z^2) \end{aligned}$$

where $\bar{A}(v)$ is a constant matrix for any given constant v. Since $f(\cdot) \in \mathcal{H}$ and, hence,$f_x(0,0)$ and $f_u(0,0)$ are zero, by continuity, $\bar{A}(v)$ remains a stable matrix for sufficiently small v, and, therefore

$$\lim_{k\to\infty} z(k) = 0$$

or

$$\lim_{k\to\infty} x(k) = x^*.$$

Thus we have shown that, after the system is stabilized by state feedback, a constant input produces a constant output for the nonlinear system Σ. Next we will examine the "nonlinear gain" of the system so that a desired constant output can be produced by an appropriate constant input. Let r be a constant reference output that $y(k)$ is to follow asymptotically. This implies that x^* must be such that

$$r = cx^* + h(x^*) = c(I-\bar{A})^{-1}bv + c\eta(v) + \bar{h}(v) \qquad (10)$$

where

$$\bar{h}(v) \triangleq h((I-\bar{A})^{-1}bv + \eta(v)).$$

Since $\eta, \bar{h} \in \mathcal{H}$, a solution for v in terms of r (in a neighborhood of the origin) to (10) exists if

$$c(I-\bar{A})^{-1}b \neq 0.$$

This corresponds to the condition that the transfer function $\bar{H}(z) = c(zI-\bar{A})^{-1}b$ does not have a zero at $z=1$, or equivalently, that the transfer function $H(z) = c(zI-A)^{-1}b$ does not have a zero at $z=1$ (since zeros are invariant under state feedback). Thus, we have the following theorem.

Theorem 6 (Set-Point Regulation): The output of the nonlinear system Σ can be regulated around a constant value r in a neighborhood of the origin, if this can be achieved for the linearized system Σ_L.

B. Periodic Sequence Tracking

The above result can be generalized to the tracking of any periodic sequence in a neighborhood of the origin.

Theorem 7 (Periodic Sequence Tracking): Given system Σ and any N-periodic sequence $y^*(k)$, the output of Σ can track $y^*(k)$ if the transfer function of Σ_L has no zero that coincides with an N^{th} root of 1.

Proof: For a detailed proof, see [16]. The essence of the proof is to examine the evolution of the system on an N-time scale: at $k=1, N+1, 2N+1, \ldots$, at $k=2, N+2, 2N+2, \ldots$, *etc.,* A periodic input to the system translates to constant inputs to the above N-time scale subsystems and, therefore, the results obtained for the case of regulation can be extended to it. ■

C. Tracking an Arbitrary Signal $y^*(k)$

We have thus far considered the regulation problem and the problem of tracking an N-periodic signal. We now proceed to consider the more general case where the output $y^*(k)$ to be followed is any arbitrary signal. However, further constraint on the structure of Σ has to be imposed before the problem can be posed precisely. This, in turn, leads to the concepts of relative degree, normal form, and zero dynamics of the system, all of which are central, both to the statement as well as the solution of the general tracking problem [11], [17], [18]. We discuss these concepts first in the case of Σ_L, before extending them to Σ.

1) Relative Degree, Normal Form, and Zero Dynamics for Σ_L: Consider the linear system (Σ_L) described by

$$\begin{aligned} x^+ &= Ax + bu \\ y &= cx. \end{aligned}$$

Let $z_1 \triangleq y = cx$. Then

$$z_1^+ = cx^+ = cAx + cbu.$$

If $cb = 0$, let $z_2 \triangleq cAx$. Then

$$z_2^+ = cA^2x + cAbu.$$

Similarly, if $cAb = 0$, let $z_3 \triangleq cA^2x$. Then

$$z_3^+ = cA^3x + cA^2bu.$$

As an example, assume that $cA^2b \neq 0$. Since $cb = 0$, and $cAb = 0$, it follows that the fact $cA^2b \neq 0$ implies that c, cA and cA^2 are linearly independent. Let T be a nonsingular matrix such that c, cA and cA^2 are its first three rows. Then, the transformation

$$z = Tx$$

will transform the system into the normal form

$$\begin{aligned} z_1^+ &= z_2 \\ z_2^+ &= z_3 \\ z_3^+ &= P\bar{z} + Q\eta + cA^2bu \\ \eta^+ &= R\bar{z} + S\eta + Ku \\ y &= z_1 \end{aligned} \qquad (11)$$

where $\bar{z} = [z_1, z_2, z_3]^T$, $\eta = [z_4, z_5, \ldots, z_n]^T$, and $z^T = [\bar{z}^T, \eta^T]$ is the state of the system, and P and Q are row vectors in $\mathbb{R}^3$ and $\mathbb{R}^{n-3}$ respectively, and $R \in \mathbb{R}^{(n-3)\times 3}$, $S \in \mathbb{R}^{(n-3)\times(n-3)}$ and $K \in \mathbb{R}^{n-3}$. Because of linearity, it can be shown that T can be chosen so that $K = 0$, and this is assumed in the following discussions.

We have assumed $cb = 0$, $cAb = 0$ and $cA^2b \neq 0$ for illustration. In a more general case we would have

$$cA^ib = 0, \quad (i = 0, i, \ldots, d-2)$$

and

$$cA^{d-1}b \neq 0.$$

In such a case, $P^T \in \mathbb{R}^d$, $Q^T \in \mathbb{R}^{n-d}$, $R \in \mathbb{R}^{(n-d)\times d}$, $S \in \mathbb{R}^{(n-d)\times(n-d)}$, $K \in \mathbb{R}^{n-d}$, and

$$z_i^+ = z_{i+1} \quad (i = 1, 2, \ldots, d-1)$$

while the equations for z_d^+ and η^+ are given by (11).

The objective is for $y(k)$ to follow a desired trajectory $y^*(k)$. Intuitively this can be achieved by setting

$$u(k) = (cA^2b)^{-1}(y^*(k+3) - P\bar{z} - Q\eta). \qquad (12)$$

However, the subsystem associated with η in Fig. 3 is driven by an arbitrary signal y^*, and can become unbounded. To pose conditions that guarantee the boundedness of η and other signals, we proceed as follows. Choose $y^*(k)$ to be identically zero.

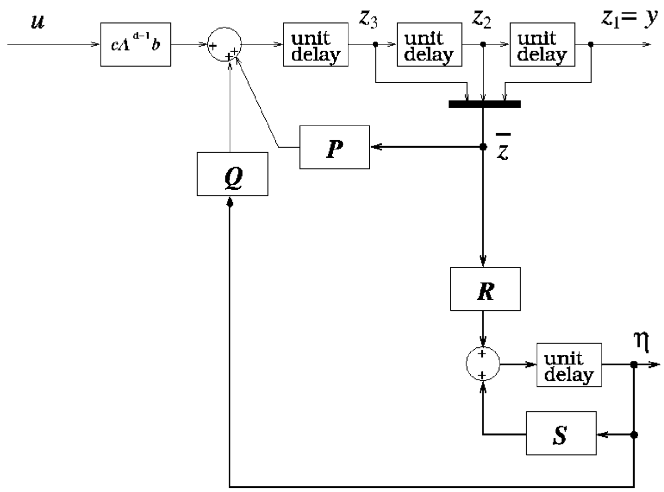

Fig. 3. Linear normal form.

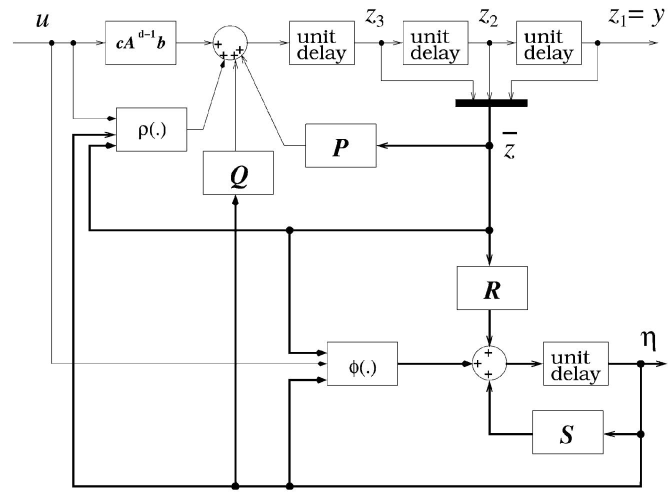

Fig. 4. Nonlinear normal form.

From (11) and the fact that $\bar{z} = 0$ and $K = 0$ we have a system evolving as

$$\eta(k+1) = S\eta(k).$$

This is called the *zero dynamics* of the linear system. If S is an asymptotically stable matrix, it follows that $\lim_{k\to\infty}\eta(k) = 0$. In such a case, even when $y^*(k)$ is nonzero but bounded, $\eta(k)$ is a bounded sequence, and so is the control input $u(k)$.

2) Relative Degree, Normal Form, and Zero Dynamics for Σ*:* The results obtained in Section III for Σ_L are global in nature. In this section, we show that similar results can be obtained, in a neighborhood of the origin, for the nonlinear system Σ.

Let Σ be described by

$$\begin{aligned} x^+ &= Ax + bu + f(x,u) \\ y &= cx + h(x). \end{aligned} \tag{13}$$

Assume that

$$cb = 0, \quad cAb = 0, \quad cA^2b \neq 0.$$

Let

$$z_1 \triangleq y = cx + h(x) \triangleq cx + p^{(0)}(x).$$

Then

$$\begin{aligned} z_1^+ &= cx^+ + p^{(0)}(x^+) \\ &= cAx + cbu + cf(x,u) + p^{(0)}(Ax + bu + f(x,u)) \\ &\triangleq cAx + \bar{p}^{(1)}(x,u) \end{aligned}$$

where $p^{(0)}(\cdot), \bar{p}^{(1)}(\cdot) \in \mathcal{H}$. If, in a neighborhood of the origin

$$\frac{\partial \bar{p}^{(1)}(x,u)}{\partial u} \equiv 0$$

i.e.,

$$\left(c + p_x^{(0)}(Ax + bu + f(x,u))\right)(b + f_u) \equiv 0 \tag{14}$$

where subscripts denote Jacobians, it follows that $z_1^+ = cAx + p^{(1)}(x)$, $p^{(1)}(\cdot) \in \mathcal{H}$. In other words, $\bar{p}^{(1)}(x,u)$ is only a function of x. A necessary (but not sufficient) condition for (14) to hold in a neighborhood of the origin is that $cb = 0$.

We now define $z_2 \triangleq cAx + p^{(1)}(x)$ to obtain

$$\begin{aligned} z_2^+ &= cA^2x + cAbu + cAf(x,u) + p^{(1)}(Ax + bu + f(x,u)) \\ &\triangleq cA^2x + \bar{p}^{(2)}(x,u) \\ &= cA^2x + p^{(2)}(x). \end{aligned}$$

The last equality holds if

$$\frac{\partial \bar{p}^{(2)}(x,u)}{\partial u} \equiv 0$$

in a neighborhood of the origin, or equivalently

$$\left(cA + p_x^{(1)}(Ax + bu + f(x,u))\right)(b + f_u) \equiv 0.$$

Again, a necessary but not sufficient condition for the above to hold is that $cAb = 0$.

Continuing in this manner, if $cA^2b \neq 0$, then

$$\left(cA^2 + p_x^{(2)}(Ax + bu + f(x,u))\right)(b + f_u) \neq 0$$

in a neighborhood of the origin, and therefore as in the linear case, a nonsingular matrix T can be chosen whose first three rows are c, cA and cA^2, respectively, and a vector $p(x)$ can be chosen whose first three elements are $p^{(0)}(x)$, $p^{(1)}(x)$, and $p^{(2)}(x)$. If

$$z = Tx + p(x)$$

is chosen as the new system of coordinates, (13) can be transformed into the *normal form*

$$\begin{aligned} z_1^+ &= z_2 \\ z_2^+ &= z_3 \\ z_3^+ &= P\bar{z} + Q\eta + cA^2bu + \rho(\bar{z},\eta,u) \\ \eta^+ &= R\bar{z} + S\eta + Ku + \phi(\bar{z},\eta,u) \\ y &= z_1 \end{aligned} \tag{15}$$

where $\bar{z} \triangleq [z_1, z_2, z_3]^T$, $\eta \triangleq [z_4, z_5, \ldots, z_n]^T$, $z^T = [\bar{z}^T, \eta^T]$ and $\rho(\cdot), \phi(\cdot) \in \mathcal{H}$. The vector z is the state of the system. A block diagram representation is shown in Fig. 4.

As in the linear case, a matrix T can be chosen so that $K = 0$ in the expression for η^+ and we will therefore assume K to be zero in the following discussions. However, η^+ still depends on u through the higher order terms in $\phi(\cdot)$.

The general case with relative degree d instead of 3 follows in a straightforward fashion.

Definition 4 (Relative Degree): The nonlinear system Σ is said to have a well-defined relative degree d if

i) its linearized system Σ_L has relative degree d:

$$\begin{aligned} cb &= 0 \\ cAb &= 0 \\ &\cdots \\ cA^{d-2}b &= 0 \\ cA^{d-1}b &\neq 0 \end{aligned}$$

and

ii) the nonlinear system Σ satisfies

$$\begin{aligned} \left(c + p_x^{(0)}(Ax + bu + f(x,u))\right)(b + f_u(x,u)) &\equiv 0 \\ \left(cA + p_x^{(1)}(Ax + bu + f(x,u))\right)(b + f_u(x,u)) &\equiv 0 \\ &\cdots \\ \left(cA^{d-2} + p_x^{(d-2)}(Ax + bu + f(x,u))\right)(b + f_u(x,u)) &\equiv 0 \end{aligned}$$

in a neighborhood of the origin, where the functions $p^{(i)}(\cdot) \in \mathcal{H}$, $i = 0, 1, \ldots, d-1$ are defined recursively as indicated previously, and subscripts denote Jacobians.

We note that in the normal form, $z_1, z_2, \ldots, z_{d-1}$ are delayed versions of z_d which is the first component of $\bar{z}$ that can be directly controlled by u. Since $\rho(\cdot) \in \mathcal{H}$, it follows that, given a desired trajectory $y^*(k+d)$ at time k, the control $u(k)$ can be determined as

$$u(k) = (cA^{d-1}b)^{-1}(y^*(k+d) - P\bar{z} - Q\eta) + g(y^*(k+r), \bar{z}, \eta) \quad (16)$$

where $g(\cdot) \in \mathcal{H}$, so that $z_1(k)$ follows $y^*(k)$. Since $y^*(k+d)$ is bounded, $z_1(k), z_2(k), \ldots, z_d(k)$ are also bounded. However, it has to be demonstrated that $\eta(k)$ is also a bounded sequence. This will assure that the entire state of the system $z(k)$ is bounded. It is only then that the control $u(k)$ can be computed using (16).

To study the behavior of $\eta(k)$, we first consider the case when $y(k) \equiv 0$ (and hence, $\bar{z}(k) \equiv 0$). The output can be forced to zero by the input

$$u_{\text{zero}}(\eta) = -(cA^{d-1}b)^{-1}Q\eta + g(0, 0, \eta).$$

Note that u_{zero} is a function only of η. Under such an input, η evolves according to

$$\eta^+ = S\eta + \phi(0, \eta, u_{\text{zero}}) \triangleq S\eta + \bar{\phi}(\eta), \quad \bar{\phi}(\cdot) \in \mathcal{H}.$$

Definition 5 (Zero Dynamics): The dynamics associated with the η part of the state vector z when the output is forced to be identically zero, i.e.

$$\eta^+ = S\eta + \bar{\phi}(\eta) \quad (17)$$

is called the zero-dynamics of the nonlinear system Σ.

A sufficient condition for the asymptotic stability of the nonlinear zero dynamics is that S is a stable matrix, or equivalently, the zero dynamics of the linearized system Σ_L is asymptotically stable.

The zero dynamics given by (17) is an autonomous nonlinear difference equation. By Malkin's theorem [19], if it is (uniformly) asymptotically stable, it is also totally stable, or, stable under sufficiently small perturbations. If we consider $y^*(k)$, the desired output signal, as a sufficiently small perturbation, the we can conclude that all signals are bounded, and that in a neighborhood of the origin, the control law (16) can be used for tracking.

With a change of coordinates, the control law can also be expressed as

$$u(k) = (cA^{d-1}b)^{-1}(y^*(k+d) - \bar{P}x) + \bar{g}(x, y^*(k+d)) \quad (18)$$

where $\bar{g}(\cdot) \in \mathcal{H}$.

Theorem 8 (General Tracking): If the nonlinear system Σ has a well defined relative degree, and if the zero dynamics of the linearized system Σ_L is asymptotically stable, then a control law (18) exists such that the output of Σ follows any desired signal of sufficiently small magnitude, and that the state only evolves in a neighborhood of the origin.

V. NARMA Model and Tracking: Input and Output Accessible

In Section IV, problems of regulation and tracking of a nonlinear system are studied under the assumption that the state vector is accessible. In practical situations, it is often the case that the state vector is not accessible while only the input and the output are accessible. This makes the study of such control problems very important. For mathematical tractability we always assume that, even when only the input and the output are accessible, the system has an underlying state space representation (3). Since it has been established that, if the linearized system Σ_L is observable, the nonlinear system Σ is also observable in a neighborhood of the origin, it follows that the state $x(k)$ of the system can be reconstructed from measurements $y(k), y(k-1), \ldots, y(k-n+1), u(k), u(k-1), \ldots, u(k-n+1)$. Thus, for the purpose of establishing existence of solutions for such problems as stabilization, regulation and tracking, we could use the results obtained in Section IV and replace $x(k)$ with its reconstructed value (for a more rigorous treatment, readers are referred to [20]). However, rather than relating all variables explicitly to the state vector $x(k)$ at every step, we prefer a representation of the system entirely in terms of its input $u(k)$ and output $y(k)$. Therefore, we will establish the nonlinear autoregressive moving average (NARMA) model of the system Σ in a neighborhood of the origin, and study the properties of this representation and the problems of regulation and tracking pertaining to it.

Before we proceed, the meaning of a (local) input–output representation of a nonlinear system has to be clarified. An input–output representation has the following properties.

i) For each initial condition x_0 in the state representation, there exists a corresponding initial condition vector in terms of past values of u and y in the input–output representation.

ii) If $x_0 \in \Omega_x$ and $u(k) \in \Omega_u, \forall\, k > 0$, then the output $y(k)$ determined by the state representation will be identical to the one determined by the input–output representation, where Ω_x and Ω_u are some neighborhoods of the origin.

As pointed out in [11], since our main analytical tool is the implicit function theorem, the results obtained for the nonlinear system Σ will be local in nature, i.e., they are valid in a neighborhood Ω of the origin. If Σ is an unstable system, then there exist input sequences which lead to unbounded output sequences, thus leaving the region of validity of the implicit function theorem. Therefore, strictly speaking, a local input–output representation will fail to be valid for an unstable Σ. However, from a practical point of view, it is usually the case that a linear controller is already in operation when we attempt nonlinear control using neural networks to improve performance, and therefore our starting point will usually be that all the signals are in the region of validity for all time. With this understanding we will treat the input–output representations of both stable and unstable systems in the same fashion.

A. The NARMA Model

From Section III it is known that if the linearized system Σ_L is observable, then the state can be reconstructed from the input and the output, i.e.

$$x(0) = W_o^{-1}\left(Y_{(0,n-1)} - PU_{(0,n-2)}\right) + \eta\left(Y_{(0,n-1)}, U_{(0,n-2)}\right)$$

where $\eta(\cdot) \in \mathcal{H}$. If we iterate the state equation in (3) to obtain $x(n-1)$, and note that the system is time-invariant, we will have for all k,

$$\begin{aligned} x(k) = {} & T_1 U_{(k-n+1,k-1)} + T_2 Y_{(k-n+1,k-1)} + Qy(k) \\ & + \tau\left(U_{(k-n+1,k-1)}, Y_{(k-n+1,k-1)}, y(k)\right) \end{aligned} \tag{19}$$

where the matrices T_1, T_2 and Q are suitably defined, and $\tau(\cdot) \in \mathcal{H}$. It follows that:

$$\begin{aligned} y(k+1) &= cx(k+1) + h(x(k+1)) \\ &= c\left(T_1 U_{(k-n+1,k)} + T_2 Y_{(k-n+1,k)}\right) \\ &\quad + g\left(Y_{(k-n+1,k)}, U_{(k-n+1,k)}\right) \\ &\triangleq \alpha^T Y_{(k-n+1,k)} + \beta^T U_{(k-n+1,k)} \\ &\quad + g(Y_{(k-n+1,k)}, U_{(k-n+1,k)}) \\ &= \alpha_0 y(k) + \cdots + \alpha_{n-1} y(k-n+1) \\ &\quad + \beta_0 u(k) + \cdots + \beta_{n-1} u(k-n+1) \\ &\quad + g\left(y(k), \ldots, y(k-n+1),\right. \\ &\qquad \left. u(k), \ldots, u(k-n+1)\right) \end{aligned} \tag{20}$$

where $g(\cdot) \in \mathcal{H}$. Thus the future value of the output, $y(k+1)$, is expressed in terms of the current and past values of the input and the output. Note that this has been achieved by only assuming that the linearized system Σ_L is observable.

Suppose Σ_L has a relative degree (delay) d, i.e., in (20), $\beta_0 = 0, \beta_1 = 0, \ldots, \beta_{d-2} = 0$, and $\beta_{d-1} \neq 0$. Then $u(k), u(k-1), \ldots, u(k-d+2)$ do not affect $y(k+1)$ through *linear* terms, but they may affect $y(k+1)$ through *nonlinear* terms in the higher order function $g(\cdot)$. Since our main analytical tools are the inverse function theorem and the implicit function theorem, we want to restrict our attention to systems that have dominant linear terms (as has been done in [21] and other papers). In the following, it is assumed that the nonlinear system Σ itself has a well-defined relative degree, as described in Section IV-C.2.

From the normal form (15) it follows that:

$$\begin{aligned} y(k+d) &= PY_{(k,k+d-1)} + Q\eta(k) + cA^{d-1}bu(k) \\ &\quad + \rho\left(Y_{(k,k+d-1)}, \eta(k), u(k)\right) \\ \rho(\cdot) &\in \mathcal{H} \end{aligned} \tag{21}$$

where

$$Y_{(k,k+d-1)} \triangleq [y(k), \ldots, y(k+d-1)] = \bar{z}(k).$$

In this representation, $u(k)$ affects $y(k+d)$ through a nonzero linear gain $cA^{d-1}b$ and a nonlinear higher order function $\rho(\cdot)$. Since η is part of the state vector, our objective is to eliminate it from the above equation and obtain an input–output representation. There are three steps involved.

Step 1: Define

$$\xi(k) \triangleq Q\eta(k) + \rho\left(Y_{(k,k+d-1)}, \eta(k), u(k)\right)$$

and (21) becomes

$$y(k+d) = PY_{(k,k+d-1)} + cA^{d-1}bu(k) + \xi(k)$$

and our objective is to express $\xi(k)$ in terms of $u(k)$ and $y(k)$, and their *past or future* values.

Step 2: Define

$$v(k) \triangleq Y_{(k,k+d-1)}$$

and, from (15), $\xi(k)$ can be considered as the output of the following system::

$$\begin{aligned} \eta(k+1) &= S\eta(k) + Rv(k) + \phi(v(k), \eta(k), u(k)) \\ \xi(k) &= Q\eta(k) + \rho(v(k), \eta(k), u(k)) \end{aligned} \tag{22}$$

where η is the state variable of dimension $\bar{n} \triangleq n - d$, and u and v are the inputs. Now we need the following:

Proposition 1: The system Σ_L in its normal form (11) is observable if and only if (Q, S) is an observable pair.

Proposition 2: If (Q, S) is an observable pair, then the system described by (22) is observable, i.e., the state η can be reconstructed from the input $[v, u]$ and the output ξ.

The aforementioned proposition is a straightforward extension of the observability result for SISO systems presented in Section III. The following notations are used for the past and present values of the input and the output

$$\begin{aligned} \Xi_{(k-\bar{n},k-1)} &\triangleq [\xi(k-\bar{n}), \xi(k-\bar{n}+1), \ldots \\ &\qquad \xi(k-1)]^T \\ U_{(k-\bar{n},k-1)} &\triangleq [u(k-\bar{n}), u(k-\bar{n}+1), \ldots \\ &\qquad u(k-1)]^T \\ V_{(k-\bar{n},k-1)} &\triangleq \left[v^T(k-\bar{n}), v^T(k-\bar{n}+1), \ldots, \right. \\ &\qquad \left. v^T(k-1)\right]^T. \end{aligned}$$

Step 3: Since

$$y(k+d) = PY_{(k,k+d-1)} + cA^{d-1}bu(k) + \xi(k)$$

it follows that:

$$\xi(k) = y(k+d) - PY_{(k,k+d-1)} - cA^{d-1}bu(k)$$

and now we can express $\Xi_{(k-\bar{n},k-l)}$ in terms of $Y_{(k-\bar{n},k+d-1)}$ and $U_{(k-\bar{n},k-1)}$. The latter two are $Y_{(k+d-n,k+d-1)}$ and $U_{(k+d-n,k-1)}$, respectively, by definition of $\bar{n}$.

Finally, it follows that:

$$y(k+d) = \bar{\alpha}^T Y_{(k+d-n,k+d-1)} + \bar{\beta}^T U_{(k+d-n,k)} + \bar{\omega}\left(Y_{(k+d-n,k+d-1)}, U_{(k+d-n,k)}\right) \quad (23)$$

where $\bar{\omega}(\cdot) \in \mathcal{H}$ and the vectors $\bar{\alpha}$ and $\bar{\beta}$ are suitably defined, with the latter having a nonzero leading element $cA^{d-1}b$.

In this input–output representation, the input $u(k)$ affects the output $y(k+d)$ through a nonzero linear term $cA^{d-1}bu(k)$. Mathematically speaking, a suitable value of $u(k)$ can be determined using the implicit function theorem for a desired value of $y(k+d)$. However, from a control point of view, such a control input $u(k)$ cannot be determined at instant k, when the outputs $y(k+1), \ldots, y(k+d-1)$ are not available. To overcome this difficulty, (23) is recursively used to replace future outputs with past outputs and results in

$$\begin{aligned} y(k+d) &= \alpha^T Y_{(k-n+1,k)} + \beta^T U_{(k-n+1,k)} \\ &\quad + \omega\left(Y_{(k-n+1,k)}, U_{(k-n+1,k)}\right) \\ &= \alpha_0 y(k) + \cdots + \alpha_{n-1} y(k-n+1) \\ &\quad + \beta_0 u(k) + \cdots + \beta_{n-1} u(k-n+1) \\ &\quad + \omega\left(y(k), \ldots, y(k-n+1),\right. \\ &\qquad \left. u(k), \ldots, u(k-n+1)\right) \end{aligned} \quad (24)$$

where $\beta_0 = cA^{d-1}b \neq 0$ and $\omega(\cdot) \in \mathcal{H}$. A diagrammatic representation of a NARMA model is shown in Fig. 5.

B. The Tracking Problem

From the discussion in Section IV, it follows that, if the desired value $y^*(k+d)$ is given at time k, a control law can be determined as

$$\begin{aligned} u(k) = \frac{1}{\beta_0}\Big(& y^*(k+d) - \alpha_0 y(k) - \cdots - \alpha_{n-1} y(k-n+1) \\ & - \beta_1 u(k-1) - \cdots - \beta_{n-1} u(k-n+1) \ ) \\ & + g\left(y(k), \ldots, y(k-n+1), y^*(k+d),\right. \\ & \left. u(k-1), \ldots, u(k-n+1)\right) \Big) \end{aligned} \quad (25)$$

where $g(\cdot) \in \mathcal{H}$, so that $y(k)$ follows $y^*(k)$.

To ensure that all signals are bounded while using the control law (25), we define zero dynamics in the input–output representation (24), which corresponds to the zero dynamics in the state representation. As discussed earlier, the state vector $\eta(k)$ of the system (22) can be uniquely determined from $\Xi_{(k,k+\bar{n}-1)}$, $U_{(k,k+\bar{n}-1)}$ and $V_{(k,k+\bar{n}-1)}$, or more explicitly, from $Y_{(k,k+n-1)}$ and $U_{(k,k+\bar{n}-1)}$. Therefore, for a *fixed* sequence of output $y(k)$, there are $\bar{n}$ values of the input that uniquely determine the state $\eta(k)$. Thus, the zero dynamics in terms of η in (17) is manifested through the evolution of $u(k)$ when the output $y(k)$ is constrained to be identically zero for all k

$$\begin{aligned} u(k) = &-\frac{1}{\beta_0}\left(\beta_1 u(k-1) + \cdots + \beta_{n-1} u(k-n+1)\right) \\ &+ g\left(0, 0, \ldots, 0, u(k-1), \ldots, u(k-n+1)\right). \end{aligned} \quad (26)$$

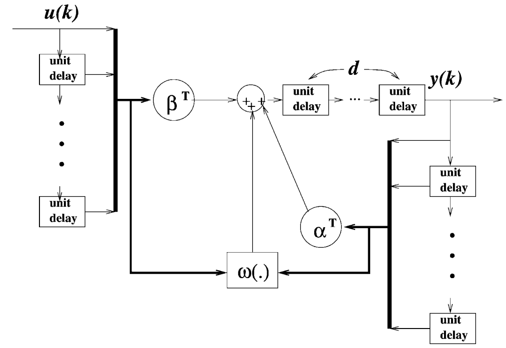

Fig. 5. The NARMA model.

The above is called the *zero dynamics in the input–output representation*. The final results can be summarized as follows.

Theorem 9 (NARMA Model): If the system Σ has a well defined relative degree, it has an input–output representation (24). And if the zero dynamics (26) is asymptotically stable, then under the control law (25), the output $y(k)$ will follow any reference signal $y^*(k)$ with a sufficiently small amplitude, while all signals in the system remain in a neighborhood of the origin.

C. Identifiers and Controllers

We have so far considered the case when the nonlinearities are known. This provides the basis for dealing with the case when the nonlinearities are unknown and the plant has to be identified and controlled adaptively. In view of the results obtained so far, the identifier and the controller can be depicted in more detail, as shown in Fig. 6.

Identifiers and controllers are usually approximated by neural networks. After collecting the input and the output of the plant, the training of an identifier can be conducted in a standard way. When the training is complete and the identifier is obtained, its *corresponding* controller, as shown in Fig. 6, can also be trained by using random numbers. Interested readers are referred to [22] for details.

VI. Conclusion

This paper presents a unified theoretical framework for the identification and control of a nonlinear discrete-time dynamical system, in which the nonlinear system is represented explicitly as a sum of its linearized component and the residual higher order functions. This representation substantially simplifies the procedure of applying the implicit function theorem to derive local properties of the nonlinear system, and reveals the role played by the linearized system in a more transparent form. Under the assumption that the linearized system is controllable and observable, it is shown that: 1) the nonlinear system is also controllable and observable in a local domain; 2) a feedback law exists to stabilize the nonlinear system locally in a finite number of steps; and 3) the nonlinear system can track a constant or a periodic sequence locally, if its linearized system can do so. With some additional assumptions, the nonlinear system is shown to have a well-defined relative degree (delay) and zero-dynamics. If the zero-dynamics of the linearized system is asymptotically stable, so is the nonlinear one, and in such a case, a control law exists for the nonlinear system to track an arbitrary reference

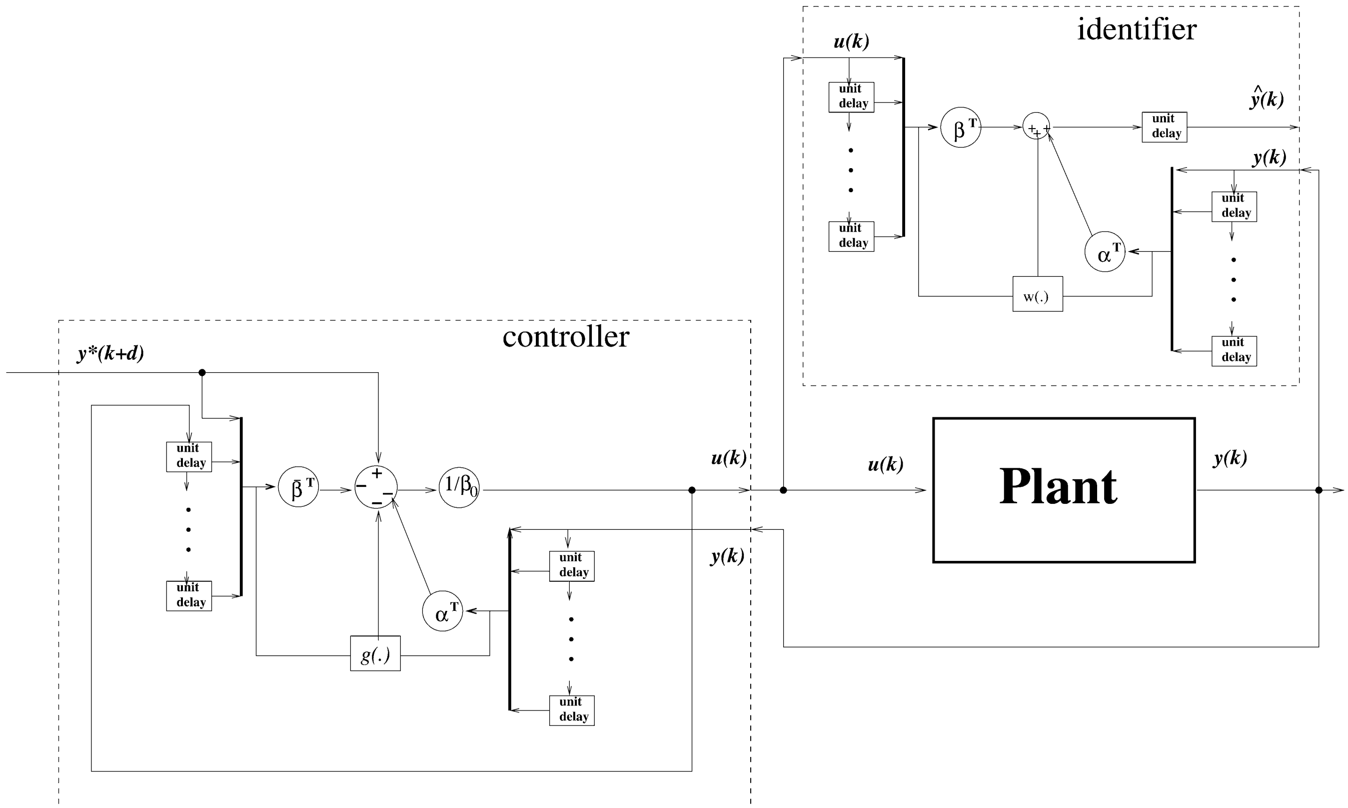

Fig. 6. Identifier and controller in more detail.

signal locally. The tracking can be achieved by using the state vector for feedback, or by using only the input and the output, in which case the NARMA model is established and utilized.

Even though many of the results have appeared in [8], [9], and particularly in [11], the main contribution of this paper is to present all the results in a unified fashion, so that the role played by the linearized system Σ_L in the analysis of the nonlinear system Σ becomes evident. It is the authors' belief that the framework presented will provide a significantly clearer understanding of the use of neural networks as identifiers and controllers for general nonlinear discrete-time dynamical systems.

APPENDIX

A. Complete Proof of Theorem 4

Proof: Let $x(n)$ be zero in (5):

$$U = -W_c^{-1}A^n x(0) + g(x(0), 0).$$

This is the (unique) control sequence of length n that brings any state $x(0)$ to the origin. Define the first element of this sequence by

$$U(0) \triangleq \Gamma x(0) + \gamma(x(0)), \quad \gamma(\cdot) \in \mathcal{H}. \tag{27}$$

We claim that the state feedback law

$$u(k) = \Gamma x(k) + \gamma(x(k)) \tag{28}$$

stabilizes the system in at most n steps. This is a consequence of the following two facts:

i) the solution (5) is unique in the neighborhood of the origin;
ii) the origin is an equilibrium state.

To see this, let the control sequence that transfers any initial state x to the origin in *exactly* n steps be denoted by $U_{(x)}$; recall that any such sequence has to satisfy (4), which has a unique solution (5) and, therefore, is uniquely determined by the initial state x. Let the first element of the sequence U be denoted by $F(U)$, the second element by $S(U)$, and the rest of the elements by $R(U)$, i.e.

$$U = [F(U), S(U), R(U)]. \tag{29}$$

Consider the case in which $U_{(x_0)}$ transfers the state x_0 to the origin in exactly n steps. Let x_1 be the state of the system after the control input $F(U_{(x_0)})$ is applied. Then we know that the remaining control sequence $[S(U_{(x_0)}), R(U_{(x_0)})]$ transfers x_1 to the origin in $(n-1)$ steps. Since the origin is an equilibrium state, it follows that the control sequence $[S(U_{(x_0)}), R(U_{(x_0)}), 0]$ transfers x_1 to the origin in exactly n steps. By uniqueness, it must follow that

$$U_{(x_1)} = [S(U_{(x_0)}), R(U_{(x_0)}), 0]$$

hence

$$S(U_{(x_0)}) = F(U_{(x_1)}) = \Gamma x_1 + \gamma(x_1)$$

as a consequence of (27) and (29). Continuing the same argument for $R(U_{(x_0)})$, we conclude that the state feedback law

$$u = \Gamma x + \gamma(x)$$

stabilizes the system in at most n steps. ∎

ACKNOWLEDGMENT

L. Chen would like to thank Dr. J. B. D. Cabrera [11], [16] for many insightful discussions.

REFERENCES

[1] K. S. Narendra and K. Parthasarathy, "Identification and control of dynamical systems using neural networks," *IEEE Trans. Neural Networks*, vol. 1, pp. 4–27, Mar. 1990.
[2] M. D. Lemmon and A. N. Michel, "Guest editorial: Special section on neural networks in control, identification, and decision making," *IEEE Trans. Automat. Contr.*, vol. 44, pp. 1993–1994, Nov. 1999.
[3] K. S. Narendra and F. L. Lewis, "Introduction to the special issue on neural network feedback control," *Automatica*, vol. 37, no. 8, pp. 1147–1148, Aug. 2001.
[4] Y. H. Kim and F. L. Lewis, *High-Level Feedback Control With Neural Networks*, Singapore: World Scientific, 1999.
[5] S. S. Ge, C. C. Hang, T. H. Lee, and T. Zhang, *Stable Adaptive Neural Network Control*. Norwell, MA: Kluwer, 2001.
[6] F. L. Lewis, S. Jagannathan, and A. Yesildirek, *Neural Network Control of Robot Manipulators and Nonlinear Systems*. London, U.K.: Taylor Francis, 1999.
[7] R. M. Sanner and J.-J. E. Slotine, "Gaussian networks for direct adaptive control," *IEEE Trans. Neural Networks*, vol. 10, pp. 898–906, 1999.
[8] A. U. Levin and K. S. Narendra, "Control of nonlinear dynamical systems using neural networks. controllability and stabilization," *IEEE Trans. Neural Networks*, vol. 4, pp. 192–206, Mar. 1993.
[9] ——, "Control of nonlinear dynamical systems using neural networks—Part II: Observability, identification, and control," *IEEE Trans. Neural Networks*, vol. 7, pp. 30–42, Jan. 1996.
[10] K. S. Narendra and S. Mukhopadhyay, "Adaptive control of nonlinear multivariable systems using neural networks," *Neural Networks*, vol. 7, no. 5, pp. 737–752, 1994.
[11] J. B. D. Cabrera and K. S. Narendra, "Issues in the application of neural networks for tracking based on inverse control," *IEEE Trans. Automat. Contr.*, vol. 44, pp. 2007–2027, Nov. 1999.
[12] S. Mukhopadhyay and K. S. Narendra, "Disturbance rejection in nonlinear systems using neural networks," *IEEE Trans. Neural Networks*, vol. 4, pp. 63–72, Jan. 1993.
[13] L. Chen and K. Narendra, "Nonlinear adaptive control using neural networks and multiple models," *Automatica, Special Issue on Neural Network Feedback Control*, vol. 37, no. 8, pp. 1245–1255, Aug. 2001.
[14] L. J. Corwin and R. H. Szczarba, *Multivariable Calculus*. New York: Marcel Dekker, 1982.
[15] B. Jakubczyk and E. Sontag, "Controllability of nonlinear discrete-time systems: A lie-algebraic approach," *SIAM J. Contr. Opt.*, pp. 1–33, 1990.
[16] J. B. D. Cabrera, "Nonlinear Arma Models and the General Tracking Problem for Discrete-Time Dynamical Systems," Ph.D. dissertation, Yale University, New Haven, CT, 1997.
[17] A. Isidori, *Nonlinear Control Systems II*, ser. Commun. Contr. Eng.. New York: Springer Verlag, 1999.
[18] S. Monaco and D. Normand-Cyrot, "Minimum-phase nonlinear discrete-time systems and feedback stabilization," in *Proc. 26th IEEE Conf. Decision and Control*, 1987, pp. 979–986.
[19] J. P. LaSalle, *Stability Theory for Difference Equations*, ser. Studies in Ordinary Differential Equations, MAA Studies in Mathematics: American Mathematical Assoc., 1977.
[20] L. Chen, "Nonlinear adaptive control of discrete-time systems using neural networks and multiple models," Ph.D. dissertation, Yale Univ., New Haven, CT, 2001.
[21] I. J. Leontaritis and S. A. Billings, "Input–output parametric models for nonlinear systems. Part I: Deterministic nonlinear systems," *Int. J. Contr.*, vol. 41, no. 2, pp. 303–328, 1985.
[22] L. Chen and K. Narendra, "On a new approach to the design of tracking controllers for nonlinear dynamical systems," in *Proc. 1999 Amer. Control Conf.*, vol. 5, 1999, pp. 3534–3538.

Lingji Chen (S'98–M'99) received the B.E. and M.E. degrees in electrical engineering from the University of Science and Technology of China, Hefei, Anhui, in 1990 and 1993, respectively, and the Ph.D. degree in electrical engineering from Yale University, New Haven, CT, in 2001.

Currently, he is a Senior Research Engineer with Scientific Systems Company Inc., Woburn, MA. His research interests include systems biology, bioinformatics, data fusion, nonlinear control, and neural networks.

Kumpati S. Narendra (S'55–M'60–SM'63–F'79–LF'97) received the B.E. degree (with honors) in electrical engineering, from Madras University, India, in 1954, and the M.S. and Ph.D. degrees from Harvard University, Cambridge, MA, in 1955 and 1959, respectively. He received an honorary M.A. degree from Yale University, New Haven, CT, in 1968, and in 1995, his alma mater, in Madras, awarded him an honorary Doctor of Science degree.

From 1961 to 1965, he was an Assistant Professor with the Division of Engineering and Applied Science at Harvard University. He joined the Department of Engineering and Applied Science of Yale University in 1965, and became Professor in 1968. He was Chairman of the Department of Electrical Engineering from 1984 to 1987, and Director of the Neuro-Engineering and Neuro-Science Center during 1995–1996. Currently, he is the Harold W. Cheel Professor of Electrical Engineering and Director of the Center for Systems Science at Yale University. He is the author of more than 175 technical articles in the area of systems theory, three books, and the editor of four others in the areas of adaptive control, stability theory, and learning. He has served on various international technical committees, and on several editorial boards of technical journals. He is currently a member of the Advisory Committee of the Institute of Advanced Engineering in Seoul, Korea, the Scientific Advisor of the Hamilton Institute in Dublin, Ireland. He has also been a consultant to more than 15 industrial research laboratories during the past 30 years.

Prof. Narendra is the recipient of numerous awards, including the Franklin V. Taylor Memorial Award (IEEE SMC Society, 1972), the George Axelby Best Paper Award (IEEE Control Systems Society, 1988), the John R. Ragazzini Education Award (American Automatic Control Council, 1990), the Outstanding Paper Award of the Neural Network Council (1991), the Neural Network Leadership Award (1994), the Bode Prize/Lecture (IEEE Control Systems Society, 1995), and the Richard E. Bellman Control Heritage Award of the AACC, for pioneering contributions to adaptive control and learning theory (June 2003). He is a Fellow of the Institute of Electrical Engineers (U.K.) and a member of Sigma Xi, and the Connecticut Academy of Science and Engineering.