INT. J. CONTROL, 1989, VOL. 49, NO. 3, 1013–1032

Representations of non-linear systems: the NARMAX model

S. CHEN† and S. A. BILLINGS†

Input–output representations of non-linear discrete-time systems are discussed. It is shown that the NARMAX (Non-linear AutoRegressive Moving Average with eXogenous inputs) model is a general and natural representation of non-linear systems and contains, as special cases, several existing non-linear models. The problem of approximating non-linear input–output systems is also addressed and several properties of non-linear models are highlighted using simple examples.

1. Introduction

An input–output model is a means of describing the input–output relationship of a system and an important question regarding the model is how to relate input to output in some straightforward way that will provide an adequate approximation to as large a class of systems as possible for a reasonable computational cost. For linear discrete-time systems, it is well known that a linear difference equation model exists that involves only a fixed and finite number of calculations at each stage if the Hankel matrix of the system has finite rank and this provides a much more concise description than the impulse response function. An analogous situation exists for non-linear discrete-time systems. The non-linear difference equation model known as the NARMAX (Non-linear AutoRegressive Moving Average with eXogenous inputs) model (Leontaritis and Billings 1985) provides a unified representation for a wide class of non-linear systems and has obvious advantages over functional series representations such as the Volterra series.

With the aim of unifying input–output models for non-linear systems, the present study shows that the NARMAX model provides a natural representation for a wide class of non-linear systems and includes several known non-linear input–output models as special cases. For practical applications, there is a need to approximate general input–output behaviour with simple models. Based on the Stone–Weierstrass theorem two models, called the output-affine and polynomial models, are shown to be suitable for this purpose. These results should provide a basis for using these two models in the identification of non-linear systems and other control implementations.

The paper is organized as follows. Section 2 reviews various non-linear input–output representations and shows that the NARMAX model provides a unified representation for finitely realizable non-linear systems. The way in which the NARMAX model provides a natural representation for real processes is illustrated in § 3 using examples and this yields further evidence of the generality of the NARMAX model. The problem of uniqueness is also briefly discussed. General non-linear input–output behaviour can be well approximated by polynomial models and this is proved in § 4. Three parametric non-linear models are compared in § 5 and it is shown that the polynomial NARMAX model has some important practical advantages. The justification for using output-affine models to approximate non-linear systems, which

Received 20 April 1988.

† Department of Control Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD, U.K.

0020-7179/89 $3.00

was presented by Sontag (1979 a) and Fliess and Normand-Cyrot (1982), is also briefly discussed. The extension to stochastic non-linear systems is presented in § 6 where an important property of the stochastic input–output model, namely m-invertibility, is briefly introduced.

Unless otherwise specifically stated, the single-input single-output case is used for notational simplicity. However, most of the discussions are valid for multi-input multi-output systems (Leontaritis and Billings 1985).

2. System representations

Consider the discrete-time time-invariant system

$$\left.\begin{aligned} x(k+1) &= g(x(k), u(k)) \\ y(k) &= h(x(k), u(k)) \end{aligned}\right\} \quad x(0) = x_0 \tag{1}$$

where $k = 1, 2, \ldots$, $x(k)$ is the state vector at k, x_0 is the initial state, $u(k)$ and $y(k)$ are the input and output, respectively, $g(\cdot)$ is the one-step ahead state transition function and $h(\cdot)$ is the output function.

The function that describes the input–output behaviour of a system is of primary importance in systems theory because this is all an external observer can see. Let $\mathbf{U}$ be the input set and $\mathbf{Y}$ be the output set with the assumption that $\mathbf{U}$ and $\mathbf{Y}$ are real Banach spaces. Denote as $\mathbf{U}^+$ the union of all $\mathbf{U}^k$, $k \geqslant 1$, where $\mathbf{U}^k$ is the set of all sequences $(u(1), \ldots, u(k))$ of length k: $\mathbf{U}^k = \{(u(1), \ldots, u(k)): u(i) \in \mathbf{U}, i = 1, \ldots, k\}$. When the system is initially at the state x_0, the response function that describes the input–output behaviour of the system is defined as $f_{x_0}: \mathbf{U}^+ \to \mathbf{Y}$. For every $k = 1, 2, \ldots$, the response function $f_{x_0|k}: \mathbf{U}^k \to \mathbf{Y}$ is a different function since the domain $\mathbf{U}^k$ is a different one. A strictly causal response function is one where, for each k, $f_{x_0|k}(u(1), \ldots, u(k))$ is independent of $u(k)$. f_{x_0} is said to be finitely realizable if and only if it has a state–space realization, (1), with the state vector $x(k)$ having a finite dimension. When $x_0 = 0$, f_{x_0} and $f_{x_0|k}$ is written as f and f_k. If x_0 is an equilibrium state, a simple variable change can transfer the system into zero-initial-state.

2.1. *Volterra series*

If each $f_{x_0|k}$ is a real analytic function the system can be represented by a Volterra series

$$y(k) = \sum_{i=0}^{\infty} h_{x_0|ki}(u(1), \ldots, u(k)) \tag{2}$$

where $h_{x_0|ki}$ is a homogeneous degree i polynomial in $u(1), \ldots, u(k)$. For a zero-initial-state or equilibrium-initial-state response function, $h_{x_0|ki}$ is simplified as h_{ki}. The Volterra series or related Wiener series represent a large class of non-linear systems, and have been extensively studied (e.g. Marmarelis and Marmarelis 1978, Schetzen 1980, Rugh 1981). The functional series expansions of Volterra or Wiener, however, map past inputs into the present output and this inevitably means that an excessive parameter set, often extending to well in excess of 500 kernal values, is required to describe even simple non-linear systems and consequently few practical applications have been reported.

2.2. *General non-linear model*

Input–output descriptions that expand the current output in terms of past inputs

and outputs provide models that represent a broad class of non-linear systems and that may avoid the difficulty of excessive parameters associated with Volterra series. The model

$$y(k) = F(y(k-1), \ldots, y(k-n_y), u(k-1), \ldots, u(k-n_u)) \tag{3}$$

where $F(\cdot)$ is some non-linear function, is about as far as one can go in terms of specifying a general finite non-linear system. Model (3) is referred to as the NARMAX (non-linear ARMAX) model by Billings and Leontaritis (1982) for its resemblance to the linear model

$$y(k) = a_0 + \sum_{i=1}^{n_y} a_i y(k-i) + \sum_{i=1}^{n_u} b_i u(k-i) \tag{4}$$

Leontaritis and Billings (1985) rigorously proved that a non-linear discrete-time time-invariant system can always be represented by the model (3) in a region around an equilibrium point subject to two sufficient conditions:

(i) the response function f of the system is finitely realizable;

(ii) a linearized model exists if the system is operated close to the chosen equilibrium point.

Notice that condition (i) simply excludes distributed parameter systems and condition (ii) implies that if the system is perturbed with a small amplitude input in the linear region around the equilibrium point, a linearized model of the system exists.

The derivation of the NARMAX model (3) is based on zero-initial-state response by Leontaritis and Billings (1985). The result can, however, be carried over to the non-zero-initial-state case using similar arguments. Notice that whereas the response functions of a system are different for different initial states, the input–output model (3) for the system will always be the same regardless of the initial state of the system provided that the system is maintained within a region around an equilibrium point. This is clarified with two examples.

Example 1

$$\left.\begin{aligned} x(k+1) &= x(k) + u(k) \\ y(k) &= x(k) + x^2(k) \end{aligned}\right\} \quad x(0) = x_0$$

The origin is an equilibrium point. The response function $f_{x_0|k}$ is

$$y(k) = f_{x_0|k}(u(1), \ldots, u(k-1), u(k)) = x_0 + \sum_{i=1}^{k-1} u(i) + \left(x_0 + \sum_{i=1}^{k-1} u(i)\right)^2$$

which is obviously different for different x_0. The NARMAX model that describes the system is

$$y(k+1) = y(k) + u(k)(1 + 4y(k))^{1/2} + u^2(k)$$

(see Leontaritis and Billings 1985) for

$$x(k) = x_0 + \sum_{i=1}^{k-1} u(i) > -0{\cdot}5$$

Example 2

$$\left.\begin{aligned} x(k+1) &= x(k) + u(k) \\ y(k) &= \exp(-x(k)) \end{aligned}\right\} \quad x(0) = x_0$$

The input–output model of the system is

$$y(k+1) = y(k)\exp(-u(k))$$

which is globally valid and is independent of initial state. The response function $f_{x_0|k}$ is obviously dependent upon initial state. Further illustrations are given later.

Several known non-linear input–output models can be thought of as special cases of the NARMAX model (3).

2.3. *Bilinear model*

A general bilinear input–output model takes the form

$$y(k) = a_0 + \sum_{i=1}^{n_y} a_i y(k-i) + \sum_{i=1}^{n_u} b_i u(k-i) + \sum_{i=1}^{n_y}\sum_{j=1}^{n_u} c_{ij} y(k-i)u(k-j) \tag{5}$$

This is obviously a simple case of the NARMAX model (3). The bilinear state space model

$$\left.\begin{aligned} x(k+1) &= Ax(k) + Bu(k) + u(k)Cx(k) \\ y(k) &= Dx(k) \end{aligned}\right\} \tag{6}$$

for example, gives rise to a bilinear input–output model

$$y(k+1) = DA(D^{\mathrm{T}}D)^{-1}D^{\mathrm{T}}y(k) + DBu(k) + u(k)DC(D^{\mathrm{T}}D)^{-1}D^{\mathrm{T}}y(k) \tag{7}$$

Bilinear system theory has been widely studied in the content of continuous-time systems where the bilinear state space model

$$\left.\begin{aligned} \frac{dx(t)}{dt} &= Ax(t) + Bu(t) + u(t)Cx(t) \\ y(t) &= Dx(t) \end{aligned}\right\} \tag{8}$$

plays an important role. This is because, roughly speaking, the set of bilinear systems is dense in the space of continuous-time systems and any continuous causal functional can be arbitrarily well approximated by bilinear systems within any bounded time interval (see for example Fliess and Normand-Cyrot 1982). Moreover, many real continuous-time processes are naturally in bilinear form. A few examples are distillation columns (España and Landau 1978), nuclear and thermal control processes (Mohler 1973).

Sampling the continuous-time bilinear system, however, produces a NARMAX model which is more complex than a discrete-time bilinear model. Assume that a zero-order-hold (ZOH) device is used, i.e. $u(t) = u(t_k)$, $t_k \leqslant t < t_{k+1}$, a fixed sampling rate is employed with a sampling period h and k is used to replace t_k. For $t \in [t_k, t_{k+1})$, from (8)

$$\left.\begin{aligned} \frac{dx(t)}{dt} &= [A + u(k)C]x(t) + Bu(k) \\ y(t) &= Dx(t) \end{aligned}\right\} \tag{9}$$

Let $t \to t_{k+1}$ and using $h = t_{k+1} - t_k$ yields

$$\left.\begin{aligned} x(k+1) &= \exp\,[[A + u(k)C]h]x(k) + \left\{\int_0^h \exp\,[[A + u(k)C](h-\tau)]B\,d\tau\right\}u(k) \\ y(k+1) &= Dx(k+1) \end{aligned}\right\} \tag{10}$$

The input–output difference equation of the system (10) is a NARMAX model

$$\begin{aligned} y(k+1) = {} & D \exp\,[[A + u(k)C]h](D^{\mathrm{T}}D)^{-1}D^{\mathrm{T}}y(k) \\ & + D\left\{\int_0^h \exp\,[[A + u(k)C](h-\tau)]B\,d\tau\right\}u(k) \end{aligned} \tag{11}$$

that cannot be expressed in the form of (7) or (5). Whatever initial state the system (10) or (8) started from, the NARMAX model for the system is always (11).

To end this subsection it is important to note that, unlike the continuous-time case, it is impossible to approximate all discrete-time systems within the class of discrete-time bilinear systems. A mathematical explanation is that the set of discrete-time bilinear systems is not closed with respect to the product operation so that the product of two discrete-time bilinear state space systems is not necessarily bilinear (Fliess and Normand-Cyrot 1982).

2.4. *Output-affine and rational models*

The response function f of a system is said to be a polynomial response function if for each k, f_k is a polynomial of finite degree in all variables, although this degree may tend to ∞ as $k \to \infty$. A polynomial response function f is said to be bounded if for all k the maximum power that any individual variable is raised to in f_k is less than a certain bound. The realization of polynomial response functions has been investigated in detail by Sontag (1979 b).

It is known that a polynomial response function f is finitely realizable if and only if it satisfies the rational difference equation (Sontag 1979 b)

$$\begin{aligned} & a(y(k-1), \ldots, y(k-r), u(k-1), \ldots, u(k-r))y(k) \\ & \qquad = b(y(k-1), \ldots, y(k-r), u(k-1), \ldots, u(k-r)) \end{aligned} \tag{12}$$

or

$$y(k) = \frac{b(y(k-1), \ldots, y(k-r), u(k-1), \ldots, u(k-r))}{a(y(k-1), \ldots, y(k-r), u(k-1), \ldots, u(k-r))} \tag{13}$$

where r is the order of the system, $a(\cdot)$ and $b(\cdot)$ are polynomials of finite degree.

Sontag (1979 b) further showed that f is a finitely realizable and bounded polynomial response function if and only if it satisfies an affine difference equation

$$\begin{aligned} a_0(u(k-1), \ldots, u(k-r))y(k) = {} & \sum_{i=1}^{r} a_i(u(k-1), \ldots, u(k-r))y(k-i) \\ & + a_{r+1}(u(k-1), \ldots, u(k-r)) \end{aligned} \tag{14}$$

or

$$y(k) = \sum_{i=1}^{r} \frac{a_i(u(k-1), \ldots, u(k-r))}{a_0(u(k-1), \ldots, u(k-r))}y(k-i) + \frac{a_{r+1}(u(k-1), \ldots, u(k-r))}{a_0(u(k-1), \ldots, u(k-r))} \tag{15}$$

where $a_i(\cdot)$, $i=0, 1, \ldots, r+1$ are polynomials of finite degree. Such a response function admits a state-affine representation

$$\left.\begin{aligned} x(k+1) &= A(u(k))x(k) + B(u(k)) \\ y(k) &= C(u(k))x(k) \end{aligned}\right\} \tag{16}$$

where $A(u(k))$, $B(u(k))$ and $[C(u(k))]^T$ are matrix and vector valued polynomials of finite degree.

Sontag (1979 a) showed how to remove the zero-initial-state assumption for systems and the above results are also satisfied for the non-zero-initial-state case. As an illustration consider Example 1 again. The output-affine model for the system is (see Leontaritis and Billings 1985)

$$\begin{aligned} u(k)y(k+2) = (u(k)+u(k+1))y(k+1) - u(k+1)y(k) + u^2(k+1)u(k) \\ + u(k+1)u^2(k) \end{aligned}$$

which can be derived independently of the initial state.

The output-affine model (15) and the rational model (13) are globally valid. The response function of the system is, however, restricted to a polynomial response. For example, the response function in Example 2 is not a polynomial function and the results for polynomial response functions do not therefore apply. By choosing the particular forms in (15) and (13) for $F(\cdot)$, it is easily seen that the NARMAX model (3) can be interpreted to contain, as special cases, the output-affine and rational models.

Fliess and Normand-Cyrot (1982) showed that on a finite-time interval and with bounded inputs, a discrete-time input–output system can be arbitrarily well approximated within the set of state-affine systems. As an example, consider the sampled continuous-time bilinear system of (10). Approximation of the system of (10) by state-affine systems involves the approximation of $\exp[[A+u(k)C]h]$ and $\int_0^h \exp[[A+u(k)C](h-\tau)]B\,d\tau$ using matrix and vector valued polynomials in $u(k)$.

3. Interpretation of the NARMAX model

The NARMAX model (3) is a natural representation for sampled non-linear continuous-time systems. One example is that the discretization of the bilinear continuous-time system produces a NARMAX model as shown in § 2.3. A few more illustrations are given in this section.

Example 3

$$\frac{dy(t)}{dt} = a_1 y(t) + a_2 y^2(t) + bu(t) \tag{17}$$

This differential equation, for instance, can be used to model a simple non-linear circuit consisting of a linear capacitor, a linear resistor and a non-linear resistor in parallel with the current source $i(t)$ (Bussgang *et al.* 1974). The non-linear differential equation relating the current excitation $i(t)$ and the voltage $v(t)$ across the capacitor is

$$i(t) = C\frac{dv(t)}{dt} + K_1 v(t) + K_2 v^2(t) \tag{18}$$

Equation (17) is obtained by identifying $i(t)$ with $u(t)$ and $v(t)$ with $y(t)$. Assume again that ZOH is used and the sampling period is h. The discretization of (17) gives rise to an integration equation

$$\int_{y(k)}^{y(k+1)} \frac{dy}{a_2y^2 + a_1y + bu(k)} = h \tag{19}$$

This integration can be completed and although the explicit input–output relationship may be very complicated it is clear that it takes the form

$$y(k+1) = F(y(k), u(k)) \tag{20}$$

Going into slightly more detail, the analytic expression of $y(k)$ depends upon the two roots of the equation

$$a_2y^2 + a_1y + bu(k) = 0$$

Obviously the value of $u(k)$ affects the nature of the two roots (i.e. real or complex) and in turn the form of non-linear function $F(\cdot)$. Therefore, (20) may not be valid globally. If the sampling period is sufficiently small, using the forward difference scheme

$$\frac{y(k+1) - y(k)}{h} \approx a_1y(k) + a_2y^2(k) + bu(k) \tag{21}$$

and an approximation of (20) is given by

$$y(k+1) = (1 + ha_1)y(k) + ha_2y^2(k) + hbu(k) \tag{22}$$

Example 4

A delignification process (Ljung 1987: 133)

$$\frac{dy(t)}{dt} = -K_1 \exp\left(-\frac{E_L}{u_1(t)}\right)[y(t)]^M[u_2(t)]^\alpha[u_3(t)]^\beta \tag{23}$$

where

$$\left.\begin{array}{ll} y(t) & \text{lignin concentration at time } t \\ u_1(t) & \text{absolute temperature at time } t \\ u_2(t) & \text{concentration of hydrogen sulfite, } [HSO_3^-] \\ u_3(t) & \text{concentration of hydrogen, } [H^+] \end{array}\right\} \tag{24}$$

and K_1, E_L, M, α and β are constants associated with the chemical reaction. It is easily seen that discretization of (23) yields

$$\int_{y(k)}^{y(k+1)} \frac{dy}{y^M} = -hK_1 \exp\left(-\frac{E_L}{u_1(k)}\right)[u_2(k)]^\alpha[u_3(k)]^\beta \tag{25}$$

and the explicit 'input–output' relationship takes the form

$$y(k+1) = F(y(k), u(k)) \tag{26}$$

where $u(k) = (u_1(k), u_2(k), u_3(k))^T$. If h is sufficiently small, an approximation of (26) can be obtained as

$$y(k+1) = y(k) - hK_1 \exp\left(-\frac{E_L}{u_1(k)}\right)[y(k)]^M[u_2(k)]^\alpha[u_3(k)]^\beta \tag{27}$$

using the forward difference scheme.

For a general differential equation

$$\frac{d^n y(t)}{dt} = \tilde{g}\left(\frac{d^{n-1}y(t)}{dt}, \ldots, \frac{dy(t)}{dt}, y(t), u(t)\right) \tag{28}$$

defining the state vector as

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_n(t) \end{bmatrix} = \begin{bmatrix} y(t) \\ \dfrac{dy(t)}{dt} \\ \vdots \\ \dfrac{d^{n-1}y(t)}{dt} \end{bmatrix} \tag{29}$$

a state–space equation is obtained

$$\left.\begin{aligned} \frac{dx(t)}{dt} &= \bar{g}(x(t), u(t)) \\ y(t) &= Dx(t) \end{aligned}\right\} \tag{30}$$

where $D = (1, 0, \ldots, 0)$. The discretization of (30) gives rise to a NARMAX model. Assume that a ZOH is used and a fixed sampling period h is employed. For $t \in [t_k, t_{k+1})$

$$x(t) = x(k) + \int_{t_k}^{t} \bar{g}(x(\tau), u(k))\, d\tau \tag{31}$$

Let $t \to t_{k+1}$, the exact discrete model of (28) is obtained

$$\left.\begin{aligned} x(k+1) &= x(k) + \int_{t_k}^{t_{k+1}} \bar{g}(x(\tau), u(k))\, d\tau \\ y(k+1) &= Dx(k+1) \end{aligned}\right\} \tag{32}$$

The exact input–output equation is a NARMAX model in the form of (3). Notice that the form of the non-linear function $F(\cdot)$ may change if the input varies over different regions. A NARMAX model may therefore be valid only in a region around some operating point.

If h is sufficiently small, an approximate difference equation model is given by

$$\left.\begin{aligned} x(k+1) &= x(k) + h\bar{g}(x(k), u(k)) \\ y(k) &= Dx(k) \end{aligned}\right\} \tag{33}$$

ferent approximation schemes may be used to obtain different approximate he stability properties of these approximate models may differ from the l (for a simple illustration, see Potts 1982). The input–output relationships ate or exact discrete models are, however, all in the form of (3) and an AX model exists in a region around some operating point even though e non-linear function $F(\cdot)$ may be very difficult to obtain.

The NARMAX model was referred to as the recursive representation of the system by Hammer (1984) where the problem of uniqueness of the representation was addressed. It is easily seen that a system has infinitely many different but input–output equivalent NARMAX models. In fact, let

$$y(k) = F(y(k-1), \ldots, y(k-n_y), u(k-1), \ldots, u(k-n_u))$$

be a NARMAX model of the system. Another NARMAX model differing from the above can be obtained as follows

$$\begin{aligned} y(k) &= F(F(y(k-2), \ldots, y(k-n_y-1), u(k-2), \ldots, u(k-n_u-1)), \\ &\quad y(k-2), \ldots, y(k-n_y), u(k-1), \ldots, u(k-n_u)) \\ &= \bar{F}[y(k-1), \ldots, y(k-(n_y+1)), u(k-1), \ldots, u(k-(n_u+1))] \end{aligned}$$

All these representations are input–output equivalent to the system and form an equivalence class $\mathbf{F}$. Within $\mathbf{F}$, there exists a minimal NARMAX model that has the smallest n_y and n_u. For example, the minimal NARMAX model for Example 2 is $y(k) = y(k-1)\exp(-u(k-1))$ where $n_y = n_u = 1$, but another equivalent model is

$$y(k) = y(k-2)\exp(-u(k-2))\exp(-u(k-1))$$

where $n_y = n_u = 2$. The minimal NARMAX model can be derived from any $F \in \mathbf{F}$ in a finite number of successive reduction steps (see Hammer 1984). By the NARMAX model of the system, we are therefore often referring to this minimal representation. Notice that the situation is similar for linear models such as (4).

4. Justification of the polynomial NARMAX model

The NARMAX model (3) is a very general model and it represents a wide class of discrete-time non-linear systems. In some cases, the non-linear form $F(\cdot)$ is known and the task of specifying the input–output relationship of the system is reduced to determining some unknown parameters. For many real sampled non-linear systems, however, their exact NARMAX models are very difficult to determine and in general the non-linear structure of $F(\cdot)$ is unknown. A means of approximating $F(\cdot)$ using some known function is therefore desired and often necessary. If the system is operated close enough to the desired operating point, a linear model of the form of (4) may be used to approximate $F(\cdot)$. The usefulness of such a linearized model, however, is restricted. Model (4) is a polynomial of degree 1 in $y(k-1), \ldots, y(k-n_y), u(k-1), \ldots, u(k-n_u)$. It is reasonable to believe that higher order polynomial functions will in general yield better approximations to the system. Indeed practical identification of several industrial systems has shown that many can be adequately modelled by polynomial NARMAX models (Billings 1986, Billings and Fadzil 1985, Billings *et al.* 1988 a, 1989). The remaining part of this section is devoted to the theoretical justification for using polynomial NARMAX models to represent non-linear systems.

Define the real Banach space $\mathbf{Z}$ as

$$\mathbf{Z} = \mathbf{Y}^{n_y} \times \mathbf{U}^{n_u} \tag{34}$$

and let $n = n_y + n_u$, $z(k) = (z_1(k), \ldots, z_n(k))^T$ where

$$z_i(k) = \begin{cases} y(k-i), & i = 1, \ldots, n_y \\ u(k-i+n_y), & i = n_y+1, \ldots, n \end{cases} \tag{35}$$

$\mathbf{C}^0(\mathbf{Z})$ will be denoted as the real Banach space of continuous functions from $\mathbf{Z}$ into $\mathbf{Y}$. Define

$$p_\theta^m : \mathbf{Z} \rightarrow \mathbf{Y} \tag{36}$$

a polynomial function of degree m

$$p_\theta^m(z(k)) = \sum_{i=0}^{m} L_\theta^i z(k) \tag{37}$$

with

$$\left.\begin{aligned} L_\theta^0 z(k) &= \theta_0 \\ L_\theta^1 z(k) &= \sum_{i=1}^{n} \theta_i z_i(k) \\ L_\theta^2 z(k) &= \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \theta_{i_1 i_2} z_{i_1}(k) z_{i_2}(k) \\ &\vdots \\ L_\theta^m z(k) &= \sum_{i_1=1}^{n} \cdots \sum_{i_m=1}^{n} \theta_{i_1 \cdots i_m} z_{i_1}(k) \cdots z_{i_m}(k) \end{aligned}\right\} \tag{38}$$

where $\theta_0, \theta_i, \theta_{i_1 i_2}, \ldots, \theta_{i_1 \ldots i_m}$ are real constants. It is obvious that the set of all polynomial functions is a subset of $\mathbf{C}^0(\mathbf{Z})$. A polynomial function p_θ^m multiplied by a real scalar ρ is done by multiplying each coefficient in (38) by ρ. The sum of two polynomial functions p_θ^m and p_α^j (assume $j \leqslant m$) is a polynomial function of degree $m = \max\{m, j\}$

$$(p_\theta^m + p_\alpha^j)(z(k)) = \sum_{i=0}^{j} (L_\theta^i + L_\alpha^i) z(k) + \sum_{i=j+1}^{m} L_\theta^i z(k) \tag{39}$$

with

$$\left.\begin{aligned} (L_\theta^0 + L_\alpha^0) z(k) &= \theta_0 + \alpha_0 \\ (L_\theta^1 + L_\alpha^1) z(k) &= \sum_{i=1}^{n} (\theta_i + \alpha_i) z_i(k) \\ &\vdots \\ (L_\theta^j + L_\alpha^j) z(k) &= \sum_{i_1=1}^{n} \cdots \sum_{i_j=1}^{n} (\theta_{i_1 \ldots i_j} + \alpha_{i_1 \ldots i_j}) z_{i_1}(k) \ldots z_{i_j}(k) \end{aligned}\right\} \tag{40}$$

The product of p_θ^m and p_α^j is a polynomial of degree $m + j$

$$(p_\theta^m \times p_\alpha^j)(z(k)) = p_\beta^{m+j}(z(k)) = \sum_{i=0}^{m+j} L_\beta^i z(k) \tag{41}$$

with

$$\left.\begin{aligned}
&L_\beta^0 z(k) = \beta_0 = \theta_0\alpha_0 \\
&L_\beta^1 z(k) = (L_\theta^0 \times L_\alpha^1 + L_\theta^1 \times L_\alpha^0)z(k) = \sum_{i=1}^{n} (\theta_0\alpha_i + \theta_i\alpha_0)z_i(k) \\
&L_\beta^2 z(k) = (L_\theta^0 \times L_\alpha^2 + L_\theta^1 \times L_\alpha^1 + L_\theta^2 \times L_\alpha^0)z(k) \\
&\qquad = \sum_{i_1=1}^{n}\sum_{i_2=1}^{n} (\theta_0\alpha_{i_1 i_2} + \theta_{i_1}\alpha_{i_2} + \theta_{i_1 i_2}\alpha_0)z_{i_1}(k)z_{i_2}(k) \\
&\quad\vdots \\
&L_\beta^{m+j} z(k) = (L_\theta^m \times L_\alpha^j)z(k) \\
&\qquad = \sum_{i_1=1}^{n}\cdots\sum_{i_m=1}^{n}\sum_{i_{m+1}=1}^{n}\cdots\sum_{i_{m+j}=1}^{n} \theta_{i_1\ldots i_m}\alpha_{i_{m+1}\ldots i_{m+j}} z_{i_1}(k)\ldots z_{i_{m+j}}(k)
\end{aligned}\right\} \quad (42)$$

It is clear that the set of all polynomial functions from $\mathbf{Z}$ into $\mathbf{Y}$ is a sub-algebra.

Theorem

Assume that $\mathbf{Z}$ is compact. Then the set of all polynomial functions from $\mathbf{Z}$ into $\mathbf{Y}$ is dense in $\mathbf{C}^0(\mathbf{Z})$.

Remark 1

If $\mathbf{Y}$ and $\mathbf{U}$ are compact subsets of some Banach spaces, the theorem is also held.

The proof is based on the Stone–Weierstrass theorem (e.g. Dieudonné 1960, Simmons 1963). Let $\mathbf{E}$ be a compact Banach space. If a sub-algebra $\mathbf{S}$ of the Banach space $\mathbf{C}^0(\mathbf{E})$ of continuous functions $\mathbf{E} \to \mathbf{R}$ separates points and contains constant functions, then $\mathbf{S}$ is dense in $\mathbf{C}^0(\mathbf{E})$.

The separation property means that for $x, y \in \mathbf{E}$, $x \neq y$, there exists $g \in \mathbf{S}$ such that $g(x) \neq g(y)$.

The set of all polynomial functions from $\mathbf{Z}$ into $\mathbf{Y}$ obviously contains constant functions (p_θ^0). To prove the separation property, take two distinct points $z(k)$ and $\bar{z}(k)$. We may assume that $z_1(k) \neq \bar{z}_1(k)$. Then a degree-1 polynomial function with a set of coefficients $\theta_0 = 0$, $\theta_1 \neq 0$ and $\theta_i = 0$, $i = 2, \ldots, n$ satisfies $p_\theta^1(z(k)) \neq p_\theta^1(\bar{z}(k))$.

The theorem implies that for an $F \in \mathbf{C}^0(\mathbf{Z})$ and given $\varepsilon > 0$, there exists a degree-m polynomial function with a set of coefficients θ such that

$$\|F(z(k)) - p_\theta^m(z(k))\| < \varepsilon, \quad \forall\, z(k) \in \mathbf{Z} \qquad (43)$$

Using polynomial NARMAX models to represent non-linear input–output systems is therefore well justified. As an illustration, consider a polynomial approximation of the NARMAX model for Example 1 given in § 2.2. Specifically, assume that

$$\mathbf{Z} = \{z(k) = (y(k), u(k)) : |y(k)| \leqslant 0{\cdot}25 \text{ and } |u(k)| \leqslant 1\}$$

and the required accuracy is $\varepsilon = 0{\cdot}04$. If a polynomial model of degree 4

$$y(k+1) = p_\theta^4(z(k)) = y(k) + u(k) + 2u(k)y(k) - 2u(k)y^2(k) + 4u(k)y^3(k) + u^2(k)$$

is used, then

$$\|F(z(k)) - p_\theta^4(z(k))\| = \|u(k)[(1 + 4y(k))^{1/2} - (1 + 2y(k) - 2y^2(k) + 4y^3(k))]\|$$

However, for $(y(k), u(k)) \in \mathbf{Z}$

$$(1+4y(k))^{1/2} = 1 + 2y(k) - 2y^2(k) + 4y^3(k) - \frac{5}{128}(4\bar{y})^4$$

where $\bar{y}$ lies between $y(k)$ and the origin. Therefore

$$\| F(z(k)) - p_\theta^4(z(k)) \| < \frac{5}{128} < \varepsilon, \quad \forall z(k) \in \mathbf{Z}$$

The linearized model cannot achieve the required accuracy. If ε is decreased, the degree of polynomial can be increased to achieve the desired accuracy.

5. Comparison of output-affine, polynomial and rational models

Three parametric input–output models, namely the output-affine, polynomial NARMAX and rational models, have been introduced and, from the point of view of formulation, these can all be considered as special forms of the general NARMAX model (3). This section presents a brief comparison of these three models.

As mentioned in § 2.4, state-affine models can be used to represent non-linear systems, particularly systems that appear naturally affine-in-the-states. For example, the exact sampled model (10) of the continuous-time bilinear state–space system is affine-in-the-states. In such a situation, the reason for using a state-affine model to approximate the system is apparent and the approximation can be achieved effectively. One way of using state-affine models to approximate the general state–space system (1) is as follows. For simplicity assume that $x_0 = 0$. First replace (1) by an approximate state–space model based on a degree-$2m$ polynomial

$$\left.\begin{aligned} x(k+1) &= \sum_{i=0}^{m}\sum_{j=0}^{m} A_{ij}\, x^{(i)}(k) u^j(k), \quad A_{00} = 0 \\ y(k) &= \sum_{i=0}^{m}\sum_{j=0}^{m} B_{ij} x^{(i)}(k) u^j(k), \quad B_{00} = 0 \end{aligned}\right\} \tag{44}$$

where $x^{(i)}(k) = x(k) \otimes \ldots \otimes x(k)$ (i terms of $x(k)$) and $\otimes$ is the Kronecker product symbol. Next define a new state vector

$$x^*(k) = \begin{bmatrix} x(k) \\ x^{(2)}(k) \\ \vdots \\ x^{(m)}(k) \end{bmatrix} \tag{45}$$

Then a state-affine model for the augmented state $x^*(k)$ can be obtained (see Rugh 1981: 254–255). Methods of constructing state-affine models from measurements have been discussed by Dang Van Mien and Normand-Cyrot (1984) and Neyran *et al.* (1987). A system admits an output-affine model if and only if it is a finite state-affine system (Sontag 1979 a). Output-affine models can therefore be employed to represent non-linear input–output systems. Identification of the output-affine model has been studied by Billings *et al.* (1988 b) and Chen and Billings (1988 a, b). Notice that the output-affine model does not contain power or cross-product terms in the outputs and this seems to be a restriction. For non-linear systems like those given in Examples 3

and 4, using output-affine models to approximate the systems may require more terms and it is not obvious how the approximation of terms like $y^2(k-1)$ is achieved. Of course, if the response function of the system is a bounded polynomial, the system can then be modelled exactly by an output-affine model. The polynomial NARMAX model is more suitable as an approximation to the general system (3), and because power terms in both the inputs and outputs are allowed more parsimonious models can be obtained.

When the system is operating close to an operating point where the linear approximation will be valid, it is desirable that the non-linear model degenerates to the linear model that is satisfied by the linearized system. A polynomial model can naturally be reduced to the linear model in such a situation. However, it is not clear how an output-affine model can achieve this. The linearized system around the origin for Example 1 is $y(k+1) = y(k) + u(k)$, and in this case the output-affine model fails completely to degenerate to this model when the system is operating close to the origin. For Example 1 it is also seen that the order of the (minimal) polynomial model is lower than that of the (minimal) output-affine model ($n_y = n_u = 1$ compared with $r = 2$). Evidently, a lower order representation is easier to implement in practice. To compute $y(k+1)$, we need to store $y(k)$ and $u(k)$ for Example 1 using the polynomial model, compared with $y(k-1)$, $u(k-1)$, $u(k)$ and $y(k)$ using the output-affine model.

It is obvious that the rational model contains the polynomial model as a special case, and it seems possible therefore that more efficient approximations of non-linear systems can be found by searching in this wider class. A prediction error estimation algorithm for rational models has been derived by Billings and Chen (1989).

For the output-affine model, from (14) or (15), it is seen that

$$a_0(u(k-1), \ldots, u(k-r)) \neq 0 \tag{46}$$

is always required in any practical application, and if $a_0(\cdot)$ is different from zero but of small magnitude, high accuracy in the execution of the algebraic operations is necessary. The rational model suffers the same drawback but the polynomial model does not. The biggest advantage of the polynomial NARMAX model over the output-affine and rational models is perhaps that it is linear in the parameters. Many linear identification results can easily be extended to the polynomial non-linear model and several combined routines of structure determination and parameter estimation have been developed (Billings and Voon 1984, 1986, Leontaritis and Billings 1988, Korenberg *et al.* 1988, Chen and Billings, 1989).

Finally, advantages of using NARMAX models rather than Volterra series are highlighted through simple examples.

Example 5

$$\left.\begin{aligned} x(k+1) &= x(k) + u^2(k) \\ y(k) &= x(k) \end{aligned}\right\} \quad x(0) = x_0$$

For this system, the input–output behaviour can be determined by the NARMAX model

$$y(k) = ay(k-1) + bu^2(k-1) \quad \textit{with} \quad a = b = 1 \quad \textit{and} \quad y(0) = x_0 \tag{47}$$

The Volterra series is

$$y(k) = \sum_{i=0}^{\infty} h_{x_0|ki}(u(1), \ldots, u(k)) = w_0(w_0) + \sum_{i_1=1}^{k-1} w_1(i_1; x_0)u(k-i_1) + \sum_{i_1=1}^{k-1}\sum_{i_2=1}^{i_1} w_2(i_1, i_2; x_0)u(k-i_1)u(k-i_2) + \ldots + \sum_{i_1=1}^{k-1}\ldots\sum_{i_m=1}^{i_{m-1}} w_m(i_1, \ldots, i_m; x_0)u(k-i_1)\ldots u(k-i_m) + \ldots \quad (48)$$

Notice that the response function $f_{x_0|k}$ is

$$y(k) = x_0 + \sum_{i=1}^{k-1} u^2(k-i) \quad (49)$$

Therefore

$$\left.\begin{aligned} w_0(x_0) &= x_0 \\ w_1(i_1; x_0) &= 0 \\ w_2(i_1, i_2; x_0) &= \delta(i_1 - i_2) \\ w_m(i_1, \ldots, i_m; x_0) &= 0. \quad m > 2 \end{aligned}\right\} \quad (50)$$

where

$$\delta(i) = \begin{cases} 1, & i = 0 \\ 0, & i \neq 0 \end{cases} \quad (51)$$

Even for such a simple non-linear system, the NARMAX model is more convenient to use than the Volterra series. If the system structure is unknown, much more computational effort is required to estimate kernels in (48) than to estimate constant parameters in (47). The necessity to use special input signals and the difficulty of interpreting and using the identification results are further disadvantages of using the Volterra series.

For example 2 in § 2.2, the exact NARMAX model is

$$y(k) = y(k-1)\exp(-u(k-1)) \quad (52)$$

Polynomial models can be used to approximate (52) to an arbitrary accuracy. If the first four terms in the infinite series

$$\exp(-x) = \sum_{i=0}^{\infty} (-1)^i \frac{x^i}{i!} \quad -\infty < x < \infty \quad (53)$$

are employed an approximate polynomial NARMAX model is obtained

$$y(k) = y(k-1) - y(k-1)u(k-1) + \tfrac{1}{2}y(k-1)u^2(k-1) - \tfrac{1}{6}y(k-1)u^3(k-1) \quad (54)$$

The response function of the system is

$$y(k) = f_{x_0|k}(u(1), \ldots, u(k)) = y(0)\prod_{i=1}^{k-1}\exp(-u(i)) = \exp(-x_0)\left(\sum_{j=0}^{\infty}\frac{(-1)^j}{j!}\left(\sum_{i=1}^{k-1}u(k-i)\right)^j\right) \quad (55)$$

The Volterra series is the same as in (48) with all the kernels $w_m(i_1, ..., i_m; x_0)$, $m \geqslant 0$ non-zero

$$\left.\begin{aligned} w_0(x_0) &= \exp(-x_0) \\ w_1(i_1; x_0) &= \exp(-x_0) H_1(i_1) \\ w_2(i_1, i_2; x_0) &= \exp(-x_0)(1 - \tfrac{1}{2}\delta(i_1 - i_2)) \\ &\vdots \end{aligned}\right\} \tag{56}$$

where

$$H_1(i) = \begin{cases} 1, & i = 0, 1, 2, ... \\ 0, & i \leqslant 0 \end{cases} \tag{57}$$

For $m > 2$, the expressions for the kernels $w_m(i_1, ..., i_m; x_0)$ become complicated. The simplicity of the NARMAX model compared with the Volterra series is obvious.

6. Non-linear stochastic input–output models

It is more appropriate to define a model in the stochastic environment if the model is to be used as a basis for the development of identification and digital controller design techniques. Define

$$\left.\begin{aligned} u^k &= (u(1), ..., u(k))^T \\ y^k &= (y(1), ..., y(k))^T \end{aligned}\right\} \tag{58}$$

Adopting the view of Leontaritis and Billings (1985) and Ljung (1987: Chap. 5) and regarding y^k as a random variable which we observe for different realizations. Then the conditional probability density function of y^k given u^k

$$\bar{f}_p(y^k | u^k) \tag{59}$$

for $k = 1, 2, ...$ specifies completely a causal stochastic system. Using Bayes's rule, it is seen that the conditional probability density function of $y(k)$ given u^k and y^{k-1}

$$\bar{f}_p(y(k) | y^{k-1}, u^k) \tag{60}$$

for $k = 1, 2, ...$ equally specifies the system. If the system is strictly causal, (60) is equivalent to

$$\bar{f}_p(y(k) | y^{k-1}, u^{k-1}) \tag{61}$$

Based on (60), we can compute the conditional mean of $y(k)$ given u^k and y^{k-1}

$$\hat{y}(k) = E[y(k) | y^{k-1}, u^k] = g_k(y^{k-1}, u^k) \tag{62}$$

and express $y(k)$ in a prediction-error or innovation form

$$y(k) = \hat{y}(k) + e(k) \tag{63}$$

where

$$e(k) = y(k) - \hat{y}(k) = y(k) - g_k(y^{k-1}, u^k) \tag{64}$$

is the prediction error or innovation at time k. In the model (63) the output is separated into two components. The part of the output that can be predicted from the

past is given by a deterministic function $g_k(y^{k-1}, u^k)$ and the unpredictable part is defined as the innovation $e(k)$. Let

$$e^k = (e(1), ..., e(k))^T \tag{65}$$

The vector e^{k-1} can be evaluated from the vectors y^{k-1} and u^{k-1} using (64) iteratively. Similarly, the vector y^{k-1} can be obtained from e^{k-1} and u^{k-1}. Therefore the information contained in (y^{k-1}, u^k) is equivalent to (e^{k-1}, u^k) and

$$\bar{f}_p(y(k)|y^{k-1}, u^k) = \bar{f}_p(y(k)|e^{k-1}, u^k) \tag{66}$$

Thus, the prediction $\hat{y}(k)$ can be given alternatively by

$$\hat{y}(k) = E[y(k)|e^{k-1}, u^k] = f_k^*(e^{k-1}, u^k) \tag{67}$$

f_k^* can be considered as the response function of a deterministic system where the input is $(u(k), e(k))^T$ and the output is $\hat{y}(k)$. Now the results for input–output representations given in § 2 can be applied to f_k^*. If, for example, f_k^* satisfies two sufficient conditions given by Leontaritis and Billings (1985) the system can be represented by the NARMAX model

$$\hat{y}(k) = F^*(\hat{y}(k-1), ..., \hat{y}(k-n_y), u(k-1), ..., u(k-n_u)) \tag{68}$$

or equivalently

$$y(k) = F(y(k-1), ..., y(k-n_y), u(k-1), ..., u(k-n_u), e(k-1), ..., e(k-n_e)) + e(k) \tag{69}$$

If f_k^* is polynomial and finitely realizable, the system can be modelled by the stochastic rational model (Billings and Chen 1989)

$$y(k) = \frac{b(y(k-1), ..., y(k-r), u(k-1), ..., u(k-r), e(k-1), ..., e(k-r))}{a(y(k-1), ..., y(k-r), u(k-1), ..., u(k-r), e(k-1), ..., e(k-r))} + e(k) \tag{70}$$

where $a(\cdot)$ and $b(\cdot)$ are polynomials of finite degree. If further f_k^* is a bounded polynomial for all k, a suitable stochastic output-affine model is (Chen and Billings 1988 a)

$$y(k) = \sum_{i=1}^{r} \frac{a_i(u(k-1), ..., u(k-r))}{a_0(u(k-1), ..., u(k-r))} y(k-i) + \frac{a_{r+1}(u(k-1), ..., u(k-r))}{a_0(u(k-1), ..., u(k-r))} + \sum_{i=1}^{r} \frac{a_{r+1+i}(u(k-1), ..., u(k-r))}{a_0(u(k-1), ..., u(k-r))} e(k-i) + e(k) \tag{71}$$

where $a_i(\cdot)$, $i = 0, 1, ..., 2r+1$ are polynomials of finite degree. Models (70) and (71) are special forms of the general model (69).

If the inputs are removed from (69), a general NARMA (non-linear ARMA) time series model is obtained

$$y(k) = F(y(k-1), ..., y(k-n_y), e(k-1), ..., e(k-n_e)) + e(k) \tag{72}$$

The polynomial AR time series model

$$y(k) = p(y(k-1), ..., y(k-n_y)) + e(k) \tag{73}$$

where $p(\cdot)$ is a polynomial of finite degree, has been criticized for being explosive (Granger and Andersen 1978, Ozaki 1985). It was argued that power terms like

$y^j(k-i)$ will almost certainly cause the model to be explosive unless the associated parameters are small and the distribution of $e(k)$ is truncated and limited in extent, and therefore, polynomial models may not be very useful in modelling non-linear time series whose underlying process is stable and non-divergent. Notice, however, that the criticisms do not carry over to control systems. Non-linear control systems are capable of being unstable and divergent unless the input is restricted to some regions. This feature is well reflected in polynomial models. Moreover, control action will influence system behaviour and often changes the nature of the system. For example

$$y(k) = 2y(k-1) + u(k-1) + e(k)$$

is an explosive linear time series if $u(k) = 0$ for all k. By introducing the feedback control

$$u(k) = -1{\cdot}5y(k)$$

the system becomes stable and behaves like

$$y(k) = 0{\cdot}5y(k-1) + e(k)$$

a stationary time series. It should be emphasized that there are time series (such as bacterial growth) that exhibit explosive behaviour and these can certainly be modelled by polynomial time series models. In practice, for stable and non-divergent, non-linear time series, polynomial models often fit the observed time series values better and produce more accurate predictions than linear models. For the ship-rolling time series given by Ozaki (1985), the polynomial model works well. Explosive behaviour may arise when using polynomial models to simulate their underlying processes under the assumption that $e(k)$ is white *gaussian* noise. If the system noise is really gaussian distributed, it may have a large value (although the probability of $e(k)$ taking such a value is very small). When this occurs, $y(k)$ may jump out of the stability region of the model, causing the divergence of future time series values. The gaussian assumption is an idealized one. The real system noise may well have limited amplitude values and therefore the time series values may in reality never be outside the stability region of the model. It seems that the question of explosive behaviour is due to a large extent on the assumptions regarding the system noise and the usefulness of polynomial time series models cannot be ruled out. This aspect will be discussed further in another publication.

One of the most important uses for time series models is to provide forecasts or predictions. In order to use the general non-linear time series model (72) for forecasting, we must be able to estimate the unobserved system noise sequence e^k from observed time series values y^k and the given model. This leads to the generalized definition of invertibility (Granger and Anderson 1978). Assume that the system is modelled exactly by (72) and all $y(k)$ are known. Let a residual sequence

$$\varepsilon^k = (\varepsilon(1), \ldots, \varepsilon(k))^T \tag{74}$$

be generated by

$$\varepsilon(k) = y(k) - F(y(k-1), \ldots, y(k-n_y), \varepsilon(k-1), \ldots, \varepsilon(k-n_e)) \tag{75}$$

with some initial conditions $\varepsilon(i), i = -n_e + 1, \ldots, 1, 0$ given. Then the model (72) is said to be invertible if

$$E[e(k) - \varepsilon(k)]^2 \to 0 \quad \text{as } k \to \infty \tag{76}$$

This definition of invertibility for time series is not very suitable for control systems because the system is often much more complex than the models that are estimated. Model-invertibility or m-invertibility has been introduced to deal with this aspect (Chen and Billings 1989). By the system is meant the data generation $\{(y(k), u(k))\}$, which is a stochastic process. The model is defined by (69) and can be parameterized with a parameter vector θ of dimension n_θ

$$y(k) = F(y(k-1), \ldots, y(k-n_y), u(k-1), \ldots, u(k-n_u), \varepsilon(k-1, \theta), \ldots, \varepsilon(k-n_e, \theta); \theta) + \varepsilon(k, \theta) \quad (77)$$

As θ ranges over D_M, a subset of $\mathbf{R}^{n_\theta}$, (77) describes the set of models. Identification, for example, consists of selecting a model within D_M that best describes the recorded data. It does not require that the system is modelled exactly within the model set. The residual or prediction error $\varepsilon(k, \theta)$ is, of course, dependent on the chosen parameter θ. For realizations of the stochastic process $\{(y(k), u(k))\}$, let two sequences $\{\varepsilon^{[i]}(k, \theta)\}$, $i = 1, 2$ be generated by

$$\varepsilon^{[i]}(k, \theta) = y(k) - F(y(k-1), \ldots, y(k-n_y), u(k-1), \ldots, u(k-n_u), \varepsilon^{[i]}(k-1, \theta), \ldots, \varepsilon^{[i]}(k-n_e, \theta); \theta) \quad (78)$$

with any two different initial conditions

$$\varepsilon^{[i]}(-n_e+1), \ldots, \varepsilon^{[i]}(0), \quad i = 1, 2 \quad (79)$$

Then the model is said to be m-invertible for θ if

$$(\varepsilon^{[1]}(k, \theta) - \varepsilon^{[2]}(k, \theta))^2 \to 0 \quad \text{as } k \to \infty \text{ w.p. } 1 \quad (80)$$

If the model is m-invertible for all $\theta \in D_M$ it is m-inevitable on D_M. Essentially the m-invertibility of the model means that the one-step-ahead predictor

$$\hat{y}(k|\theta) = g_k(y^{k-1}, u^k; \theta) \quad (81)$$

is exponentially stable. Convergence analysis of recursive estimators for linear models critically depends upon the stability of the predictors (Ljung and Söderström 1983) and this is also true for non-linear models (Chen and Billings 1989).

If the non-linear function $F(\cdot)$ in (77) is a polynomial function, in order to guarantee m-invertibility, the model should be linear in the prediction errors (Chen and Billings 1989). The model can however be non-linear in the inputs and outputs. This is illustrated by an example.

Example 6 The underlying process is a Wiener system

$$\left.\begin{aligned} w(k) &= 0{\cdot}8w(k-1) + 0{\cdot}4u(k-1) \\ y(k) &= w(k) + w^3(k) + e(k) \end{aligned}\right\} \quad (82)$$

excited by a uniformly distributed input with mean 0·2 and the amplitude range from −0·8 to 1·2. The system noise $e(k)$ is a gaussian sequence of zero mean with variance 0·01. A model was identified as (Billings and Voon 1986)

$$\begin{aligned} y(k) = {} & 0{\cdot}7578y(k-1) + 0{\cdot}3891u(k-1) - 0{\cdot}03723y^2(k-1) \\ & + 0{\cdot}3794y(k-1)u(k-1) + 0{\cdot}0684u^2(k-1) + 0{\cdot}1216y(k-1)u^2(k-1) \\ & + 0{\cdot}0633u^3(k-1) - 0{\cdot}739\varepsilon(k-1) - 0{\cdot}368u(k-1)\varepsilon(k-1) + \varepsilon(k) \end{aligned} \quad (83)$$

A sufficient condition for (83) to be m-invertible is $E[0{\cdot}739 + 0{\cdot}368u(k)]^2 < 1$ (Chen and Billings 1989). A simple computation shows $E[0{\cdot}739 + 0{\cdot}368u(k)]^2 = 0{\cdot}7055$ which indicates that the model is indeed m-invertible for the particular input signal used. It is important to emphasize that the restriction of being linear in the prediction errors is a restriction on the polynomial model not the data generation. The actual system output can contain non-linear terms of the system noise.

It is easily seen that the stochastic output-affine model (71) is naturally linear in the prediction errors. For the stochastic rational model, the restriction of being linear in the prediction errors can be removed but analysis becomes more complicated. Notice that for modelling time series, a non-explosive rational model can always be achieved by a suitable choice of the degrees for polynomials $a(\cdot)$ and $b(\cdot)$

7. Conclusions

A unified approach has been adopted in modelling non-linear discrete-time systems. It has been shown that the NARMAX model is a general input–output representation for finitely realizable systems and it includes, as special cases, linear, bilinear, output-affine and rational models. It has been demonstrated that sampling real continuous-time systems naturally produces NARMAX models and several properties of input–output models have been discussed and illustrated using simple examples.

Acknowledgment

This work is supported by the Science and Engineering Research Council (Grant Ref. GR/D/30587).

References

Billings, S. A., 1986, Introduction to nonlinear systems analysis and identification, *Signal Processing for Control*, edited by K. Godfrey and P. Jones (Berlin: Springer-Verlag), pp. 261–294.

Billings, S. A., and Chen, S., 1989, *Int. J. Systems Sci.*, **20**, 467.

Billings, S. A., Chen, S., and Backhouse, R. J., 1989, The identification of linear and nonlinear models of a turbocharged automotive diesel engine. *Mech. Systems Signal Process.*, to be published.

Billings, S. A., and Fadzil, M. B., 1985, The practical identification of systems with nonlinearities. *Proc. 7th IFAC Symp. on Identification and System Parameter Estimation*, York, U.K., pp. 155–160.

Billings, S. A., Fadzil, M. B., Sulley, J., and Johnson, P. M., 1988 a, Identification of a nonlinear difference equation model of an industrial diesel generator. *Mech. Systems Signal Process.*, **2**, 59.

Billings, S. A., Korenberg, M. J., and Chen, S., 1988 b, *Int. J. Systems Sci.*, **19**, 1559.

Billings, S. A., and Leontaritis, I. J., 1982, Parameter estimation techniques for nonlinear systems. *Proc. 6th IFAC Symposium on Identification and System Parameter Estimation*, Washington, DC, pp. 505–510.

Billings, S. A., and Voon, W. S. F., 1984, *Int. J. Systems Sci.*, **15**, 601.

Billings, S. A., and Voon, W. S. F., 1986, *Int. J. Control*, **44**, 803.

Bussgang, J. J., Ehrman, L., and Graham, J. W., 1974, *Proc. Inst. elect. electron. Engrs*, **62**, 1088.

Chen, S., and Billings, S. A., 1988 a, *Int. J. Control*, **47**, 309; 1988 b, Recursive maximum likelihood identification of a nonlinear output-affine model. *Ibid.*, **48**, 1605; 1989, A recursive prediction error parameter estimator for nonlinear models. *Ibid.*, **49**, 569.

DANG VAN MIEN, H., and NORMAND-CYROT, D., 1984, *Automatica*, **20**, 175.
DIEUDONNÉ, J., 1960, *Foundations of Modern Analysis* (New York: Academic Press).
ESPAÑA, M., and LANDAU, I. D., 1978, *Automatica*, **14**, 345.
FLIESS, M., and NORMAND-CYROT, D., 1982, On the approximation of nonlinear systems by some simple state-space models. *Proc. 6th IFAC Symposium on Identification and System Parameter Estimation*, Washington, DC, pp. 511–514.
GRANGER, C. W. J., and ANDERSEN, A. P., 1978, *An Introduction to Bilinear Time Series Models.* Göttingen: Vandenhoeck and Ruprecht).
HAMMER, J., 1984, *Int. J. Control*, **40**, 1.
KORENBERG, M. J., BILLINGS, S. A., LIU, Y. P., and MCILROY, P. J., 1988, *Int. J. Control*, **48**, 193.
LEONTARITIS, I. J., and BILLINGS, S. A., 1985, *Int. J. Control*, **41**, 303; 1988, *Int. J. Systems Sci.*, **19**, 519.
LJUNG, L., 1987, *System Identification: Theory and User* (Englewood Cliffs, NJ: Prentice Hall).
LJUNG, L., and SÖDERSTRÖM, T., 1983, *Theory and Practice of Recursive Identification* (Cambridge, Mass: MIT Press).
MARMARELIS, P. Z., and MARMARELIS, V. Z., 1978, *Analysis of Physiological Systems: The White-Noise Approach* (New York: Plenum Press).
MOHLER, R. R., 1973, *Bilinear Control Processes* (New York: Academic Press).
NEYRAN, B., THOMASSET, D., and DUFOUR, J., 1987, *Proc. Instn elect. Engrs*, **134**, Pt D, 89.
OZAKI, T., 1985, Non-linear time series models and dynamical systems, *Handbook of Statistics*, Vol. 5, *Time Series in the Time Domain*, edited by E. J. Hannan, P. R. Krishnaiah and M. M. Rao (Amsterdam: North-Holland), pp. 25–83.
POTTS, R. B., 1982, *Nonlinear Anal. Theory, Methods Applic.* **6**, 659.
RUGH, W. J., 1981, *Nonlinear System Theory: The Volterra/Wiener Approach* (Baltimore, Md: Johns Hopkins University Press).
SCHETZEN, M., 1980, *The Volterra and Wiener Theories of Nonlinear Systems* (New York: John Wiley).
SIMMONS, G. F., 1963, *Introduction to Topology and Modern Analysis* (New York: McGraw-Hill).
SONTAG, E. D., 1979 a, *I.E.E.E. Trans. Circuits Systems*, **26**, 342; 1979 b, *Polynomial response maps. Lecture Notes in Control and Information Sciences* 13 (Berlin: Springer-Verlag).