Piece-Wise Deterministic Signals

Åström, Karl Johan

Published in:
Time Series

1980

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Åström, K. J. (1980). Piece-Wise Deterministic Signals. In O. D. Anderson (Ed.), *Time Series* North-Holland.

Total number of authors:
1

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

O.D. Anderson, ed., TIME SERIES

PIECE-WISE DETERMINISTIC SIGNALS

K. J. Åström

Department of Automatic Control
Lund Institute of Technology
S-220 07 Lund 7, Sweden

There are cases where it is unnatural to model disturbances as time series. Command signals and large upsets in control systems are typical examples. The notion of piece-wise deterministic signals is introduced to capture the essential features of such signals. A formal treatment of the signals and their prediction theory is given.

INTRODUCTION

The problem of predicting the future value of a signal based on past values occur in many different disciplines. To develop a suitable theory it is necessary to have mathematical models of the signal. A suitable model should allow for some degree of regularity. To be realistic the model should, however, not allow the signal to be predicted exactly. This means, for example, that analytic functions are not suitable because such functions can be predicted exactly if their values on an arbitrarily small interval are known. Stochastic processes and time series have been used very successfully as signal models. There are, however, many situations where it is unnatural to model signals as random processes. Command signals (set points) in industrial control systems are typical examples. The set point is normally kept constant over long periods with occasional changes. There is no particular pattern to the changes. They may appear regularly or irregularly. The amplitudes of the changes may vary substantially. Disturbances in industrial processes is another type of signal. Over periods of time the disturbances are small or even negligible. There may, however, be periods when large upsets occur. The amplitude of the disturbance during an upset may then be substantially larger than during the normal operation. The character of the signal may also be different.

In classical control theory disturbances are often described as steps, ramps, and sinusoids, or more generally as signals, which are the solutions of initial value problems for linear constant coefficient differential equations. In this paper a new class of signals is introduced. These signals share properties both with deterministic signals generated by difference equations and with stochastic processes of the ARMA type. The ARMA processes can be thought of as generated from dynamical systems with white noise inputs. The piece-wise deterministic signals can be regarded in a similar way. The input signal which drives the system is, however, a deterministic signal instead of a random process. It is zero over long intervals and different from zero only at isolated points. The points where the signal is nonzero need not to be known apriori. Because of the similarity to ARMA processes, the formalism for predicting such processes can be applied to piece-wise deterministic signals with only small modifications.

Innovations representations play a central role in the theory of stochastic processes [1]. In such representations a signal is represented as the output of a dynamical system whose input is a sequence of independent (or uncorrelated) identically distributed random variables. Similarly in deterministic control theory it has been common to model disturbances as signals which are solutions to ordinary difference or differential equations. Typical examples are stepfunctions and sinusoids. The piece-wise deterministic signals share properties both with stochastic processes and with deterministic signals. They are described as solutions to ordinary difference equations over certain time intervals but they change in an unpredictable way at isolated points. The piece-wise deterministic signals are therefore also a special type of splines [2]. The piece-wise deterministic signals are also related to the shot-noise model of random processes [3]. In the shot-noise model the random process is generated by sending a random impulse train through a dynamical system. Similar processes are also discussed in [4].

The paper is organized as follows. Piece-wise deterministic signals with polynomial generators are first introduced. The signals are defined and the prediction theory is developed. An extension to signals with rational generators is then given. State space description are introduced and signals which are a sum of an ARMA process and piece-wise deterministic signals are finally discussed.

PIECE-WISE DETERMINISTIC SIGNALS WITH POLYNOMIAL GENERATORS

Definitions

A formal definition will now be given. Only discrete time signals are considered. Therefore let time be the set of integers. Furthermore introduce a subset

$$T_i = \{\ldots, t_{-1}, t_0, t_1, \ldots\}$$

of the integers such that

$$\min_j (t_{j+1} - t_j) = \ell > 1.$$

The elements of the set T_i are obviously isolated. The spacing between the points is at least ℓ. Furthermore let T_r be the complement of T_i with respect to all integers.

DEFINITION 1. Let $A(d)$ be a monic polynomial of degree $n < \ell$ in the backward shift operator d. A signal y is called *piece-wise deterministic* of degree n and index ℓ with polynomial generator if

$$A(d)\, y(t) = 0 \quad \text{if} \quad t \in T_r \tag{2.1}$$

and

$$A(d)\, y(t) \neq 0 \quad \text{if} \quad t \in T_i. \; \square \tag{2.2}$$

The polynomial $A(d)$ is called the *generator* of the signal. The set T_r is called the set of *regular points* and T_i is called the set of *irregular points*. The properties of the set of irregular points are not important as long as it is assumed that the distance between two irregular points is larger than ℓ. When constructing predictors it will be assumed that the set T_i is unknown. The predictors obtained will then be independent of T_i.

In the early literature on stochastic processes the word *completely deterministic* was used to describe a stochastic process such that

$$A(d)\ y = 0 \quad \forall\ t.$$

See [5]. This motivates the chosen terminology.

In analogy with the terminology for random processes the signal ν defined by

$$\nu(t) = A(d)\ y(t) \tag{2.3}$$

is called the *innovation* of the signal y. The signal y can thus be thought of as being generated by feeding the innovation sequence through a filter with the transfer function $1/A(d)$. Notice that the innovations associated with a piece-wise deterministic signal are different from zero only at the irregular points.

The generator of a piece-wise deterministic signal is unique. To see this, let the set of irregular points be given and assume that a signal y has two generators A_1 and A_2. Let ν_1 and ν_2 denote the corresponding innovations. Equation (2.3) gives

$$\nu_1 = A_1 y = (A_1/A_2)\ \nu_2,$$

since ν_1 and ν_2 are zero except at T_i and A_1 and A_2 monic A_1 equals A_2. Some examples of piece-wise deterministic signals will now be given.

EXAMPLE 1. A piece-wise constant signal has the generator

$$A = 1 - d.$$

The set of irregular points are all the points where the signal changes level. □

EXAMPLE 2. A piece-wise linear signal has the generator

$$A = 1 - 2d + d^2.$$

The set of irregular points are all points such that the change of slope is immediately to the left of the points. See Figure 1. □

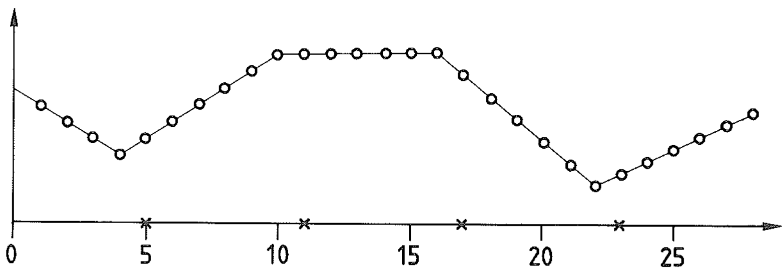

Figure 1
A piece-wise linear signal and its set of irregular points.

EXAMPLE 3. A piece-wise sinusoidal signal with period $2\pi/\omega$ has the generator

$$A = 1 - 2d \cos \omega + d^2. \quad \square$$

Prediction

Predictors for piece-wise deterministic signals will now be constructed. The choice of criteria will first be discussed. Predictors which are optimal with respect to the chosen criterion are then given.

Since the signals satisfy the deterministic equation (2.1) at the regular points it may be expected that the signals can be predicted exactly at those points. Furthermore there will always be an error when the signal is predicted at an irregular point. Since no specific assumptions are made on the changes at the irregular points, it is not natural to use criteria like mean square error etc. Instead it will be attempted to find predictors which brings the prediction error to zero as quickly as possible after an irregular point. The following result then holds.

THEOREM 1. Consider a piece-wise deterministic signal with polynomial generator A. Let F and G be polynomials which are the unique solutions to

$$1 = F(d)\ A(d) + d^k\ G(d) \tag{2.4}$$

such that

$$\deg F(d) < k.$$

Assume that

$$k < \ell. \tag{2.5}$$

The k-step predictor of y given by

$$\hat{y}(t|t-k) = G(d)\ y(t-k) \tag{2.6}$$

brings the prediction error to zero k steps after each irregular point. The prediction error is

$$e(t) = y(t) - \hat{y}(t|t-k) = F(d)\ \nu(t), \tag{2.7}$$

where ν is the innovation of the signal y.

Proof. It follows from (2.1), (2.4), and (2.5) that

$$[1 - d^k G(d)]\ y(t) = F(d)\ A(d)\ y(t) = 0$$

for

$$t = t_i + k,\ t_i + k + 1,\ \ldots,\ t_{i+1} - 1, \tag{2.8}$$

where $\{t_i\}$ is the set of irregular points. The predictor (2.6) thus predicts y(t) exactly in the interval (2.8). Since there will always be a prediction error at the irregular points and their k-1 right successors, the predictor (2.6) is optimal in the sense that it brings the prediction error to zero as quickly as possible. The prediction error is

$$e(t|t-k) = y(t) - \hat{y}(t|t-k) = [1 - d^k G(d)]y(t) = F(d)\ A(d)\ y(t),$$

where the last equality follows from (2.4). The formula (2.7) then follows from (2.3). $\square$

Remark 1. Notice that the predictor (2.6) is a moving average of the signal. □

Remark 2. Notice that the predictor (2.6) is also optimal in the sense of least squares. □

Piece-wise deterministic signals with polynomial generators share many properties with autoregressive processes. The formulas for the predictor (2.6) and the prediction error (2.7) are identical. See [6]. The main difference is that the innovations for autoregressive processes are a sequence of independent (or uncorrelated) random variables, while the innovations for piece-wise deterministic signals are zero except at the irregular points. This has as a consequence that the prediction error for piece-wise deterministic signals will be zero over certain intervals while the prediction errors for autoregressive processes will be a moving average of white noise.

Simple examples of predictors are given in the following examples.

EXAMPLE 4. A piece-wise constant signal has the generator $A = 1 - d$. Simple calculations give

$$F(d) = 1 + d + \ldots + d^{k-1}$$

$$G(d) = 1.$$

The predictor is thus

$$\hat{y}(t|t-k) = y(t-k). \ \square$$

EXAMPLE 5. A piece-wise linear signal has the generator $A = 1 - d + d^2$. Simple calculations give

$$F(d) = 1 + d + \ldots + kd^{k-1}$$

$$G(d) = k + 1 - kd.$$

The predictor is thus

$$\hat{y}(t|t-k) = y(t-k) + k[y(t-k) - y(t-k-1)]. \ \square$$

The problem of predicting a piece-wise deterministic signal can be formulated as the problem of finding a causal operator R such that

$$Ry(t) = 0.$$

It follows from Definition 1 that $R = A$ makes $Ry = 0$ except at the irregular points. Since

$$Ry(t) = \frac{R}{A}\,\nu(t)$$

and $\nu(t_i) \neq 0$, it follows that the signal y can not be predicted exactly by a rational predictor of order less than ℓ. If the order of the operator is larger than ℓ it may be possible to predict the signal. This is illustrated by the following example.

EXAMPLE 6. Consider a square wave with period $2p$. The signal can be predicted using the predictor

$$\hat{y}(t+1|t) = y(t). \tag{2.9}$$

This predictor gives the correct prediction except at those points where the square wave changes level. The square wave can thus be regarded as a piece-wise deterministic signal with generator $Q = 1 - d$.

The square wave can, however, also be predicted exactly by the predictor

$$\hat{y}(t+1|t) = y(t) - y(t-p+1) + y(t-p) \tag{2.10}$$

which requires that $p+1$ past values of the signal are stored. The square wave can thus be regarded in two different ways. If it is considered as a piece-wise constant signal it can be predicted by the simple predictor (2.9) which does not require any storage of past data. The predictor (2.9) gives an error each time the signal changes. An exact predictor (2.10) can be obtained by considering the square wave as a periodic signal. The exact predictor (2.10) is, however, complex in the sense that many past values must be stored if the period p is large. Whether it is worthwhile to use the fact that the square wave is periodic and not just piece-wise constant will thus depend on the circumstances. Notice also that the simple predictor is more robust. The exact predictor requires that the signal is an exact square wave with known period. □

PIECE-WISE DETERMINISTIC SIGNALS WITH RATIONAL GENERATORS

The class of signals introduced in the preceding section will now be generalized. Rational functions will be used as generators instead of polynomials. The signal class will first be defined and the prediction theory is then given.

Definitions

The sets of regular and irregular points are defined as before. Introduce

DEFINITION 2. Let $A(d)$ and $C(d)$ be polynomials of degrees $n < \ell$ in the backward shift operator d. A signal y is called piece-wise deterministic with rational generator if

$$\frac{A(d)}{C(d)}\, y(t) = 0, \qquad \text{if} \quad t \in T_r, \tag{3.1}$$

and

$$\frac{A(d)}{C(d)}\, y(t) \neq 0, \qquad \text{if} \quad t \in T_i. \; \square \tag{3.2}$$

The rational function $A(d)/C(d)$ is called the generator of the signal. As before the signal ν defined by

$$\nu(t) = \frac{A(d)}{C(d)}\, y(t) \tag{3.3}$$

is called the *innovation* of the signal. The signal y can also be represented as

$$y(t) = \frac{C(d)}{A(d)}\, \nu(t),$$

where ν is the innovation. It follows from the definition that the innovation is zero at all regular points and different from zero only at the irregular points. Notice that the irregular points are isolated and that they have a minimum distance ℓ. Piece-wise deterministic signals with rational generators share many properties with mixed autoregressive moving average stochastic processes. Compare the references [6], [7], and [8].

Prediction

Some notation is needed to describe the predictor. A polynomial $A(d)$ is called *stable* if all its zeros are strictly outside the unit disc. A key problem in the prediction is to calculate the innovations from the measurements. For signals with polynomial generators this is easy because it follows from (2.3) that the innovation is simply a moving average of the observations. It also follows from (2.3) that the value $\nu(t_i)$ of the innovation at time t_i can be computed from the value of the signal at time t_i and n past values. The problem is more involved for signals with rational generators. It follows from equation (3.3) that the innovations are obtained as a solution of a difference equation. The value $\nu(t_i)$ of the innovation at time t_i then depends on the value of the signal at time t_i, all past values of y, and on the initial conditions for the difference equation (3.3).

Instead of calculating the innovation ν from the signal y it is then possible to calculate the signal

$$\eta(t) = C(d)\ \nu(t) = A(d)\ y(t).$$

The signal η is, however, different from zero at the irregular points and their deg C successors. To determine the signal η it is thus necessary to wait deg C time units after each irregular point. When η is known it is possible to predict y exactly until next irregular point occur. A predictor for signals with rational generators is given by

THEOREM 2. Consider a piece-wise deterministic signal with rational generator A/C. Let F and G be polynomials which are the unique solution of the equation

$$C(d) = A(d)\ F(d) + d^k\ C(d)\ G(d) \tag{3.4}$$

such that

$$\deg F < k + \deg C. \tag{3.5}$$

Assume that

$$k < \ell - \deg C. \tag{3.6}$$

Then a k-step predictor of y which brings the prediction error to zero in $k+\deg C$ steps after each irregular point is given by

$$\hat{y}(t|t-k) = G(d)\ y(t-k). \tag{3.7}$$

The prediction error is

$$e(t) = y(t) - \hat{y}(t|t-k) = F(d)\ \nu(t) \tag{3.8}$$

where ν is the innovation of the signal y.

Proof. It follows from (3.3) and (3.4) that

$$F\ \nu(t) = \frac{AF}{C}\ y(t) = [1-d^kG]\ y(t) = 0 \tag{3.9}$$

for

$$t = t_i + \deg F,\ t_i + \deg F+1,\ \ldots,\ t_{i+1} - 1, \tag{3.10}$$

where $\{t_i\}$ are the irregular points. The predictor (3.7) thus predicts y exactly in the interval (3.10). It follows from (3.6) that the interval (3.10) is not empty. The formula for the predictor error follows from (3.9). □

Remark. Notice that the predictor (3.7) is a moving average. Since the polynomial C divides the polynomial F, the predictor polynomial G is identical to the polynomial G in Theorem 1. □

Steady State Predictors

It will now be shown that it is possible to obtain other predictors where the prediction errors goes to zero quicker. These predictors are, however, dynamical systems and will thus require initial conditions. In steady state the initial conditions are unimportant.

If the polynomial $C(d)$ is stable the innovation ν can be computed from y using the stable difference equation (3.3). Since the difference equation is stable the influence of the initial conditions will decrease exponentially as time increases. In steady state the initial conditions are thus unimportant. A straightforward extension of the arguments in the proof of Theorem 2 shows that the steady-state predictor is given by

$$\hat{y}(t|t-k) = \frac{G(d)}{C(d)} y(t-k) \tag{3.11}$$

and that the prediction error is

$$y(t) - \hat{y}(t|t-k) = F(d)\ \nu(t). \tag{3.12}$$

The polynomials F and G are the unique solutions to the equation

$$C(d) = A(d)\ F(d) + d^k\ G(d) \tag{3.13}$$

such that

$$\deg F < k. \tag{3.14}$$

Notice that the predictor (3.11) is a dynamical system and that initial conditions are required. Since C is stable the initial conditions are unimportant in steady state.

To handle cases when the polynomial $C(d)$ also has zeros in the unit disc the polynomial C is factored as

$$C(d) = C^+(d)\ C^-(d) \tag{3.15}$$

where all the zeros of $C^+(d)$ are strictly outside the unit disc and all zeros of $C^-(d)$ are inside the unit disc or on the unit circle. The factors C^+ and C^- are called the *stable* and *unstable factors*, respectively. If C has an unstable factor it is not possible to generate the innovations from the measurements in real time. The signal

$$\eta(t) = C^-(d)\ \nu(t) = \frac{A(d)}{C^+(d)}\ y(t)$$

can be generated in real time from $y(t)$. The signal $\eta(t)$ is however different from zero at the irregular point and their $\deg C^-$ successors. The steady-state predictor for signals with rational generators is given by

THEOREM 3. Consider a piece-wise deterministic signal with rational generator A/C. Let F and G be polynomials which are the unique solution to the equation

$$C(d) = A(d)\ F(d) + d^k\ C^-(d)\ G(d) \tag{3.16}$$

such that

$$\deg F < k + \deg C^-. \tag{3.17}$$

Assume that

$$k < \ell - \deg C^-. \tag{3.18}$$

The k-step predictor of y given by

$$\hat{y}(t|t-k) = \frac{G(d)}{C^+(d)}\ y(t-k) \tag{3.19}$$

is then the *stable steady-state* predictor which brings the prediction error to zero in deg F steps after each irregular point. The prediction error is

$$e(t) = y(t) - \hat{y}(t|t-k) = F(d)\ \nu(t) \tag{3.20}$$

where ν is the innovation of the signal y.

Proof. The proof is analogous to the proof of Theorem 2. It follows from (3.3) and (3.16) that

$$F\ \nu(t) = \frac{FA}{C}\ y(t) = \left[1 - d^k\ \frac{G}{C^+}\right]\ y(t).$$

The predictor is thus given by (3.19) and the prediction error by (3.20). □

Remark. Notice that the predictor is a dynamical system whose characteristic polynomial is equal to the stable factor $C^+(d)$. □

Mean Square Predictors

For signals with polynomial generators the minimum time predictor was the same as the mean square predictor when this exists. For signals with rational generators the predictors are however different. The notion of reciprocal polynomial is necessary to describe the the mean square predictor. If A(d) is the polynomial

$$A(d) = a_0 + a_1 d + \ldots + a_n d^n$$

then the reciprocal polynomial is given by

$$A_*(d) = a_0 d^n + a_1 d^{n-1} + \ldots + a_n.$$

Further introduce

$$A_*^- = (A^-)_*$$

Notice that this is different from $(A_*)^-$.

Let the prediction error be ε. The mean square predictor minimizes the criterion

$$\lim_{N\to\infty} \frac{1}{N} \sum_{t=1}^{N} \varepsilon^2(t).$$

This predictor is given by

THEOREM 4. Consider a piece-wise deterministic signal with rational generator A/C. Assume that the polynomial C has no zeros on the unit circle. Factor C into its stable and unstable factors as

$$C(d) = C^+(d)\ C^-(d). \tag{3.21}$$

Let F and G be polynomials which are the unique solution to the equation

$$C^+(d)\ C_*^-(d) = A(d)\ F(d) + d^k\ G(d)\ C^-(d) \tag{3.22}$$

such that

$$\deg F < k + \deg C^-. \tag{3.23}$$

The steady-state mean square predictor of y is then given by

$$\hat{y}(t|t-k) = \frac{G(d)}{C^+(d)\ C_*^-(d)}\ y(t-k) \tag{3.24}$$

and the prediction error is

$$y(t) - \hat{y}(t|t-k) = \frac{C^-(d)}{C_*^-(d)}\ \nu(t).\ \square \tag{3.25}$$

Remark. Notice that the predictor (3.24) is a dynamical system with the characteristic polynomial $C^+C_*^-$. □

Multivariable Extensions

The results can be extended to multivariable signals by considering signal models of the type

$$A(d)\ y(t) = C(d)\ \nu(t),$$

where the signal y and the innovation are vectors of the same dimension and A(d) and C(d) are matrix polynomials in the delay operator.

STATE SPACE MODELS

Dynamical systems can be described by internal or external models. The external models only describe the input-output properties while the internal models give a detailed account of the internal couplings in the system. It is of course useful to consider a problem in different ways. The discussion of piece-wise deterministic signals in the previous sections was based on external models. Piece-wise deterministic signals will now be described using internal models.

Definitions

Let ν be a signal which is zero at the regular points T_r and different from zero only at the irregular points T_i. As before it is assumed that the irregular points are isolated with a minimal spacing ℓ. The signal ν may, however, now take values in R^p. Consider the signal

$$\begin{cases} x(t+1) = Ax(t) + B\nu(t+1) \\ y(t) = Cx(t) \end{cases} \tag{4.1}$$

where x(t) is an n-vector and y(t) is an m-vector. Equation (4.1) describes a piece-wise deterministic signal as the output y of a dynamical system driven by the innovation ν. The model (4.1) is an internal description. The corresponding external description is given by

$$y(t) = C[qI - A]^{-1} B\nu(t+1), \tag{4.2}$$

where q is the forward shift operator. When the signal y and the innovation ν are scalars (4.2) is equivalent to (3.3).

Prediction

It will now be shown how the problem of predicting a piece-wise deterministic signal can be solved using the state space formalism. The results are closely related to Kalman filters [9] and Luenberger observers [10]. As for systems with rational generators there are several different cases that are of interest. It is assumed that the following condition holds:

$$\operatorname{rank} \begin{pmatrix} CA \\ CA^2 \\ CA^n \end{pmatrix} = n. \tag{4.3}$$

A predictor for the signal (4.1) is given by

THEOREM 5. Consider a piece-wise deterministic signal generated by the observable model (4.1). Let K be a matrix such that all eigenvalues of [I-KC]A are zero. Assume that

$$\ell > k + n. \tag{4.4}$$

The predictor

$$\hat{y}(t|t-k) = CA^k\hat{x}(t-k), \tag{4.5}$$

where $\hat{x}$ is given by the difference equation

$$\hat{x}(t+1) = [I-KC]A\hat{x}(t) + Ky(t+1), \tag{4.6}$$

gives a prediction error which is zero in the intervals

$$t_i + n + k \leq t < t_{i+1}, \tag{4.7}$$

where $\{t_i\}$ are the irregular points.

Proof. Introduce

$$\tilde{x}(t) = x(t) - \hat{x}(t).$$

It follows from (4.1) and (4.6) that

$$\tilde{x}(t+1) = [I-KC][A\tilde{x}(t) + B\nu(t+1)].$$

Since all eigenvalues of the matrix [I-KC]A are zero it follows that

$$\{[I-KC]A\}^n = 0. \tag{4.8}$$

Furthermore, since $\nu(t)$ is different from zero only at the irregular points, it follows that

$$\tilde{x}(t) = 0 \quad \text{for} \quad t_i + n \leq t < t_{i+1}.$$

The prediction error $\tilde{y}$

$$\tilde{y}(t) = y(t) - \hat{y}(t|t-k) = CA^k\tilde{x}(t-k) + C \sum_{i=1}^{k} A^{i-1}B\nu(t-i+1)$$

is thus zero on the set (4.7) which is not empty because of (4.4). □

Remark 1. Notice that because of (4.8) the predictor (4.6) can be written as

$$\hat{x}(t) = \{I - d[I-KC]A\}^{-1} Ky(t) = P(d)\ y(t),$$

where $P(d)$ is a polynomial. The predictor is thus a polynomial operator. □

Remark 2. Notice that in special cases the prediction error may be zero in fewer than $n+k$ steps after the irregular points. □

Steady State Predictors

It is possible to obtain predictors which are more efficient than those given by Theorem 5 in the sense that the prediction errors go to zero quicker. These predictors will, however, be dynamical systems which can not be characterized as polynomial operators. Initial conditions are thus important. If the difference equations are stable the influence of the initial conditions will be negligible as time increases. There are many different possibilities. A typical result is given by

THEOREM 6. Consider a piece-wise deterministic signal y generated by the observable model (4.1). Assume that $m = p$, that all eigenvalues of the matrix

$$[I - B(CB)^{-1}C]\,A \tag{4.9}$$

are inside the unit disc and that

$$\ell > n. \tag{4.10}$$

The predictor

$$\hat{y}(t|t-k) = CA^k\ \hat{x}(t-k), \tag{4.11}$$

where $\hat{x}$ is given by the difference equation

$$\hat{x}(t+1) = A\hat{x}(t) + B(CB)^{-1}[y(t+1) - CA\hat{x}(t)], \tag{4.12}$$

gives a prediction error which in the *steady state* is zero in the interval

$$t_i + k \leq t < t_{i+1}, \tag{4.13}$$

where $\{t_i\}$ are the irregular points.

Proof. Introduce

$$\tilde{x}(t) = x(t) - \hat{x}(t).$$

It follows from (4.1) and (4.12) that

$$\begin{aligned}\tilde{x}(t+1) &= [I - B(CB)^{-1}C]\ [A\tilde{x}(t) + B\nu(t+1)] = \\ &= [I - B(CB)^{-1}C]\ A\tilde{x}(t).\end{aligned} \tag{4.14}$$

Since the matrix (4.9) has all eigenvalues inside the unit disc it follows from (4.14) that $x(t) \to 0$ as $t \to \infty$. Consider now a time t between two irregular points. Then

$$\tilde{y}(t|t-k) = y(t) - \hat{y}(t|t-k) = CA^k\tilde{x}(t-k) + C\sum_{i=1}^{k} A^{i-1}B\nu(t-i+1).$$

For large t it thus follows that $\tilde{y}(t|t-k) = 0$ in the interval (4.13). □

Remark. Equation (4.12) has the same form as a Kalman filter. The filter gain is $K = B(CB)^{-1}$. Notice that the filter gain reflects how the disturbances enter the system. □

MIXED PROCESSES

Piece-wise deterministic signals and ARMA processes are simple to deal with. Both capture certain aspects of real disturbances. It is thus natural to combine the models to obtain more realistic models of real signals. One possibility is simply to add a piece-wise deterministic signal and an ARMA signal. The prediction theory for mixed signals of this type is unfortunately quite complicated. In this section it is indicated how simple suboptimal predictors for mixed signals can be constructed. An example indicates that the simple predictors have interesting properties.

A Signal Model

Let y_1 be a piece-wise deterministic signal with polynomial generator A_1. The signal y_1 can be represented as

$$A_1(d)\ y_1(t) = \nu(t), \tag{5.1}$$

where

$$A_1(d) = 1 + a_1 d + \ldots + a_n d^n$$

and

$$\nu(t) = 0, \quad t \in T_r,$$
$$\nu(t) \neq 0, \quad t \in T_i.$$

Furthermore let y_2 be a mixed autoregressive moving average process represented by

$$A_2(d)\ y_2(t) = C_2(d)\ \varepsilon(t), \tag{5.2}$$

where $\{\varepsilon(t)\}$ is a sequence of independent equally distributed random variables. Consider the signal

$$y(t) = y_1(t) + y_2(t). \tag{5.3}$$

Prediction

Both the formulation and the solution of an exact prediction problem for the signal (5.3) is complicated. It will therefore be attempted to postulate a predictor for the signal (5.3) and to explore its properties. If the model (5.3) should be reasonable the innovations ν of the piece-wise deterministic signal should be larger than the innovations of the ARMA process. If this is not the case the piece-wise deterministic signal could probably be absorbed in an ARMA model. A predictor for y can heuristically be formulated as follows. For simplicity the discussion is limited to one-step prediction.

The predictors for the signals y_1 and y_2 are given by

$$\hat{y}_1(t|t-1) = G_1(d)\ y_1(t-1)$$

and

$$\hat{y}_2(t|t-1) = G_2(d)/C_2(d)\ y_2(t-1),$$

where

$$G_1(d) = [1 - A_1(d)]/d$$
$$G_2(d) = [C_2(d) - A_2(d)]/d.$$

These predictors can, however, not be realized since the signals y_1 and y_2 can not be measured separately. Assume for a moment that approximate predictors for y_1 and y_2 can be constructed. Let the residual be defined by

$$e(t) = y(t) - \hat{y}_1(t) - \hat{y}_2(t).$$

If $|e(t)| < a$ it is assumed that $\nu(t) = 0$. The predictions are then updated as

$$\hat{y}_1(t+1) = G_1(d)\ \hat{y}_1(t)$$
$$\hat{y}_2(t+1) = G_2(d)/C_2(d)\ [y(t) - \hat{y}_1(t)].$$

If $|e(t)| > a$ it is instead assumed that $\nu(t) = e(t)$ and the predictions are updated as

$$\hat{y}_1(t) = y(t) - \hat{y}_2(t)$$
$$\hat{y}_1(t+1) = G_1(d)\ \hat{y}_1(t)$$
$$\hat{y}_2(t+1) = G_2(d)/C_2(d)\ [y(t) - \hat{y}_1(t)].$$

The selection of the test level a is a crucial part of this predictor. A simple example illustrates that the proposed predictor has interesting prospects.

Example

Consider a mixed signal characterized by

$$A_1(d) = 1 - d$$
$$A_2(d) = 1 - 1.8\,d + 0.8\,d^2$$
$$C_2(d) = 1 + 0.8\,d + 0.8\,d^2.$$

A realization of the signal is shown in Figure 2. The prediction errors are also shown in this figure. Figure 3 shows the signal and its prediction on an expanded scale. The figures show that the predictor succeeds quite well in following both the jumps and the random fluctuations.

CONCLUSIONS

Random processes like time series can be thought of as being generated by dynamical systems whose inputs are sequences of independent random variables. Piece-wise deterministic signals can similarly be thought of as outputs of dynamical systems whose inputs are zero except at isolated points. These points are not known apriori. Many properties of time series and piece-wise deterministic signals are characterized by the dynamical systems which generates them. In particular the predictors for the signals are uniquely given by the systems which generates the signals. In this paper different classes of piece-wise deterministic signals have been discussed and

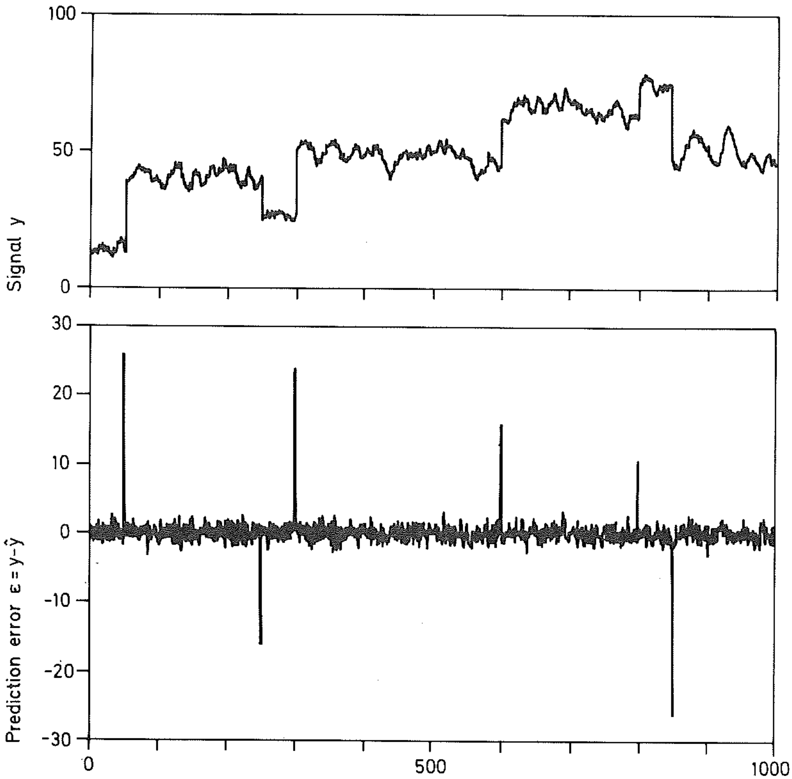

Figure 2
A realization of the signal and its prediction error.

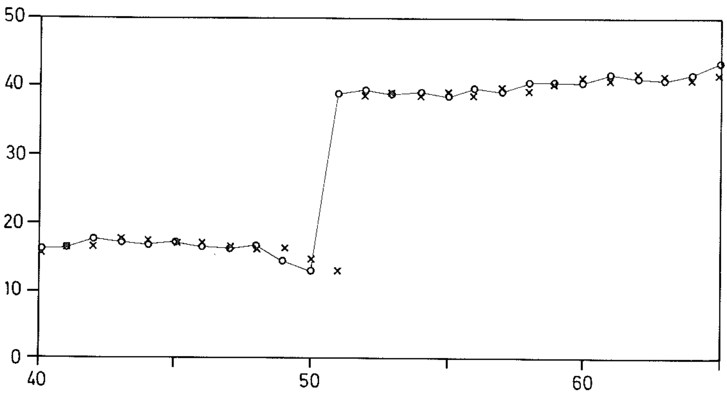

Figure 3
Signal and prediction.
The signal is denoted by o and the prediction by x.

predictors have been constructed. A natural criterion for prediction is to bring the prediction error to zero in a short time after the occurrence of a nonzero input. It has been shown that the prediction problem for piece-wise deterministic signals leads to mathematical problems which are very similar to mean square prediction problems for time series. The systems which generates the signals have been characterized both by external and internal models. This leads to slightly different approaches. Processes which are mixes of random processes and piece-wise deterministic signals have also been discussed briefly. There are several problem which merit further studies. There are some additional details to be worked out for the multivariable problem. This is straightforward. The mixed processes appear interesting. The simple example discussed indicates that useful predictors for real signals could be obtained by this approach. It would also be of interest to bring in probabilistic descriptions of the irregular points. The processes obtained will then be closely related to branching processes [11]. The parameter estimation problems for piece-wise deterministic signals are also of interest.

ACKNOWLEDGEMENTS

The support from Swedish Board of Technical Development under contract 78-3763 is gratefully acknowledged. I also wish to thank Mrs Eva Dagnegård for her careful typing of the manuscript and the camera-ready original.

REFERENCES

[1] Kailath, T.: The innovations approach to detection and estimation theory. Proc. IEEE 58 (1970) 680-695.

[2] Ahlberg, J.H., Nilson, E.N., and Walsh, J.L.: The Theory of Splines and Their Applications (Academic Press, New York, 1967).

[3] Rice, S.O.: Mathematical analysis of random noise. Bell System Tech. J. 23 (1944) 282-332 and 24 (1945) 46-156.

[4] Bartlett, M.S.: An Introduction to Stochastic Processes with Special Reference to Methods and Applications (University Press, Cambridge, 1961).

[5] Wold, H.: A Study in the Analysis of Stationary Time Series (Almqvist & Wiksell, Stockholm, 1938).

[6] Åström, K.J.: Introduction to Stochastic Control Theory (Academic Press, New York, 1970).

[7] Åström, K.J.: Stochastic control problems. In Coppel, W.A. (ed.) Mathematical Control Theory, Lecture Notes in Mathematics, Vol. 680 (Springer Verlag, Berlin, 1978).

[8] Box, G.E.P. and Jenkins, G.M.: Time series analysis forecasting and control. Holden-Day, San Francisco (1970).

[9] Kalman, R.E.: A new approach to linear filtering and prediction problems. J. Basic. Eng. 82 (1960) 34-45.

[10] Luenberger, D.G.: Observers for multivariable systems. IEEE Trans. AC-11 (1966) 190-191.

[11] Harris, T.E.: The Theory of Branching Processes (Springer Verlag, Berlin, 1963).