Step Response of a Second-Order Digital Filter With Two's Complement Arithmetic

Bingo Wing-Kuen Ling, Peter Kwong-Shun Tam, *Member, IEEE*, and Xinghuo Yu, *Senior Member, IEEE*

***Abstract*—It is well known that the autonomous response of a second-order digital filter with two's complement arithmetic may exhibit chaotic behaviors. In this paper, results of the step-response case are presented. Despite the presence of the overflow nonlinearity, it is found that the step-response behaviors can be related to some corresponding autonomous-response behaviors by means of an appropriate affine transformation. Based on this method, some differences between the step response and the autonomous response are explored. The effects of the filter parameter and input step size on the trajectory behaviors are presented. Some previous necessary conditions for the trajectory behaviors, initial conditions and symbolic sequences are extended and strengthened to become necessary and sufficient conditions. Based on these necessary and sufficient conditions, some counter-intuitive results are reported. For example, it is found that for some sets of filter parameter values, the system may exhibit the type I trajectory even when a large input step size is applied and overflow occurs. On the other hand, for some sets of filter parameter values, the system will not give the type I trajectory for any small input step size, no matter what the initial conditions are.**

***Index Terms*—Affine transformation, autonomous response, chaotic behavior, second-order digital filter with two's complement arithmetic, symbolic sequences.**

I. Introduction

MANY practical higher order digital filters can be realized by first-order digital filters and second-order digital filters in cascade and/or parallel realizations. When such a digital filter is implemented using a fixed-point microprocessor with a two's complement arithmetic for the addition operation, the physically realized filter is a nonlinear discrete-time system. Due to the nonlinearity, the dynamics of such a system may be quite complex, and chaotic behaviors may occur [1].

Many researchers had studied the chaotic behaviors of an autonomous system [1]–[10]. Some trajectory equations and sets of initial conditions corresponding to some types of trajectories are characterized in [1]. The admissibility of symbolic sequences is studied in [1], [4], [6] and [7]. The analysis is extended by considering all the real values of the filter parameter a in [8], and all the real values of the filter parameters a and b in [9]. The periodic behaviors of the symbolic sequences are discussed in [1] and [10]. A saturation-type adder overflow is analyzed in [2]. Some chaotic behaviors of a finite-state machine are presented in [5], and those of a third-order autonomous system are reported in [3].

However, practically, various types of input signals are usually applied. As a result, we need to analyze forced-input systems in addition to autonomous systems. In this paper, we concentrate on the analysis of systems with step inputs.

In Section II, we will briefly describe the system using the notations employed in the existing literature [1]–[10]. In Section III, analytical and simulation results on the system with step inputs are presented. Finally, a conclusion is presented in Section IV.

II. System Description

The system in [1]–[10] is defined as follows.

Assume a second-order digital filter can be represented by a state–space model

$$\mathbf{x}(k+1) = \begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = F(\mathbf{x}(k)). \tag{1}$$

If it can be realized by a direct form representation, then, the system can be further represented as

$$\mathbf{x}(k+1) = \begin{bmatrix} x_2(k) \\ f(b \cdot x_1(k) + a \cdot x_2(k) + u(k)) \end{bmatrix} \tag{2}$$

where a and b are the filter parameters, $u(k)$ is the input signal, $x_1(k)$ and $x_2(k)$ are the state variables, and f is the nonlinearity due to the use of two's complement arithmetic.

The nonlinearity f can be modeled as

$$f(\nu) = \nu - 2 \cdot n \tag{3}$$

such that

$$2 \cdot n - 1 \le \nu < 2 \cdot n + 1 \tag{4}$$

and $n \in Z$. Hence, the system can be represented as

$$\mathbf{x}(k+1) = \begin{bmatrix} x_2(k) \\ b \cdot x_1(k) + a \cdot x_2(k) + u(k) + 2 \cdot s(k) \end{bmatrix} \tag{5}$$

$$= \mathbf{A} \cdot \mathbf{x}(k) + \mathbf{B} \cdot u(k) + \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s(k) \tag{6}$$

for $k \ge 0$, where

$$\begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} \in I^2 \equiv \left\{ \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} : -1 \le x_1(k) < 1, -1 \le x_2(k) < 1 \right\} \tag{7}$$

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ b & a \end{bmatrix} \tag{8}$$

Manuscript received May 8, 2002; revised October 29, 2002. This work was supported by The Hong Kong Polytechnic University. The work of X. Yu was supported by the City University of Hong Kong.

The authors are with the Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong, ROC.

Digital Object Identifier 10.1109/TCSI.2003.809814

$$\mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{9}$$

and

$$s(k) \in \{-m, \ldots, -1, 0, 1, \ldots, m\} \tag{10}$$

in which m is the minimum integer satisfying

$$-2 \cdot m - 1 \le b \cdot x_1(k) + a \cdot x_2(k) + u(k) < 2 \cdot m + 1. \tag{11}$$

Since $s(k)$ is an element in the discrete set $\{-m, \ldots, -1, 0, 1, \ldots, m\}$, the values of $s(k)$ can be viewed as symbols and $s(k)$ is called a symbolic sequence.

In this paper, we only consider the case when the filter is marginally stable, that is

$$b = -1 \tag{12}$$

$$|a| \le 2 \tag{13}$$

$$u(k) = c \tag{14}$$

for $k \ge 0$, and $c \in \Re$.

It is worth noting that once the initial condition $\mathbf{x}(0)$, the filter parameters, and the input signal are given, the state variables and the symbolic sequences are uniquely defined by (5)–(14).

For the system defined above, it is reported in [1], that, for the autonomous response, there are three types of trajectories on the phase portrait, namely type I, II, and III trajectories, respectively, depending on the initial conditions. The type I–III trajectories are defined as trajectories that give a single rotated ellipse, some rotated, and translated ellipses, and an elliptical fractal pattern on the phase portrait, respectively [1]. In this paper, we investigate solutions to the following problems. Do we have similar trajectory behaviors with step inputs for the system in the presence of overflow nonlinearity? What are the effects of the parameter a and input step size on the trajectory behaviors? What are the relationships among the trajectory behaviors, symbolic sequences and the initial conditions, when overflow occurs?

III. Analytical and Simulation Results

It is obvious that, in general, the method of affine transformation cannot relate the step-response behaviors and autonomous-response behaviors in a straightforward and simplistic manner if the system is nonlinear. However, we will show, in this section, that the step-response behaviors can be related to the autonomous-response behaviors in an explicit manner even overflow nonlinearity occurs. Based on the method of affine transformation, some differences between the step response and the autonomous response are explored. The effects of the filter parameter a and input step size c, on the trajectory behaviors, are discussed. Some previous necessary conditions for the trajectory behaviors, initial conditions and symbolic sequences are extended and strengthened to become necessary and sufficient conditions. Based on these necessary and sufficient conditions, some counter-intuitive results are reported.

A. Type I Trajectory

For the type I trajectory, there is a single ellipse shown on the phase portrait. The set of initial conditions for the type I trajectory is an elliptic region on the phase portrait. In this section, some results related to the symbolic sequences and initial conditions are explored. The set of the filter parameter a and the input step size c that gives the type I trajectory is also determined. The details are explained by the following lemma.

Lemma 1: Let

$$\cos\theta = \frac{a}{2} \tag{15}$$

$$\mathbf{T} = \begin{bmatrix} 1 & 0 \\ \cos\theta & \sin\theta \end{bmatrix} \tag{16}$$

$$\widehat{\mathbf{A}} = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \tag{17}$$

$$\mathbf{x}^* = \frac{c}{2-a} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{18}$$

and

$$\widehat{\mathbf{x}}(k) = \begin{bmatrix} \hat{x}_1(k) \\ \hat{x}_2(k) \end{bmatrix} = \mathbf{T}^{-1} \cdot \left(\mathbf{x}(k) - \frac{c + 2 \cdot s_0}{c} \cdot \mathbf{x}^* \right) \tag{19}$$

where

$$s_0 \equiv s(0). \tag{20}$$

Then, the following three statements are equivalents:

i) $\widehat{\mathbf{x}}(k+1) = \widehat{\mathbf{A}} \cdot \widehat{\mathbf{x}}(k), \qquad \text{for } k \ge 0.$

ii) $s(k) = s_0, \qquad \text{for } k \ge 0.$

iii) $\mathbf{x}(0) \in \left\{ \mathbf{x}(0) \colon \left\| \mathbf{T}^{-1} \cdot \left(\mathbf{x}(0) - \frac{c + 2 \cdot s_0}{c} \cdot \mathbf{x}^* \right) \right\| \le 1 - \frac{|c + 2 \cdot s_0|}{2-a} \right\}.$

Proof: The proof can be easily obtained by the use of an affine transformation. The key steps are shown in the Appendix. ■

1) Trajectory Pattern: Since $\widehat{\mathbf{A}}$ is a rotation matrix, $\widehat{\mathbf{x}}(k+1) = \widehat{\mathbf{A}} \cdot \widehat{\mathbf{x}}(k)$ for $k \ge 0$ corresponds to a circular trajectory with center at the origin and radius

$$\|\widehat{\mathbf{x}}(0)\| = \left\| \mathbf{T}^{-1} \cdot \left(\mathbf{x}(0) - \frac{c + 2 \cdot s_0}{c} \cdot \mathbf{x}^* \right) \right\|.$$

By the transformation in Lemma 1, that is

$$\widehat{\mathbf{x}}(k) = \mathbf{T}^{-1} \cdot \left(\mathbf{x}(k) - \frac{c + 2 \cdot s_0}{c} \cdot \mathbf{x}^* \right)$$

the trajectory of $\mathbf{x}(k)$ is an elliptical orbit with the center of the ellipse at $((c + 2 \cdot s_0)/c) \cdot \mathbf{x}^*$. Since

$$\frac{c + 2 \cdot s_0}{c} \cdot \mathbf{x}^* = \frac{c + 2 \cdot s_0}{2-a} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

the center is on the diagonal line $x_2 = x_1$.

Compared to that of the autonomous response, the center of the ellipse is shifted by the vector $(c/(2-a)) \cdot \left[\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right]$, which depends on the input step size and the filter parameter a.

Fig. 1 shows the phase portrait of such a system when $a = -1.5$ and $c = 1$ with different initial conditions. Fig. 1(a) shows the trajectory when $\mathbf{x}(0) = \left[\begin{smallmatrix} -0.8 \\ -0.9 \end{smallmatrix}\right]$. Fig. 1(b) shows the corresponding symbolic sequence, $s(k) = -2$ for $k \ge 0$. Fig. 1(c) shows the trajectory when $\mathbf{x}(0) = \left[\begin{smallmatrix} 0.2 \\ -0.8 \end{smallmatrix}\right]$. Fig. 1(d)

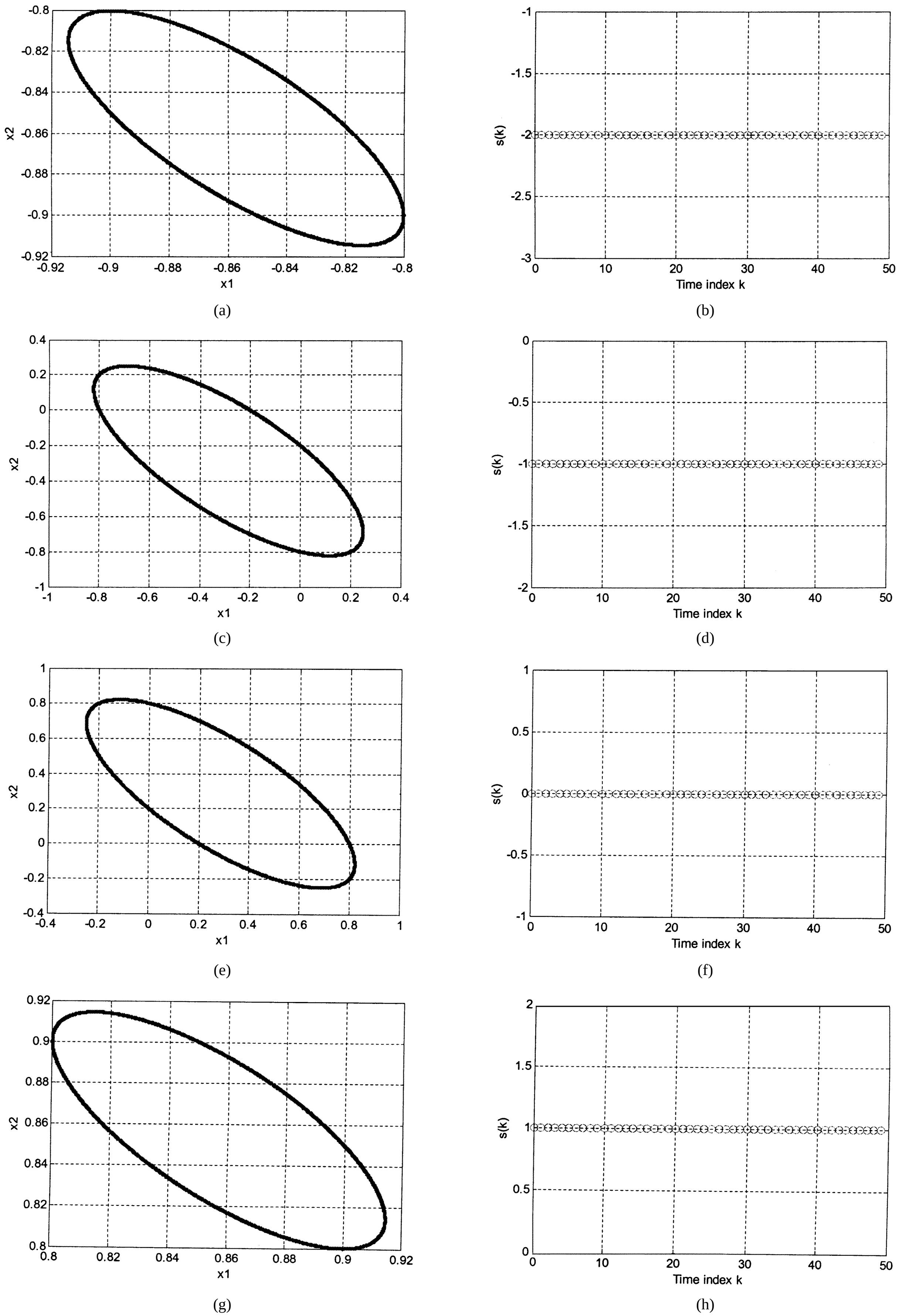


Fig. 1. Phase portrait and the symbolic sequences for the type I trajectories.

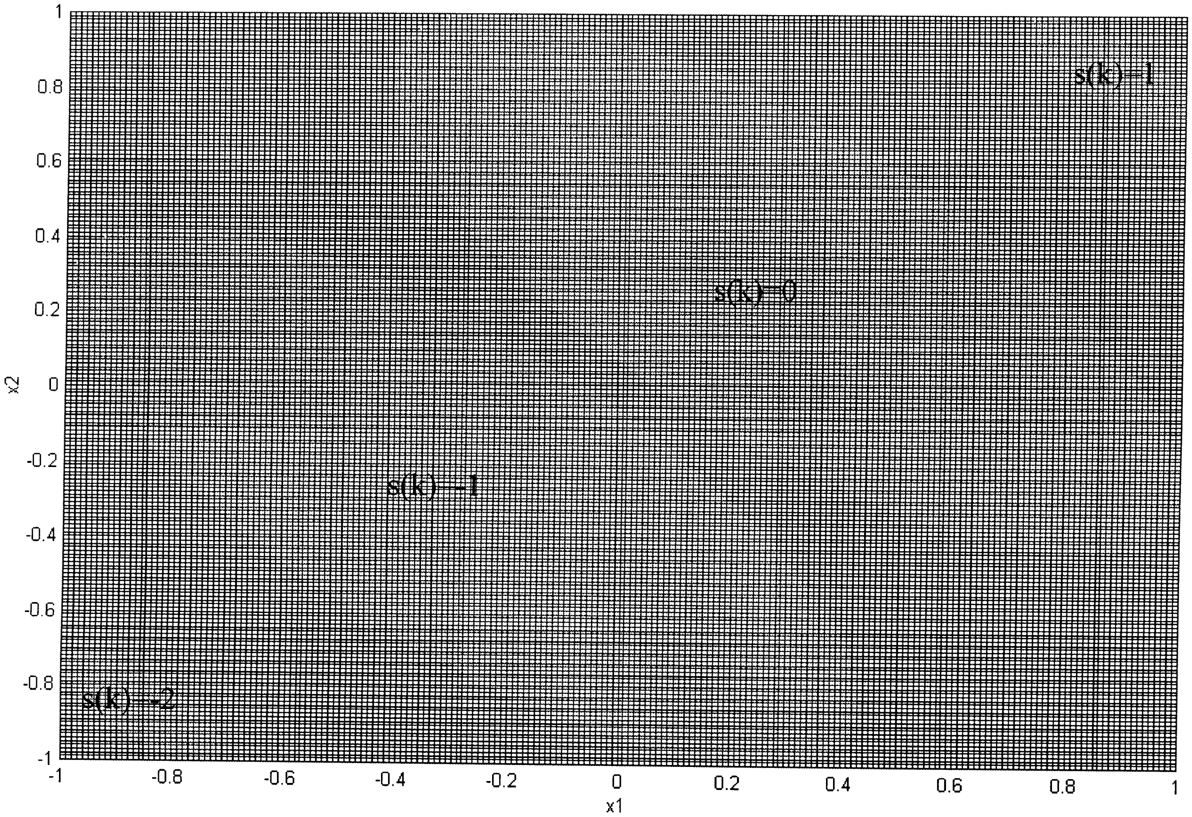


Fig. 2. Set of initial conditions that gives the type I trajectory.

shows the corresponding symbolic sequence, $s(k) = -1$ for $k \geq 0$. Fig. 1(e) shows the trajectory when $\mathbf{x}(0) = \begin{bmatrix} -0.2 \\ 0.8 \end{bmatrix}$. Fig. 1(f) shows the corresponding symbolic sequence, $s(k) = 0$ for $k \geq 0$. Fig. 1(g) shows the trajectory when $\mathbf{x}(0) = \begin{bmatrix} 0.9 \\ 0.8 \end{bmatrix}$. Fig. 1(h) shows the corresponding symbolic sequence, $s(k) = 1$ for $k \geq 0$.

The significance of Lemma 1 is that some previous necessary conditions for the trajectory behaviors, initial conditions, and symbolic sequences, are extended and strengthened to become necessary and sufficient conditions. We will have a detailed discussion on the relationship to the set of initial conditions in Section III-A-3, but we would like to mention in this section, that a system will give a type I trajectory if and only if the symbolic sequence is constant at any time instant $k \geq 0$. Based on this necessary and sufficient condition, some counter-intuitive results are found. Lemma 1 implies that the system may also give the type I trajectory even when overflow occurs, that is, $s(k) \neq 0$ for some $k \in \mathbb{Z}^+ \bigcup \{0\}$. As examples, when $s(k) = -2$ for $k \geq 0$, $s(k) = -1$ for $k \geq 0$, or $s(k) = 1$ for $k \geq 0$, overflow does occur, but the system still gives the type I trajectory, as illustrated in Fig. 1. This is a counter-intuitive phenomenon that has not been reported before.

2) Periodicity of the State Vector: In [1] and [10], it is reported that even if a single ellipse is exhibited on the phase portrait, the state variables may not be periodic. Although this is not a new result, we discuss them at this point, briefly, for completeness, by using the results in Lemma 1 and applying the approach in [11] to obtain some partial results in [1], [10] in an easy way. Since $\widehat{\mathbf{A}}$ is a rotation matrix, $\mathbf{x}(k)$ is periodic if and only if θ is a rational multiple of π [1], [10], [11]. It is worth stating here that the frequency sprectrum of $x_1(k)$ and $x_2(k)$ consists of impulses located at the dc frequency and at the natural frequency of the digital filter, no matter $\mathbf{x}(k)$ is periodic or not (It is one of various differences between continuous-time systems and discrete-time systems [12]). For the symbolic sequences, since $s(k) = s_0$ for $k \geq 0$, the frequency sprectrum of $s(k)$ consists of an impulse located at the dc frequency only.

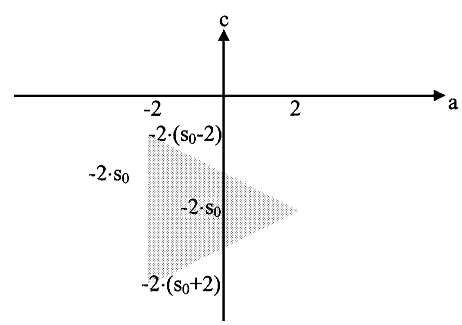


Fig. 3. Possible values of a and c for a fixed s_0.

3) Set of Initial Conditions for the Type I Trajectory: The set of initial conditions corresponding to the type I trajectory consists of rotated and translated elliptical regions. For each value of s_0, the region is characterized by a single rotated and translated ellipse with center at $((c + 2 \cdot s_0)/c) \cdot \mathbf{x}^*$. It is interesting to note that these centers are the same as that of the trajectory described in Section III-A-1, and so these centers are also on the diagonal line $x_2 = x_1$. Compared to the autonomous response,

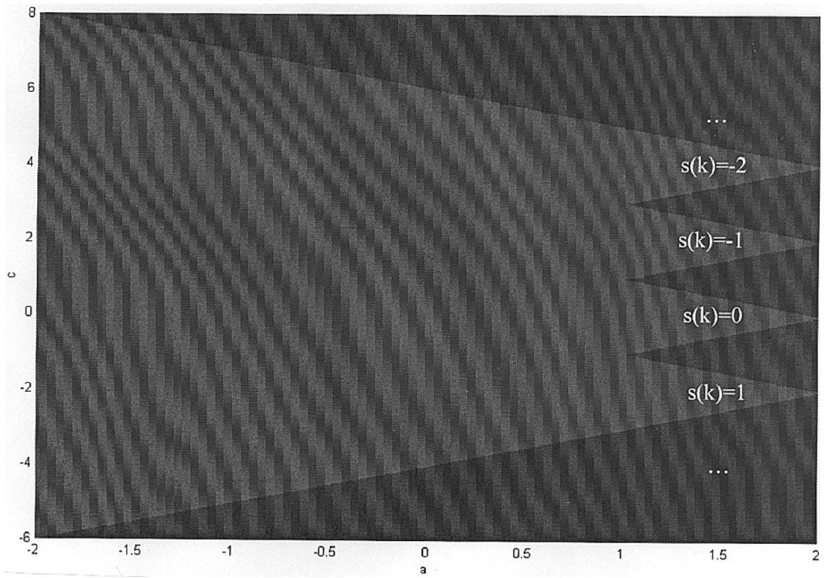


Fig. 4. Parameter space for the type I trajectory.

the set of admissible initial conditions is shifted by the vector $(c/(2-a)) \cdot \left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$, which also depends on the input step size and the filter parameter a.

Moreover, by transforming these ellipses to circles, the radii of the circles are $1 - (|c + 2 \cdot s_0|/(2 - a))$. For the autonomous-response case, the radii of the circles are $1, 1 - (2/(2 - a)), 1 - (4/(2 - a)), \ldots$, depending on the values of s_0. The corresponding radii of the circles for the step-response case are $\ldots, 1 - (|c + 2|/(2 - a))$, $1 - (|c|/(2 - a))$, $1 - (|c - 2|/(2 - a)), \ldots$. Comparing the two sequences, we can conclude that the sizes of the ellipses for the step-response case are smaller than those for the autonomous-response case.

Fig. 2 shows the set of initial conditions that gives the type I trajectory when $a = -1.5$ and $c = 1$.

4) Set of Filter Parameter and Input Step Size Values for the Type I Trajectory: The size of the elliptical region of the set of initial conditions depend on $1 - (|c + 2 \cdot s_0|/(2 - a))$, which should of course be greater than zero. This implies that $2 - a \geq |c + 2 \cdot s_0|$. For a given s_0, the possible values of a and c are in a translated triangle as shown in Fig. 3.

By combining the results for the different values of s_0, that is, combining those triangles together, we have the parameter space for the type I trajectory being characterized as shown in Fig. 4.

Note that

$$1 - \frac{|c + 2 \cdot s_0|}{2 - a} \geq 0 \;\Rightarrow\; \frac{c + 2 \cdot s_0}{c} \cdot \mathbf{x}^* \in I^2.$$

So, if a and c are in the parameter space shown in Fig. 4, then, the centers of the trajectories are in I^2 automatically.

The parameter space for the type I trajectory includes the line $c = 0$. So, for any arbitrary value of $a \in \{a\colon |a| \leq 2\}$, there exist some initial conditions such that the autonomous system will give the type I trajectory.

A counter-intuitive phenomenon can be derived from Lemma 1. The parameter space also includes the points with large values of c. It means that the system will also give the type I trajectory, even if the input step size is so large that overflow always occurs. On the other hand, there are some values of a and c which are not in the parameter space shown in Fig. 4. This region includes the case for very small values of c. This implies that the corresponding system can never give the type I trajectory even though the input step size tends to a value very close to zero, no matter what the initial conditions are. Under this condition, the system will give either the type II or type III trajectory, depending on the values of the initial conditions.

B. Type II Trajectory

Now we extend the analysis to the type II trajectory. For the type I trajectory, by the method of affine transformation, the step-response behaviors can be readily related to the autonomous-response behaviors. However, it is not the case for the type II trajectory. The method of affine transformation need to be modified, as discussed below.

1) Trajectory Pattern: It is found that there are some ellipses on the phase portrait for the step response, and the centers of these ellipses on the phase portrait are just shifted versions of that of the autonomous response. The shift depends only on filter parameter a and input step size c. However, the shifts of the different ellipses are different. The shifts also depend on the periodicity of the symbolic sequences. The detail analytical results are explained in the following lemma.

Lemma 2: For a second-order digital filter, $b \neq 0$. Otherwise, the system is a first-order system with delays. So, θ is not an integer multiple of π. Hence, we have

i) $|\mathbf{I} - \mathbf{A}^M| \neq 0$, and so $(\mathbf{I} - \mathbf{A}^M)^{-1}$ exists.

ii) By defining

$$\mathbf{x}_0^* \equiv \left(\mathbf{I} - \mathbf{A}^M\right)^{-1} \cdot \left(\sum_{j=0}^{M-1} \mathbf{A}^j \cdot \mathbf{B} \cdot c + \sum_{j=0}^{M-1} \mathbf{A}^{M-1-j} \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s(j)\right) \tag{21}$$

$$\mathbf{x}_{i+1}^* \equiv \mathbf{A} \cdot x_i^* + \mathbf{B} \cdot c + \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s(i) \tag{22}$$

for $i = 0, 1, \ldots, M-2$

$$\widehat{\mathbf{x}}_i(k) = \mathbf{T}^{-1} \cdot (\mathbf{x}(k \cdot M + i) - \mathbf{x}_i^*) \tag{23}$$

for $k \geq 0$ and $i = 0, 1, \ldots, M-1$, we have $\exists M$ such that $s(M \cdot k + i) = s(i)$ for $k \geq 0$ and $i = 0, 1, \ldots, M-1$ *if and only if* $\widehat{\mathbf{x}}_i(k+1) = \widehat{\mathbf{A}}^M \cdot \widehat{\mathbf{x}}_i(k)$ for $k \geq 0$ and $i = 0, 1, \ldots, M-1$.

Proof:

i) The proof is obvious [1], [10].

ii) For the *necessity* part, by definition

$$\mathbf{x}_0^* = \left(\mathbf{I} - \mathbf{A}^M\right)^{-1} \cdot \left(\sum_{j=0}^{M-1} \mathbf{A}^j \cdot \mathbf{B} \cdot c + \sum_{j=0}^{M-1} \mathbf{A}^{M-1-j} \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s(j)\right) \tag{24}$$

$$\Rightarrow \mathbf{x}_0^* = \mathbf{A}^M \cdot \mathbf{x}_0^* + \sum_{j=0}^{M-1} \mathbf{A}^j \cdot \mathbf{B} \cdot c + \sum_{j=0}^{M-1} \mathbf{A}^{M-1-j} \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s(j). \tag{25}$$

Since

$$\mathbf{x}_{i+1}^* = \mathbf{A} \cdot \mathbf{x}_i^* + \mathbf{B} \cdot c + \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s(i) \tag{26}$$

for $i = 0, 1, \ldots, M-2$, we have

$$\begin{aligned} &\mathbf{A} \cdot \mathbf{x}_{M-1}^* + \mathbf{B} \cdot c + \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s(M-1) \\ &= \mathbf{A}^M \cdot \mathbf{x}_0^* \\ &\quad + \sum_{j=0}^{M-1} \mathbf{A}^j \cdot \mathbf{B} \cdot c + \sum_{j=0}^{M-1} \mathbf{A}^{M-1-j} \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s(j) \\ &= \mathbf{x}_0^*. \end{aligned} \tag{27}$$

If

$$s(M \cdot k + i) = s(i) \tag{28}$$

for $k \geq 0$ and $i = 0, 1, \ldots, M-1$, we have

$$\mathbf{x}_i^* = \mathbf{A}^M \cdot \mathbf{x}_i^* + \sum_{j=0}^{M-1} \mathbf{A}^j \cdot \mathbf{B} \cdot c + \sum_{j=0}^{M-1} \mathbf{A}^{M-1-j} \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s(i+j) \tag{29}$$

for $i = 0, 1, \ldots, M-1$

$$\Rightarrow \sum_{j=0}^{M-1} \mathbf{A}^j \cdot \mathbf{B} \cdot c + \sum_{j=0}^{M-1} \mathbf{A}^{M-1-j} \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s(k \cdot M + i + j) = \left(\mathbf{I} - \mathbf{A}^M\right) \cdot \mathbf{x}_i^* \tag{30}$$

for $i = 0, 1, \ldots, M-1$.

Since

$$\mathbf{x}((k+1) \cdot M + i) = \mathbf{A}^M \cdot \mathbf{x}(k \cdot M + i) + \sum_{j=0}^{M-1} \mathbf{A}^j \cdot \mathbf{B} \cdot c + \sum_{j=0}^{M-1} \mathbf{A}^{M-1-j} \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s(k \cdot M + i + j) \tag{31}$$

for $k \geq 0$ and $i = 0, 1, \ldots, M-1$

$$\Rightarrow \mathbf{x}((k+1) \cdot M + i) = \mathbf{A}^M \cdot \mathbf{x}(k \cdot M + i) + (\mathbf{I} - \mathbf{A}^M) \cdot \mathbf{x}_i^* \tag{32}$$

for $k \geq 0$ and $i = 0, 1, \ldots, M-1$.

Since

$$\begin{aligned} \widehat{\mathbf{x}}_i(k) &= \mathbf{T}^{-1} \cdot (\mathbf{x}(k \cdot M + i) - \mathbf{x}_i^*) \Rightarrow \mathbf{x}(k \cdot M + i) \\ &= \mathbf{T} \cdot \widehat{\mathbf{x}}_i(k) + \mathbf{x}_i^* \end{aligned} \tag{33}$$

for $k \geq 0$ and $i = 0, 1, \ldots, M-1$, we have

$$\begin{aligned} \mathbf{T} \cdot \widehat{\mathbf{x}}_i(k+1) + \mathbf{x}_i^* &= \mathbf{A}^M \cdot (\mathbf{T} \cdot \widehat{\mathbf{x}}_i(k) \\ &\quad + \mathbf{x}_i^*) + (\mathbf{I} - \mathbf{A}^M) \cdot \mathbf{x}_i^* \\ &\Rightarrow \widehat{\mathbf{x}}_i(k+1) = \mathbf{T}^{-1} \cdot \mathbf{A}^M \cdot \mathbf{T} \cdot \widehat{\mathbf{x}}_i(k) \end{aligned} \tag{34}$$

for $k \geq 0$ and $i = 0, 1, \ldots, M-1$.

Since $\mathbf{A} = \mathbf{T} \cdot \widehat{\mathbf{A}} \cdot \mathbf{T}^{-1}$, we have

$$\widehat{\mathbf{x}}_i(k+1) = \widehat{\mathbf{A}}^M \cdot \widehat{\mathbf{x}}_i(k) \tag{35}$$

for $k \geq 0$ and $i = 0, 1, \ldots, M_1$, and this proves the *necessity* part.

For the *sufficiency* part, if

$$\widehat{\mathbf{x}}_i(k) = \mathbf{T}^{-1} \cdot (\mathbf{x}(k \cdot M + i) - \mathbf{x}_i^*) \tag{36}$$

and

$$\widehat{\mathbf{x}}_i(k+1) = \widehat{\mathbf{A}}^M \cdot \widehat{\mathbf{x}}_i(k) \tag{37}$$

for $k \geq 0$ and $i = 0, 1, \ldots, M-1$, then

$$\begin{aligned} \widehat{\mathbf{x}}_i(k+1) &= \mathbf{T}^{-1} \cdot (\mathbf{x}(k \cdot M + M + i) - \mathbf{x}_i^*) \\ &= \widehat{\mathbf{A}}^M \cdot \mathbf{T}^{-1} \cdot (\mathbf{x}(k \cdot M + i) - \mathbf{x}_i^*) \end{aligned} \tag{38}$$

for $k \geq 0$ and $i = 0, 1, \ldots, M-1$

$$\Rightarrow \mathbf{x}(k \cdot M + M + i) = \mathbf{A}^M \cdot \mathbf{x}(k \cdot M + i) + \left(\mathbf{I} - \mathbf{A}^M\right) \cdot \mathbf{x}_i^* \tag{39}$$

for $k \geq 0$ and $i = 0, 1, \ldots, M-1$.

But

$$\begin{aligned} \mathbf{x}(k \cdot M + M + \subset) &= \mathbf{A}^M \cdot \mathbf{x}(k \cdot M + i) \\ &\quad + \sum_{j=0}^{M-1} \mathbf{A}^j \cdot \mathbf{B} \cdot c \\ &\quad + \sum_{j=0}^{M-1} \mathbf{A}^{M-1-j} \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s(k \cdot M + i + j) \end{aligned} \tag{40}$$

for $k \geq 0$ and $i = 0, 1, \ldots, M-1$

$$\begin{aligned} &\Rightarrow \left(\mathbf{I} - \mathbf{A}^M\right) \cdot \mathbf{x}_i^* - \sum_{j=0}^{M-1} \mathbf{A}^j \cdot \mathbf{B} \cdot c \\ &= \sum_{j=0}^{M-1} \mathbf{A}^{M-1-j} \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s(k \cdot M + i + j) \end{aligned} \tag{41}$$

for $i = 0, 1, \ldots, M-1$, and $k \geq 0$, which

$$\begin{aligned} &\Rightarrow \left(\mathbf{I} - \mathbf{A}^M\right) \cdot \mathbf{x}_{i+1}^* - \sum_{j=0}^{M-1} \mathbf{A}^j \cdot \mathbf{B} \cdot c \\ &= \sum_{j=0}^{M-1} \mathbf{A}^{M-1-j} \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s(k \cdot M + i + 1 + j) \end{aligned} \tag{42}$$

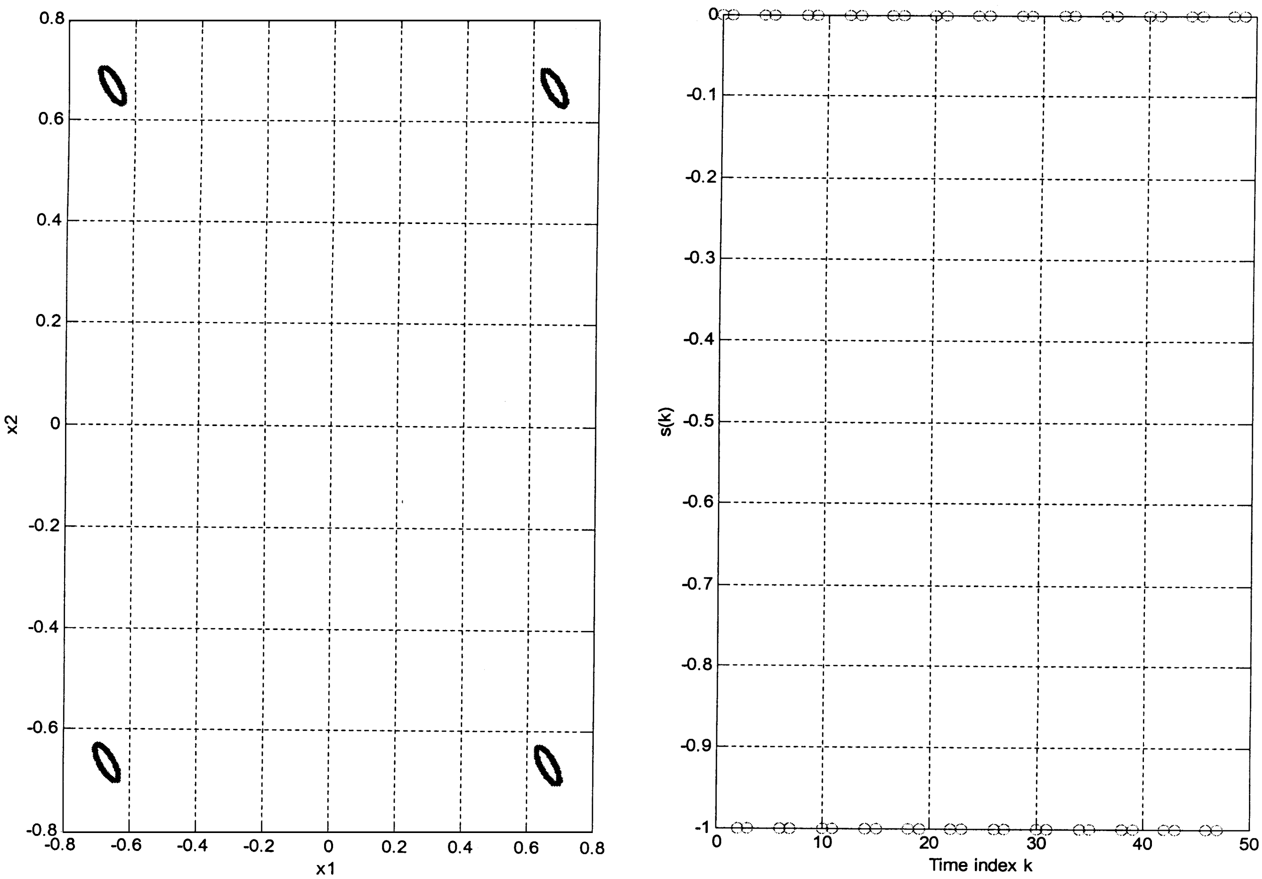


Fig. 5. Phase portrait and the corresponding symbolic sequence for the type II trajectory.

for $i = 0, 1, \ldots, M-2$, and $k \geq 0$

$$\therefore \left(\mathbf{I} - \mathbf{A}^M\right) \cdot \mathbf{x}_{i+1}^* - \sum_{j=0}^{M-1} \mathbf{A}^j \cdot \mathbf{B} \cdot c - \mathbf{A} \cdot \left(\left(\mathbf{I} - \mathbf{A}^M\right) \cdot \mathbf{x}_i^* - \sum_{j=0}^{M-1} \mathbf{A}^j \cdot \mathbf{B} \cdot c \right) = \sum_{j=0}^{M-1} \mathbf{A}^{M-1-j} \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s(k \cdot M + i + 1 + j) - \mathbf{A} \cdot \left(\sum_{j=0}^{M-1} \mathbf{A}^{M-1-j} \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s(k \cdot M + i + j) \right) \quad (43)$$

for $i = 0, 1, \ldots, M-2$, and $k \geq 0$

$$\Rightarrow \left(\mathbf{I} - \mathbf{A}^M\right) \cdot \left(\mathbf{A} \cdot \mathbf{x}_i^* + \mathbf{B} \cdot c + \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s(i) \right) - \mathbf{A} \cdot \left(\mathbf{I} - \mathbf{A}^M\right) \cdot \mathbf{x}_i^* + \left(\mathbf{A}^M - \mathbf{I}\right) \cdot \mathbf{B} \cdot c = \left(\mathbf{I} \cdot s\left(k \cdot M + M + i\right) - \mathbf{A}^M \cdot s(k \cdot M + i)\right) \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} \quad (44)$$

for $k \geq 0$ and $i = 0, 1, \ldots, M-2$

$$\Rightarrow \left(\mathbf{I} - \mathbf{A}^M\right) \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s(i) = \left(\mathbf{I} \cdot s(k \cdot M + M + i) - \mathbf{A}^M \cdot s(k \cdot M + i)\right) \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} \quad (45)$$

for $k \geq 0$ and $i = 0, 1, \ldots, M-2$.

When $k = 0$, we have

$$\left(\mathbf{I} - \mathbf{A}^M\right) \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s(i) = \left(\mathbf{I} \cdot s(M + i) - \mathbf{A}^M \cdot s(i)\right) \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} \quad (46)$$

for $i = 0, 1, \ldots, M-2$

$$\Rightarrow \; s(i) = s(M + i) \quad (47)$$

for $i = 0, 1, \ldots, M-2$.

Similarly, we can prove that

$$s(i) = s(M + i) \quad (48)$$

for $i = M-1$.

When $k = 1$, since $s(i) = s(M+i)$ for $i = 0, 1, \ldots M-1$, we have $s(i) = s(2 \cdot M + i)$, for $i = 0, 1, \ldots, M-1$. Similarly, we have $s(i) = s(k \cdot M + i)$ for $k \geq 0$ and $i = 0, 1, \ldots, M-1$. This proves the *sufficiency* part, completing the proof. ■

Since $\widehat{\mathbf{x}}_i(k+1) = \widehat{\mathbf{A}}^M \cdot \widehat{\mathbf{x}}_i(k)$ for $k \geq 0$ corresponds to a circular trajectory for each $i = 0, 1, \ldots, M-1$, there are M circles with centers at the origin and radii $\|\widehat{\mathbf{x}}_i(0)\| = \|\mathbf{T}^{-1} \cdot (\mathbf{x}(i) - \mathbf{x}_i^*)\|$ for $i = 0, 1, \ldots, M-1$ on the phase portrait when plotting the trajectory of $\widehat{\mathbf{x}}_i(k)$ for $i = 0, 1, \ldots, M-1$. Similarly, by applying these M different transformations $\widehat{\mathbf{x}}_i(k) = \mathbf{T}^{-1} \cdot (\mathbf{x}(k \cdot M + i) - \mathbf{x}_i^*)$ for $i = 0, 1, \ldots, M-1$, there are M rotated and translated ellipses on the phase portrait when plotting the trajectory of $\mathbf{x}(k)$. The centers of these ellipses are at $\mathbf{x}_i^*$ for $i = 0, 1, \ldots, M-1$.

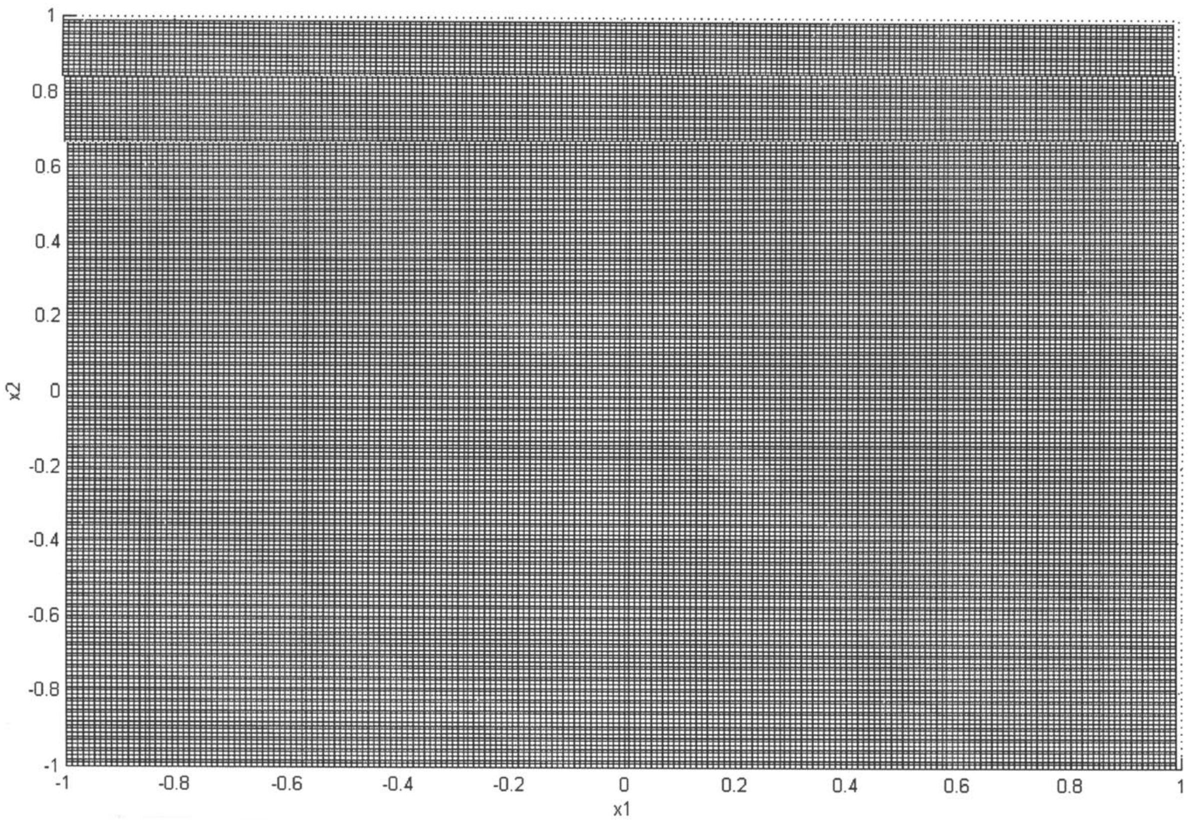


Fig. 6. Set of initial conditions that give the type II trajectory.

It is interesting to note that when $c=0$

$$\mathbf{x}_0^* = \left(\mathbf{I} - \mathbf{A}^M\right)^{-1} \cdot \sum_{j=0}^{M-1} \mathbf{A}^{M-1-j} \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s(j)$$

$$\mathbf{x}_{i+1}^* = \mathbf{A} \cdot x_i^* + \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s(i)$$

for $i = 0, 1, \ldots, M-2$. Compared to $c \neq 0$, $\mathbf{x}_0^*$ is shifted by $(\mathbf{I} - \mathbf{A}^M)^{-1} \cdot \sum_{j=0}^{M-1} \mathbf{A}^j \cdot \mathbf{B} \cdot c$, and $\mathbf{x}_i^*$ is shifted by

$$\mathbf{A}^i \cdot \left(\mathbf{I} - \mathbf{A}^M\right)^{-1} \cdot \sum_{j=0}^{M-1} \mathbf{A}^j \cdot \mathbf{B} \cdot c + \sum_{j=0}^{i-1} \mathbf{A}^{i-1-j} \cdot \mathbf{B} \cdot c$$

for $i = 1, \ldots, M-1$. Since different values of i correspond to different shift values, the centers of the ellipses are shifted in different positions. Moreover, the shift also depends on the periodicity of the symbolic sequences, which cannot be predicted by the simple affine transformation used in Lemma 1.

Fig. 5(a) shows the phase portrait of such a system with $a = -1.5$ and $c = 1$ at the initial condition $\mathbf{x}(0) = \left[\begin{smallmatrix} -0.7 \\ 0.7 \end{smallmatrix}\right]$. Fig. 5(b) shows the corresponding symbolic sequence. In this particular case, the period of the symbolic sequence is four and the trajectory consists of four ellipses with centers at $\mathbf{x}_i^*$ for $i = 0, 1, 2, 3$.

The significance of the second part of Lemma 2 is to give a necessary and sufficient condition relating the symbolic sequence and the type II trajectory. A system gives a type II trajectory if and only if the symbolic sequence is periodic with period M.

Furthermore, the second part of Lemma 2 is a generalization of the first part of Lemma 1. By substituting $M = 1$ into the second part of Lemma 2, $s(k)$ becomes a constant for $k \geq 0$, and we have $\mathbf{x}_i^* = ((c + 2 \cdot s_0)/c) \cdot \mathbf{x}^*$. This gives the same results given by the first part of Lemma 1.

2) Periodicity of State Vector: For the type I trajectory, $\mathbf{x}(k)$ is periodic if and only if θ is a rational multiple of π. For the type II trajectory, this is also a necessary and sufficient condition for the state variables to be periodic. However, the frequency sprectrum of $x_1(k)$ and $x_2(k)$ consists of impulses located at the harmonic frequencies of the symbolic sequences, that is, $(2 \cdot \pi \cdot j)/M$ for $j = 1, \ldots, M-1$, and at the natural frequency of the digital filter, no matter $\mathbf{x}(k)$ is periodic or not [12]. For the symbolic sequences, since it is periodic with period M, the frequency sprectrum of $s(k)$ consists of impulses located at its harmonic frequencies only.

3) Set of Initial Condition for the Type II Trajectory: The set of initial conditions that gives type II trajectory is given by the following lemma.

Lemma 3: $\exists M$ such that $s(M \cdot k + i) = s(i)$ for $k \geq 0$ and $i = 0, 1, \ldots, M-1$ *if and only if*

$$\mathbf{x}(0) \in \left\{\mathbf{x}(0) \colon \left\|\mathbf{T}^{-1} \cdot (\mathbf{x}(i) - \mathbf{x}_i^*)\right\| \leq 1 - \|\mathbf{x}_i^*\|_\infty\right\}$$

for $i = 0, 1, \ldots, M-1$.

Proof: For the *necessity* part, if $\exists M$ such that $s(M \cdot k + i) = s(i)$, for $k \geq 0$ and $i = 0, 1, \ldots, M-1$, then, by Lemma 2, we have $\widehat{\mathbf{x}}_i(k+1) = \widehat{\mathbf{A}}^M \cdot \widehat{\mathbf{x}}_i(k)$, for $k \geq 0$, and $i = 0, 1, \ldots, M-1$, where $\widehat{\mathbf{x}}_i(k) = \mathbf{T}^{-1} \cdot (\mathbf{x}(k \cdot M + i) - \mathbf{x}_i^*)$ for $k \geq 0$, and $i = 0, 1, \ldots, M-1$. Since the phase portrait of $\widehat{\mathbf{x}}_i(k+1) = \widehat{\mathbf{A}}^M \cdot \widehat{\mathbf{x}}_i(k)$ for $i = 0, 1, \ldots, M-1$ is a circle with radius $\|\widehat{\mathbf{x}}_i(0)\| = \|\mathbf{T}^{-1}\cdot(\mathbf{x}(i)-\mathbf{x}_i^*)\|$, for $i = 0, 1, \ldots, M-1$, we can let

$$\widehat{\mathbf{x}}_i(k) = \left\|\mathbf{T}^{-1}\cdot(\mathbf{x}(i)-\mathbf{x}_i^*)\right\|\cdot\begin{bmatrix}\cos\alpha(k)\\ \sin\alpha(k)\end{bmatrix} \tag{49}$$

for $k \geq 0$ and $i = 0, 1, \ldots, M-1$

$$\widehat{\mathbf{x}}_i(k) = \mathbf{T}^{-1}\cdot(\mathbf{x}(k\cdot M+i)-\mathbf{x}_i^*) \Rightarrow \mathbf{x}(k\cdot M+i) = \mathbf{T}\cdot\widehat{\mathbf{x}}_i(k)+\mathbf{x}_i^* \tag{50}$$

for $k \geq 0$ and $i = 0, 1, \ldots, M-1$

$$\Rightarrow \mathbf{x}(k\cdot M+i) = \left\|\mathbf{T}^{-1}\cdot(\mathbf{x}(i)-\mathbf{x}_i^*)\right\|\cdot\begin{bmatrix}\cos\alpha(k)\\ \cos(\theta-\alpha(k))\end{bmatrix}+\mathbf{x}_i^* \tag{51}$$

for $k \geq 0$ and $i = 0, 1, \ldots, M-1$.

Since $\mathbf{x}(k) \in I^2$, we have $|x_1(k)| \leq 1$ and $|x_2(k)| \leq 1$. Hence, we have

$$\left\|\mathbf{T}^{-1}\cdot(\mathbf{x}(i)-\mathbf{x}_i^*)\right\| + \|\mathbf{x}_i^*\|_\infty \leq 1 \Rightarrow \left\|\mathbf{T}^{-1}\cdot(\mathbf{x}(i)-\mathbf{x}_i^*)\right\| \leq 1-\|\mathbf{x}_i^*\|_\infty \tag{52}$$

for $i = 0, 1, \ldots, M-1$. This proves the *necessity* part.

For the *sufficiency* part, since

$$\widehat{\mathbf{x}}_i(k) = \mathbf{T}^{-1}\cdot(\mathbf{x}(k\cdot M+i)-\mathbf{x}_i^*) \tag{53}$$

for $k \geq 0$ and $i = 0, 1, \ldots, M-1$, we have

$$\widehat{\mathbf{x}}_i(k+1) = \mathbf{T}^{-1}\cdot(\mathbf{x}((k+1)\cdot M+i)-\mathbf{x}_i^*) \tag{54}$$

for $k \geq 0$ and $i = 0, 1, \ldots, M-1$.

$$\widehat{\mathbf{x}}_i(k+1) = \mathbf{T}^{-1}\cdot\left(\mathbf{A}^M\cdot\mathbf{x}(k\cdot M+i)+\sum_{j=0}^{M-1}\mathbf{A}^j\cdot\mathbf{B}\cdot c + \sum_{j=0}^{M-1}\mathbf{A}^{M-1-j}\cdot\begin{bmatrix}0\\2\end{bmatrix}\cdot s(k\cdot M+i+j)-\mathbf{x}_i^*\right) \tag{55}$$

for $k \geq 0$ and $i = 0, 1, \ldots, M-1$

$$\begin{aligned}\widehat{\mathbf{x}}_i(k+1) &= \widehat{\mathbf{A}}^M\cdot\widehat{\mathbf{x}}_i(k)+\mathbf{T}^{-1}\cdot\left(\mathbf{A}^M-\mathbf{I}\right)\cdot\mathbf{x}_i^*+\mathbf{T}^{-1}\\ &\quad\cdot\sum_{j=0}^{M-1}\mathbf{A}^j\cdot\mathbf{B}\cdot c+\mathbf{T}^{-1}\cdot\sum_{j=0}^{M-1}\mathbf{A}^{M-1-j}\cdot\begin{bmatrix}0\\2\end{bmatrix}\\ &\quad\cdot s(k\cdot M+i+j)\end{aligned} \tag{56}$$

for $k \geq 0$ and $i = 0, 1, \ldots, M-1$.

Since

$$\mathbf{x}_i^* = \mathbf{A}^i\cdot\mathbf{x}_0^*+\sum_{j=0}^{i-1}\mathbf{A}^{i-1-j}\cdot\mathbf{B}\cdot c+\sum_{j=0}^{i-1}\mathbf{A}^{i-1-j}\cdot\begin{bmatrix}0\\2\end{bmatrix}\cdot s(j) \tag{57}$$

for $i = 0, 1, \ldots, M-1$, we have

$$\left(\mathbf{I}-\mathbf{A}^M\right)\cdot\mathbf{x}_i^* = \left(\mathbf{I}-\mathbf{A}^M\right)\cdot\left(\mathbf{A}^i\cdot\mathbf{x}_0^*+\sum_{j=0}^{i-1}\mathbf{A}^{i-1-j}\cdot\mathbf{B}\cdot c + \sum_{j=0}^{i-1}\mathbf{A}^{i-1-j}\cdot\begin{bmatrix}0\\2\end{bmatrix}\cdot s(j)\right) \tag{58}$$

for $i = 0, 1, \ldots, M-1$.

Since

$$\left(\mathbf{I}-\mathbf{A}^M\right)\cdot\mathbf{x}_0^* = \sum_{j=0}^{M-1}\mathbf{A}^j\cdot\mathbf{B}\cdot c+\sum_{j=0}^{M-1}\mathbf{A}^{M-1-j}\cdot\begin{bmatrix}0\\2\end{bmatrix}\cdot s(j) \tag{59}$$

we have

$$\begin{aligned}\left(\mathbf{I}-\mathbf{A}^M\right)\cdot\mathbf{x}_i^* &= \sum_{j=0}^{M-1}\mathbf{A}^j\cdot\mathbf{B}\cdot c+\sum_{j=0}^{M-1-i}\mathbf{A}^{M-1-j}\cdot\begin{bmatrix}0\\2\end{bmatrix}\\ &\quad\cdot s(i+j)+\sum_{j=M-i}^{M-1}\mathbf{A}^{M-1-j}\cdot\begin{bmatrix}0\\2\end{bmatrix}\cdot s(i+j-M)\end{aligned} \tag{60}$$

for $i = 1, \ldots, M-1$, and

$$\begin{aligned}\widehat{\mathbf{x}}_i(k+1) &= \widehat{\mathbf{A}}^M\cdot\widehat{\mathbf{x}}_i(k)+\mathbf{T}^{-1}\cdot\sum_{j=0}^{M-1}\mathbf{A}^{M-1-j}\cdot\begin{bmatrix}0\\2\end{bmatrix}\\ &\quad\cdot s(k\cdot M+i+j)-\mathbf{T}^{-1}\cdot\sum_{j=0}^{M-1-i}\mathbf{A}^{M-1-j}\cdot\begin{bmatrix}0\\2\end{bmatrix}\\ &\quad\cdot s(i+j)-\mathbf{T}^{-1}\cdot\sum_{j=M-i}^{M-1}\mathbf{A}^{M-1-j}\cdot\begin{bmatrix}0\\2\end{bmatrix}\cdot s(i+j-M)\end{aligned} \tag{61}$$

for $k \geq 0$ and $i = 0, 1, \ldots, M-1$.

When $k = 0$, we have

$$\widehat{\mathbf{x}}_i(1) = \widehat{\mathbf{A}}^M\cdot\widehat{\mathbf{x}}_i(0)+\mathbf{T}^{-1}\cdot\left(\sum_{j=M-i}^{M-1}\mathbf{A}^{M-1-j}\cdot\begin{bmatrix}0\\2\end{bmatrix}\cdot(s(i+j)-s(i+j-M))\right) \tag{62}$$

for $i = 1, \ldots, M-1$.

If

$$\mathbf{x}(0) \in \left\{\mathbf{x}(0)\colon \left\|\mathbf{T}^{-1}\cdot(\mathbf{x}(i)-\mathbf{x}_i^*)\right\| \leq 1-\|\mathbf{x}_i^*\|_\infty\right\} \tag{63}$$

for $i = 1, \ldots, M-1$, then

$$\|\hat{\mathbf{x}}_i(0)\| \leq 1-\|\mathbf{x}_i^*\|_\infty \tag{64}$$

for $i = 0, \ldots, M-1$.

Since $\|\hat{\mathbf{A}}\| = 1$, by taking $i = 1$, we have

$$s(M) = s(0). \tag{65}$$

Similarly, by taking $i = 2$ and $s(M) = s(0)$, we have $s(M+1) = s(1)$. As a result, we have

$$s(M+i) = s(i) \tag{66}$$

for $i = 0, \ldots, M-2$.

When $k = 1$ and $i = 0$, we have

$$\widehat{\mathbf{x}}_0(2) = \widehat{\mathbf{A}}^M\cdot\widehat{\mathbf{x}}_0(1)+\mathbf{T}^{-1}\cdot\sum_{j=0}^{M-1}\mathbf{A}^{M-1-j}\cdot\begin{bmatrix}0\\2\end{bmatrix}\cdot(s(M+j)-s(j)). \tag{67}$$

Since $s(M+i) = s(i)$ for $i = 0, \ldots, M-2$, so

$$s(2\cdot M-1) = s(M-1). \tag{68}$$

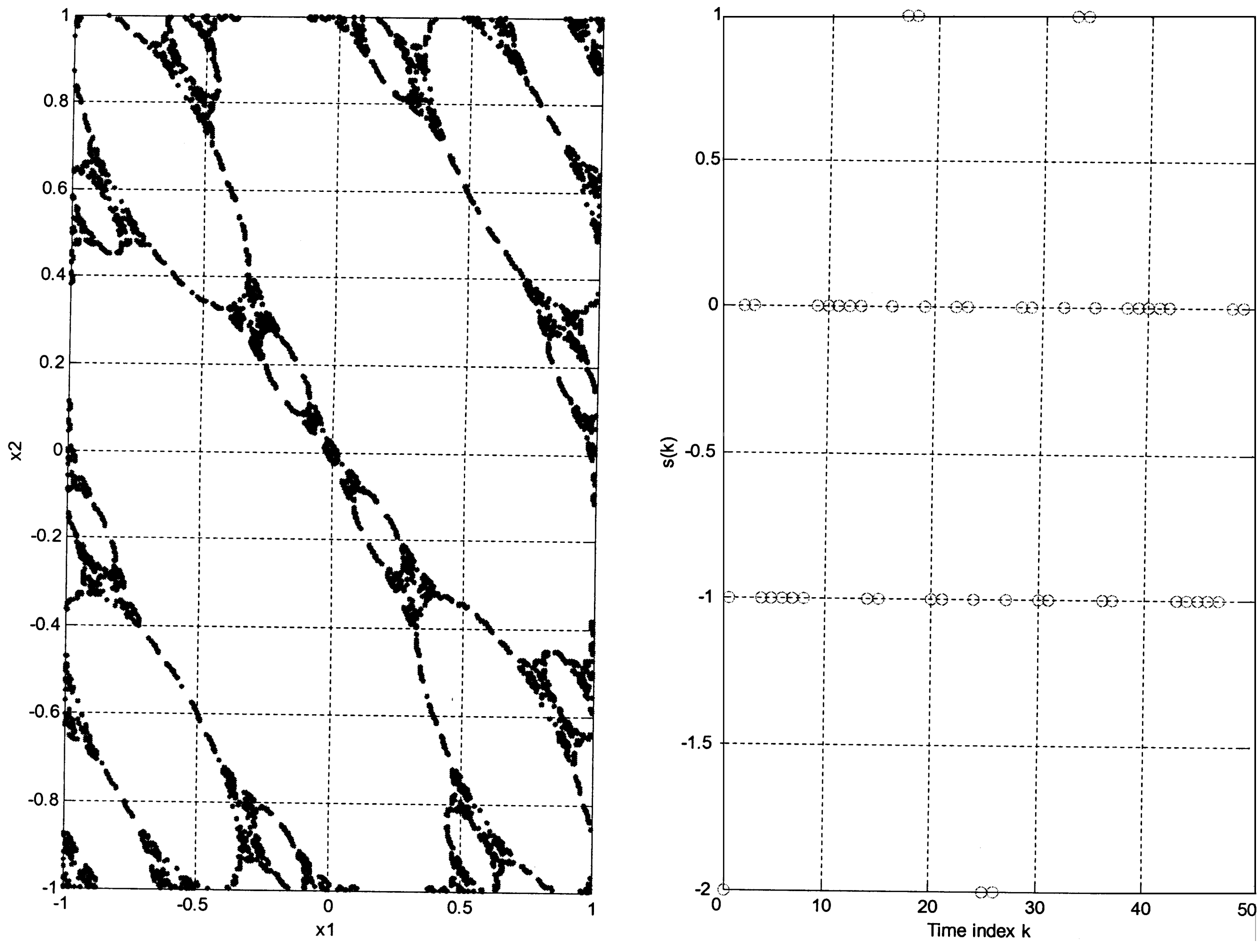


Fig. 7. Phase portrait and the corresponding symbolic sequence for the type III trajectory.

Since $\|\hat{\mathbf{x}}_i(0)\| \leq 1 - \|\mathbf{x}_i^*\|_\infty$ for $i = 0, \ldots, M-1$, we have $\|\hat{\mathbf{x}}_i(1)\| \leq 1 - \|\mathbf{x}_i^*\|_\infty$ for $i = 0, \ldots, M-1$, and $s(M+i) = s(i)$ for $i = 0, \ldots, M-1$, by mathematical induction, we conclude that $s(M \cdot k + i) = s(i)$ for $k \geq 0$ and $i = 0, \ldots, M-1$. Hence, this proves the *sufficiency* part, completing the proof. ■

Lemma 3 implies that the set of initial conditions corresponding to the type II trajectory consists of rotated and translated elliptical regions. For each periodic symbolic sequence $s(M \cdot k + i) = s(i)$ for $i = 0, \ldots, M-1$, we have a corresponding single rotated and translated ellipse with center at $\mathbf{x}_i^*$ for $i = 0, \ldots, M-1$. Overally, the centers are the same as that of the trajectory described in Lemma 2. Hence, comparing to that of the autonomous response, the centers are shifted to different positions, and the shifts depend on filter parameter a, input step size, and the periodicity of the symbolic sequences. By transforming these ellipses to circles, the radii of these circles are $1 - \|\mathbf{x}_i^*\|_\infty$ for $i = 0, \ldots, M-1$.

Fig. 6 shows the set of initial conditions that give the type II trajectory when $a = -1.5$ and $c = 1$. It can be seen from the figure that the set of initial conditions is in the elliptical regions with centers at $\mathbf{x}_i^*$ for $i = 0, \ldots, M-1$.

We can consider Lemma 3 as a generalization of the second part of Lemma 1. By substituting $M = 1$ into Lemma 1, we have the results given by second part of Lemma 1.

C. Type III Trajectory

Now, we proceed to the type III trajectory.

Remark 1: Based on our intensive simulations, we observe that the system may give an elliptical fractal pattern of trajectory if the system does not give the type I or type II trajectories.

Fig. 7(a) shows the phase portrait of such a system with $a = -1.5$ and $c = 1$ at the initial condition $\mathbf{x}(0) = \begin{bmatrix} -0.9 \\ -0.98 \end{bmatrix}$. Fig. 7(b) shows the corresponding symbolic sequence. In this particular case, there is an elliptical fractal pattern shown on the phase portrait and the symbolic sequences are aperiodic.

Remark 2: The symbolic sequences and the state variables are, in general, aperiodic. Hence, the frequency spectrum of the state variables and the symbolic sequences are continuous.

Remark 3: Let $D_M = \{\mathbf{x}(0) \colon \|\mathbf{T}^{-1} \cdot (\mathbf{x}(i) - \mathbf{x}_i^*)\| \leq 1 - \|\mathbf{x}_i^*\|_\infty$ and $s(i) = s(i+M)\}$ for $M \in Z \backslash \{0\}$ and $i = 0, \ldots, M-1$. Define $D \equiv I^2 \backslash \bigcap_{\forall M} D_M$. Then, the set of initial conditions that may give the type III trajectory is D. That is, the system with initial condition in the rest of the space after taking away the regions indicated in Figs. 2 and 6 may give rise to an elliptical fractal pattern of trajectory.

Fig. 8 shows the set of initial conditions that will give the type III trajectory when $a = -1.5$ and $c = 1$.

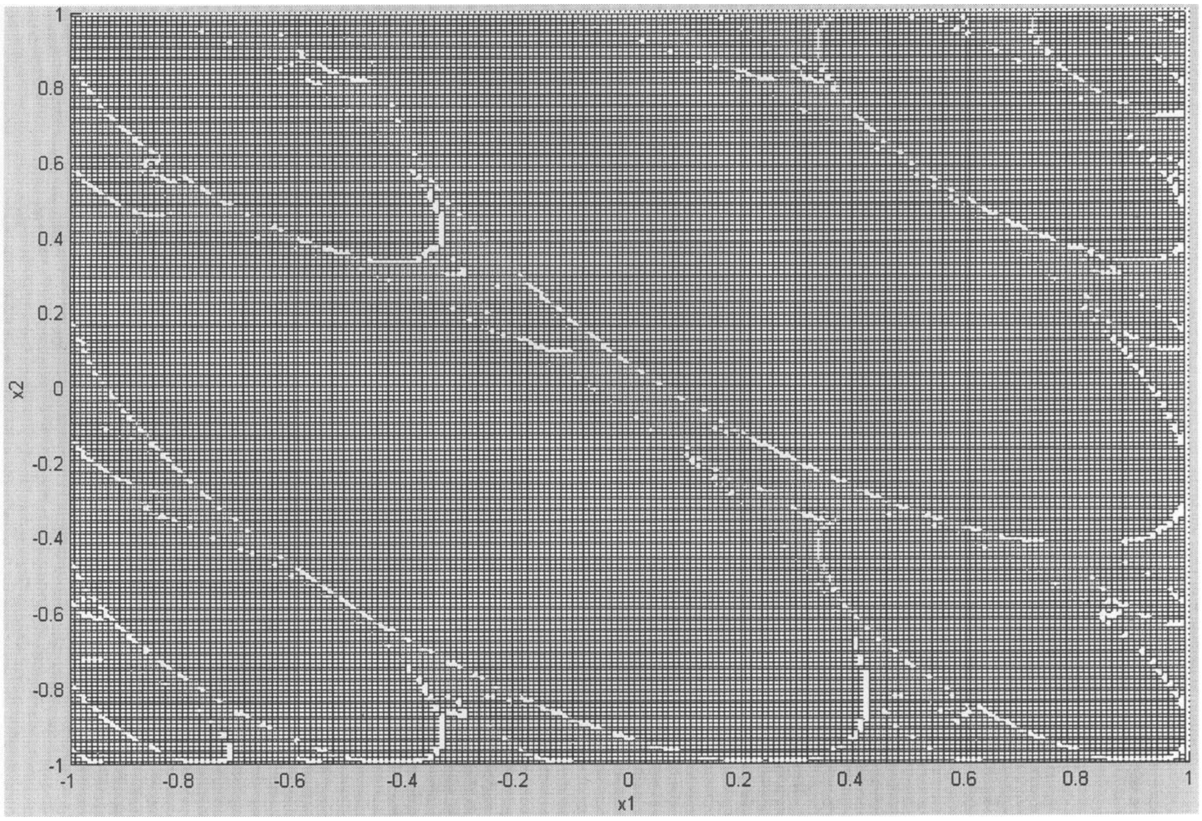


Fig. 8. Set of initial conditions that gives the type III trajectory.

IV. CONCLUSION

The main focus of this paper is the analysis of second-order digital filters with two's complement arithmetic, when there are step inputs. Even though in the presence of the overflow nonlinearity, the step-response behaviors can be related to some corresponding autonomous-response behaviors by means of affine transformations. Based on this method, some differences between the step response and autonomous response are explored.

For the type I trajectory, there is a single ellipse on the phase portrait where the center of the ellipse is located on the diagonal line. Compared to that of the autonomous response, the ellipse is shifted. The shift depends on the filter parameter and the input step size. For the type II trajectory, we have similar results. There are some ellipses on the phase portrait and these ellipses are the shifted versions of the autonomous-response case. However, different ellipses are shifted in different positions, and the shifts also depend on the periodicity of the symbolic sequences. For the rest of the space of initial conditions, the system with step input may exhibit the type III trajectory with an elliptical fractal pattern shown on the phase portrait.

Besides, we have found the sets of initial conditions for different types of trajectory. The sets of initial conditions for the type I and type II trajectories are elliptical regions on the phase portrait, while the centers of these ellipses are the centers of their elliptical trajectories. The sizes of those ellipses for the type I trajectory are smaller compared to that of the autonomous case. For the type III trajectory, the set of initial conditions is the unit square minus the sets of initial conditions that give the type I or the type II trajectory.

Some previous necessary conditions for the trajectory behaviors, initial conditions, and symbolic sequences are extended and strengthened to become necessary and sufficient conditions. It is proved in this paper, that the system gives the type I trajectory if and only if the symbolic sequences are constant at any time instant. This implies that the system will give the type I trajectory even when overflow occurs. The system gives the type II trajectory if and only if the symbolic sequences are periodic with period greater than one. In general, the results obtained in the type I trajectory can be obtained by setting the period equal to one into the results obtained in the type II trajectory. For the type III trajectory, the symbolic sequences are aperiodic.

The filter parameter values and the input step sizes that might give the type I trajectory are found, and some counter-intuitive results are reported. It is interesting to note that for some filter parameters, the system may give the type I trajectory even when a large input step size is applied. On the other hand, the system will not give the type I trajectory for some values of filter parameter and any arbitrarily small input step size no matter what the initial conditions are. Under this situation, the system will give either the type II or the type III trajectory, depending on the initial conditions.

APPENDIX

This Appendix is to prove Lemma 1. The proof can be divided into two parts. The first part is to prove $\widehat{\mathbf{x}}(k+1) = \widehat{\mathbf{A}} \cdot \widehat{\mathbf{x}}(k)$

for $k \geq 0$ if and only if $s(k) = s_0$ for $k \geq 0$. And the second part is to prove $s(k) = s_0$ for $k \geq 0$ if and only if

$$\mathbf{x}(0) \in \left\{ \mathbf{x}(0) \colon \left\| \mathbf{T}^{-1} \cdot \left(\mathbf{x}(0) - \frac{c + 2 \cdot s_0}{c} \cdot \mathbf{x}^* \right) \right\| \leq 1 - \frac{|c + 2 \cdot s_0|}{2 - a} \right\}.$$

For the *necessity* of the first part, since $\mathbf{x}(k+1) = \mathbf{A} \cdot \mathbf{x}(k) + \mathbf{B} \cdot u(k) + \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s(k)$

$$u(k) = c, \ \text{ and } \ s(k) = s_0, \qquad \text{for } k \geq 0, \ \mathbf{B} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{69}$$

we have

$$\mathbf{x}(k+1) = \mathbf{A} \cdot \mathbf{x}(k) + (c + 2 \cdot s_0) \cdot \mathbf{B} \tag{70}$$

$$\widehat{\mathbf{x}}(k+1) = \mathbf{T}^{-1} \cdot \left(\mathbf{x}(k+1) - \frac{c + 2 \cdot s_0}{c} \cdot \mathbf{x}^* \right) \tag{71}$$

$$\Rightarrow \widehat{\mathbf{x}}(k+1) = \mathbf{T}^{-1} \cdot \left(\mathbf{A} \cdot \mathbf{x}(k) + (c + 2 \cdot s_0) \cdot \mathbf{B} - \frac{c + 2 \cdot s_0}{c} \cdot \mathbf{x}^* \right). \tag{72}$$

Since $\mathbf{A} = \mathbf{T} \cdot \widehat{\mathbf{A}} \cdot \mathbf{T}^{-1}$ and $\mathbf{B} \cdot c = (\mathbf{I} - \mathbf{A}) \cdot \mathbf{x}^*$, we have

$$\widehat{\mathbf{x}}(k+1) = \widehat{\mathbf{A}} \cdot \widehat{\mathbf{x}}(k) \tag{73}$$

and this proves the *necessity* of the first part.

For the *sufficiency* of the first part, if

$$\widehat{\mathbf{x}}(k+1) = \widehat{\mathbf{A}} \cdot \widehat{\mathbf{x}}(k) \tag{74}$$

then

$$\mathbf{T}^{-1} \cdot \left(\mathbf{x}(k+1) - \frac{c + 2 \cdot s_0}{c} \cdot \mathbf{x}^* \right) = \widehat{\mathbf{A}} \cdot \mathbf{T}^{-1} \cdot \left(\mathbf{x}(k) - \frac{c + 2 \cdot s_0}{c} \cdot \mathbf{x}^* \right) \tag{75}$$

$$\Rightarrow \mathbf{x}(k+1) = \mathbf{A} \cdot \mathbf{x}(k) + \mathbf{B} \cdot c + \begin{bmatrix} 0 \\ 2 \end{bmatrix} \cdot s_0 \tag{76}$$

$$\Rightarrow s(k) = s_0 \tag{77}$$

for $k \geq 0$. This proves the *sufficiency* of the first part.

For the *necessity* of the second part, when $s(k) = s_0$ for $k \geq 0$, from the above, we have $\widehat{\mathbf{x}}(k+1) = \widehat{\mathbf{A}} \cdot \widehat{\mathbf{x}}(k)$ for $k \geq 0$, where

$$\widehat{\mathbf{x}}(k) = \mathbf{T}^{-1} \cdot \left(\mathbf{x}(k) - \frac{c + 2 \cdot s_0}{c} \cdot \mathbf{x}^* \right).$$

Since the phase portrait of $\widehat{\mathbf{x}}(k+1) = \widehat{\mathbf{A}} \cdot \widehat{\mathbf{x}}(k)$ is a circle with radius

$$\left\| \mathbf{T}^{-1} \cdot \left(\mathbf{x}(0) - \frac{c + 2 \cdot s_0}{c} \cdot \mathbf{x}^* \right) \right\|$$

we can let

$$\widehat{\mathbf{x}}(k) = \left\| \mathbf{T}^{-1} \cdot \left(\mathbf{x}(0) - \frac{c + 2 \cdot s_0}{c} \cdot \mathbf{x}^* \right) \right\| \cdot \begin{bmatrix} \cos \phi(k) \\ \sin \phi(k) \end{bmatrix} \tag{78}$$

$$\Rightarrow \mathbf{x}(k) = \left\| \mathbf{T}^{-1} \cdot \left(\mathbf{x}(0) - \frac{c + 2 \cdot s_0}{c} \cdot \mathbf{x}^* \right) \right\| \cdot \begin{bmatrix} \cos \phi(k) \\ \cos(\theta - \phi(k)) \end{bmatrix} + \frac{c + 2 \cdot s_0}{2 - a} \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}. \tag{79}$$

Since $\mathbf{x}(k) \in I^2$, that is

$$|x_1(k)| \leq 1 \tag{80}$$

and

$$|x_2(k)| \leq 1 \tag{81}$$

we have

$$\left| \left\| \mathbf{T}^{-1} \cdot \left(\mathbf{x}(0) - \frac{c + 2 \cdot s_0}{c} \cdot \mathbf{x}^* \right) \right\| \cdot \cos(\phi(k)) + \frac{c + 2 \cdot s_0}{2 - a} \right| \leq 1 \tag{82}$$

$$\Rightarrow \left\| \mathbf{T}^{-1} \cdot \left(\mathbf{x}(0) - \frac{c + 2 \cdot s_0}{c} \cdot \mathbf{x}^* \right) \right\| \leq 1 - \frac{|c + 2 \cdot s_0|}{2 - a} \tag{83}$$

and this proves the *necessity* of the second part.

For the *sufficiency* of the second part, since

$$\widehat{\mathbf{x}}(k) = \mathbf{T}^{-1} \cdot \left(\mathbf{x}(k) - \frac{c + 2 \cdot s_0}{c} \cdot \mathbf{x}^* \right)$$

$\mathbf{A} = \mathbf{T} \cdot \widehat{\mathbf{A}} \cdot \mathbf{T}^{-1}$ and

$$\mathbf{B} \cdot c = (\mathbf{I} - \mathbf{A}) \cdot \mathbf{x}^* \tag{84}$$

we have

$$\widehat{\mathbf{x}}(k+1) = \widehat{\mathbf{A}} \cdot \widehat{\mathbf{x}}(k) + \mathbf{T}^{-1} \cdot (s(k) - s_0) \cdot \begin{bmatrix} 0 \\ 2 \end{bmatrix}. \tag{85}$$

If

$$\mathbf{x}(0) \in \left\{ \mathbf{x}(0) \colon \left\| \mathbf{T}^{-1} \cdot \left(\mathbf{x}(0) - \frac{c + 2 \cdot s_0}{c} \cdot \mathbf{x}^* \right) \right\| \leq 1 - \frac{|c + 2 \cdot s_0|}{2 - a} \right\}$$

then

$$\|\hat{\mathbf{x}}(0)\| \leq 1 - \frac{|c + 2 \cdot s_0|}{2 - a}.$$

Since $\|\widehat{\mathbf{A}}\| = 1$, we have $s(0) = s_0$ and

$$\|\hat{\mathbf{x}}(1)\| = \left\| \widehat{\mathbf{A}} \cdot \widehat{\mathbf{x}}(0) \right\| \leq 1 - \frac{|c + 2 \cdot s_0|}{2 - a}.$$

Assume that

$$\mathbf{x}(k) \in \left\{ \mathbf{x}(k) \colon \left\| \mathbf{T}^{-1} \cdot \left(\mathbf{x}(k) - \frac{c + 2 \cdot s_0}{c} \cdot \mathbf{x}^* \right) \right\| \leq 1 - \frac{|c + 2 \cdot s_0|}{2 - a} \right\}$$

then, using an approach similar to that for $\mathbf{x}(0)$, we can show that $s(k) = s_0$ and

$$\|\hat{\mathbf{x}}(k+1)\| = \left\| \widehat{\mathbf{A}} \cdot \widehat{\mathbf{x}}(k) \right\| \leq 1 - \frac{|c + 2 \cdot s_0|}{2 - a}. \tag{86}$$

Hence, by mathematical induction, we conclude that $s(k) = s_0$ for $k \geq 0$, and this proves the *sufficiency* of the second part, and completing the proof. ■

REFERENCES

[1] L. O. Chua and T. Lin, "Chaos in digital filters," *IEEE Trans. Circuits Syst.*, vol. 35, pp. 648–658, June 1988.

[2] Z. Galias and M. J. Ogorzalek, "Bifurcation phenomena in second-order digital filter with saturation-type adder overflow characteristic," *IEEE Trans. Circuits Syst.*, vol. 37, pp. 1068–1070, Aug. 1990.

[3] L. O. Chua and T. Lin, "Chaos and fractals from third-order digital filters," *Int. J. Circuit Theory Applicat.*, vol. 18, pp. 241–255, 1990.

[4] ——, "Fractal pattern of second-order nonlinear digital filters: A new symbolic analysis," *Int. J. Circuit Theory Applicat.*, vol. 18, pp. 541–550, 1990.

[5] T. Lin and L. O. Chua, "On chaos of digital filters in the real world," *IEEE Trans. Circuits Syst.*, vol. 38, pp. 557–558, May 1991.

[6] Z. Galias and M. J. Ogorzalek, "On symbolic dynamics of a chaotic second-order digital filter," *Int. J. Circuit Theory Applicat.*, vol. 20, pp. 401–409, 1992.

[7] C. W. Wu and L. O. Chua, "Properties of admissible symbolic sequences in a second-order digital filter with overflow nonlinearity," *Int. J. Circuit Theory Applicat.*, vol. 21, pp. 299–307, 1993.

[8] L. Kocarev and L. O. Chua, "On chaos in digital filters: Case $b = -1$," *IEEE Trans. Circuits Syst. II*, vol. 40, pp. 404–407, June 1993.

[9] L. Kocarev, C. W. Wu, and L. O. Chua, "Complex behavior in digital filters with overflow nonlinearity: Analytical results," *IEEE Trans. Circuits Syst. II*, vol. 43, pp. 234–246, Mar. 1996.

[10] X. Yu and Z. Galias, "Periodic behaviors in a digital filter with two's complement arithmetic," *IEEE Trans. Circuits Syst. I*, vol. 48, pp. 1177–1190, Oct. 2001.

[11] R. L. Devaney, *An Introduction to Chaotic Dynamical Systems*. Reading, MA: Addison-Wesley, 1986.

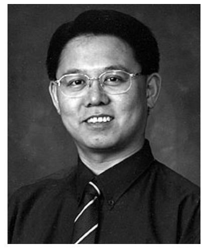
[12] A. L. Oppenheim and A. S. Willsky, *Signals and Systems*. Englewood Cliffs, NJ: Prentice-Hall, 1997.

Bingo Wing-Kuen Ling received the Bachelor's and Master's degrees in electronic engineering from the the Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, in 1997 and 2000, respectively, and is currently working toward the doctoral degree in electronic and information engineering, at the Hong Kong Polytechnic University, Kowloon, Hong Kong, in the area of chaos.

His research interests include filter bank and wavelet theory, image processing, and chaos.

Peter Kwong-Shun Tam (S'74–M'76) received the B.E., M.E., and Ph.D. degrees from the University of Newcastle, Newcastle, Australia, in 1971, 1973, and 1976, respectively, all in electrical engineering.

From 1967 to 1980, he held a number of industrial and academic positions in Australia. In 1980, he joined the Hong Kong Polytechnic University, Kowloon, Hong Kong, as a Senior Lecturer, where he is currently an Associate Professor in the Department of Electronic and Information Engineering. His research interests include signal processing, automatic control, fuzzy systems, and neural networks. He has participated in the organization of a number of symposiums and conferences.



Xinghuo Yu (M'91–SM'96) received the B.Sc. (EEE) and M.Sc. (EEE) from the University of Science and Technology, Heifei, China, in 1982 and 1985, respectively, and the Ph.D. degree from South-East University, Nanjing, China, in 1988.

From 1987 to 1989, he was a Research Fellow with the Institute of Automation, Chinese Academy of Sciences, Beijing, China. From 1989 to 1991, he was a Postdoctoral Fellow with the Applied Mathematics Department, University of Adelaide, Adelaide, Australia. From 1991 to 2002, he was with Central Queensland University, Rockhampton, Australia where, before he left, he was Professor of Intelligent Systems, and the Associate Dean (Research) of the Faculty of Informatics and Communication. Since March 2002, he has been with the Royal Melbourne Institute of Technology, Melbourne, Australia, where, from March 2002 to January 2003 he was a Professor, Deputy of School, and Director for Software and Networks, School of Electrical and Computer Engineering, and has been Associate Dean Research & Development of the Faculty of Engineering since January 2003. He has also held Visiting Professor positions in the City University of Hong Kong, Hong Kong, and Bogazici University, Istanbul, Turkey. He is a Guest Professor at Harbin Institute of Technology, Harbin, China, Huazhong University of Science and Technology, Wuhan, China, and South-East University. His research interests include chaos and chaos control, antichaos, soft computing, and applications, and nonlinear control. He has published over 200 refereed papers in technical journals, books and conference proceedings. He has also coedited five research books.

Prof. Yu serves as an Associate Editor of IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS —PART I, and is on the Editorial Board of the *International Journal of Applied Mathematics and Computer Science*. He has been on program committees of many international conferences, and was the General Chair of the 6th IEEE International Workshop on Variable Structure Systems held in December 2000 in Australia. He was the sole recipient of the 1995 Central Queensland University Vice Chancellor's Award for Research. He is also a Fellow of Institution of Engineers Australia. He has recently been conferred an Emeritus Professor of Central Queensland University Australia.