A Complete Stability Analysis of Planar Discrete-Time Linear Systems Under Saturation

Tingshu Hu and Zongli Lin, *Senior Member, IEEE*

***Abstract*—A complete stability analysis is performed on a planar discrete-time system of the form $x(k+1) = \mathrm{sat}(Ax(k))$, where A is a Schur stable matrix and sat is the saturation function. Necessary and sufficient conditions for the system to be globally asymptotically stable are given. In the process of establishing these conditions, the behaviors of the trajectories are examined in detail.**

***Index Terms*—Limit trajectories, neural networks, saturation, stability.**

I. Introduction

DYNAMICAL systems with saturation nonlinearities arise frequently in neural networks, analogue circuits and control systems (see, for example, [9], [5], [2], [6] and the references therein). In this paper, we consider systems of the following form:

$$x(k+1) = \mathrm{sat}(Ax(k)), \qquad x \in \mathbf{R}^n \tag{1}$$

where sat $\mathbf{R}^n \to \mathbf{R}^n$ is the standard saturation function. With a slight abuse of notation, we use the same symbol to denote both the vector saturation function and the scalar saturation function, i.e., if $v \in \mathbf{R}^n$, then sat $(v) = [\mathrm{sat}(v_1), \mathrm{sat}\,(v_2), \ldots, \mathrm{sat}(v_n)]^T$ and

$$\mathrm{sat}(v_i) = \begin{cases} -1, & \text{if } v_i < -1 \\ v_i, & \text{if } -1 \le v_i \le 1 \\ 1, & \text{if } v_i > 1. \end{cases} \tag{2}$$

Systems of the form (1) and their continuous-time counterparts mainly arise in neural networks and in digital filters.

As with any dynamical system, stability of these systems is of primary concern and has been heavily studied in the literature for a long period of time (see, for example, [1], [7]–[11] and the references therein). As seen in the literature, the stability analysis of such systems are highly nontrivial even for the planar case. For the continuous-time counterpart of (1), only until recently have the necessary and sufficient conditions for global asymptotic stability (GAS) been established for the planar case [4]. For the planar discrete-time system of the form (1), to the best of our knowledge, no necessary and sufficient conditions have been known, although various sufficient conditions are available [9], [11]. This paper attempts to carry out a complete stability analysis of planar systems of the form (1). In particular, necessary and sufficient conditions for the system to be GAS will be identified. In the process of establishing these conditions, the behaviors of the trajectories are examined in detail.

This work is motivated by our recent result [4] on the planar continuous-time system

$$\dot{x} = \mathrm{sat}(Ax), \qquad x \in \mathbf{R}^2. \tag{3}$$

However, the two systems (1) and (3) behave quite differently even though they have a similar description. First of all, (3) operates on the entire plane while (1) operates only on the unit square. The trajectories of (3) do not intersect each other but the connected trajectory of (1) [by connecting $x(k)$ and $x(k+1)$] can intersect itself. The limit trajectories of (3) must be periodic but a limit trajectory of (1) need not be. Finally, it is known that in the stability analysis for nonlinear systems, many more tools are available for continuous-time systems than for discrete-time systems.

We will start our investigation of the planar system (1) by characterizing some general properties of its limit trajectories. An important feature is that a nontrivial limit trajectory can only intersect two opposite pair of boundaries of the unit square and it cannot have intersections with both of the neighboring boundaries. This result turns our attention to a simpler system which has only one saturated state

$$x(k+1) = \begin{bmatrix} a_{11}x_1(k) + a_{12}x_2(k) \\ \mathrm{sat}\,(a_{21}x_1(k) + a_{22}x_2(k)) \end{bmatrix}. \tag{4}$$

For this simpler system, we will establish a relation between the present intersection of a trajectory with the lines $x_2 = \pm 1$ and the next one in terms of a set of points on the line $x_2 = 1$. The relation is discontinuous but piecewise linear. The set of points are the places where the discontinuity occurs. Some attractive properties about these points and the relation between the next intersection and the present one are revealed. These properties help us to establish the condition for the system (4) to be GAS and to characterize an interval on the line $x_2 = 1$ from which the trajectories of (4) will converge to the origin. This in turn leads to our final result on the necessary and sufficient conditions for the GAS of a planar system of the form (1).

This paper is organized as follows. In Section II, we give the necessary and sufficient conditions for the GAS of the planar system in the form of (1). An example is also given to help interpret these conditions. These conditions are established in Sections III–V. In the process of establishing these conditions, the intricate properties of the system trajectories are also revealed. In particular, Section III reveals some general properties of the

Manuscript received July 6, 2000; revised December 23, 2000. This work was supported in part by the US Office of Naval Research Young Investigator Program under Grant N00014-99-1-0670. This paper was recommended by Associate Editor P. K. Rajan.

The authors are with the Department of Electrical Engineering, University of Virginia, Charlottesville, VA 22903 USA.

Publisher Item Identifier S 1057-7122(01)04290-8.

possible limit trajectories of the system which help us to exclude the existence of limit trajectories under the condition of the main theorem and focus our attention to the simpler system with one saturated state. Section IV investigates system (4) and gives a necessary and sufficient condition for the system to be GAS. Section V proves the main result of this paper. Finally, a brief concluding remark is made in Section VI.

II. Main Results

Consider the following system:

$$x(k+1) = \operatorname{sat}(Ax(k)), \qquad x \in \mathbf{R}^2 \tag{5}$$

where $A = \left[\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right]$ and sat: $\mathbf{R}^2 \to \mathbf{R}^2$ is the saturation function, i.e., if $v \in \mathbf{R}^2$, then $\operatorname{sat}(v) = \left[\begin{smallmatrix} \operatorname{sat}(v_1) \\ \operatorname{sat}(v_2) \end{smallmatrix}\right]$ and sat$(\cdot)$ is as defined by (2).

Given an initial state $x(0) = x_0$, denote the trajectory of the system (5) that passes through x_0 at $k = 0$ as $\psi(k, x_0)$. In this paper, we only consider the positive trajectories. Hence, throughout the paper, $k \geq 0$.

Definition 2.1: The system (5) is said to be stable at its equilibrium $x_e = 0$ if, for any $\varepsilon > 0$, there exists a $\delta > 0$ such that, $\|\psi(k, x_0)\| \leq \varepsilon$, for all $k \geq 0$ and $\|x_0\| \leq \delta$. It is said to be globally asymptotically stable (GAS) if $x_e = 0$ is a stable equilibrium and $\lim_{k\to\infty} \psi(k, x_0) = 0$ for all $x_0 \in \mathbf{R}^2$. Also, it is said to be locally asymptotically stable if it is stable and $\lim_{k\to\infty} \psi(k, x_0) = 0$ for all x_0 in a neighborhood U_0 of $x_e = 0$.

The system is GAS only if it is locally asymptotically stable, which is equivalent to that A has eigenvalues inside the unit circle. In this case, A is said to be Schur stable, or simply stable. In this paper, we assume that A is stable. Denote the closed unit square as $\mathbf{S}$ and its boundary as $\partial\mathbf{S}$. It is easy to see that no matter where $x(0)$ is, we always have $x(1) \in \mathbf{S}$. Hence, the global asymptotic stability is equivalent to $\lim_{k\to\infty} \psi(k, x_0) = 0$ for all $x_0 \in \mathbf{S}$. The main result of this paper is presented as follows:

Theorem 2.1: The system (5) is globally asymptotically stable if and only if A is stable and there exists no $x_0 \in \partial\mathbf{S}$ and $N > 0$ such that $\psi(N, x_0) = \pm x_0$ and $\psi(k, x_0) = A^k x_0 \in \mathbf{S}$ for all $k < N$.

If $\psi(k, x_0) = A^k x_0 \in \mathbf{S}$ for all $k < N$, then $\psi(N, x_0) = \operatorname{sat}(A^N x_0)$. Hence, this theorem can be interpreted as follows. Assume that A is stable, then the system (5) is GAS if and only if none of the following statements are true.

1) There exist $N \geq 1$ and $d_1, d_2 \geq 0$ such that

$$A^N \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \pm \begin{bmatrix} 1+d_1 \\ 1+d_2 \end{bmatrix} \quad \text{and} \quad A^k \begin{bmatrix} 1 \\ 1 \end{bmatrix} \in \mathbf{S} \qquad \forall k < N.$$

2) There exist $N \geq 1$ and $d_1, d_2 \geq 0$ such that

$$A^N \begin{bmatrix} -1 \\ 1 \end{bmatrix} = \pm \begin{bmatrix} -1-d_1 \\ 1+d_2 \end{bmatrix}$$

and

$$A^k \begin{bmatrix} -1 \\ 1 \end{bmatrix} \in \mathbf{S} \qquad \forall k < N.$$

3) There exist $N \geq 1$, $d_2 > 0$ and $x_1 \in (-1, 1)$ such that

$$A^N \begin{bmatrix} x_1 \\ 1 \end{bmatrix} = \pm \begin{bmatrix} x_1 \\ 1+d_2 \end{bmatrix}$$

and

$$A^k \begin{bmatrix} x_1 \\ 1 \end{bmatrix} \in \mathbf{S} \qquad \forall k < N.$$

4) There exist $N \geq 1$, $d_1 > 0$ and $x_2 \in (-1, 1)$ such that

$$A^N \begin{bmatrix} 1 \\ x_2 \end{bmatrix} = \pm \begin{bmatrix} 1+d_1 \\ x_2 \end{bmatrix}$$

and

$$A^k \begin{bmatrix} 1 \\ x_2 \end{bmatrix} \in \mathbf{S} \qquad \forall k < N.$$

Each of the above conditions implies that there is a simple periodic trajectory that starts at some x_0 with period N or $2N$. The trajectory stays inside $\mathbf{S}$ as that of the corresponding linear system for the first $N-1$ steps, and when the linear trajectory goes out of $\mathbf{S}$ at step N, the saturation function makes $\psi(N, x_0) = \operatorname{sat}(A^N x_0)$ return exactly at x_0 or $-x_0$. These conditions can be verified. Since A is stable, there exists an integer N_0 such that $A^k x_0 \in \mathbf{S}$ for all $k > N_0$ and all $x_0 \in \partial\mathbf{S}$. Hence, it suffices to check the four conditions only for $N < N_0$.

Conditions 1) and 2) are very easy to check. As to 3) or 4), for each N, at most two $x_1's$ can be solved from

$$A^N \begin{bmatrix} x_1 \\ 1 \end{bmatrix} = \pm \begin{bmatrix} x_1 \\ 1+d_2 \end{bmatrix}.$$

To see this, denote the elements of A^N as $(A^N)_{ij}$, $i, j = 1, 2$. Then from 3), we have

$$(A^N)_{11} x_1 + (A^N)_{12} = \pm x_1. \tag{6}$$

If $(A^N)_{11} \neq \pm 1$, then there are two x_1's that satisfy (6). If $(A^N)_{11} = \pm 1$, we must have $(A^N)_{12} \neq 0$. Otherwise A^N would have an eigenvalue ± 1, which is impossible since A is stable. In this case, (6) has only one solution. It remains to check if $d_2 > 0$ and

$$A^k \begin{bmatrix} x_1 \\ 1 \end{bmatrix} \in \mathbf{S} \qquad \forall k < N.$$

In the process of proving Theorem 2.1, we will develop a more efficient method to check the conditions.

Example 2.1: Consider (5) with

$$A = \begin{bmatrix} 1.5840 & -1.3990 \\ 3.9702 & -2.9038 \end{bmatrix}.$$

The following results are presented with accuracy up to four decimal digits. There are two points on $\partial\mathbf{S}$ that satisfy condition 3), one with

$$x_1 = \frac{(A^3)_{12}}{1 - (A^3)_{11}} = 0.7308$$

and the other with

$$x_1 = \frac{(A^2)_{12}}{-1 - (A^2)_{11}} = 0.9208.$$

But there are four periodic trajectories as listed

$$1)\quad \begin{bmatrix}0.7308\\1.0000\end{bmatrix}\quad \begin{bmatrix}-0.2414\\-0.0023\end{bmatrix}\quad \begin{bmatrix}-0.3791\\-0.9516\end{bmatrix}\quad \begin{bmatrix}0.7308\\1.0000\end{bmatrix}$$

$$2)\quad \begin{bmatrix}0.9028\\1.0000\end{bmatrix}\quad \begin{bmatrix}0.0310\\0.6804\end{bmatrix}\quad \begin{bmatrix}-0.9028\\-1.0000\end{bmatrix}\quad \begin{bmatrix}-0.0310\\-0.6804\end{bmatrix}\quad \begin{bmatrix}0.9028\\1.0000\end{bmatrix}$$

$$3)\quad \begin{bmatrix}0.7424\\1.0000\end{bmatrix}\quad \begin{bmatrix}-0.2230\\0.0438\end{bmatrix}\quad \begin{bmatrix}-0.4145\\-1.0000\end{bmatrix}\quad \begin{bmatrix}0.7424\\1.0000\end{bmatrix}$$

$$4)\quad \begin{bmatrix}1.0000\\1.0000\end{bmatrix}\quad \begin{bmatrix}0.1850\\1.0000\end{bmatrix}\quad \begin{bmatrix}-1.0000\\-1.0000\end{bmatrix}\quad \begin{bmatrix}-0.1850\\-1.0000\end{bmatrix}\quad \begin{bmatrix}1.0000\\1.0000\end{bmatrix}.$$

In the third periodic trajectory, the first coordinate $x_1 = 0.7424$ of the initial state is computed from

$$x_1 = \frac{a_{12} - a_{11}(A^2)_{12}}{a_{11}(A^2)_{11} - 1} = 0.7424.$$

It should be noted that 4) is the only stable periodic trajectory.

As we can see from the example, there are other kinds of periodic trajectories than what are inferred by the conditions 1)–4), e.g., trajectories 3) and 4). There may also be trajectories that neither are periodic nor converge to the origin. We will prove in the subsequent sections that if none of the conditions 1)–4) is true, then there exist no nonconvergent trajectory of any kind.

III. Limit Trajectories

To prove that (5) is GAS, we need to show that the only limit point of any trajectory is the origin. It is known that A being stable alone is not sufficient to guarantee the GAS of the system. Actually, it is well-known [9] that the system may have stationary points other than the origin; there may be periodic trajectories and even trajectories that neither are periodic, nor converge to a stationary point. In this section, we are going to characterize some general properties of the nonconvergent trajectories. These properties will facilitate us to exclude the existence of such nonconvergent trajectories under the condition of Theorem 2.1.

Since every trajectory is bounded by the unit square, there exists a set of points such that the trajectory will go arbitrarily close to them infinitely many times.

Definition 3.1: For a given $x_0 \in \mathbf{R}^2$, a point $x^* \in \mathbf{R}^2$ is called a (positive) limit point of the trajectory $\psi(k, x_0)$ if there exists a subsequence of $\psi(k, x_0)$, $\psi(k_i, x_0)$, $i = 1, 2, \ldots$, such that $\lim_{i\to\infty} \psi(k_i, x_0) = x^*$. The set of all such limit points is called the limit set of the trajectory. We denote this limit set as $\Gamma(x_0)$.

Since the function sat(Ax) is continuous in x, if a trajectory $\psi(k, x_0)$ returns arbitrarily close to $x \in \Gamma(x_0)$, it will also return arbitrarily close to sat(Ax). We state this property in the following lemma.

Lemma 3.1: If $y_0 \in \Gamma(x_0)$, then $\psi(k, y_0) \in \Gamma(x_0)$ for all $k \geq 0$. Given any $\varepsilon > 0$ (arbitrarily small) and any integer $N > 0$ (arbitrarily large), there exists an integer $K_0 > 0$ such that

$$|\psi(k + K_0, x_0) - \psi(k, y_0)|_\infty < \varepsilon \qquad \forall\, k \leq N.$$

Because of Lemma 3.1, for $y_0 \in \Gamma(x_0)$, $\psi(k, y_0)$ is called a limit trajectory of $\psi(k, x_0)$. It is periodic if and only if $\Gamma(x_0)$ has finite number of elements.

The following notation is defined for simplicity. Denote

$$L_h = \left\{ \begin{bmatrix}x_1\\1\end{bmatrix} : x_1 \in (-1, 1)\right\}$$

$$L_v = \left\{ \begin{bmatrix}1\\x_2\end{bmatrix} : x_2 \in (-1, 1)\right\}.$$

We see that L_h and $-L_h$ are the two horizontal sides of $\mathbf{S}$, and L_v and $-L_v$ are the two vertical sides of $\mathbf{S}$. Notice that they do not include the four vertices of the unit square. Also, denote $v_1 = \left[\begin{smallmatrix}1\\1\end{smallmatrix}\right]$, $v_2 = \left[\begin{smallmatrix}-1\\1\end{smallmatrix}\right]$ as the two upper vertices of the square.

Let y_0 be a limit point of some trajectory and for simplicity, let $y_k = \psi(k, y_0)$. Denote $Y = \{\pm y_k \colon k \geq 0\}$ and $AY = \{\pm A y_k \colon k \geq 0\}$. Clearly, Y must have an intersection with the boundary of the unit square. If $Y \cap L_h$ is not empty, define

$$\gamma_1 = \inf\left\{x_1 \colon \begin{bmatrix}x_1\\1\end{bmatrix} \in Y \cap (L_h \cup \{v_1, v_2\})\right\}$$

and

$$\gamma_2 = \sup\left\{x_1 \colon \begin{bmatrix}x_1\\1\end{bmatrix} \in Y \cap (L_h \cup \{v_1, v_2\})\right\}.$$

If $Y \cap L_v$ is not empty, define

$$\gamma_3 = \sup\left\{x_2 \colon \begin{bmatrix}1\\x_2\end{bmatrix} \in Y \cap (L_v \cup \{v_1, -v_2\})\right\}$$

and

$$\gamma_4 = \inf\left\{x_2 \colon \begin{bmatrix}1\\x_2\end{bmatrix} \in Y \cap (L_v \cup \{v_1, -v_2\})\right\}.$$

The following proposition shows that a limit trajectory can only intersect one opposite pair of the sides of the unit square, not both of the neighboring pair. This result will reduce our problem to a much simpler one.

Proposition 3.1: Let y_0 be a limit point of some trajectory.

1) If $y_0 \in L_h$, then $\psi(k, y_0)$ will not touch L_v or $-L_v$ for all $k \geq 0$. Moreover, $\psi(k, y_0)$ will stay inside the strip

$$\left\{ \begin{bmatrix}x_1\\x_2\end{bmatrix} : |x_1| \leq \max\{|\gamma_1|, |\gamma_2|\}\right\}.$$

2) If $y_0 \in L_v$, then $\psi(k, y_0)$ will not touch L_h or $-L_h$ and will stay inside the strip

$$\left\{ \begin{bmatrix}x_1\\x_2\end{bmatrix} : |x_2| \leq \max\{|\gamma_3|, |\gamma_4|\}\right\}.$$

3) The set Y cannot include both v_1 and v_2.

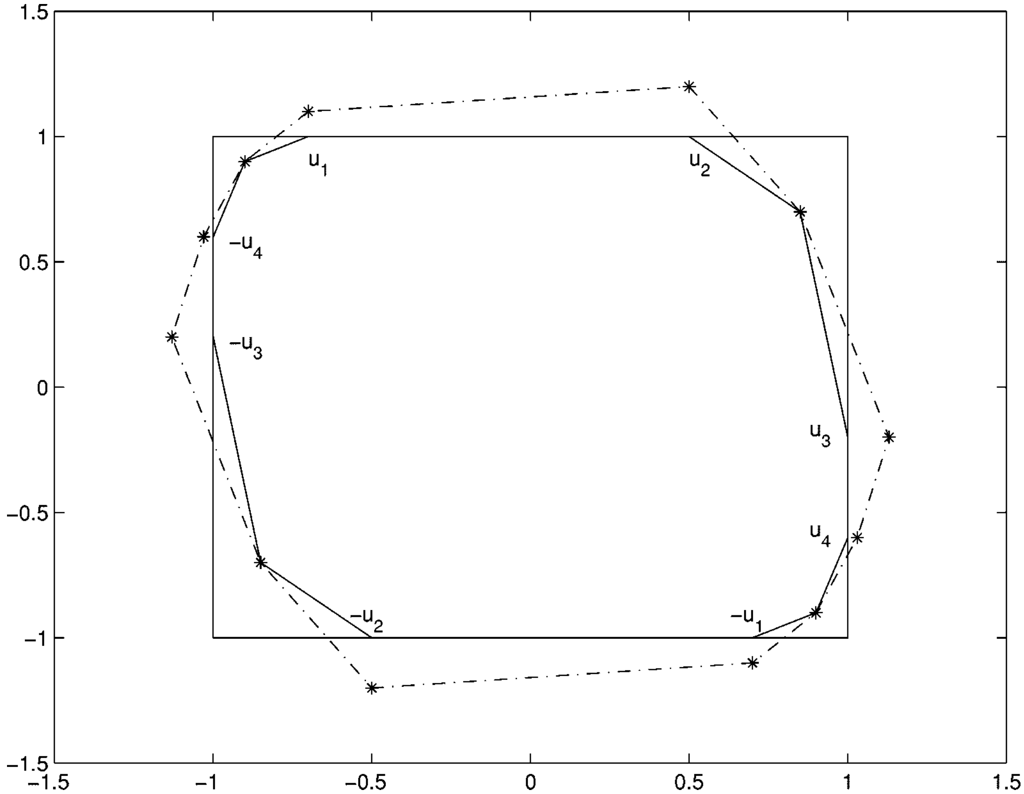

Fig. 1. Illustration for the proof of Proposition 3.1.

Proof: The proof is built up on a simple geometric fact. Let X be a set in $\mathbf{R}^2$ and let AX be the image of X under the linear map $x \rightarrow Ax$. Then, the area of AX equals to the area of X times $|\det(A)|$.

1)We first assume that Y contains a finite number of elements, i.e., $\psi(N, y_0) = y_0$ for some N. Suppose on the contrary that the trajectory will touch L_v or $-L_v$ at some step. The main idea of the proof is to show that the area of the convex hull of AY is no less than that of Y, which contradicts the fact that $|\det(A)| < 1$.

Since Y contains points on both L_h and L_v, γ_i, $i = 1, 2, 3, 4$, are all defined.

If y_k is in the interior of the unit square, then $y_k = Ay_{k-1}$; if $y_k \in L_h$, then $y_k = \mathrm{sat}(Ay_{k-1})$ and

$$Ay_{k-1} = \begin{bmatrix} x_1 \\ 1+d \end{bmatrix}$$

for some $|x_1| < 1$ and $d \geq 0$ [note that $y_0 = y_N = \mathrm{sat}(Ay_{N-1})$]; if $y_k \in L_v$, then

$$Ay_{k-1} = \begin{bmatrix} 1+d \\ x_2 \end{bmatrix}$$

for some $|x_2| < 1$ and $d \geq 0$. If $y_k = v_1$ (or v_2), then $y_k = \mathrm{sat}(Ay_{k-1})$ and

$$Ay_{k-1} = \begin{bmatrix} 1+d_1 \\ 1+d_2 \end{bmatrix} \quad \left(\text{or } \begin{bmatrix} -1-d_1 \\ 1+d_2 \end{bmatrix}\right)$$

for some $d_1, d_2 \geq 0$. Hence, AY contains all the elements of Y which are in the interior of $\mathbf{S}$, and for those y_k on the boundary of $\mathbf{S}$, if $y_k \in L_h$, there is a point in AY that is just above y_k (on the same vertical line) and if $y_k \in L_v$, then there is a point in AY that is just to the right of y_k (on the same horizontal line).

Denote the areas of the convex hulls of Y and AY as $\mathcal{A}(Y)$ and $\mathcal{A}(AY)$, respectively. Also, let

$$u_1 = \begin{bmatrix} \gamma_1 \\ 1 \end{bmatrix} \quad u_2 = \begin{bmatrix} \gamma_2 \\ 1 \end{bmatrix}$$
$$u_3 = \begin{bmatrix} 1 \\ \gamma_3 \end{bmatrix} \quad u_4 = \begin{bmatrix} 1 \\ \gamma_4 \end{bmatrix}$$

as shown in Fig. 1. In the figure, the points marked with "$*$" belong to AY, the polygon with dash–dotted boundary is the convex hull of AY and the polygon with vertices $\pm u_i$, $i = 1, 2, 3, 4$, and some points in the interior of $\mathbf{S}$ is the convex hull of Y. Since there is at least one point in Y that is to the left of u_1, one to the right of u_2, one above u_3 and one below u_4, the convex hull of Y is a subset of the convex hull of AY. (This may not be true if u_1 is the leftmost point in Y, or if u_2 is the rightmost). It follows that $\mathcal{A}(AY) \geq \mathcal{A}(Y)$. This is a contradiction since $\mathcal{A}(AY) = |\det(A)|\mathcal{A}(Y)$ and $|\det(A)| < 1$.

If, on the contrary, Y has a point outside of the strip

$$\left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : |x_1| < \max\{|\gamma_1|, |\gamma_2|\} \right\}$$

then, there will be a point in Y that is to the left of u_1 (or on the same vertical line with u_1), and a point to the right of u_2 (or on the same horizontal line with u_2). In this case, we also have $\mathcal{A}(AY) \geq \mathcal{A}(Y)$, which is a contradiction.

Now we extend the result to the case that Y has infinite many elements. Also suppose on the contrary that the trajectory will touch L_v, $-L_v$ or go outside of the strip at some step. By

Lemma 3.1, for any $\varepsilon > 0$ and any integer $N \geq 1$, there exists a $K_0 > 0$ such that

$$|\psi(k+K_0, x_0) - \psi(k, y_0)|_\infty < \varepsilon$$

for all $k \leq N$ and in particular

$$|\psi(K_0, x_0) - y_0|_\infty < \varepsilon.$$

So, the trajectory $\psi(k+K_0, x_0)$, $k \geq 0$, will also touch (or almost touch) L_v, $-L_v$, or go outside of the strip. Since y_0 is a limit point of $\psi(k+K_0, x_0)$, there exists a $K_1 > 0$ such that

$$|\psi(K_1+K_0, x_0) - y_0|_\infty < \varepsilon.$$

Define

$$Z(\varepsilon) = \{\psi(k+K_0, x_0) \colon 0 \leq k \leq K_1\}$$

and

$$AZ(\varepsilon) = \{A\psi(k+K_0, x_0) \colon 0 \leq k \leq K_1\}.$$

Using similar arguments as in the finite element case, we can show that

$$|\det(A)| = \frac{\mathcal{A}(AZ(\varepsilon))}{\mathcal{A}(Z(\varepsilon))} \geq 1 - O(\varepsilon).$$

Letting $\varepsilon \to 0$, we obtain $|\det(A)| \geq 1$, which is a contradiction.

2) Similar to 1).

3) If, on the contrary, Y contains both v_1 and v_2, then the convex hull of Y is $\mathbf{S}$. Also, AY contains a point

$$Ay_j = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad x_1 \leq -1 \quad x_2 \geq 1$$

and a point

$$Ay_k = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \qquad x_1 \geq 1 \quad x_2 \geq 1$$

hence, the convex hull of AY contains $\mathbf{S}$. This also leads to $\mathcal{A}(AY) \geq \mathcal{A}(Y)$, a contradiction. $\square$

IV. Systems with One Saturated State

Now, we are clear from Proposition 3.1 that if there is any limit trajectory, it can intersect only one opposite pair of the sides of the unit square, either $(L_h, -L_h)$, or $(L_v, -L_v)$, not both of them. So, we only need to investigate the possibility that a limit trajectory only intersects $\pm L_h$. The other possibility that it only intersects $\pm L_v$ is similar. For this reason, we consider the following system:

$$x(k+1) = \begin{bmatrix} a_{11}x_1(k) + a_{12}x_2(k) \\ \mathrm{sat}\,(a_{21}x_1(k) + a_{22}x_2(k)) \end{bmatrix} := \mathrm{sat}_2\,(Ax(k)). \tag{7}$$

Assume that $A = \left[\begin{smallmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{smallmatrix}\right]$ is stable. If $a_{21} = 0$ or $a_{12} = 0$, it is easy to see that both systems (5) and (7) are GAS and none of the conditions 1)–4) following Theorem 2.1 can be true. So we assume in the following that $a_{21}, a_{12} \neq 0$.

The terms *GAS*, *limit point* and *limit trajectory* for (5) are extended to (7) in a natural way.

For a given initial state $x(0) = x_0$, denote the trajectory of the system (7) as $\psi_2(k, x_0)$. Denote the line $x_2 = 1$ as L_h^e, the line $x_2 = -1$ as $-L_h^e$ and the region between these two lines (including $\pm L_h^e$) as $\mathbf{S}^e$. We will show later that (7) has nontrivial limit trajectory in $\mathbf{S}$ if and only if (5) has nontrivial limit trajectory that intersects $\pm L_h$. In the sequel, when we say "limit trajectory," we mean a limit trajectory other than the trivial one at the origin.

In this section, we study the GAS of the system (7) and will also determine a subset in L_h^e which is free of limit points. Our investigation will be based on the study of the linear system

$$x(k+1) = Ax(k). \tag{8}$$

For a stable continuous-time linear planar system, if a trajectory stays in $\mathbf{S}^e$ for a whole cycle [$\angle x(t)$ increases or decreases by 2π], then $x(t)$ will be in $\mathbf{S}^e$ for all $t > 0$. But, for the discrete-time linear planar system (8), a trajectory might go out of $\mathbf{S}^e$ after staying within $\mathbf{S}^e$ for several cycles. In the continuous-time case, the trajectories never intersect but in the discrete-time case, the connected trajectory [by connecting $x(k)$ and $x(k+1)$] can intersect itself. These facts make the analysis much more complicated than the continuous-time system as discussed in [1], [3], [4] and [10].

A simple one or two point periodic trajectory can be formed if $A\left[\begin{smallmatrix} x_1 \\ 1 \end{smallmatrix}\right] = \pm\left[\begin{smallmatrix} x_1 \\ 1+d \end{smallmatrix}\right]$ for some $d > 0$. An N or $2N$ point periodic trajectory will be formed if $A^N\left[\begin{smallmatrix} x_1 \\ 1 \end{smallmatrix}\right] = \pm\left[\begin{smallmatrix} x_1 \\ 1+d \end{smallmatrix}\right]$, $d > 0$ and $A^k\left[\begin{smallmatrix} x_1 \\ 1 \end{smallmatrix}\right] \in \mathbf{S}^e$, for all $k < N$.

Proposition 4.1: The system (7) is GAS if and only if A is stable and the following statement is not true for any $x_1 \in \mathbf{R}$: There exist an integer $N > 0$ and a real number $d > 0$ such that

$$A^N \begin{bmatrix} x_1 \\ 1 \end{bmatrix} = \pm \begin{bmatrix} x_1 \\ 1+d \end{bmatrix}$$

and

$$A^k \begin{bmatrix} x_1 \\ 1 \end{bmatrix} \in \mathbf{S}^e \qquad \forall k < N. \tag{9}$$

Let $\alpha_s = \min\{|x_1| \colon x_1 \text{ satisfies (9)}\}$, then no limit trajectory can exist completely within the strip

$$\left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : |x_1| < \alpha_s \right\}.$$

.

Remark: If (9) is true for some x_1, then there will be a stationary point or periodic trajectory such as $x_0, Ax_0, \ldots, A^{N-1}x_0$, $\mathrm{sat}_2(A^N x_0) = x_0, Ax_0, \ldots$. There may also exist other kind of limit trajectories. Proposition 4.1 says that if there is no simple periodic trajectory as inferred by (9), there will be no limit trajectory of any kind (except the one at the origin).

To prove Proposition 4.1, we need to establish the relation between the next intersection of a trajectory with $\pm L_h^e$ and the present one.

For $x_0 \in L_h^e$, suppose that $\psi_2(k, x_0)$ will intersect $\pm L_h^e$ at $k = k_i$, $i = 1, 2, \ldots$, with $0 < k_1 < k_2 < \ldots$. Since the trajectory can be switched to $-\psi_2(k, x_0)$ at any k without changing its convergence property, we assume for simplicity that all the intersections $\psi_2(k_i, x_0)$ are in L_h^e (If not so, just multiply it with -1). Denote

$$x_0^1 = \psi_2(k_1, x_0) \quad x_0^2 = \psi_2(k_2, x_0) \quad \ldots.$$

We call x_0, x_0^1 and x_0^2 the first, the second and the third intersections, respectively. We also call x_0 and x_0^1 the present and the next intersections.

Clearly, x_0^1 is uniquely determined by x_0. We also see that the relation $x_0 \to x_0^1$ is a map from L_h^e to itself. To study the GAS of the system (7), it suffices to characterize the relation between x_0 and x_0^1. Through this relation, we can show that if (9) is not true for any x_1, then for every $x_0 \in L_h^e$, the intersections x_0^1, $x_0^2, \ldots$ will move closer and closer toward an interval, and all the trajectories starting from this interval will not touch the lines $\pm L_h^e$ and will converge to the origin.

Let $x_0 \in L_h^e$. The next intersection of $\psi_2(k, x_0)$ with L_h^e occurs at step k_1 if

$$\left|[0 \quad 1]A^{k_1}x_0\right| \geq 1$$

and

$$\left|[0 \quad 1]A^{k}x_0\right| < 1 \qquad \forall\, k < k_1.$$

The next intersection is $x_0^1 = \psi_2(k_1, x_0) = \text{sat}_2(A^{k_1}x_0)$ [or $-\text{sat}_2(A^{k_1}x_0)$]. Since for different $x_0 \in L_h^e$, the number of steps for the trajectories to return to $\pm L_h^e$, i.e., the number k_1 as defined above, is different, we see that the relation between x_0 and x_0^1 must be discontinuous.

We will first determine an interval on L_h^e from which a trajectory will not intersect $\pm L_h^e$ again (no x_0^1) and will converge to the origin.

Since A is stable, there exists a positive definite matrix P such that

$$A^TPA - P < 0.$$

Define the Lyapunov function as

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

then for every $x \in \mathbf{R}^2$, $V(A^kx) < V(x)$ for all $k > 1$.

Given a real number $\rho > 0$, denote the Lyapunov level set as

$$\mathcal{E}(\rho) := \{x \in \mathbf{R}^2 \colon x^TPx \leq \rho\}.$$

Let ρ_c be such that $\mathcal{E}(\rho_c) \subset \mathbf{S}^e$ and $\mathcal{E}(\rho_c)$ just touches $\pm L_h^e$. In this case, $\mathcal{E}(\rho_c)$ has only one intersection with L_h^e. Let this intersection be

$$p_c = \begin{bmatrix} \alpha_c \\ 1 \end{bmatrix}.$$

If $x_0 = p_c$, then the linear trajectory A^kx_0 will be inside $\mathcal{E}(\rho_c) \subset \mathbf{S}^e$. Hence, $\psi_2(k, x_0) = \text{sat}_2(A^kx_0) = A^kx_0$ for all $k > 0$ and will converge to the origin. Since

$$A^TPA - P < 0$$

there exists an interval around p_c in L_h^e, of nonzero length, such that for every x_0 in this interval, $\psi_2(k, x_0) = A^kx_0$, $k \geq 1$, will never touch $\pm L_h^e$ and will converge to the origin.

Here we will use a simple way to denote a line segment. Given two points $p_1, p_2 \in \mathbf{R}^2$, denote

$$[p_1, p_2] := \{\lambda p_1 + (1-\lambda)p_2 \colon 0 \leq \lambda \leq 1\}$$

and similarly

$$\begin{aligned} (p_1, p_2] &= [p_1, p_2] \backslash \{p_1\} \\ [p_1, p_2) &= [p_1, p_2] \backslash \{p_2\} \\ (p_1, p_2) &= [p_1, p_2] \backslash \{p_1, p_2\}. \end{aligned}$$

Define

$$\alpha_0 := \min\left\{\alpha < \alpha_c \colon A^k \begin{bmatrix} \alpha \\ 1 \end{bmatrix} \in \mathbf{S}^e \qquad \forall\, k \geq 0\right\}$$

and

$$\beta_0 := \max\left\{\beta > \alpha_c \colon A^k \begin{bmatrix} \beta \\ 1 \end{bmatrix} \in \mathbf{S}^e \qquad \forall\, k \geq 0\right\}.$$

Since $a_{21} \neq 0$, the line $AL_h^e := \{Ax \colon x \in L_h^e\}$ has intersections with both L_h^e and $-L_h^e$, so there exist points on both sides of p_c which will be mapped out of $\mathbf{S}^e$ under A. Hence α_0 and β_0 are finite numbers. Now, let

$$p_0 = \begin{bmatrix} \alpha_0 \\ 1 \end{bmatrix} \quad q_0 = \begin{bmatrix} \beta_0 \\ 1 \end{bmatrix}$$

then, for all $x_0 \in [p_0, q_0]$, $\psi_2(k, x_0) = A^kx_0$ will converge to the origin. Because of the extremal nature in the definition of α_0 and β_0, we must have $A^kp_0 \in \pm L_h^e$ for some k, otherwise α_0 would not be the minimum of the set. Therefore, define

$$m_0 := \min\{k \colon A^kp_0 \in \pm L_h^e\}$$

and similarly

$$n_0 := \min\{k \colon A^kq_0 \in \pm L_h^e\}.$$

If $m_0 > 1$, then by definition

$$\left|[0 \quad 1]A^kp_0\right| < 1 \qquad \forall\, k < m_0$$

and by continuity, there exists a neighborhood of p_0 such that $|[0\ 1]A^kx_0| < 1$, $\forall\, k < m_0$ for all x_0 in this neighborhood. Because of this, we can define

$$\alpha_1 := \min\left\{\alpha < \alpha_0 \colon A^k \begin{bmatrix} \alpha \\ 1 \end{bmatrix} \in \mathbf{S}^e \qquad \forall\, k < m_0\right\}.$$

If $n_0 > 1$, then define

$$\beta_1 := \max\left\{\beta > \beta_0 \colon A^k \begin{bmatrix} \beta \\ 1 \end{bmatrix} \in \mathbf{S}^e, \qquad \forall\, k < n_0\right\}.$$

Also, because $a_{21} \neq 0$, α_1 and β_1 are finite. Let

$$p_1 = \begin{bmatrix} \alpha_1 \\ 1 \end{bmatrix} \quad q_1 = \begin{bmatrix} \beta_1 \\ 1 \end{bmatrix}.$$

It follows from the extremal nature in the definition of α_1 and β_1 that there exists a $k < m_0$ such that $A^kp_1 \in \pm L_h^e$, so we define

$$m_1 := \min\{k \colon A^kp_1 \in \pm L_h^e\}$$

and similarly

$$n_1 := \min\{k: A^k q_1 \in \pm L_h^e\}.$$

For simplicity, we denote $\left[\begin{smallmatrix}-\infty\\1\end{smallmatrix}\right]$ as p_∞ and $\left[\begin{smallmatrix}\infty\\1\end{smallmatrix}\right]$ as q_∞. Implied by the definitions are the following:

$$p_1 \in (p_\infty, p_0) \quad q_1 \in (q_0, q_\infty),$$

and

$$m_1 < m_0 \quad n_1 < n_0.$$

Inductively, if $m_{i-1} > 1$, then define

$$\alpha_i := \min\left\{\alpha < \alpha_{i-1}: A^k \begin{bmatrix}\alpha\\1\end{bmatrix} \in \mathbf{S}^e \qquad \forall\, k < m_{i-1}\right\}$$

and if $n_{j-1} > 1$, define

$$\beta_j := \max\left\{\beta > \beta_{j-1}: A^k \begin{bmatrix}\beta\\1\end{bmatrix} \in \mathbf{S}^e \qquad \forall\, k < n_{j-1}\right\}.$$

Let

$$p_i = \begin{bmatrix}\alpha_i\\1\end{bmatrix} \quad q_j = \begin{bmatrix}\beta_j\\1\end{bmatrix}$$

and

$$m_i := \min\{k: A^k p_i \in \pm L_h^e\} \quad n_j := \min\{k: A^k q_j \in \pm L_h^e\}.$$

Then

$$\begin{aligned} &p_i \in (p_\infty, p_{i-1}) \quad q_j \in (q_{j-1}, q_\infty),\\ &m_i < m_{i-1} \quad n_j < n_{j-1}. \end{aligned}$$

The induction procedure ends if both $m_i = 1$ and $n_j = 1$. We denote the maximum index of i as I and the maximum index of j as J. As an immediate consequence of these definitions, we have

$$\begin{aligned} \alpha_I < \alpha_{I-1} < \cdots < \alpha_1 < \alpha_0 < \alpha_c < \beta_0 < \beta_1 < \cdots\\ < \beta_{J-1} < \beta_J \end{aligned}$$

and

$$\begin{aligned} 1 = m_I < m_{I-1} < \cdots < m_1 < m_0\\ 1 = n_J < n_{J-1} < \cdots < n_1 < n_0. \end{aligned}$$

We claim that this set of p_i, $i = 0, 1, 2, \ldots, I$, and q_j, $j = 0, 1, \ldots, J$, forms exactly the set of points where discontinuity occurs on the relation between the next intersection of a trajectory with $\pm L_h^e$ and the present one.

Lemma 4.1:

1) If $x_0 \in [p_0, q_0]$, then $\psi_2(k, x_0) = A^k x_0$ will be inside $\mathbf{S}^e$ for all $k > 0$ and will converge to the origin.
2) If $x_0 \in (p_{i+1}, p_i]$, then the next intersection of $\psi_2(k, x_0)$ with $\pm L_h^e$ is $\psi_2(m_i, x_0) = \mathrm{sat}_2(A^{m_i} x_0)$; If $x_0 \in (p_\infty, p_I]$, then the next intersection is $\psi_2(1, x_0) = \mathrm{sat}_2(A x_0)$. Moreover, $A^{m_i} p_i \in \pm L_h^e$ and $A^{m_i} x_0 \notin \mathbf{S}^e$ for all $x_0 \in (p_\infty, p_i)$.
3) If $x_0 \in [q_j, q_{j+1})$, then the next intersection of $\psi_2(k, x_0)$ with $\pm L_h^e$ is $\psi_2(n_j, x_0) = \mathrm{sat}_2(A^{n_j} x_0)$; If $x_0 \in [q_J, q_\infty)$, then the next intersection is $\psi_2(1, x_0) = \mathrm{sat}_2(A x_0)$. Moreover, $A^{n_j} q_j \in \pm L_h^e$ and $A^{n_j} x_0 \notin \mathbf{S}^e$ for all $x_0 \in (q_j, q_\infty)$.
4) $|[0\ 1]A^k p_{i+1}| \le 1$ for all $k < m_i$ and $|[0\ 1]A^{m_i} p_{i+1}| > 1$; $|[0\ 1]A^k q_{j+1}| \le 1$ for all $k < n_j$ and $|[0\ 1]A^{n_j} q_{j+1}| > 1$.

Proof: 1) This is a direct consequence of the definition of p_0 and q_0.

2) From the definition of m_i, $|[0\ 1]A^k p_i| < 1$ for all $k < m_i$ and $|[0\ 1]A^{m_i} p_i| = 1$. Since $A^{m_i} L_h^e$ is a straight line and $A^{m_i} p_c$ is in the interior of $\mathbf{S}^e$, we have $|[0\ 1]A^{m_i} x_0| > 1$ for all $x_0 \in (p_\infty, p_i)$ (Note that p_c is to the right of p_i).

On the other hand, since

$$\alpha_{i+1} = \min\left\{\alpha < \alpha_i: A^k \begin{bmatrix}\alpha\\1\end{bmatrix} \in \mathbf{S}^e \qquad \forall\, k < m_i\right\}$$

we have $|[0\ 1]A^k p_{i+1}| \le 1$ for all $k < m_i$. Also since $A^k p_c$ is in the interior of $\mathbf{S}^e$, we have $|[0\ 1]A^k x_0| < 1$ for all $k < m_i$ and for all $x_0 \in (p_{i+1}, p_c)$.

Combining the above arguments we have, for all $x_0 \in (p_{i+1}, p_i]$, $|[0\ 1]A^{m_i} x_0| \ge 1$ and $|[0\ 1]A^k x_0| < 1$ for all $k < m_i$. This means that the next intersection with $\pm L_h^e$ is $\psi_2(m_i, x_0) = \mathrm{sat}_2(A^{m_i} x_0)$.

3) Similar to 2).

4) This is contained in the proof of 2). □

It is obvious that $\mathrm{sat}_2(A^{m_i} x)$ is a continuous function of x. Lemma 4.1 2) implies that for all $x_0 \in (p_{i+1}, p_i]$, the second coordinate of $\mathrm{sat}_2(A^{m_i} x_0)$, $[0\ 1]\mathrm{sat}_2(A^{m_i} x_0)$, is the constant 1 or -1, while the first coordinate remains linear on x_0. Similarly, for all $x_0 \in [q_j, q_{j+1})$, the second coordinate of $\mathrm{sat}_2(A^{n_j} x_0)$ is the constant 1 or -1 and the first coordinate is linear on x_0. Same relation holds for $x_0 \in (p_\infty, p_I]$ and $x_0 \in [q_I, q_\infty)$.

We will provide an easy way to compute p_i and q_j after revealing more properties about this set of points. In fact, the following properties will lead directly to the proof of Proposition 4.1. For $x_0 \in L_h^e$, the next intersection of $\psi_2(k, x_0)$ with $\pm L_h^e$ can be on L_h^e or on $-L_h^e$. For simplicity, we will assume that the next intersection is on L_h^e, otherwise we can replace the state $x(k)$ at the intersection with $-x(k)$, noting that we can multiply the state at any step with -1 without changing the convergence rate of a trajectory. Hence in the following, when we say that $x \in [p_i, q_j]$, we mean $x \in \pm[p_i, q_j]$; and when we say that $x \in \pm L_h^e$ is to the left (or right) of p_i, it could also be that x is to the right (or left) of $-p_i$.

Denote $p_i^1 = \psi_2(m_i, p_i) = A^{m_i} p_i$ and $q_j^1 = \psi_2(n_j, q_j) = A^{n_j} q_j$. We see that p_i^1 is the second intersection of $\psi_2(k, p_i)$ with $\pm L_h^e$ (the first one is p_i), and q_j^1 is the second intersection of $\psi_2(k, q_j)$ with $\pm L_h^e$.

Lemma 4.2:

1) $p_0^1, q_0^1 \in [p_0, q_0]$;
2) If $p_i^1 \in (q_{j-1}, q_j]$, then $m_{i-1} = m_i + n_{j-1}$; if $q_j^1 \in [p_i, p_{i-1})$, then $n_{j-1} = n_j + m_{i-1}$;
3) For $i \ge 1$, $p_i^1 \in (q_0, q_\infty)$ and for $j \ge 1$, $q_j^1 \in (p_\infty, p_0)$;
4) $p_i^1 \in (p_{i-1}^1, q_\infty)$, and $q_j^1 \in (p_\infty, q_{j-1}^1)$;
5) For $i, j \ge 1$, p_i^1 and q_j^1 cannot be both in $[p_i, q_j]$, nor both outside of $[p_i, q_j]$, i.e., there must be one of them inside $[p_i, q_j]$ and the other one outside of the interval.

Proof: First, we give a simple property arising from the Lyapunov function $V(x)$. Since $V(x)$ is a convex function and p_c takes the minimum value from all $x \in L_h^e$, we have, if both

s_1 and s_2 are to the left of p_c and s_1 is to the left of s_2, then $V(s_1) > V(s_2)$; if both s_1 and s_2 are to the right of p_c and s_1 is to the right of s_2, then $V(s_1) > V(s_2)$.

1)Clearly, p_0^1 cannot be to the left of p_0, otherwise we would have $V(p_0^1) = V(A^{n_0} p_0) > V(p_0)$. Suppose on the contrary that $p_0^1 \in (q_0, q_\infty)$, then by Lemma 4.1 3), we would have $|[0\ 1]A^{n_0} A^{m_0} p_0| > 1$. A contradiction to the definition of p_0. Similar argument holds for q_0.

2) From Lemma 4.1 4), the first time $A^k p_i$ goes out of $\mathbf{S}^e$ is at $k = m_{i-1}$. And by Lemma 4.1 3) and 4), for $x_0 \in (q_{j-1}, q_j]$, the first time $A^k x_0$ goes out of $\mathbf{S}^e$ is $k = n_{j-1}$. Since $p_i^1 = A^{m_i} p_i \in (q_{j-1}, q_j]$ and $A^k p_i \in \mathbf{S}^e$ for all $k < m_i$, we have $m_{i-1} = m_i + n_{j-1}$. Similarly, for q_j, we have $n_{j-1} = n_j + m_{i-1}$.

3) Similar to 1), p_i^1 cannot be to the left of p_i. Suppose on the contrary, that $p_i^1 \in [p_0, q_0]$, then $A^k p_i$ never goes out of $\mathbf{S}^e$. This is a contradiction since $A^{n_0} p_i \notin \mathbf{S}^e$. Also, suppose on the contrary that $p_i^1 \in [p_{i-l}, p_{i-l-1})$, $l \geq 0$, then similar to the argument in 2), we would have $m_{i-1} = m_i + m_{i-l-1}$ and hence $m_{i-1} > m_{i-l-1}$. This is a contradiction since m_i decreases as i is increased. So, we must have $p_i^1 \in (q_0, q_\infty)$ and similarly, $q_j^1 \in (p_\infty, p_0)$.

4) Since p_i is to the left of p_{i-1}, we have

$$V(p_i) > V(p_{i-1}) > V(p_c).$$

By 3), p_i^1 is to the right of p_c, this implies

$$V(p_c) < V(p_i^1) = V(A^{m_i} p_i)$$

and hence

$$V(A^{m_i} p_c) < V(p_c) < V(A^{m_i} p_i). \tag{10}$$

Since $p_{i-1} \in (p_i, p_c)$, the point $A^{m_i} p_{i-1}$ is on the line between $A^{m_i} p_c$ and $A^{m_i} p_i$. Also, since the function $V(x)$ is convex, it follows from (10) that

$$V(A^{m_i} p_{i-1}) < \max\{V(A^{m_i} p_c), V(A^{m_i} p_i)\} = V(A^{m_i} p_i).$$

Since $m_{i-1} > m_i$, we have $V(A^{m_{i-1}} p_{i-1}) < V(A^{m_i} p_{i-1})$ and

$$\begin{aligned} V(p_{i-1}^1) &= V(A^{m_{i-1}} p_{i-1}) < V(A^{m_i} p_{i-1}) \\ &< V(A^{m_i} p_i) = V(p_i^1). \end{aligned} \tag{11}$$

By 3), p_{i-1}^1 and p_i^1 are both to the right of p_c, hence the inequality (11) implies that p_i^1 is to the right of p_{i-1}^1, i.e., $p_i^1 \in (p_{i-1}^1, q_\infty)$.

Similarly, we have $q_j^1 \in (p_\infty, q_{j-1}^1)$.

5) Suppose that both $q_j^1, p_i^1 \in [p_i, q_j]$, then by 2) and 3), we have

$$n_{j-1} = n_j + m_{i-k1} \quad m_{i-1} = m_i + n_{j-k2}$$

where $k_1, k_2 \geq 1$. Since $m_{i-1} \leq m_{i-k1}$ and $n_{j-1} \leq n_{j-k2}$, it follows that

$$m_{i-1} < n_{j-1} \quad n_{j-1} < m_{i-1}$$

which is a contradiction.

On the other hand, suppose that both $q_j^1, p_i^1 \notin [p_i, q_j]$, then we must have q_j^1 to the left of p_i and p_i^1 to the right of q_j. Hence

$$V(q_j^1) > V(p_i) \quad V(p_i^1) > V(q_j).$$

Recall that $q_j^1 = A^{n_j} q_j$ and $p_i^1 = A^{m_i} p_i$, it follows that

$$V(q_j^1) > V(p_i) > V(p_i^1) > V(q_j) > V(q_j^1)$$

which is also a contradiction. □

From 1), 3), and 5) of Lemma 4.2, we can see that the only pair of p_i and q_j such that $p_i^1, q_j^1 \in [p_i, q_j]$ is p_0 and q_0. This fact can be used to generate the points p_i, $0, 1, \ldots, I$, and q_j, $j = 0, 1, \ldots, J$. Although it is possible to determine these points directly from the definition, it is hard to derive a computationally efficient method to generate the points from inside to outside, i.e., from p_0, q_0 to p_I, q_J. In the following, we provide an iterative method based on the properties in Lemmas 4.1 and 4.2 to generate the points from outside to inside, i.e., from p_I, q_J to p_0, q_0 and use the unique property that $p_0^1, q_0^1 \in [p_0, q_0]$ as a sign to stop the iteration.

`Algorithm for Generating` p_i, q_j, m_i, n_j `and` p_i^1, q_j^1

Step 1 Set $ii = 1$. Get the two intersections of AL_h^e with $\pm L_h^e$. They are p_I^1 and $-q_J^1$ (or $-p_I^1$ and q_J^1). Denote the line segment of AL_h^e between L_h^e and $-L_h^e$ as L_1. Multiply the two end points of L_1 from left with A^{-1}, then we get p_I and q_J. Clearly, the one to the left of p_c is p_I and the one to its right is q_J. If $p_I^1, q_J^1 \in [p_I, q_J]$, then $I = J = 0$ and stop the algorithm.

Step 2 $ii = ii + 1$. Check if AL_{ii-1} has intersections with $\pm L_h^e$. If not, let $L_{ii} = AL_{ii-1}$ and repeat this step. If there is, then cut off the part of AL_{ii-1} that is outside of $\mathbf{S}^e$ and let the remaining part be L_{ii}. The cut-off place is one of $\pm p_i^1$ and $\pm q_j^1$. Multiply the cut-off place from left with A^{-ii}. The result is so far the innermost p_i if it is to the left of p_c, or the innermost q_j if to the right of p_c. In the mean time, we also obtain $m_i = ii$ and\or $n_j = ii$. Let the innermost pair be p_i, q_j, if $p_i^1, q_j^1 \in [p_i, q_j]$, then we must have $i = j = 0$ and stop the algorithm since all the p_i, q_j have been computed. If $p_i^1, q_j^1 \in [p_i, q_j]$ is not true, then repeat this step.

We see from the above algorithm that the number of iterations equals $\max\{m_0, n_0\}$.

Item 4) in Lemma 4.2 shows that $[p_{i-1}, p_{i-1}^1] \subset [p_i, p_i^1]$ and $[q_{j-1}^1, q_{j-1}] \subset [q_j^1, q_j]$. Item 5) shows that either we have $[p_i, p_i^1] \subset [q_j^1, q_j]$, or $[q_j^1, q_j] \subset [p_i, p_i^1]$. Item 3) shows that all these intervals must include $[p_0, q_0]$. In summary, the facts in Lemma 4.2 jointly show that the intervals $[p_i, p_i^1]$ and $[q_j^1, q_j]$, $i = 0, 1, \ldots, I$, $j = 0, 1, \ldots J$ are ordered by inclusion. We can draw a figure for easy understanding of Lemma 4.2. If we draw arcs from p_i to p_i^1 and arcs from q_j to q_j^1, then these arcs

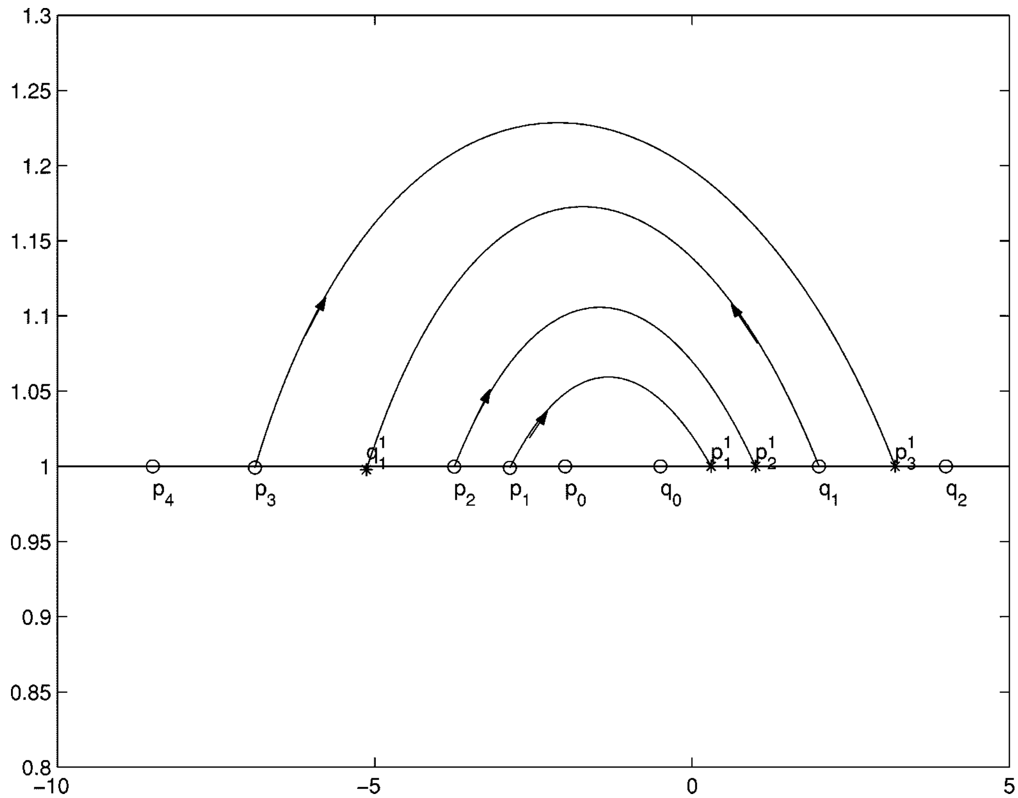

Fig. 2. Illustration for Lemma 4.2.

can be made not to intersect each other (see Fig. 2). In the figure, we have

$$[p_0, q_0] \subset [p_1, p_1^1] \subset [p_2, p_2^1] \subset [q_1^1, q_1] \subset [p_3, p_3^1].$$

Let $p_i^2 = \psi_2(m_{i-1}, p_i) = \mathrm{sat}_2(A^{m_{i-1}}p_i)$, then p_i^2 is the third intersection of $\psi_2(k, p_i)$ with $\pm L_h^e$. By Lemma 4.1 4), we know that m_{i-1} is the smallest integer k such that $A^k p_i$ is outside of $\mathbf{S}^e$. Also let $q_j^2 = \psi_2(n_{j-1}, q_j) = \mathrm{sat}_2(A^{n_{j-1}}q_j)$.

Lemma 4.3: Suppose that (9) is not true for any $x_1 \in \mathbf{R}$, then $p_i^2 \in (p_i, p_i^1)$ and $q_j^2 \in (q_j^1, q_j)$.

Proof: Suppose $p_i^1 \in (q_j, q_{j+1}]$, then by Lemma 4.2 2), the smallest k for $A^k p_i$ to go out of $\mathbf{S}^e$ is $k = m_{i-1} = m_i + n_j$. So $p_i^2 = \mathrm{sat}_2(A^{m_i+n_j}p_i) = \mathrm{sat}_2(A^{n_j}p_i^1)$. If on the contrary that $p_i^2 \in [p_i^1, q_\infty)$ (to the right of p_i^1), since $q_j^1 = A^{n_j}q_j \in \pm L_h^e$ is to the left of q_j, there must be a point $x \in (q_j, p_i^1) \subset (q_j, q_{j+1}]$ such that $A^{n_j}x$ is right above x, i.e., $\mathrm{sat}_2(A^{n_j}x) = x$. By Lemma 4.1 3), the next intersection of $\psi_2(k, x)$ with $\pm L_h^e$ is $\mathrm{sat}_2(A^{n_j}x)$, so we must have $A^k x \in \mathbf{S}^e$ for all $k < n_j$ and there exists x_1 such that (9) is true. A contradiction. On the other hand, if $p_i^2 = \mathrm{sat}_2(A^{m_{i-1}}p_i) \in (p_\infty, p_i]$ (to the left of p_i), since $p_{i-1}^1 = A^{m_{i-1}}p_{i-1}$ is to the right of p_{i-1}, there must be a point $x \in [p_i, p_{i-1})$ such that $A^{m_{i-1}}x$ is right above x. Similar to the former case, we have a contradiction. Therefore, $p_i^2 \in (p_i, p_i^1)$, and similarly, $q_j^2 \in (q_j^1, q_j)$. □

This lemma says that if a trajectory starts from p_i or q_j, its third intersection with $\pm L_h^e$ will be closer to the central interval $[p_0, q_0]$ than the first intersection or the second one. We will show in the next lemma that this property can be actually extended to all $x \in L_h^e$.

Lemma 4.4: Assume that the condition (9) is not true for any $x_1 \in \mathbf{R}$. Given $x(0) = x_0 \in L_h^e$. Let x_0^1 and x_0^2 be the second and the third intersection of the trajectory $\psi_2(k, x_0)$ with L_h^e, (if the intersections are on $-L_h^e$, then get symmetric projections on L_h^e). If $x_0 \in (p_\infty, p_0]$, then $x_0^1 \in (x_0, q_\infty)$ and one of the following must be true.

1) $x_0^1 \in (p_0, q_0)$ and there is no third intersection x_0^2;
2) $x_0^1 \in (x_0, p_0]$;
3) $x_0^1 \in [q_0, q_\infty)$ and $x_0^2 \in (x_0, x_0^1)$.

Similarly, if $x_0 \in [q_0, q_\infty)$, then $x_0^1 \in (p_\infty, x_0)$ and one of the following must be true.

4) $x_0^1 \in (p_0, q_0)$ and there is no third intersection x_0^2;
5) $x_0^1 \in [q_0, x_0)$;
6) $x_0^1 \in (p_\infty, p_0]$ and $x_0^2 \in (x_0^1, x_0)$.

Also, if $x_0 \in (p_i, p_i^1)$ [or $x_0 \in (q_j^1, q_j)$], then x_0^1, x_0^2 and the subsequent intersections will all be in the interval (p_i, p_i^1) [or (q_j^1, q_j)]. Furthermore, for any $x_0 \in L_h^e$, there is a finite k_1 such that $\psi_2(k_1, x_0) \in (p_0, q_0)$. After that, $\psi_2(k, x_0)$ will have no more intersection with L_h^e and will converge to the origin.

Proof: Lemma 4.2 says that all the segments $[p_i, p_i^1]$ and $[q_j^1, q_j]$ are ordered by inclusion. We will prove the result of this lemma from the innermost segment to the outermost with an inductive procedure. Without loss of generality, assume that $[p_1, p_1^1]$ is the innermost segment (except for $[p_0, q_0]$), then we must have $p_1^1 \in (q_0, q_1]$ and $p_1^2 = \mathrm{sat}_2(A^{n_0}p_1^1) = \mathrm{sat}_2(A^{m_0}p_1)$.

By Lemma 4.3, $p_1^2 \in (p_1, p_1^1)$. There are three possibilities.

Case 1—$p_1^2 \in [q_0, p_1^1)$: (See Fig. 3.) For $x_0 \in [q_0, p_1^1)$, $x_0^1 = \mathrm{sat}_2(A^{n_0}x_0)$ by Lemma 4.1 3). Since $q_0^1 = \mathrm{sat}_2(A^{n_0}q_0)$ and $p_1^2 = \mathrm{sat}(A^{n_0}p_1^1)$ are to the left of q_0 and p_1^1 respectively,

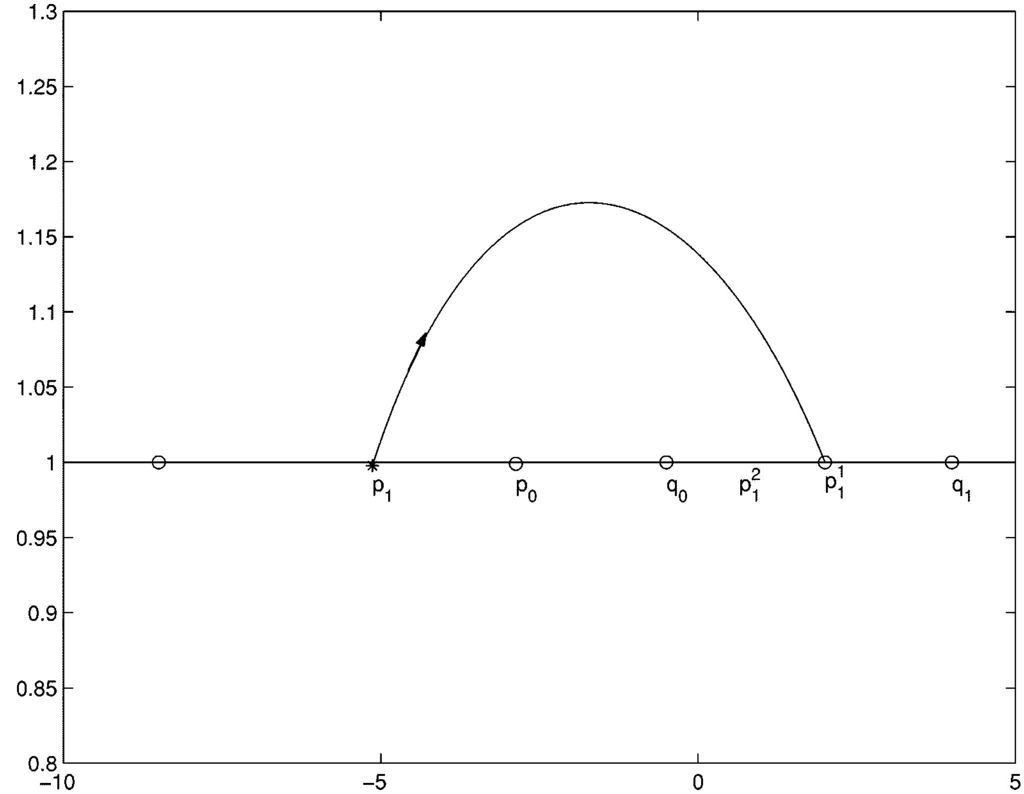

Fig. 3. Illustration for the proof of Lemma 4.4: Case 1.

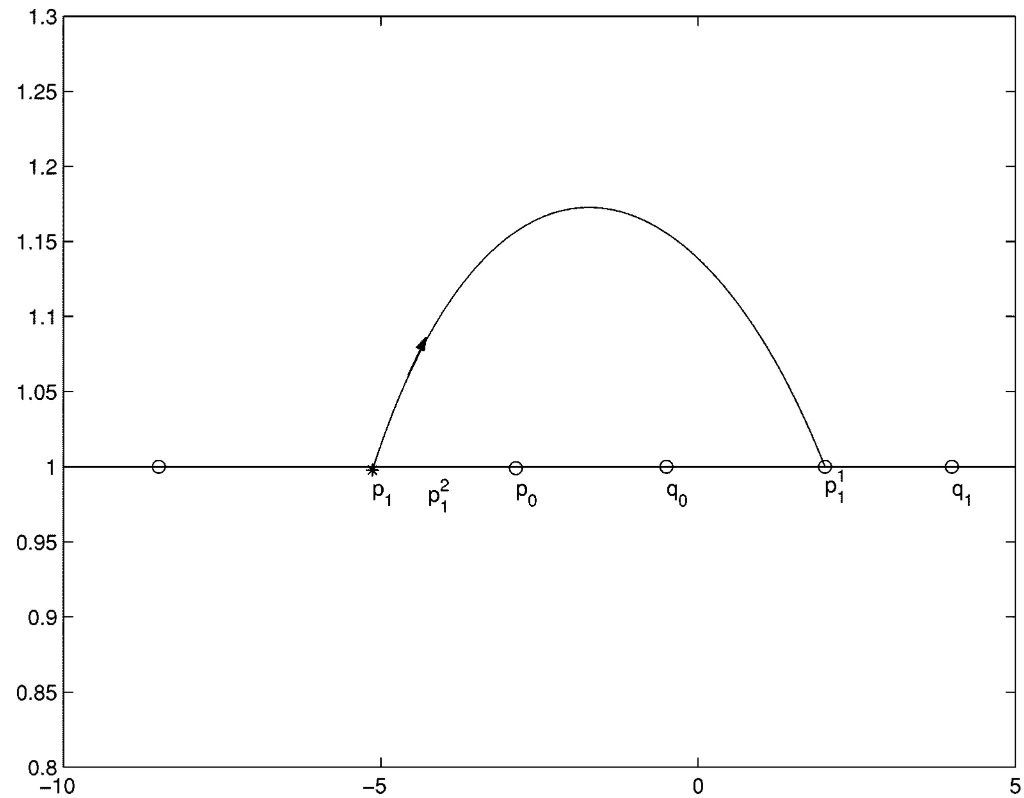

Fig. 4. Illustration for the proof of Lemma 4.4: Case 2.

x_0^1 must be to the left of x_0 and also, $x_0^1 \in [q_0^1, p_1^2)$. So we have $x_1^1 \in (q_0^1, x_0)$. This belongs to 4) or 5) of the lemma. If it is 4), then there will be no more intersection; If it is 5), then with the same argument, we have $x_0^2 \in (q_0^1, x_0^1), \ldots$. Moreover, the subsequent intersections will fall between p_0 and q_0 in a finite number of steps since there is no x_1 satisfying (9).

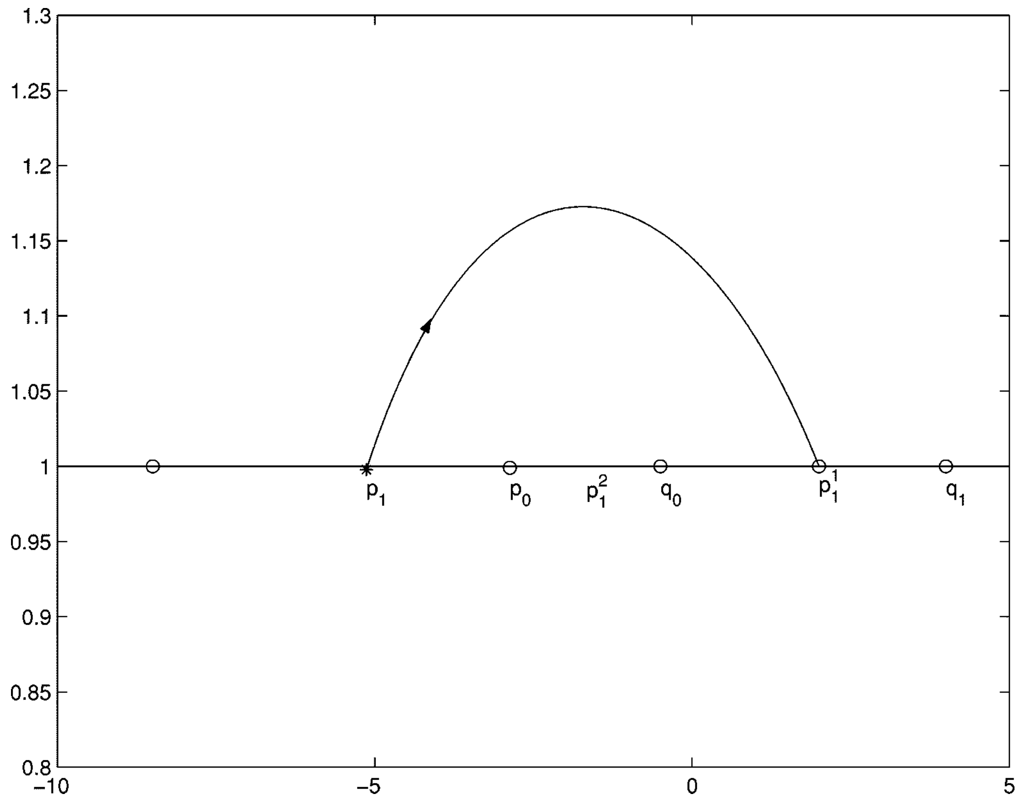

Fig. 5. Illustration for the proof of Lemma 4.4: Case 3.

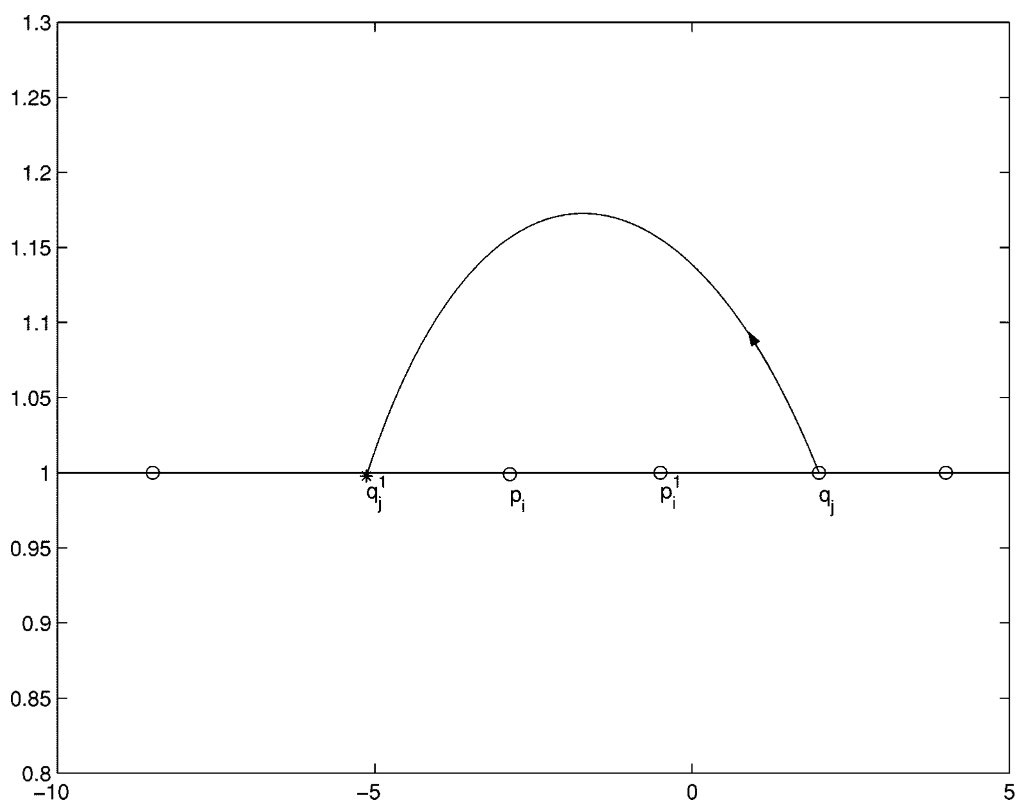

Fig. 6. Illustration for the proof of Lemma 4.4.

For $x_0 \in (p_1, p_0]$, $x_0^1 = \mathrm{sat}_2(A^{m_0}x_0) \in [p_0^1, p_1^2)$. If $x_0^1 \in (p_0, q_0)$, then we get 1) of the lemma. If $x_0^1 \in [q_0, p_1^2)$, then by the argument in the previous paragraph, we must have $x_0^2 \in [q_0^1, x_0^1) \subset (x_0, x_0^1)$ and we get 3) of the lemma.

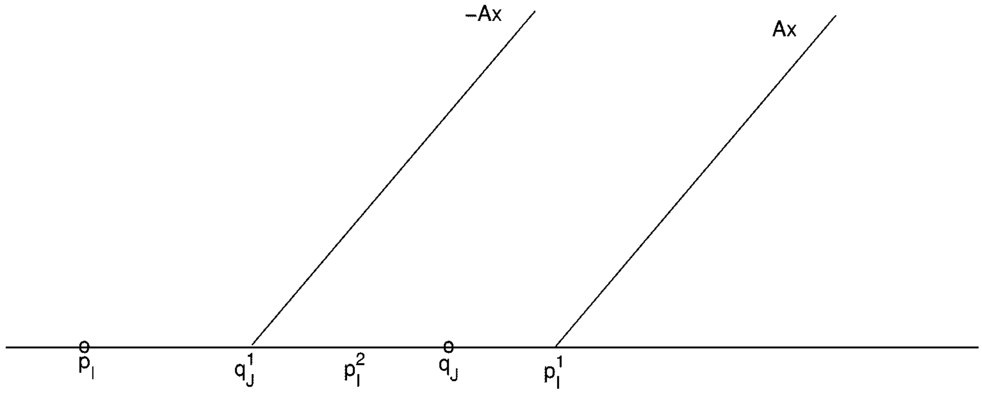

Fig. 7. Illustration for the proof of Lemma 4.4: Case i.

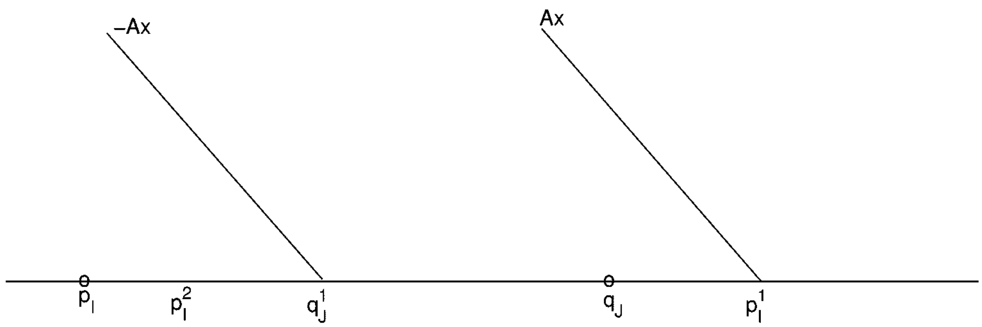

Fig. 8. Illustration for the proof of Lemma 4.4: Case ii.

Case 2—$p_1^2 \in (p_1, p_0]$: (See Fig. 4.) For $x_0 \in (p_1, p_0]$, $x_0^1 = \mathrm{sat}_2(A^{m_0}x_0) \in (p_1^2, p_0^1]$. Since p_1^2 and p_0^1 are to the right of p_1 and p_0 respectively, x_0^1 must also be to the right of x_0, i.e., $x_0^1 \in (x_0, p_0^1]$. If $x_0^1 \in (x_0, p_0)$, then we get 2) and the subsequent intersections, if any, will move rightward until falling between p_0 and q_0; if $x_0^1 \in (p_0, p_0^1) \subset (p_0, q_0)$, then we get 1).

For $x_0 \in [q_0, p_1^1)$, $x_0^1 = \mathrm{sat}_2(A^{n_0}x_0) \in (p_1^2, q_0^1)$. If $x_0^1 \in [p_0, q_0^1) \subset (p_0, q_0)$ then we obtain 4). If $x_0^1 \in (p_1^2, p_0]$, then the argument in the foregoing paragraph applies and we have $x_0^2 \in (x_0^1, p_0^1] \subset (x_0^1, x_0)$, which belongs to 6).

Case 3—$p_1^2 \in (p_0, q_0)$: (See Fig. 5.) For $x_0 \in (p_1, p_0]$, we have $x_0^1 \in (p_1^2, p_0^1] \subset (p_0, q_0)$, which belongs to 1). For $x_0 \in [q_0, p_1^1)$, we have $x_0^1 \in (p_1^2, q_0^1] \subset (p_0, q_0)$, which belongs to 4).

So far, we have shown that one of 1)–6) holds for all $x_0 \in [p_1, p_1^1]$. And in each of the above three cases, we see that for all $x_0 \in (p_1, p_1^1)$, x_0^1, x_0^2 and the subsequent intersections are all in (p_1, p_1^1) and will fall between p_0 and q_0 in a finite number of steps.

Next, we assume that these properties hold for all $x_0 \in [p_i, p_i^1]$ and the next segment which includes $[p_i, p_i^1]$ is $[q_j^1, q_j]$ (see Fig. 6). We also have three cases: $q_j^2 \in [p_i^1, q_j)$, $q_j^2 \in (q_j^1, p_i]$ and $q_j^2 \in (p_i, p_i^1)$. By treating the segment $[p_i, p_i^1]$ as $[p_0, q_0]$ in the proof for $[p_1, p_1^1]$, we can use the same argument to show that one of 1)–6) holds for all $x_0 \in (q_j^1, p_i] \cup [p_i^1, , q_j)$. Moreover, for all $x_0 \in (q_j^1, p_i] \cup [p_i^1, , q_j)$, the intersections will move toward $[p_i, p_i^1]$ and fall between $[p_i, p_i^1]$ in a finite number of steps.

Now, suppose that $[p_I, p_I^1]$ is the outermost segment. By induction, we have obtained the properties in the lemma for all $x_0 \in [p_I, p_I^1]$ and we would like to extend the properties to the whole line L_h^e.

Recall that $m_I = n_J = 1$, so $p_I^1 = Ap_I$ and $p_I^2 = \mathrm{sat}_2(A^2p_I) = \mathrm{sat}_2(Ap_I^1)$. The line AL_h^e will actually intersect with $\pm L_h^e$ at p_I^1 and $-q_J^1$ (or $-p_I^1$ and q_J^1). Assume that p_I^1 is on L_h^e. Then the ray $\{Ax\colon x \in [q_J, q_\infty)\}$ is below the line $-L_h^e$. Get a symmetric projection of this ray as $\{-Ax\colon x \in [q_J, q_\infty)\}$. Then the two rays $\{-Ax\colon x \in [q_J, q_\infty)\}$ and $\{Ax\colon x \in (p_\infty, p_I]\}$ are parallel and are both above the line L_h^e (see Figs. 7 and 8). Here, we have two cases.

Case i: The rays have a positive slope (see Fig. 7). Since $p_I^1 \in [q_J, q_\infty)$, we have $p_I^2 \in [q_J^1, q_\infty)$ and by Lemma 4.3, $p_I^2 \in (p_I, p_I^1)$. So, $p_I^2 \in (q_J^1, p_I^1)$.

For $x_0 \in (p_I^1, q_\infty)$, since both $q_J^1 = -\mathrm{sat}_2(Aq_J)$ and $p_I^2 = -\mathrm{sat}_2(Ap_I^1)$ are to the left of q_J and p_I^1, respectively, and since there exists no $x \in L_h^e$ such that $-\mathrm{sat}_2(Ax) = x$, we must have, $x_0^1 = -\mathrm{sat}_2(Ax_0)$ to the left of x_0 and $x_0^1 \in (p_I^2, x_0)$. If $x_0^1 \in (p_0, q_0)$, then we obtain 4). If $x_0^1 \in [q_0, x_0)$, then we obtain 5). If $x_0^1 \in (q_J^1, p_0) \subset (q_J^1, q_J)$, then by the established properties on $[q_J^1, q_J]$, we must have $x_0^2 \in (x_0^1, q_J)$ to the left of p_I^1 and hence to the left of x_0. Therefore, $x_0^2 \in (x_0^1, x_0)$ and we get 6).

For $x_0 \in (p_\infty, p_I)$, $x_0^1 = \mathrm{sat}_2(Ax)$ is to the right of p_I^1, so the properties for $x_0 \in (p_I^1, p_\infty)$ applies. Also note that x_0^2 is to the right of p_I^2. Hence $x_0^2 \in (x_0, x_0^1)$ and we obtain 3).

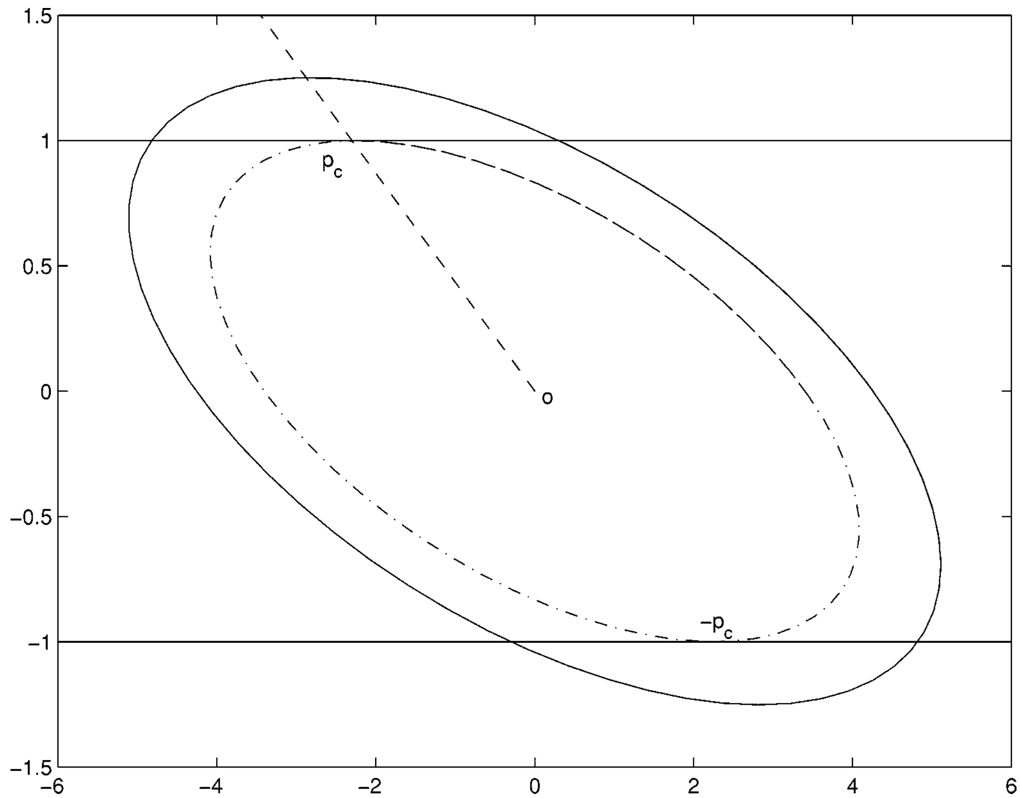

Fig. 9. Illustration for the proof of Lemma 4.5.

Case ii: The rays have a negative slope (see Fig. 8). In this case, $p_I^2 \in (p_I,\, q_J^1)$.

For $x_0 \in (p_\infty,\, p_I)$, since $p_I^1 = Ap_I$ is to the right of p_I and there exist no $x \in L_h^e$ such that $\mathrm{sat}_2(Ax) = x$, we must have $x_0^1 = \mathrm{sat}_2(Ax_0)$ to the right of x_0, in particular, $x_0 \in (x_0,\, p_I^1)$. If $x_0^1 \in (x_0,\, p_0]$, then we obtain 2). If $x_0^1 \in (p_0,\, q_0)$, then we have 1). If $x_0^1 \in (q_0,\, p_I^1)$, then by using the established property in the interval $(p_I,\, p_I^1)$, we have $x_0^2 \in (p_I,\, x_0^1) \subset (x_0,\, x_0^1)$ and we obtain 3).

For $x_0 \in (p_I^1,\, q_\infty)$, $x_0^1 = -\mathrm{sat}_2(Ax)$ is to the right of p_I^2. By applying the property for x_0 in $(p_\infty,\, p_I)$ and $(p_I,\, p_I^1)$, we have $x_0^2 \in (x_0^1,\, x_0)$, which belongs to f).

Similar to the argument for the interval $(p_1,\, p_1^1)$, it can be shown that the intersections will fall between $(p_I,\, p_I^1)$ in a finite number of steps for all $x_0 \notin (p_I,\, p_I^1)$.

In summary, the intersections of a trajectory $\psi_2(k,\, x_0)$ with the lines $\pm L_h^e$ will move from the outer intervals to the inner intervals until falling into $(p_0,\, q_0)$ in a finite number of steps. After that, it will not touch the lines $\pm L_h^e$ and will converge to the origin. □

Next, we suppose that the condition (9) is true for some $x_1 \in \mathbf{R}$. We would like to determine an interval in L_h^e such that a trajectory starting from this interval will converge to the origin.

Recall that p_c is defined to be the unique intersection of the Lyapunov ellipsoid $\mathcal{E}(\rho_c)$ with the line L_h^e (see Fig. 9). Also, α_c is the first coordinate of p_c, i.e., $p_c = \left[\begin{smallmatrix}\alpha_c\\ 1\end{smallmatrix}\right]$.

Lemma 4.5: Assume that $\alpha_c \le 0$. If there exist an integer $N > 0$ and a $d > 0$ such that $A^N\left[\begin{smallmatrix}x_1\\ 1\end{smallmatrix}\right] = \pm\left[\begin{smallmatrix}x_1\\ 1+d\end{smallmatrix}\right]$, then we must have $x_1 < \alpha_c < 0$.

Proof: When $\alpha_c \le 0$, an ellipsoid $\mathcal{E}(\rho)$ takes the shape in Fig. 9. Each ellipsoid $\mathcal{E}(\rho)$ has an intersection with the ray that starts from the origin and passes through p_c. This intersection is the highest point in the ellipsoid. Since $\alpha_c \le 0$, it can be seen that

$$V\left(\begin{bmatrix}x_1\\ 1\end{bmatrix}\right) < V\left(\pm\begin{bmatrix}x_1\\ 1+d\end{bmatrix}\right) \qquad \forall x_1 \ge \alpha_c, \quad d > 0.$$

Since

$$V\left(A^N\begin{bmatrix}x_1\\ 1\end{bmatrix}\right) < V\left(\begin{bmatrix}x_1\\ 1\end{bmatrix}\right)$$

it is impossible to have $A^N\left[\begin{smallmatrix}x_1\\ 1\end{smallmatrix}\right] = \pm\left[\begin{smallmatrix}x_1\\ 1+d\end{smallmatrix}\right]$ for any $x_1 \ge \alpha_c$. □

Lemma 4.6: Let

$$\alpha_s = \min\{|x_1|\colon x_1 \text{ satisfies (9)}\}.$$

Case 1) $\alpha_c \le 0$. Let

$$p_s = \begin{bmatrix}-\alpha_s\\ 1\end{bmatrix}$$

then $p_s^1 = p_s \in (p_\infty,\, p_0)$. Suppose that $p_s \in [p_{i+1},\, p_i)$. Then, for every $x_0 \in (p_s,\, p_i^1]$, the trajectory $\psi_2(k,\, x_0)$ will converge to the origin;

Case 2) $\alpha_c > 0$. Let

$$p_s = \begin{bmatrix}\alpha_s\\ 1\end{bmatrix}$$

then $p_s^1 = p_s \in (q_0,\, q_\infty)$. Suppose that $p_s \in (q_j,\, q_{j+1}]$. Then, for every $x_0 \in [q_j^1,\, p_s]$, the trajectory $\psi_2(k,\, x_0)$ will converge to the origin.

In both cases, no limit trajectory can be formed completely inside the strip

$$\left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : |x_1| < \alpha_s \right\}.$$

Proof: We only prove Case 1. Since $\alpha_c \le 0$, by Lemma 4.5, there is no $x_1 \ge \alpha_c$ satisfying (9), so we have $-\alpha_s < \alpha_c < 0$ and p_s must be between p_{i+1} and p_i for some i, noting that p_s cannot be in $[p_0, q_0]$ by Lemma 4.1 1). Following the iterative procedure in the proof of Lemma 4.4, we can show that for all $x_0 \in [p_i, p_i^1]$, the trajectory will converge to the origin. Now we consider a point between p_s and p_i. For x_0 in this interval, the next intersection of the trajectory with the lines $\pm L_h^e$ is $x_0^1 = \mathrm{sat}_2(A^{m_i} x_0)$. Since $\mathrm{sat}_2(A^{m_i} p_s) = p_s$ (or $-p_s$) and $p_i^1 = \mathrm{sat}_2(A^{m_i} p_i)$ is to the right of p_i, we must have $x_0^1 \in (x_0, p_i^1)$, and the subsequent intersections will move rightward and fall between p_i and p_i^1 in a finite number of steps. Therefore, the trajectory $\psi_2(k, x_0)$ will converge to the origin.

Now, consider $x_0 \in (p_i^1, q_\infty) \subset (p_c, q_\infty)$. Let k_1 be the minimal integer such that $A^{k_1} x_0$ goes out of $\mathbf{S}^e$, then by the shape of the Lyapunov ellipsoid, the point $A^{k_1} x_0$ must be to the left of x_0 (or to the right of $-x_0$ if $A^k x_0$ is below the line $-L_h^e$), otherwise we would have $V(A^{k_1} x_0) > V(x_0)$, which is impossible. Hence, $x_0^1 = \mathrm{sat}_2(A^{k_1} x_0)$ must be to the left of x_0, and the subsequent intersections either fall between p_s and p_i^1 at a finite step, or go to the left of p_s. This shows that no limit trajectory can be formed completely to the right of p_s and symmetrically, to the left of $-p_s$. Hence, no limit trajectory can be formed completely inside the strip

$$\left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : |x_1| < \alpha_s \right\}.$$

□

Proof of Proposition 4.1: It follows immediately from Lemmas 4.4 and 4.6. □

The number α_s and the point $p_s = \left[\begin{smallmatrix} \pm\alpha_s \\ 1 \end{smallmatrix}\right]$ can be easily computed by applying Lemma 4.1. Actually, all the x_1 satisfying (9) can be determined. Let $x_0 = \left[\begin{smallmatrix} x_1 \\ 1 \end{smallmatrix}\right]$, then x_1 satisfies (9) for some N if and only if x_0 satisfies

$$A^N x_0 \notin \mathbf{S}^e \quad A^k x_0 \in \mathbf{S}^e \qquad \forall\, k < N$$

and

$$\mathrm{sat}_2(A^N x_0) = x_0. \tag{12}$$

Assume that $\alpha_c < 0$, then by Lemma 4.5, we only need to check if there is such an x_0 in the interval (p_∞, p_0). Clearly, no x_0 in $[p_0, q_0]$ satisfies (12) by Lemma 4.1 1). So we need to check over the intervals $[p_{i+1}, p_i)$ with i increased from 0 to $I - 1$ and the interval (p_∞, p_I).

Consider a point x_0 in the interval $[p_{i+1}, p_i)$. By Lemma 4.1 2), the smallest integer N for $A^N x_0 \notin \mathbf{S}^e$ is $N = m_i$. By Lemma 4.2 3), $A^{m_i} p_i = p_i^1 \in (q_0, q_\infty)$ is to the right of p_i. So there exists $x_0 \in [p_{i+1}, p_i)$ satisfying (12) if and only if $A^{m_i} p_{i+1}$ is to the left of p_{i+1}, i.e.,

$$\mathrm{sat}_2(A^{m_i} p_{i+1}) \in (p_\infty, p_{i+1}]. \tag{13}$$

If this is true, then x_1, the first coordinate of x_0, can be solved from

$$[1 \quad 0] A^{m_i} \begin{bmatrix} x_1 \\ 1 \end{bmatrix} = \pm x_1. \tag{14}$$

In summary, we have the following

Algorithm for Determining All the x_1 Satisfying Condition (9): Assume $\alpha_c \le 0$. Initially set $i = 0$.

Step 1) $i = i + 1$. If (13) is satisfied, then compute x_1 from (14). Repeat this step until $i = I - 1$.

Step 2) Solve

$$[1 \quad 0] A \begin{bmatrix} x_1 \\ 1 \end{bmatrix} = \pm x_1$$

for x_1, if $x_1 \in (-\infty, \alpha_I)$, then x_1 satisfies (9) with $N = 1$.

V. Proof of the Main Results

Now, we turn back to the system (5),

$$x(k+1) = \mathrm{sat}(Ax(k)). \tag{15}$$

For easy reference, we restate Theorem 2.1 as follows.

Theorem 5.1: The system (15) is globally asymptotically stable if and only if A is stable and none of the following statements are true.

1) There exists an $N \ge 1$ such that

$$\mathrm{sat}(A^N v_1)) = \pm v_1 \quad \text{and} \quad A^k v_1 \in \mathbf{S} \qquad \forall\, k < N.$$

2) There exists an $N \ge 1$ such that

$$\mathrm{sat}(A^N v_2)) = \pm v_2 \quad \text{and} \quad A^k v_2 \in \mathbf{S} \qquad \forall\, k < N.$$

3) There exists an $x_1 \in (-1, 1)$ and an $N \ge 1$ such that

$$\mathrm{sat}\left(A^N \begin{bmatrix} x_1 \\ 1 \end{bmatrix}\right) = \pm \begin{bmatrix} x_1 \\ 1 \end{bmatrix}$$

and

$$A^k \begin{bmatrix} x_1 \\ 1 \end{bmatrix} \in \mathbf{S} \qquad \forall\, k < N.$$

4) There exists an $x_2 \in (-1, 1)$ and an $N \ge 1$ such that

$$\mathrm{sat}\left(A^N \begin{bmatrix} 1 \\ x_2 \end{bmatrix}\right) = \pm \begin{bmatrix} 1 \\ x_2 \end{bmatrix}$$

and

$$A^k \begin{bmatrix} 1 \\ x_2 \end{bmatrix} \in \mathbf{S} \qquad \forall\, k < N.$$

Proof: We will exclude the possibility of the existence of limit trajectories (except for the trivial one at the origin) under the condition that none of statements 1)–4) in the theorem is true. In the following, when we say a limit trajectory, we mean a nontrivial one other than the origin. Clearly, every limit trajectory must include at least one point on the boundary of the unit square, i.e., a point in the set $\pm(L_h \cup L_v \cup \{v_1, v_2\})$. By Proposition 3.1, we know that a limit trajectory cannot have points in both $\pm L_h$ and $\pm L_v$. So we have two possibilities here, limit trajectories including points in $\pm(L_h \cup \{v_1, v_2\})$, and those including points in $\pm(L_v \cup \{v_1, v_2\})$. Because of the similarity, we only exclude the first possibility under the condition that none of 1)–3) is true, the second possibility can be excluded under the condition that none of 1), 2) and 4) is true.

For a given initial state x_0, we denote the trajectory of the system (15) as $\psi(k, x_0)$ and the trajectory of (7) as $\psi_2(k, x_0)$.

Clearly, if $x_1 \in (-1, 1)$ satisfies 3), then this x_1 also satisfies (9). On the other hand, suppose that there is some x_1 that satisfies (9). Let p_s be as defined in Lemma 4.6 for the system (7) [if there is no x_1 that satisfies (9), then we can assume that

$$p_s = \begin{bmatrix} \pm\infty \\ 1 \end{bmatrix}$$

and the following argument also goes through]. Note that, if there is some $x_1 \in \mathbf{R}$, $|x_1| \le 1$, that satisfies (9), i.e.,

$$A^N \begin{bmatrix} x_1 \\ 1 \end{bmatrix} = \pm \begin{bmatrix} x_1 \\ 1+d \end{bmatrix}$$

and

$$\left| [0 \quad 1] A^k \begin{bmatrix} x_1 \\ 1 \end{bmatrix} \right| \le 1 \qquad \forall k < N$$

we must also have

$$\left| [1 \quad 0] A^k \begin{bmatrix} x_1 \\ 1 \end{bmatrix} \right| \le |x_1| \qquad \forall k < N$$

which indicates that x_1 satisfies 3). Otherwise, as in the proof of Proposition 3.1, the area of the convex hull of the set

$$\{\pm x_0, \pm Ax_0, \ldots, \pm A^{N-1} x_0\}$$

would be less than the area of the convex hull of the set

$$\{\pm Ax_0, \pm A^2 x_0, \ldots, \pm A^N x_0\}.$$

This would be a contradiction to the fact that $|\det(A)| < 1$.

Hence, if no x_1 satisfies 3), then p_s must be outside of $\mathbf{S}$. By Proposition 4.1, no limit trajectory of (7) can lie completely inside the strip

$$\left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : |x_1| < \alpha_s \right\}.$$

It follows that no limit trajectory of (15) can lie completely between $-L_v$ and L_v. Therefore, no limit trajectory of (15) can include only boundary points in $\pm L_h$. On the other hand, if a limit trajectory include only boundary points $\pm v_1$ (or $\pm v_2$, note that, by Proposition 3.1, no limit trajectory can include both $\pm v_1$ and $\pm v_2$), then 1) or 2) must be true, which contradicts our assumption. In short, if there is a limit trajectory that include points in $\pm(L_h \cup \{v_1, v_2\})$, it must include at least one point on $\pm L_h$ and one on $\pm v_1$ (or $\pm v_2$). Here, we assume that it includes v_2.

Let us consider the trajectories $\psi(k, v_2)$ and $\psi_2(k, v_2)$. Suppose that $\psi(k, v_2)$ has an intersection with $\pm L_h$ but does not include v_1 and any point in $\pm L_v$, we conclude that $\psi(k, v_2) = \psi_2(k, v_2)$ will converge to the origin. The argument goes as follows.

Let k_0 be the smallest k such that $\psi(k, v_2)$ intersects $\pm L_h$. Denote $v_2^1 = \psi(k_0, v_2)$. Since 2) is not true, k_0 must also be the smallest k such that

$$\left| [0 \quad 1] A^k v_2 \right| \ge 1.$$

So, we have $\psi_2(k, v_2) = \psi(k, v_2)$ for all $k \le k_0$. Here, we have two cases.

Case 1—$\alpha_c \le 0$: In this case, p_s is to the left of v_2. Since $v_2^1 = \mathrm{sat}(A^{k_0} v_2) = \mathrm{sat}_2(A^{k_0} v_2)$ goes to the right of v_2, by Lemma 4.5, v_2 must be to the left of p_0. It follows that $v_2 \in (p_s, p_i^1]$, where $(p_s, p_i^1]$ is the interval in Lemma 4.6 2). Hence, $\psi_2(k, v_2)$ will converge to the origin. Moreover, the subsequent intersections of $\psi_2(k, v_2)$ with $\pm L_h$ are between v_2 and v_2^1. Since $\psi(k, v_2)$ does not touch $\pm L_v$, we must have $\psi(k, v_2) = \psi_2(k, v_2)$ and hence $\psi(k, v_2)$ will also converge to the origin.

Case 2—$\alpha_c > 0$: In this case p_s is to the right of v_1. By the assumption that $\psi(k, v_2)$ does not include v_1, the intersections of $\psi_2(k, v_2)$ with $\pm L_h$ will stay to the left of v_1 (or to the right of $-v_1$). Since $\alpha_c > 0$, by Lemma 4.5, the intersections will move rightward until falling on $[q_j^1, p_s)$, where $[q_j^1, p_s)$ is the interval in Lemma 4.6 3). Similar to Case 1, we have that $\psi_2(k, v_2)$ converges to the origin and $\psi(k, v_2) = \psi_2(k, v_2)$.

So far, we have excluded the possibility that a limit trajectory includes any point in the set $\pm(L_h \cup \{v_1, v_2\})$. The possibility that a limit trajectory includes any point in the set $\pm(L_v \cup \{v_1, v_2\})$ can be excluded in a similar way. Thus, there exists no limit trajectory of any kind and the system (15) must be globally asymptotically stable. □

Here we provide a simple method to check the conditions 3) and 4) of Theorem 5.1 based on the algorithm to determine all the x_1 satisfying (9) and hence p_s in the previous section. From the proof of Theorem 5.1, we see that 3) is true if and only if $p_s \in \mathbf{S}$. To check 4), we can exchange x_1 and x_2, i.e., use a state transformation $y = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right] x$. The system (15) is then equivalent to

$$y(k+1) = \mathrm{sat}\left(\overline{A} y(k)\right) \tag{16}$$

where $\overline{A} = \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right] A \left[\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right]$. The condition 4) for the system (15) is equivalent to the condition 3) for the system (16).

VI. Conclusions

We gave a complete stability analysis of a planar discrete-time linear system under saturation. The analysis involves intricate investigation on the intersections of the trajectories with the lines $x_1 = \pm 1$ and $x_2 = \pm 1$. Our main result provides a necessary and sufficient condition for such a system to be globally asymptotically stable.

References

[1] F. Albertini and D. D'Alessandro, "Asymptotic stability of continuous-time systems with saturation nonlinearities," *Syst. Control Lett.*, vol. 29, no. 3, pp. 175–180, 1996.

[2] D. S. Bernstein and A. N. Michel, "A chronological bibliography on saturating actuators," *Int. J. Robust Nonlinear Control*, vol. 5, no. 5, pp. 375–380, 1995.

[3] L. Hou and A. N. Michel, "Asymptotic stability of systems with saturation constraints," *IEEE Trans. Automat. Contr.*, vol. 43, pp. 1148–1154, Aug. 1998.

[4] T. Hu and Z. Lin, "A complete stability analysis of planar linear systems under saturation," *IEEE Trans. Circuits Syst. I*, vol. 47, pp. 498–512, Apr. 2000.

[5] L. Jin, P. N. Nikiforuk, and M. M. Gupta, "Absolute stability conditions for discrete-time recurrent neural networks," *IEEE Trans. Neural Networks*, vol. 5, pp. 954–964, Nov. 1994.

[6] Z. Lin, *Low Gain Feedback*, ser. Lecture Notes in Control and Information Sciences. London, U.K.: Springer, 1998, vol. 240.

[7] D. Liu and A. N. Michel, "Asymptotic stability of systems operating on a closed hypercube," *Syst. Control Lett.*, vol. 19, no. 4, pp. 281–285, 1992.

[8] ——, "Sparsely interconnected neural networks for associative memories with applications to cellular neural networks," *IEEE Trans. Circuits Syst.*, vol. 41, pp. 295–307, Apr. 1994.

[9] ——, *Dynamical Systems with Saturation Nonlinearities*, ser. Lecture Notes in Control and Information Sciences. London: Springer, 1994, vol. 195.

[10] R. Mantri, A. Saberi, and V. Venkatasubramanian, "Stability analysis of continuous time planar systems with state saturation nonlinearity," *IEEE Trans. Circuits Syst. I*, vol. 45, pp. 989–993, Sept. 1998.

[11] J. H. F. Ritzerfeld, "A condition for the overflow stability of second-order digital filters that is satisfied by all scaled state-space structures using saturation," *IEEE Trans. Circuits Syst.*, vol. 36, pp. 1049–1057, Aug. 1989.

Tingshu Hu was born in Sichuan, China in 1966. She received the B.S. and M.S. degrees in electrical engineering from Shanghai Jiao Tong University, Shanghai, China, in 1985 and 1988, respectively. She is currently working toward the Ph.D degree in electrical engineering, at the University of Virginia, Charlottesville.

Her research interests include systems with saturation nonlinearities and robust control theory, areas in which she has published several papers. She is also the co-author (with Zongli Lin) of the book *Control Systems with Actuator Saturation: Analysis and Design* (Birkhäuser, Boston, 2001).

Ms. Hu is an Associate Editor on the Conference Editorial Board of the IEEE Control Systems Society.

Zongli Lin (S'89–M'90–SM'98) was born in Fuqing, Fujian, China, in 1964. He received the B.S. degree in mathematics and computer science from Amoy University, Xiamen, China, in 1983, the Master of Engineering degree in automatic control from the Chinese Academy of Space Technology, Beijing, China, in 1989, and the Ph.D. degree in electrical and computer engineering, from Washington State University, Pullman, WA, in May 1994.

From July 1983 to July 1986, Dr. Lin worked as a Control Engineer at Chinese Academy of Space Technology. In January 1994, he joined the Department of Applied Mathematics and Statistics, State University of New York, Stoney Brook as a Visiting Assistant Professor, where he became an Assistant Professor in September 1994. Since July 1997, he has been an Assistant Professor in electrical engineering at University of Virginia, Charlottesville. His current research interests include nonlinear control, robust control, and control of systems with saturating actuators. In these areas he has published several papers. He is also the author of the recent book, *Low Gain Feedback* (Springer-Verlag, London, 1998) and a co-author (with Tingshu Hu) of the book *Control Systems with Actuator Saturation: Analysis and Design* (Birkhäuser, Boston, 2001).

Dr. Lin has been an Associate Editor on the Conference Editorial Board of the IEEE Control Systems Society and currently serves as an Associate Editor of IEEE TRANSACTIONS ON AUTOMATIC CONTROL. He is the recipient of a U.S. Office of Naval Research Young Investigator Award.