Proceedings of the 36th
Conference on Decision & Control
San Diego, California USA • December 1997

The Controllability and Stabilization of Unstable LTI Systems with Input Saturation [1]

Tingshu Hu[2] Li Qiu[3] Zongli Lin[4]

Abstract

In this paper, we study the controllable region of a general unstable continuous-time LTI system with input saturation and we also study the stabilization of such a system by a saturated linear state feedback. We give simple exact descriptions of the controllable regions for certain classes of unstable systems. The study on stabilization is quite preliminary. We only deal with anti-stable planar systems. We conjecture that for such a system its asymptotic stability region (domain of attraction) under a saturated linear state feedback can be easily obtained from a stable limit cycle of its time-reversed system. We conjecture with convincing arguments that for such a system a saturated linear state feedback can be designed so that the asymptotic stability region is arbitrarily close to its controllable region.

1 Introduction

Recently there is a renewed interest in the control of systems with bounded inputs. Great progress has been made in the past few years. In the continuous-time setting, most of the existing work deals with systems that have no poles on the open right half of the complex plane. (We will call such systems semistable systems.)

It is well-known, see for example [6, 10, 11], that a semistable LTI system controllable in the usual sense is globally controllable with bounded inputs. Based on this fact, extensive literature is devoted to the control of semistable systems using bounded control. In [12] and [13], nonlinear globally asymptotically stabilizing feedback laws were designed. Later, linear saturated state feedback laws were constructed so that the closed-loop system is asymptotically stable within any prescribed bounded region, see, e.g., [4, 5, 9]. In these papers, the feedback gains are kept small so that within a prescribed region of state, the control signal will not exceed the saturation level. It was also recognized that if the feedback is designed by the LQ method, then the feedback can be amplified by any positive gain without affecting the stability region. This positive gain is then utilized to improve other performance of the system, see [3, 9]. In [7], a nonlinear feedback is designed to guarantee globally asymptotic stability and $\mathcal{L}_2$ BIBO stability.

For strictly unstable systems that have poles on the open right half of the complex plane, however, the existing results are quite limited. It is known that such systems are not globaly controllable and hence cannot be globally stabilized in any way. Just as the controllability result of [6, 10, 11] paved the way for the development of stabilization theory for semistable systems with bounded inputs, the control of strictly unstable systems requires the exact descriptions of their controllable region. It is shown in [2] that there exists a nice seperation result concerning the controllable region of a strictly unstable system. Suppose a strictly unstable system is decomposed into the sum of a semistable subsystem and an anti-stable subsystem, then the controllable region of the whole system is the Cartesian product of the controllable region of the semistable subsystem, which is the whole state space of this subsystem, and that of the anti-stable subsystem, which is a bounded convex open set. Because of this, it suffices to study the controllability of anti-stable systems.

The stabilization of a strictly unstable system is a much harder issue. Even the analysis problem of decribing the asymptotic stability region (domain of contraction) of a closed loop system is not sufficiently addressed. Although a stability region of the closed-loop system can be estimated, even some performance can be guaranteed within this region, it is not clear whether this region is too conservative or not, nor is it clear how to enlarge this region or to make it meet the performance requirements.

In this paper, we first study the controllable region of a continuous time anti-stable system with saturated inputs, i.e., the input bound is given by an ∞-norm bound. For anti-stable systems with only real poles and second order anti-stable systems with complex eigenvalues, simple formulas for the boundaries of the controllable regions are obtained. These formulas provides interesting geometric insigts which are useful in the study of stabilization of such systems.

The study of stabilization in this paper focuses on planar anti-stable systems. The reason for this is the symplicity of such systems. More general systems are left for future study. For planar anti-stable systems, we conjecture that the asymptotic stability region under a saturated linear state feedback can be easily obtained. We also conjecture that a saturated linear state feedback can be designed so that the asymptotic stability region of the closed-loop system is arbitrarily close to the controllable region of the open loop system.

[1]This research was supported by the Hong Kong Research Grants Council.

[2]Department of Automation, Shanghai Jiao Tong University, Shanghai 200030, P.R. China.

[3]Department of Electrical & Electronic Engineering, Hong Kong University of Science & Technology, Clear Water Bay, Kowloon, Hong Kong.

[4]Department of Electrical Engineering, University of Virginia, Charlottesville, VA 22903, USA.

0-7803-3970-8/97 $10.00 © 1997 IEEE

2 Preliminaries and Notation

Consider LTI system

$$\dot{x}(t) = Ax(t) + Bu(t) \tag{1}$$

where $x(t) \in \mathbf{R}^n$ is the state and $u(t) \in \mathbf{R}^m$ is the control. A control signal u is said to be *admissible* if $\|u(t)\|_\infty \leq 1$ for all $t \geq 0$. In this paper, we are interested in the control of system (1) by using admissible controls. Our first concern is the set of states that can be steered to the origin by an admissible control.

Definition 1 *Consider system (1).*

(a) *A state x_0 is said to be controllable at a given time T if there exists an admissible control u such that the state trajectory x of the system satisfies $x(0) = x_0$ and $x(T) = 0$. The set of all states controllable at T is called the controllable region of the system at T and is denoted by $\mathcal{C}(T)$.*

(b) *A state x_0 is said to be controllable if it is controllable at some $T < \infty$. The set of all controllable states is called the controllable region of the system and is denoted by $\mathcal{C}$.*

Let $\langle A|B \rangle$ be the controllable subspace of the pair (A, B). Since the controllable region of (1) has to be a subset of $\langle A|B \rangle$, it can be given by that of its controllable subsystem. Hence we assume in the following, without loss of generality, that (A, B) is controllable.

Proposition 1 *Assume that (A, B) is controllable.*

(a) *If A is semi-stable, then $\mathcal{C} = \mathbf{R}^n$.*

(b) *If A is anti-stable, then $\mathcal{C}$ is a bounded convex open set containing the origin.*

(c) *If $A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$ with $A_1 \in \mathbf{R}^{n_1 \times n_1}$ being anti-stable and $A_2 \in \mathbf{R}^{n_2 \times n_2}$ being semi-stable, and B is partitioned as $\begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$ accordingly, then $\mathcal{C} = \mathcal{C}_1 \times \mathbf{R}^{n_2}$ where $\mathcal{C}_1$ is the controllable region of the anti-stable system (A_1, B_1).*

Statement (a) is well-known [6, 10, 11]. Statements (b) and (c) are proved in [2]. Because of this proposition, we can concentrate on the study of controllable regions of anti-stable systems. For such systems $\mathcal{C}$ can be approximated by $\mathcal{C}(T)$ for sufficiently large T. To make this formal let us introduce the Hausdorff distance between bounded subsets of $\mathbf{R}^n$. Let $\mathcal{S}_1, \mathcal{S}_2$ be two bounded subsets of $\mathbf{R}^n$. Then their Hausdorff distance is defined as:

$$d(\mathcal{S}_1, \mathcal{S}_2) := \max\{\vec{d}(\mathcal{S}_1, \mathcal{S}_2), \vec{d}(\mathcal{S}_2, \mathcal{S}_1)\}$$

where

$$\vec{d}(\mathcal{S}_1, \mathcal{S}_2) = \sup_{x_1 \in \mathcal{S}_1} \inf_{x_2 \in \mathcal{S}_2} \|x_1 - x_2\|.$$

Here the vector norm used is arbitrary. With the Hausdorff distance, we have

$$\lim_{T \to \infty} d(\mathcal{C}(T), \mathcal{C}) = 0.$$

If we restrict ourselves to the compact subsets of $\mathbf{R}^n$, then the Hausdorff distance is a metric. Under this metric, we clearly have $\lim_{T \to \infty} \mathcal{C}(T) = \overline{\mathcal{C}}$, the closure of $\mathcal{C}$, and $\lim_{T \to \infty} \partial\mathcal{C}(T) = \partial\mathcal{C}$. Here we use "$\partial$" to denote the boundary of a set. In Section 3, we will derive methods for generating the boundary of $\mathcal{C}$.

If $B = \begin{bmatrix} b_1 & \cdots & b_m \end{bmatrix}$ and the controllable region of the system (A, b_i), $i = 1, \cdots, m$, is $\mathcal{C}_i$, then $\mathcal{C} = \sum_{i=1}^m \mathcal{C}_i$. Hence in the study of the controllable regions we can further assume without loss of generality that $m = 1$.

In summary, we will assume in the following that (A, B) is controllable, A is anti-stable, and $m = 1$.

In Section 4, we study the stabilization of a system with input saturation. Our ultimate purpose is to show that a satuarted linear feedback can be designed so that the stability region (domain of attraction) of the closed loop system is arbitrarily close to the controllable region of the systems. We will in Section 4 start with the stabilization of anti-stable systems since the stabilization of semistable systems has become a rather mature topic [4, 5, 7, 9, 12, 13]. For technical reasons, we further restrict our study to second order anti-stable systems. The results in Section 4 are preliminary and in some cases speculatory. Further study is underway.

In many situations, it may be more convenient to study the controllability of a system through the reachability of its time-reversed system. For a nonlinear system

$$\dot{x} = f(x, u) \tag{2}$$

its time reversed system is

$$\dot{z} = -f(z, v) \tag{3}$$

It is easy to see that $x(t)$ solves (2) with $x(0) = x_0, x(t_1) = x_1$, and certain u if and only if $z(t) = x(t_1 - t)$ solves (3) with $z(0) = x_1$, $z(t_1) = x_0$, and $v(t) = u(t_1 - t)$. The two systems have the same curves as trajectories, but traversed in opposite directions.

Consider the time reversed system of (1):

$$\dot{z}(t) = -Az(t) - Bv(t). \tag{4}$$

Definition 2 *Consider system (4).*

(a) *A state z_T is said to be reachable at a given time T if there exists an admissible control v such that the state trajectory z of system (4) satisfies $z(0) = 0$ and $z(T) = z_T$. The set of all states reachable at T is called the reachable region of system (4) at T and is denoted by $\mathcal{R}(T)$.*

(b) *A state z is said to be reachable if it is reachable at some $T < \infty$. The set of all reachable states is called the reachable region of system (4) and is denoted by $\mathcal{R}$.*

It is a known result that $\mathcal{C}(T)$ and $\mathcal{C}$ of (1) are the same as $\mathcal{R}(T)$ and $\mathcal{R}$ of (4), see for example [6]. To avoid confusion, we will reserve the notation x, u, $\mathcal{C}(T)$, and $\mathcal{C}$ for the original system (1), and reserve z, v, $\mathcal{R}(T)$, and $\mathcal{R}$ for the time-reversed system (4).

3 Controllable Regions

3.1 Description of the controllable region via extremal control

In this section, we consider the controllable regions $\mathcal{C}(T)$ and $\mathcal{C}$ of system (1) via the study of the reachable regions $\mathcal{R}(T)$ and $\mathcal{R}$ of system (4). Here we assume A is anti-stable, (A, B) is controllable, and $m = 1$. Since B is now a column vector, we rename it as b for convenience.

Definition 3 *A control v is said to be extremal on $[0,T]$ if the response $z(t)$ of system (4) lies on $\partial\mathcal{R}(t)$ for all $t \in [0,T]$.*

Lemma 1 *Let $z_T \in \partial\mathcal{R}(T)$, and v be a control that steers the state from the origin to z_T at time T, then v is extremal on $[0,T]$.*

Lemma 2 ([6, p. 62]) *A control v is extremal on $[0,T]$ for system (4) if and only if there is a nonzero vector $c \in \mathbf{R}^n$ such that $v(t) = \text{sgn}(c'e^{At}b)$ for $t \in [0,T]$.*

From Lemma 2, the set of extremal controls can be written as:

$$\mathcal{E}(T) := \{v(t) = \text{sgn}(c'e^{At}b), t \in [0,T] : \|c\|_2 = 1.\}.$$

Combining Lemmas 1 and 2 gives

$$\partial\mathcal{R}(T) = \left\{-\int_0^T e^{-A(T-\tau)}b\,\text{sgn}(c'e^{A\tau}b)\mathrm{d}\tau : \|c\|_2 = 1\right\}.$$

This shows that $\partial\mathcal{R}(T)$ can be determined from the surface of a unit ball in $\mathbf{R}^n$. In the following two sections, we will further simplify this for some special cases.

3.2 Systems with only real eigenvalues

It follows from e.g. [6], that if A has only real eigenvalues and $c \neq 0$, then $c'e^{At}b$ has at most $n-1$ zeros. This implies that an extremal control can have at most $n-1$ switches. We will show that the converse is also true. That is, for system (4), any bang-bang control with $n-1$ or less switches is an extremal control. In this way, $\partial\mathcal{R}(T)$ and $\partial\mathcal{R}$ can be described in a simple manner.

Lemma 3 : *For system (4), assume that A has only real eigenvalues, then*

(a) *an extremal control has at most $n-1$ switches;*

(b) *any bang-bang control with $n-1$ or less switches is an extremal control.*

By Lemma 3, the set of extremal controls on $[0,T]$ can be described as follows:

$$\mathcal{E}(T) = \left\{\pm v : v(t) = \begin{cases} 1, & 0 \le t < t_1 \\ (-1)^i, & t_i \le t < t_{i+1} \\ (-1)^{n-1}, & t_{n-1} \le t \le T \end{cases},\right. \quad \left. 0 \le t_1 \le t_2 \le \cdots \le t_{n-1} \le T \right\}.$$

Here we allow $t_i = t_{i+1}$ and $t_{n-1} = T$. Hence $\mathcal{E}(T)$ consists of all bang-bang controls on $[0,T]$ with $n-1$ or less switches.

For $v \in \mathcal{E}(T)$, with switches $0 \le t_1 \le t_2 \le \cdots \le t_{n-1} \le T$, the state of system (4) at T is,

$$z(T) = \pm\left[e^{-AT} + 2\sum_{i=1}^{n-1}(-1)^i e^{-A(T-t_i)} + (-1)^n I\right]A^{-1}b.$$

This shows that

$$\partial\mathcal{R}(T) = \left\{\pm\left[e^{-AT} + 2\sum_{i=1}^{n-1}(-1)^i e^{-A(T-t_i)} + (-1)^n I\right]A^{-1}b : \right. \quad \left. 0 \le t_1 \le t_2 \le \cdots \le t_{n-1} \le T \right\}.$$

If we let $\tau_i = T - t_i, i = 1, \cdots, n-1$, then $\partial\mathcal{R}(T)$ can be rewritten as

$$\partial\mathcal{R}(T) = \left\{\pm\left[e^{-AT} + 2\sum_{i=1}^{n-1}(-1)^i e^{-A\tau_i} + (-1)^n I\right]A^{-1}b : \right. \quad \left. T \ge \tau_1 \ge \tau_2 \ge \cdots \ge \tau_{n-1} \ge 0 \right\}.$$

By letting T go to infinity, we get the following theorem.

Theorem 1

$$\partial\mathcal{C} = \partial\mathcal{R} = \left\{\pm\left[2\sum_{i=1}^{n-1}(-1)^i e^{-A\tau_i} + (-1)^n I\right] A^{-1}b : \right. \quad \left. \infty \ge \tau_1 \ge \cdots \ge \tau_{n-1} \ge 0 \right\}.$$

By using this theorem, the boundary of the controllable region of system (1) can be easily plotted at least for the low dimensional cases. For example, when $n = 2$, we have

$$\partial\mathcal{C} = \left\{\pm(-2e^{-A\tau} + I)A^{-1}b : 0 \le \tau \le \infty\right\},$$

and when $n = 3$, we have

$$\partial\mathcal{C} = \left\{\pm(-2e^{-A\tau_1} + 2e^{-A\tau_2} - I)A^{-1}b : 0 \le \tau_2 \le \tau_1 \le \infty\right\}.$$

Now we give another interpretation of $\partial\mathcal{C}$ and $\partial\mathcal{R}$ which will be used later. Let $z_e^+ = -A^{-1}b$ and $z_e^- = A^{-1}b$. Clearly z_e^+ is the equilibrium point of system (4) with control $v(t) \equiv 1$ and z_e^- with $v(t) \equiv -1$. Assume that the system has initial state z_e^+ and the following bang-bang control is applied:

$$v(t) = \begin{cases} -1 & , \quad 0 \le t < t_1 \\ (-1)^{i+1} & , \quad t_i \le t < t_{i+1} \\ (-1)^{n-1} & , \quad t_{n-2} \le t \le T \end{cases}$$

Then it is easy to verify that the state at T is

$$z(T) = \left[2\sum_{i=1}^{n-1}(-1)^i e^{-A(T-t_{i-1})} + (-1)^n I\right] A^{-1}b.$$

where $t_0 = 0$.

Clearly $z(T) \in \partial\mathcal{C}$ from Theorem 1. On the other hand, any element of $\partial\mathcal{C}$ can be reached by applying a bang-bang control with $n-2$ or less switches to system (4) from the initial state z_e^+ or z_e^-. Thus we see that that $\partial\mathcal{C} = \partial\mathcal{R}$ has two branches. The first branch consists of trajectories of (4) when the initial state is z_e^+ and the input is a bang-bang control with $n-2$ or less switchings. The second branch consists of the trajectories of (4) when the initial state is z_e^- and the input is a bang-bang control with $n-2$ or less switchings. In particular, if $n = 2$, then $\partial\mathcal{C} = \partial\mathcal{R}$ can be formed by the trajectory of (4) starting from z_e^+ with $v(t) = -1$ and the trajectory from z_e^- with $v(t) = 1$. The first trajectory approaches z_e^- as $t \to \infty$ and the second trajectory approaches z_e^+ as $t \to \infty$, so the two trajectories form a closed curve. If $n = 3$, then one half of $\partial\mathcal{C} = \partial\mathcal{R}$ can be formed by the trajectories of (4) starting from z_e^+ with the first control being -1 and one swicth at any time to 1. So the trajectories go toward z_e^- first and then turn back toward z_e^+. The other half is just symmetric to the first half. See the example in section 3.4.

3.3 Second-order systems with complex eigenvalues

In this subsection, we consider the special case when A is 2×2 with a pair of complex antistable eigenvalues. In this case, the original system (1) can be assumed, without loss of generality, to have the following form

$$\dot{x}(t) = Ax(t) + bu(t) = \begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} x(t) + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u(t) \quad (5)$$

and the corresponding time-reversed system

$$\dot{z}(t) = -\begin{bmatrix} \alpha & -\beta \\ \beta & \alpha \end{bmatrix} z(t) - \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} v(t). \quad (6)$$

It can be shown that

$$\{c'e^{At}b : \|c\|_2 = 1\} = \left\{\sqrt{b_1^2 + b_2^2}\sin(\beta t + \theta)e^{\alpha t} : \theta \in [0, 2\pi]\right\}.$$

Hence the set of extremal controls is

$$\mathcal{E}(T) = \{v(t) = \mathrm{sgn}[\sin(\beta t + \theta)], t \in [0, T] : \ \theta \in [0, 2\pi]\}.$$

Therefore, we get

$$\partial\mathcal{R}(T) = \left\{-\int_0^T e^{-A(T-\tau)}b\,\mathrm{sgn}[\sin(\beta\tau + \theta)]\mathrm{d}\tau : \theta \in [0, 2\pi]\right\}.$$

In the following, we show that $\partial\mathcal{R}(T)$ approaches the steady state phase plot of (6) with a square wave input, a limit cycle Γ, as $T \to \infty$.

Denote $T_p = \frac{\pi}{\beta}$,

$$z_s^+ = -(I + e^{-AT_p})^{-1}(I - e^{-AT_p})A^{-1}b, \ z_s^- = -z_s^+ \quad (7)$$

and

$$\Gamma^+ = \{e^{-At}z_s^- - (I - e^{-At})A^{-1}b : \ t \in [0, T_p]\} \quad (8)$$

$$\Gamma^- = \{e^{-At}z_s^+ + (I - e^{-At})A^{-1}b : \ t \in [0, T_p]\} \quad (9)$$

Let $v^*(t) = \mathrm{sgn}[\sin(\beta t)]$, then $v^*(\cdot)$ is a bang-bang control with the length of each switch being T_p and the first control being 1. In other words, $v^*(\cdot)$ is a $2T_p$ periodic square wave starting with 1. Denote $z^*(t)$ as the time response of (6) with initial value z_s^- and control $v^*(\cdot)$, that is

$$z^*(t) = e^{-At}z_s^- - \int_0^t e^{-A(t-\tau)}b\,\mathrm{sgn}[\sin(\beta\tau)]\mathrm{d}\tau.$$

Then it is easy to verify that $z^*(t)$ starts from z_s^-, goes along Γ^+, reaches z_s^+ at $t = T_p$, then goes along Γ^- and returns z_s^- at $t = 2T_p$. This process is repeated with periodic $2T_p$. This shows that $\Gamma = \Gamma^+ \cup \Gamma^-$ forms a closed curve. To be exact,

$$\Gamma = \{z^*(T + t) : t \in [0, 2T_p]\} = \left\{z^*(T + \frac{\theta}{\beta}) : \theta \in [0, 2\pi]\right\}$$

for any $T \geq 0$.

Since the zero input response of system (6) is exponentially stable, it follows that for any initial state, under control $v(t) = \mathrm{sgn}[\sin(\beta t)]$, the time response $z(t)$ approaches $z^*(t)$ as $t \to \infty$. Hence Γ is in fact the periodic orbit generated by applying a periodic bang-bang control with period $2T_p$ to system (6).

Theorem 2 : $\partial\mathcal{C} = \partial\mathcal{R} = \Gamma$.

3.4 An example

A third-order system is described by (1) with

$$A = \begin{bmatrix} 0.2 & 1 & 0 \\ 0 & 0.2 & 0 \\ 0 & 0 & 0.4 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$z_e^+ = -A^{-1}B = \begin{bmatrix} 20 \\ -5 \\ -2 \end{bmatrix}$, $z_e^- = -z_e^+$. In Figure 1, the solid curves are trajectories of the time-reversed system starting from z_e^+ by applying a bang-bang control with one switch and the first control is negative. The trajectories leave z_e^+ and go toward z_e^- at first, then turn back toward z_e^+. The dashed curves are symmetric to the solid curves. The boundary of the controllable region, $\partial\mathcal{C}$ is formed by these trajectories.

The controllable regions for some second-order systems are plotted in the next section as a comparation to the stability regions.

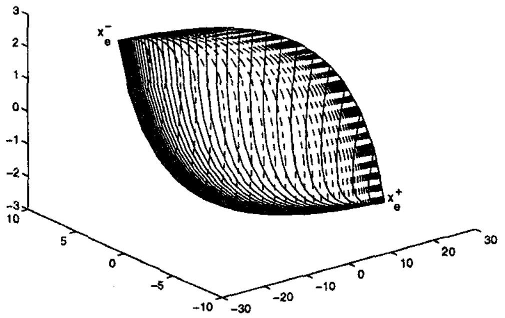

Figure 1: $\partial\mathcal{C}$ of a third-order system

4 Stabilization with saturated linear feedbacks

For an LTI plant controlled by a linear state feedback, local stability implies global stability. However, this is usually not the case in the presence of input saturation. Consider the open loop system

$$\dot{x}(t) = Ax(t) + bu(t) \tag{10}$$

with admissible control $|u(t)| \le 1$. A saturated linear state feedback is given by $u = \sigma(fx)$, where $f \in \mathbf{R}^{1\times n}$ is the feedback gain and $\sigma(\cdot)$ is the saturation function

$$\sigma(s) = \begin{cases} 1 & , \quad s \ge 1 \\ s & , \quad |s| < 1 \\ -1 & , \quad s \le -1. \end{cases}$$

Such a feedback is said to be stabilizing if $A + bf$ is stable. With a saturated linear state feedback applied, the closed loop system is

$$\dot{x}(t) = Ax(t) + b\sigma[fx(t)]. \tag{11}$$

Denote the state transition map of (11) by $\phi : (t, x_0) \mapsto x(t)$. The asymptotic stability region (domain of attraction) $\mathcal{S}$ of (11) is defined by

$$\mathcal{S} = \left\{ x_0 \in \mathbf{R}^n : \lim_{t\to\infty} \phi(t, x_0) = 0 \right\}.$$

Obviously, $\mathcal{S}$ must lie within the controllable region $\mathcal{C}$ of system (10). Therefore, a design problem is to choose the state feedback gain so that $\mathcal{S}$ is close to $\mathcal{C}$.

This seemingly simple task is actually quite non-trivial, even for semistable systems. In the past few years, extensive research has been reported on the stabilization of semistable plant using saturated linear feedback, e.g. [4, 5, 9, 13]. The problem for general unstable systems is much harder. In this section, we will only deal with antistable planar systems as a starting point.

The study in this section is not quite conclusive. We will present two interesting conjectures and some supporting arguments.

4.1 The asymptotic stability region under a given feedback

Consider system (10). Assume that $A \in \mathbf{R}^{2\times 2}$ and A is antistable. In this section, we analyze the asymptotic stability region of (11). In [1], it was shown that the boundary of $\mathcal{S}$, denoted by $\partial\mathcal{S}$, is a closed orbit, but no method to find this closed orbit is provided. Generally, only a subset of $\mathcal{S}$ lying between $fx = 1$ and $fx = -1$ is detected from some Lyapunov function, see [3] for example. Here, we will show that $\partial\mathcal{S}$ is a limit cycle of (11). This limit cycle can be easily detected from the time reversed-system of (11).

Consider the time-reversed system of (11):

$$\dot{z}(t) = -Az(t) - b\sigma[fz(t)] \tag{12}$$

Since (12) has only one equilibrium point [1], all the limit cycles of (12) are totally ordered by enclosement. Let P be a positive matrix such that $(A + bf)'P + P(A + bf)$ is negative definite and since $\{z \in \mathbf{R}^2 : -1 < fz < 1\}$ is an open neighborhood of the origin, it must contain

$$\mathcal{Q}_0 := \{z \in \mathbf{R}^2 : z'Pz \le r_0\}$$

for some $r_0 > 0$. Denote the state transition map of (12) by $\psi : (t, z_0) \mapsto z(t)$.

Theorem 3 *Consider system (12).*

(a) *For every $z_0 \ne 0$, $\psi(t, z_0)$ converges to a limit cycle as $t \to \infty$.*

(b) *There exists a smallest limit cycle Γ_m and a largest limit cycle Γ_M with $\Gamma_m \subset \mathbf{R}^2 \setminus \mathcal{Q}_0$ and $\Gamma_M \subset \mathcal{C}$.*

This theorem shows that Γ_m is stable from inside and Γ_M is stable from outside. The original system (11) has the same limit cycles as the time-reversed system (12). For (11), however, Γ_m is unstable from inside and Γ_M is unstable from outside. Because of Theorem 3 (a), we have $\partial\mathcal{S} = \Gamma_m$.

Of all the examples that have been tested, it holds that $\Gamma_M = \Gamma_m$. This leads to the following conjecture.

Conjecture 1 *Each of system (11) and system (12) has only one limit cycle.*

If this conjecture is true, then this limit cycle is equal to the boundary $\partial\mathcal{S}$ of the asymptotic stability region of (11). However, the limit cycle is an unstable one of system (11). It is difficult to determine it from system (11) directly. On the contrary, it is very easy to determine this limit cycle from the time-reversed system (12) since now the limit cycle becomes a stable one. The method is as follows: choose any $z_0 \ne 0$, then the trajectory $\psi(t, z_0)$ converges to $\partial\mathcal{S}$ as $t \to \infty$. See Figure 2, where the solid curve and the dashed curve are generated from different initial states. The straight lines are $fz = 1$ and $fz = -1$.

We are unable to prove Conjecture 1 at the present time. However, the describing function analysis suggests that it is true.

4.2 On enlarging the stability region

It is obvious that the stability region of (11) must lie within the controllable region of (10), i.e., $\mathcal{S} \subset \mathcal{C}$. By

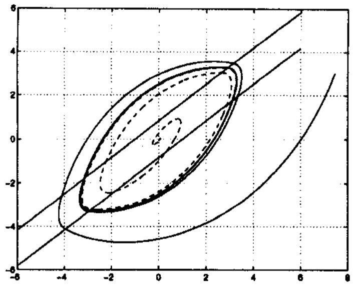

Figure 2: Determine $\partial\mathcal{S}$ from the limit cycle

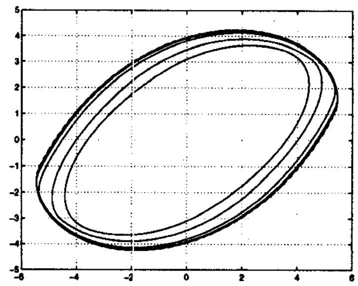

Figure 3: Stability regions under different feedbacks

comparing their boundary behavior, we can show that to make $\partial\mathcal{S}$ close to $\partial\mathcal{C}$, the two straight lines $fz = 1$ and $fz = -1$ must be close to each other and be parallel to the line between z_e^+ and z_e^-(or z_s^+ and z_s^-). All f such that $fz = 1$ is parallel to the line between z_e^+ and z_e^-(or z_s^+ and z_s^-) can be written as

$$f = kb'A'^{-1}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}, \quad k \neq 0$$

As $|k| \to \infty$, the distance between $fz = 1$ and $fz = -1$ will approach to zero.

Let $f_0 = -b'A'^{-1}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$. It's now clear that to make $\partial\mathcal{S}$ close to $\partial\mathcal{C}$, f must be chosen as $f = kf_0$ with large $|k|$. Now the question becomes if sufficiently large k can be chosen so that $A + kbf_0$ is stable.

Claim 1 *There exists $k_0 > 0$, such that $A + kbf_0$ is stable for all $k > k_0$ or $k < -k_0$.*

The above argument show that if we let $f = kb'A'^{-1}\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ and set k be arbitrarily large, then it is possible that $\mathcal{S}$ will be arbitrarily close to $\mathcal{C}$. This leads to the following conjecture.

Conjecture 2 *For every $\epsilon > 0$, there exists f such that $d(\mathcal{S}, \mathcal{C}) < \epsilon$.*

In the rest of this subsection, we illustrate the idea using one example.

Example 1: Let $A = \begin{bmatrix} 0.6 & -0.8 \\ 0.8 & 0.6 \end{bmatrix}$ and $b = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$. Then $f_0 = \begin{bmatrix} -0.8 & 4.4 \end{bmatrix}$. In Figure 3, the boundaries of the asymptotic stability regions corresponding to different $f = -kf_0$, $k = 0.08, 0.1, 0.2, 0.4, 0.8$ are plotted from the inner to the outer. The region do become biger for greater k. The outmost dashed curve is $\partial\mathcal{C}$. When $k = 0.8$, it can be seen that $\partial\mathcal{S}$ is very close to $\partial\mathcal{C}$. This shows that larger k generates larger asymptotic stability regions, which is quite different from the view that smaller feedback tends to generate larger stability region.

References

[1] J. Alvarez, R. Suarez and J. Alvarez,"Planar linear systems with single saturated feedback", *Systems & Control Letters*, vol. 20, pp. 319-326, 1993.

[2] O. Hájek, *Control Theory in the Plane*, Springer-Verlag, 1991

[3] P.-O. Gutman and P. Hagander, "A new design of constrained controllers for linear systems", *IEEE Trans. Automat. Contr.*, vol. AC-30, pp. 22-33, 1985.

[4] Z. Lin and A. Saberi, "Semi-global exponential stabilization of linear systems subject to 'input saturation' via linear feedbacks", *Systems & Control Letters*, vol. 21, pp. 225-239, 1993.

[5] Z. Lin and A. Saberi, "A semi-global low-and-high design for linear systems with input saturation: stabilization and disturbance rejection", *Int. J. of Robust and Nonlinear Control*, vol. 5, pp. 381-398, 1995.

[6] J. Macki and M. Strauss, *Introduction to Optimal Control*, Springer-Verlag, 1982.

[7] A. Megretski,"L_2 BIBO output feedback stabilization with saturated control", *13th Triennial IFAC World Congress*, vol. D, pp. 435-440, 1996.

[8] L. Pandolfi, "Linear control systems: controllability with constrained controls", *J. Optimiz. Theory Appl.*, vol. 19, pp. 577-585, 1976.

[9] A. Saberi, Z. Lin and A. R. Teel, "Control of linear systems with saturating actuators", *IEEE Trans. Automat. Contr.* vol. AC-41, pp. 368-378, 1996.

[10] W. E. Schmitendorf and B. R. Barmish, "Null controllability of linear systems with constrained controls", *SIAM J. Control and Optimization*, vol. 18, pp. 327-345, 1980.

[11] E. D. Sontag, "An algebraic approach to bounded controllability of linear systems", *Int. J. Control*, vol. 39, pp. 181-188, 1984.

[12] H. J. Sussmann, E. D. Sontag, and Y. Yang, "A general result on the stabilization of linear systems using bounded controls", *IEEE Trans. Automat. Contr.*, vol. AC-39, pp. 2411-2425, 1994.

[13] A. R. Teel, "Global stabilization and restricted tracking for multiple integrators with bounded controls", *System & Control Letters*, vol. 18, pp. 165-171, 1992.