GLOBAL STABILIZATION OF LINEAR SYSTEMS
WITH BOUNDED FEEDBACK

BY YUDI YANG

A dissertation submitted to the
Graduate School—New Brunswick
Rutgers, The State University of New Jersey
in conjunction with
The Graduate School of Biomedical Sciences
the University of Medicine and Dentistry of New Jersey
in partial fulfillment of the requirements for the
Joint Degree of Doctor of Philosophy
Graduate Program in Mathematics

Written under the direction of
Héctor J. Sussmann

and approved by

_________________________________________

_________________________________________

_________________________________________

_________________________________________

New Brunswick, New Jersey
October, 1993
ABSTRACT OF THE DISSERTATION

Global Stabilization of Linear Systems
with Bounded Feedback

by Yudi Yang, Ph.D.

Dissertation Director: Héctor J. Sussmann

There are two parts in the thesis. The first part deals with the problem of global stabilization of linear systems with bounded feedback. A linear system can be globally stabilized by a bounded feedback if and only if all eigenvalues of the system matrix have nonpositive real parts and all eigenvalues of the uncontrollable part have strictly negative real parts. If a linear system satisfies this condition, then we can provide an algorithm to find a bounded stabilizing feedback. The design employs linear combinations and compositions of linear functions and saturations. We also show that if we use only a saturated linear feedback, then a linear system cannot be globally stabilized in general. Some applications, such as output feedback stabilization, the stabilization of cascade systems, and the stabilization of flight control, are also presented here.

The second part deals with questions of global stabilizability of nonlinear systems. Based on the use of control-Lyapunov functions, we obtain a class of stabilizing feedback laws. A sufficient condition for such feedbacks to be continuously differentiable is presented. We then apply this condition to a wide class of two- and three-dimensional systems, extending some recent results on stabilization.
Acknowledgements

During my research, I received much help from my advisor, Professor Héctor J. Sussmann. Without his help, this thesis could not have been finished. I cannot sufficiently thank him for his invaluable help during the past years.

I also had many discussions with Professor Eduardo D. Sontag, and received constant suggestions from him during the research. I would like to thank him, especially for his introductory courses on systems and control theory.

I would like to thank the US Air Force and NSF for partially supporting my graduate study under grants AFOSR-91-0343, AFOSR-91-0346, DMS-8902994, and DMS92-02554.

A special thanks goes to my wife for her encouragement and understanding in the past three years.
Dedication

This thesis is dedicated to my parents in China.
Table of Contents

Abstract ................................................................. ii
Acknowledgements ....................................................... iii
Dedication ................................................................. iv

1. Introduction .......................................................... 1
   1.1. Part One .......................................................... 1
   1.2. Part Two .......................................................... 5
   1.3. The Organization ................................................ 7

2. Existence Results .................................................... 9
   2.1. Preliminary Definitions and Results ......................... 9
   2.2. Some Properties of Analytic Functions ..................... 12
   2.3. The Proof of Theorem 2.1 .................................... 17

3. The Naive Design .................................................... 31
   3.1. Some Special Cases ............................................ 31
   3.2. A Negative Result .............................................. 33
   3.3. The proof of Theorem 3.1 .................................... 34

4. A Design Using Hidden Layers ..................................... 48
   4.1. Constructions of Feedback .................................... 48
   4.2. Technical lemmas .............................................. 51
   4.3. The Proof of Theorem 4.1 .................................... 71
   4.4. An Algorithm .................................................. 79
   4.5. Multiple Integrators .......................................... 87
5. Applications ......................................................... 91
   5.1. Output Feedback ............................................. 91
   5.2. Cascaded Systems ........................................... 93
   5.3. F-8 Longitudinal Flight Control ......................... 95
      5.3.1. The Model .............................................. 96
      5.3.2. Trim Conditions ....................................... 98
      5.3.3. Controller Design .................................... 99
      5.3.4. Stability of the Linear System ...................... 101
      5.3.5. Simulation Results ................................... 106

6. Discrete-Time Linear Systems ................................. 112
   6.1. Statement of the Main Results ........................... 112
   6.2. Preliminaries .............................................. 114
   6.3. The Proof of Theorem 6.1 ................................. 121
   6.4. The Proof of Theorem 6.2 ................................. 125

7. Nonlinear Affine Systems ...................................... 126
   7.1. General Theory ............................................ 126
   7.2. Desingularizing Functions ............................... 133
   7.3. Applications to Planar Systems ......................... 136
   7.4. Applications to Three-dimensional Systems .......... 146

References .......................................................... 151

Vita ................................................................. 154
Chapter 1

Introduction

There are two parts in the thesis. In the first part we completely solve the problem on global stabilization of linear systems with bounded feedback. In the second part we present some results on global stabilization of nonlinear affine systems with continuously differentiable feedback.

1.1 Part One

We consider linear time-invariant continuous-time systems

\[
\Sigma: \quad \dot{x} = Ax + Bu,
\]

(1.1.1)

where \( A \in \mathbb{R}^{n \times n} \), \( B \in \mathbb{R}^{n \times m} \) for some integers \( n \) (the dimension of the system) and \( m \) (the number of inputs), and for which the control values \( u(t) \) are restricted to lie in a bounded set \( U \subseteq \mathbb{R}^m \). We assume that \( U \) contains zero in its interior.

The study of such systems is motivated by the possibility of actuator saturation or constraints on actuators, reflected sometimes also in bounds on available power supplies or rate limits. These systems cannot be naturally dealt with within the context of standard (algebraic) linear control theory but are ubiquitous in control applications. To quote the recent textbook [23] (page 171): "saturation is probably the most commonly encountered nonlinearity in control engineering." Mathematically, control questions become nontrivial, since only control values in \( U \) are allowed into the underlying linear system.

We will present results on global stabilization of systems of this form. Precisely, we want to find a bounded vector-valued function \( k : \mathbb{R}^n \to \mathbb{R}^n \) such that the closed-loop system of (1.1.1) with the feedback \( u = k(x) \) has the origin as a global asymptotically
stable equilibrium. Of course, this cannot always be achieved for general linear systems, even for controllable ones. For instance, if the feedback is required to be bounded by 1, then no trajectory of the one-dimensional open-loop system \( \dot{x} = x + u \) starting in the region \( \{ x : |x| > 1 \} \) can approach zero. So, we need to determine conditions under which the system (1.1.1) is globally asymptotically stabilizable.

The theory of controllability of linear systems with bounded controls is a well-studied topic; see e.g. the fundamental paper [21], as well as the different, more algebraic approach discussed in [24]. In these two references it was proved that such asymptotic null-controllability of the system (that is, the existence of open-loop controls that steer each state to the origin, in the limit as \( t \to \infty \)) is equivalent to the following pair of algebraic conditions:

(a) all eigenvalues of \( A \) have nonpositive real part, and

(b) all eigenvalues of the uncontrollable part of \( \Sigma \) have strictly negative real parts

(that is, the pair \( (A, B) \) is stabilizable in the ordinary sense).

Note that under Conditions (a) and (b) there may very well be nontrivial Jordan blocks corresponding to critical eigenvalues, so that the system \( \dot{x} = Ax \) need not be asymptotically stable, or even Lyapunov-stable. This is what makes the problem interesting and allows inclusion of examples of practical importance such as systems involving integrators.

The first work on global stabilization of linear systems with bounded feedback appeared in [28]. In that paper it was shown that the above two conditions are equivalent to the global stabilizability of (1.1.1) with bounded smooth feedback. Using the same technique as in [28], we can even show that under Conditions (a) and (b), there exists a bounded analytic feedback that stabilizes (1.1.1). The proof is by induction. The induction starts with presenting an explicit stabilizing feedback for diagonalizable \( A \) of (1.1.1). Then, if \( A \) is not diagonalizable, we obtain the system matrix by adding Jordan blocks one by one to a diagonalizable system matrix without changing its eigenvalues and use the inductive hypothesis each time we add a Jordan block. Since the resulting feedback relies on many submanifolds defined by integrals corresponding to some system
trajectories, this approach is not constructive.

In very special cases, including all one- and two-dimensional systems, stabilization is possible by simply using a saturated linear feedback law of the type:

\[ u = \sigma(Fx), \quad (1.1.2) \]

where \( F \) is an \( m \times n \) matrix and \( \sigma \) is a function that computes a saturation in each coordinate of the vector \( Fx \), for instance, \( u_i = \tanh((Fx)_i) \) or \( u_i = \text{sat}((Fx)_i) \) where,

\[ \text{sat}(s) = \text{sign}(s) \min\{|s|, 1\}. \quad (1.1.3) \]

As stated above, this is also possible in the case of systems for which the Jordan form of \( A \) has no off-diagonal ones, i.e., neutrally stable systems. (See [15] and [22].) Thus it is natural to ask if such simple control laws as that of Equation (1.1.2) can also be used for more general systems. This was negatively answered in a paper by A.T. Fuller as far back as the late 1960’s. He showed in [13] that already for triple integrators such saturated linear feedback is not sufficient, provided that certain assumptions are satisfied by the saturation \( \sigma \). In [31] we gave an independent proof of such a negative result, which applies to basically arbitrary \( \sigma \)’s. Our result shows that if \( \sigma : \mathbb{R} \rightarrow \mathbb{R} \) is a locally Lipschitz function for which both limits \( \lim_{s \to \pm \infty} \sigma(s) \) exist and are nonzero, then there is no stabilizing feedback of the form (1.1.2) for \( n \)-dimensional integrators with \( n \geq 3 \).

The fact that linear feedback laws when saturated can lead to instability has motivated a large amount of research. See for instance [17] and [18], and references there, for estimates of the size of the regions of attraction that result when using linear saturated controllers. Rather than working with linear saturated control laws \( u = \sigma(Fx) \) and trying to show that they are globally stabilizing, or trying to estimate their domains of attraction, we allow more general bounded (and hence necessarily nonlinear) laws. Since saturated linear feedbacks suffice for up to two dimensions, in higher dimensions it is natural to look for other simple control rules, for instance, employing linear combinations and compositions of saturation nonlinearities. In the language of neural networks, one wants control laws that are implementable by feedforward nets using “hidden layers” rather than the “perceptrons” represented by (1.1.2). Motivated in part by [28]
and [31], Andrew Teel showed in [32] how, in the particular case of single-input multiple integrators, such combinations of saturations are indeed sufficient to obtain stabilizing feedback controllers. In [30] we obtained a general solution of the same type, for the full case treated in [28]. The approach is explicit and constructive. Our solution is inspired by the techniques introduced in [32] for the particular case treated there, but the details are far more complicated, due to the possibilities of having both multiple inputs and (perhaps multiple) purely imaginary eigenvalues and to the need to deal with arbitrary saturations.

Our result was first announced in [29] and [40], where we considered a very special type of feedback for which the saturations are exactly linear near 0. When a system has a pure imaginary eigenvalue, a saturation with three different slopes may be needed. Later, in [30], we extended the result by allowing essentially arbitrary saturations. The only conditions imposed on the saturation functions $\sigma$ are that

(i) $\sigma$ is locally Lipschitz,

(ii) $s\sigma(s) > 0$ whenever $s \neq 0$,

(iii) $\sigma$ is differentiable at 0 and $\sigma'(0) > 0$, and

(iv) $\liminf_{|s| \to \infty} |\sigma(s)| > 0$.

So, mathematically, the results in [30] show that one can use analytic functions to implement feedback laws (the results in [40] and [32] would not give this) and, from an engineering point of view, they insure that rather general components could be employed, subject only to mild conditions which are robustly satisfied. In the terminology of current "artificial neural networks" technology, our results allow the implementation of feedback controllers using very general types of activation (neuron characteristic) functions.

In addition to the design that employs linear combinations and compositions of saturation nonlinearities as constructed in [32], we also exhibit a different design that uses linear combinations of saturated linear functions. (In neural network terms, it involves a "single hidden layer net.") So, from the above point of view, we have completely
solved the problem of global stabilization of linear systems with bounded feedback on both the theoretical level and the algorithmic level. As applications of our results, we studied the output stabilization problem and the stabilization of cascaded systems. Those results were already shown in [40], [30]. In particular, following the general philosophy outlined in [29, 40, 30], we developed in detail in [39] an explicit design for the linearized equations of longitudinal flight control for an F-8 aircraft and tested the resulting controller—via simulations—on the original nonlinear model. In the process of working out this example, we were able to obtain, in certain particular cases, bounds tighter than those of [40, 30]. With these improved bounds, better performance can be achieved.

1.2 Part Two

The purpose of this part is to seek continuously differentiable feedbacks for nonlinear affine systems of the type

$$\Sigma: \dot{x} = f(x) + g(x)u,$$

where $x \in \mathbb{R}^n$, $u$ is a scalar input, and $f, g \in C^1(\mathbb{R}^n)$, such that the origin is a globally asymptotically stable equilibrium of the resulting closed-loop systems. There are many articles dealing with the stabilization problem for $\Sigma$. (See the references listed below.) Particularly, in [1] Artstein proved that if there exists a control-Lyapunov function $V$ of a nonlinear affine system, then there is a feedback that stabilizes the system. Sontag, in [26], improved upon this result by giving a simple explicit formula for such a feedback in terms of directional derivatives of $V$. This feedback has the same regularity as $L_f V$ and $L_g V$ outside the origin but generally fails to be continuously differentiable at the origin. The interest in Part Two is in studying the existence of a stabilizing feedback which is continuously differentiable at the origin.

We first show that the feedback law given in [26] can be replaced by a class of functions expressed in terms of $L_f V$ and $L_g V$, and that when $L_f V$ and $L_g V$ are homogeneous with respect to a one-parameter group of dilations, some of these feedbacks are continuously differentiable at the origin. We then present some applications to the
stabilization of two- and three-dimensional systems of the type

\[
\begin{align*}
\dot{x} &= f(x, y), \\
\dot{y} &= u.
\end{align*}
\]

A well-known result on the stabilization of this composite system is based on the existence of a smooth or differentiable stabilizing feedback for \( \dot{x} = f(x, u) \). But, our results show that even if the system \( \dot{x} = f(x, u) \) cannot be stabilized by a continuously differentiable feedback, it may still be possible to stabilize the composite system \( \dot{x} = f(x, y) \), \( \dot{y} = u \) by a continuously differentiable feedback. The approach pursued here relies on a "desingularizing function" as defined in [20] whose use will be illustrated below. Since we use a constructive proof, such a feedback can be exactly computed.

Our general approach is applied to particular classes of two- and three-dimensional systems. For example, we consider two-dimensional polynomial systems of the type

\[
\begin{align*}
\dot{x} &= \sum_{i \geq 0, m_0 - j, \geq 0} c_{0j} x^{m_0 - j} y^{n_0 + j}, \\
\dot{y} &= u,
\end{align*}
\] (1.2.1)

where \( i > 0, j > 0, m_0 \geq 2, n_0 \geq 0, \) and \( c_0 \neq 0 \) (e.g., \( \dot{x} = x^2 - 3xy^2 + y^4, \dot{y} = u \)). Our result in this case is that (1.2.1) can be globally asymptotically stabilized by a continuously differentiable feedback if and only if there exist \((x_1, y_1)\) and \((x_2, y_2)\) with \( x_1 > 0, x_2 < 0 \) such that \( f(x_1, y_1) < 0 \) and \( f(x_2, y_2) > 0 \), where \( f(x, y) = \sum_{i \geq 0, m_0 - j, \geq 0} c_{0j} x^{m_0 - j} y^{n_0 + j} \). This generalizes the result obtained in [12] for homogeneous planar systems.

We also analyze the stabilizability with continuously differentiable feedback of the system

\[
\begin{align*}
\dot{x} &= (x - y^5)(x^2 + |y|^k), \\
\dot{y} &= u.
\end{align*}
\] (1.2.2)

We show that (1.2.2) is stabilizable with continuously differentiable feedback if \( k > 4 \), and (1.2.2) is not stabilizable with continuously differentiable feedback if \( k < 4 \).

As a third example, we consider three-dimensional polynomial systems of the type

\[
\begin{align*}
\dot{x} &= f(x, y), \\
\dot{y} &= g(x, y) + z^q, \\
\dot{z} &= u,
\end{align*}
\] (1.2.3)
where both $f$ and $g$ are homogeneous polynomials of degree $p$, and $p, q$ are odd numbers.

We show that if $p > 1$, then (1.2.3) can be globally asymptotically stabilized by a continuously differentiable feedback if and only if there exists $(x_1, y_1)$ with $x_1 > 0$ such that $f(x_1, y_1) < 0$. When $p = 1$ there is a similar result on the stabilization with a continuous feedback. When $f(x, y) = x^p$ and $g(x, y) = 0$, the system (1.2.3) reduces

$$\begin{align*}
\dot{x} &= y^p, \\
\dot{y} &= z^q, \\
\dot{z} &= u,
\end{align*}$$

(1.2.4)

where $p, q$ are odd numbers. So, in particular for (1.2.4), we are able to extend the results of [5] and [11], where only the case of $p = q$ was considered.

1.3 The Organization

The thesis is organized as follows. Chapters 2 through 6 discuss the problem of global stabilization of linear systems with bounded feedback. In Chapter 2 we develop some conditions that are equivalent to the global stabilizability of linear systems with bounded feedback using a nonconstructive approach.

In Chapter 3 we present a simple design, based on linear saturated feedback, for some special cases such as double integrators and systems whose system matrices are diagonalable. We also show that the linear saturated feedback design does not work for general linear systems.

In Chapter 4 we exhibit a constructive design and a precise algorithm to find a bounded stabilizing feedback for a given linear system when such a feedback exists. Our analysis is for general linear systems. To study a special system, we can either apply the algorithm to the system or just follow the general philosophy of the approach and thus obtain a simpler feedback. So, we also exhibit a particular design that applies to scalar-input multiple integrators.

In Chapter 5 we present the applications to output feedback stabilization, the stabilization of cascade systems, and F-8 aircraft control.
Finally, in Chapter 6 we discuss the global stabilizability of linear discrete-time systems with bounded feedback. We present a constructive approach for general linear discrete-time systems.

The second part of the thesis is included in Chapter 7, which is totally independent from the first six chapters. Section 7.1 presents the main results of the general theory. We describe a class of stabilizing feedback laws that use a control-Lyapunov function design and give a sufficient condition for feedbacks in the class to be continuously differentiable. In Section 7.2 we introduce desingularizing functions and state a simple, easy to apply version of Lemma 1 of [20]. Finally, in Sections 7.3 and 7.4 we apply the theory of Sections 7.1 and 7.2 to some examples of two- and three-dimensional systems.
Chapter 2

Existence Results

This chapter develops some conditions that are equivalent to the global stabilizability of linear systems with bounded feedback.

2.1 Preliminary Definitions and Results

**Definition 2.1.1** Let $\Sigma$ be a finite-dimensional linear system $\dot{x} = Ax + Bu$, $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$. Let $k : \mathbb{R}^n \to \mathbb{R}^m$ be a locally Lipschitz function. Then we say that $k$ stabilizes $\Sigma$ if $0$ is a globally asymptotically stable (GAS) equilibrium of the closed-loop system $\dot{x} = Ax + Bk(x)$.

We use $\Sigma_k$ to denote the closed-loop system: $\dot{x} = Ax + Bk(x)$. If $k(x) = Kx$ for all $x \in \mathbb{R}^n$, where $K$ is a matrix of an appropriate size, we use $\Sigma_K$ to denote the closed-loop system.

**Definition 2.1.2** A linear system $\Sigma$ is asymptotically controllable (AC) if for every initial condition $\bar{x} \in \mathbb{R}^n$ there exists a bounded measurable open-loop control $u : [0, \infty) \to \mathbb{R}^m$ with the property that the trajectory $t \to x(t)$ of $\Sigma$ that corresponds to $u(\cdot)$ and satisfies $x(0) = \bar{x}$ is such that $\lim_{t \to \infty} x(t) = 0$.

- If there is a fixed constant $C$ such that for every $\bar{x}$ the control $u(\cdot)$ can be chosen in such a way that $\|u(t)\| \leq C$ for all $t \geq 0$, then we say that $\Sigma$ is asymptotically controllable with bounded control (ACBC).

- If in addition $C$ can be chosen arbitrarily small, then we say that $\Sigma$ is asymptotically controllable with small controls (ACSC).  


Definition 2.1.3 Let $\Sigma$ be a linear system. Then $\Sigma$ is:

- **bounded feedback stabilizable (BFS)** if there exists a bounded locally Lipschitz feedback $k$ that stabilizes $\Sigma$.

- **small feedback stabilizable (SFS)** if for every $\varepsilon > 0$ there exists a stabilizing feedback $k$ for $\Sigma$ such that $\|k(x)\| \leq \varepsilon$ for all $x$, i.e. $\|k\|_\infty \leq \varepsilon$. □

Definition 2.1.4 Let $\mathcal{F}$ be any normed space of feedback controls. Let $\Sigma$ be a linear system. Then $\Sigma$ is:

- **$\mathcal{F}$-stabilizable** if there is a stabilizing feedback $k \in \mathcal{F}$ for $\Sigma$.

- **small norm $\mathcal{F}$-stabilizable** if for every $\varepsilon > 0$ there is a stabilizing feedback $k \in \mathcal{F}$ for $\Sigma$ with $\mathcal{F}$-norm $< \varepsilon$. □

If $\mathcal{F}$ is the space of all linear feedbacks $x \to Kx \overset{\text{def}}{=} k_K(x)$, with the norm $\|k_K\| = \|K\|$, then we will use the expressions "linearly stabilizable" and "linearly stabilizable with small norm".

Definition 2.1.5 For an $n$-tuple $\alpha = (\alpha_1, \ldots, \alpha_n)$ of nonnegative integers, define $|\alpha| = \alpha_1 + \cdots + \alpha_n$, and let $D^\alpha$ denote $D_1^{\alpha_1} \cdots D_n^{\alpha_n}$, where $D_i^{\alpha_i} = \frac{\partial^{\alpha_i}}{\partial x_i^{\alpha_i}}$. Let $\mathcal{F}^w_{n,m,r}$ denote the space of all analytic functions $k : \mathbb{R}^n \to \mathbb{R}^m$ such that

$$\|k\|_{C^r} = \sup\{||D^\alpha k(x)|| : x \in \mathbb{R}^n, 0 \leq |\alpha| \leq r\}$$

is finite. Then a linear system is **analytically stabilizable with small $C^r$-norm** if it is small-norm $\mathcal{F}^w_{n,m,r}$-stabilizable. □

For an $n$-dimensional linear system

$$\Sigma : \dot{x} = Ax + Bu, \quad x \in \mathbb{R}^n, \quad u \in \mathbb{R}^m, \quad (2.1.1)$$

$\mathbb{R}^n$ has a direct sum decomposition $\mathbb{R}^n = E_s \oplus E_c \oplus E_u$, where $AE_s \subseteq E_s, AE_c \subseteq E_c, AE_u \subseteq E_u$, such that the eigenvalues of the restrictions $A_s, A_c, A_u$ of $A$ to $E_s, E_c, E_u$,
$E_u$ are contained, respectively, in $\{ s : \Re s < 0 \}$, $\{ s : \Re s = 0 \}$, $\{ s : \Re s > 0 \}$. If we write $x \in \mathbb{R}^n$ as $(x_s, x_c, x_u)$, and $Bu = (B_s u, B_c u, B_u u)$, then the system equations for $\Sigma$ are written in the form

$$\dot{x}_s = A_s x_s + B_s u, \quad x_s \in E_s,$$

$$\dot{x}_c = A_c x_c + B_c u, \quad x_c \in E_c,$$

$$\dot{x}_u = A_u x_u + B_u u, \quad x_u \in E_u.$$

These three systems will be denoted, respectively, by $\Sigma_s$, $\Sigma_c$ and $\Sigma_u$, and referred to as the stable, critical and unstable parts of $\Sigma$. We will write $\Sigma_u = 0$, $\Sigma_s = 0$ if $E_u = \{0\}$ or $E_s = \{0\}$.

**Theorem 2.1** Let $\Sigma$ be a finite-dimensional linear system. Then the following conditions are equivalent:

1. $\Sigma_u = 0$ and $\Sigma_c$ is controllable,

2. $\Sigma$ is ACBC,

3. $\Sigma$ is ACSC,

4. $\Sigma$ is BFS,

5. $\Sigma$ is SFS,

6. $\Sigma$ is analytically stabilizable with small $C^1$-norm,

7. $\Sigma$ is linearly stabilizable with small norm.

We remark that (6) implies that there exists an analytic stabilizing feedback which is globally Lipschitz. This property is needed to deal with output feedback stabilization; see [28]. In Chapter 4 we will give another equivalent condition, namely, $\Sigma$ is analytically stabilizable with small $C^0$-norm, and the resulting closed-loop system has the small-input small-state property. As will be shown in chapter 5, the small-input small-state property makes output stabilization considerably simpler.
2.2 Some Properties of Analytic Functions

The most delicate part of the proof of Theorem 2.1 is the implication (1) \( \Rightarrow \) (6). Before we prove Theorem 2.1, we need some properties of analytic functions.

**Lemma 2.2.1** Let \( f : \mathbb{R}^n \rightarrow \mathbb{R}^n \) and \( g : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n \) be two real analytic functions such that

(i) \( f(0) = 0 \),

(ii) all eigenvalues of the linearization of \( f(x) \) at the origin have negative real parts,

(iii) \( |g(z,t)| \leq C|z| \) for \( z \) in a neighborhood of \( 0 \) in \( \mathbb{R}^n \), where \( C \) is a constant.

Then there exists a neighborhood \( U \) of the origin in \( \mathbb{R}^n \) such that the function \( \xi : U \rightarrow \mathbb{R}^n \) given by

\[
\xi(z) = \int_0^\infty g(\gamma(z,t), t) dt
\]

is well-defined and analytic, where \( t \rightarrow \gamma(z,t) \) denotes the trajectory of \( \dot{\phi} = f(\phi) \) starting at \( z \).

**Proof.** We consider the differential equation \( \dot{z} = f(z) \) for \( z \in \mathbb{C}^n \). For \( z \) near \( 0 \) in \( \mathbb{C}^n \), let \( \xi(z) \) be defined by the expression (2.2.1), where \( t \rightarrow \gamma(z,t) \) is the trajectory of \( \dot{\phi} = f(\phi) \) starting at \( z \). Note that the system \( \dot{z} = f(z) \) is locally asymptotically stable because its linearization at the origin is stable. Therefore, \( \gamma(z,t) \) exists for all \( t \geq 0 \) provided that \( z \) is sufficiently close to \( 0 \). We will show that the integral (2.2.1) converges when \( z \) is near \( 0 \) and \( \xi \) is analytic at the origin. Since the restriction of \( \xi \) to \( \mathbb{R}^n \) is a real-valued function, it follows that \( \xi \) is analytic at the origin.

First, we need to give a precise meaning to the formula for \( \xi(z) \) for \( z \in \mathbb{C}^n \). Note that \( f : \mathbb{R}^n \rightarrow \mathbb{R}^n \) is analytic at \( 0 \). Using the power series of \( f \) at \( 0 \), we can extend \( f \) to a neighborhood \( \mathcal{O}_1 \) of the origin in \( \mathbb{C}^n \). Similarly, since \( g : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n \) is an analytic function, given \( (z,t) \in \mathbb{R}^{n+1} \), we can use the Taylor's series of \( g \) at \( (z,t) \) to extend \( g \) to a neighborhood of \( (z,t) \) in \( \mathbb{C}^{n+1} \). Therefore, we can extend \( g \) to an open set \( \mathcal{O}_2 \) in \( \mathbb{C}^{n+1} \), and \( \mathbb{R}^{n+1} \in \mathcal{O}_2 \). So, if \( t \rightarrow \gamma(z,t) \), \( z \in \mathbb{C}^n \), is a trajectory of \( \dot{\phi} = f(\phi) \) such that
γ(z,t) ∈ O₁ and (γ(z,t), t) ∈ O₂ for all t ≥ 0 and \( \int_0^\infty g(\gamma(x,t), t)dt \) converges, then \( \tilde{\xi}(z) \) is well-defined.

Let A denote the linearization matrix of \( f(x) \) at the origin. From the assumption, we know that all eigenvalues of A have negative real parts. Therefore there exist two positive constants \( K \) and \( \sigma \) such that \( |e^{At}| ≤ Ke^{-\sigma t} \) for \( t ≥ 0 \). Write \( f(z) = Az + h(z) \) for \( z \) near 0 in \( \mathbb{C}^n \). Then \( h \) is analytic at the origin in \( \mathbb{C}^n \), and \( h(z) = O(|z|^2) \) for \( z \) near the origin. So there is a \( \delta > 0 \), such that \( |h(z)| ≤ \frac{\sigma}{2K}|z| \) for \( |z| < \delta \).

Let \( z \in \mathbb{C}^n \) be such that \( |z| < \delta \). Then for \( T > 0 \), if \( γ(z,t) \) exists for \( 0 ≤ t < T \), we have

\[
γ(z,t) = e^{At}z + \int_0^t e^{A(t-s)}h(γ(z,s))ds.
\]

It follows that if \( |γ(z,t)| < \delta \) for \( 0 ≤ t < T \), then

\[
|γ(z,t)| ≤ Ke^{-\sigma t}|z| + K \int_0^t e^{-\sigma(t-s)}|h(γ(z,s))|ds ≤ Ke^{-\sigma t}|z| + \frac{\sigma}{2} \int_0^t e^{-\sigma(t-s)}|γ(z,s)|ds.
\]

Multiplying \( e^{\sigma t} \) to the above inequality, we get

\[
e^{\sigma t}|γ(z,t)| ≤ K|z| + \frac{\sigma}{2} \int_0^t e^{\sigma s}|γ(z,s)|ds.
\]

From Grownwall's inequality, we conclude that

\[
e^{\sigma t}|γ(z,t)| ≤ K|z|e^{\frac{\sigma}{2}t}.
\]

Therefore, if \( γ(z,t) \) exists for \( 0 < t < T \), where \( T > 0 \), and \( |γ(z,t)| < \delta \), then

\[
|γ(z,t)| ≤ K|z|e^{-\frac{\sigma}{2}t}. \tag{2.2.2}
\]

Now let \( z \in \mathbb{C}^n \) be such that \( |z| < \min\{\delta, \frac{\delta}{K}\} \). It is obvious that there is a solution for small \( t \) such that \( |γ(z,t)| < \delta \). Therefore, (2.2.2) holds for such \( t \). Let \( T \) be the infimum \( t \) such that (2.2.2) fails. Then for \( t \in [0,T) \), we have \( |γ(z,t)| ≤ K|z| < \delta \).

Therefore \( \limsup_{t→T} |γ(z,t)| ≤ K|z| < \delta \). If \( T < ∞ \), then the trajectory can be extended at \( (γ(z,t),T) \) and consequently (2.2.2) holds for some \( t > T \), which contradicts with the assumption that (2.2.2) fails for \( t > T \). Thus \( T = ∞ \).

To summarize, we have shown that for \( z \in B_δ = \{z ∈ \mathbb{C}^n : |z| < δ, |z| < δ/K\} \), \( γ(z,t) \) exists for all \( t \) and (2.2.2) is satisfied. Since \( f \) is analytic on \( \{z ∈ \mathbb{C}^n : |z| < δ\} \),
it follows that the solutions of $\dot{\phi} = f(\phi)$ starting from $B_\delta$ analytically depend on $t$ and the initial values (see [6], Theorem 7.2, for instance). So, for $z \in B_\delta$, $\gamma(z, t)$ is analytic in $z$ and $t \ (\geq 0)$.

Recall that

$$|g(z, t)| \leq C|z|$$

(2.2.3)

for $z$ near the origin in $\mathbb{C}^n$. Let $\delta$ be sufficiently small such that (2.2.3) is satisfied for $z \in B_{K\delta}$ and (2.2.2) is satisfied for $z \in B_{\delta}$. Then for $z \in B_{\delta}$ we have

$$|g(\gamma(z, t), t)| \leq C|\gamma(z, t)| \leq CK|z|e^{-\frac{\pi}{2}t}.$$ 

Thus the integral $\int_0^\infty g(\gamma(z, t), t)dt$ converges. Write $\xi(z) = \sum_{n=0}^\infty u_n(z)$, where $u_n(z) = \int_0^{n+1} g(\gamma(z, t), t)dt$. Note that $|u_n(z)| \leq \int_0^{n+1} CK|z|e^{-\frac{\pi}{2}t}dt \leq \frac{2CK}{\pi}e^{-\frac{\pi}{2}n}$. It follows that $\sum_{n=0}^\infty u_n(z)$ converges uniformly on $z \in B_{\delta}$ by the dominated convergence theorem. Since every $u_n$ is analytic on $B_{\delta}$, we conclude that $\xi$ is analytic on $B_{\delta}$. □

**Remark 2.2.2** It is obvious that if we replace (2.2.1) by

$$\xi(z) = \int_0^\infty g(\gamma(z, t), t + T)dt,$$

where $T$ is a positive constant, then the conclusion of Lemma 2.2.1 is still true. In particular, we can find a neighborhood of the origin so that $\xi$ is analytic there. From the proof of the lemma, we also see that the neighborhood can be chosen independently of $T$, because $B_\delta$ depends only on the value of $K$ and the region where (2.2.3) is true in the proof. We will use this fact in the proof of the next lemma. □

**Lemma 2.2.3** Let $f : \mathbb{R}^n \to \mathbb{R}^n$ and $g : \mathbb{R}^{n+1} \to \mathbb{R}^n$ be two real analytic functions and satisfy the conditions (i)-(iii) of Lemma 2.2.1. Suppose that the origin is a globally asymptotically stable equilibrium of the system $\dot{x} = f(x)$. Let $t \to \gamma(x, t)$ denote the trajectory of $\dot{\phi} = f(\phi)$ with $\gamma(x, 0) = x$. Then the function $\xi$ defined by

$$\xi(z) = \int_0^\infty g(\gamma(z, t), t)dt$$

is analytic on $\mathbb{R}^n$.
Proof. From the assumptions of the lemma, we know that for any \( x \in \mathbb{R}^n \), \( \gamma(x, t) \) converges to zero exponentially as \( t \to \infty \). Since \( g(x, t) \) is bounded by \( C|x| \) when \( |x| \) is small, it follows that \( g(\gamma(x, t), t) \) also converges to zero exponentially as \( t \to \infty \). So \( \xi(x) \) is well-defined for every \( x \in \mathbb{R}^n \).

Let \( T \) be an arbitrary positive number and use \( x_T \) to denote \( \gamma(x, T) \). Let \( \xi_T(x) \) and \( \xi^T(x) \) denote the following two integrals:

\[
\xi_T(x) = \int_0^T g(\gamma(x, t), t) \, dt,
\]

\[
\xi^T(x) = \int_T^\infty g(\gamma(x, t), t) \, dt = \int_0^\infty g(\gamma(x_T, t), T + t) \, dt.
\]

Then \( \xi(x) = \xi_T(x) + \xi^T(x) \). From Remark 2.2.2, we know that there is a neighborhood \( B \) of the origin, which is independent of \( T \), such that \( \tilde{\xi}^T(x) = \int_0^\infty g(\gamma(x, t), T + t) \, dt \) is analytic in \( B \).

Let \( x_0 \) be an arbitrary point in \( \mathbb{R}^n \). Since \( \gamma(x_0, t) \) approaches zero as \( t \to \infty \), it follows that, for sufficiently large \( T \), \( \gamma(x_0, T) \in B \). We conclude that there is a neighborhood \( B_{x_0} \) of \( x_0 \) such that \( \gamma(x, T) \in B \) for \( x \in B_{x_0} \). (See Theorem 4.1 in [7].)

It is clear that \( \gamma(\cdot, T) : B_{x_0} \to B \) is analytic at \( x_0 \). So \( \xi_T(\cdot) = \tilde{\xi}^T(\gamma(\cdot, T)) \) is analytic at \( x_0 \) for sufficiently large \( T \). Also, since \( g \) and \( \gamma : \mathbb{R}^{n+1} \to \mathbb{R}^n \) are analytic functions, it is obvious that \( \xi_T : \mathbb{R}^n \to \mathbb{R}^n \) is analytic for any \( T > 0 \). Therefore, by choosing sufficiently large \( T \), we see that \( \xi = \xi_T + \xi^T \) is analytic at \( x_0 \). \( \square \)

Lemma 2.2.4 Suppose \( h \) is an increasing function from \( \mathbb{R}^+ \) to \( \mathbb{R}^+ \). Then there exists an analytic function \( \varphi \) from \( \mathbb{R} \) to \( \mathbb{R}^+ \) such that \( \varphi(x) \leq h(|x|)^{-1} \) and \( |\varphi(x)| \leq h(|x|)^{-2} \).

Proof. We first prove that there exists a positive entire function \( h_1 : \mathbb{R} \to \mathbb{R}^+ \) such that \( h(|x|) \leq h_1(x) \).

Let \( h_1(x) \) be a power series in the form \( \sum_{i=1}^\infty c_i x^{2k_i} \), where \( c_i, k_i, i = 1, 2, \ldots \), are given inductively below:

- \( c_1 = h(2), k_1 = 0; \)
- if \( c_i, k_i \) are defined for \( i = 1, 2, \ldots, n - 1 \), then let \( k_n \) be an integer greater than \( k_{n-1} \) satisfying \( n^{k_n} > h(n + 1) \), and set \( c_n = n^{-k_n} \).
To see that $\sum_{i=1}^{\infty} c_i x^{2k_i}$ converges for all $x \in \mathbb{C}$, we write the series as $\sum_{k=0}^{\infty} a_n x^n$. Then

for $n = 2k_i$, we have $a_n^{\frac{1}{n}} = c_i^{\frac{1}{2k_i}} = i^{\frac{1}{2}}$; otherwise, $a_n^{\frac{1}{n}} = 0$. So, $a_n^{\frac{1}{n}} \to 0$ as $n \to \infty$. From Hadamard’s theorem about the convergence radius of series, we see that $\sum c_i x^{2k_i}$ converges for all $x \in \mathbb{C}$.

Now given any $x$, suppose $n \leq |x| < n + 1$. Then

$$h_1(x) \geq h_1(n) \geq c_n n^{2k_n} = n^k > h(n + 1) \geq h(|x|).$$

So $h_1$ is an entire function as required.

Let $h_2(x) = h_1(x) + e^x$. Then $h_2(x) \to \infty$ as $x \to +\infty$, and $h_2'(0) = 1$. Write $h_2(0) = a$ ($> 0$). Then $h_2^{-1}$ (the inverse of $h_2$) is an increasing analytic function from $(a, +\infty)$ to $\mathbb{R}^+$. Notice that $h_2'$ is also an increasing analytic function from $(0, +\infty)$ to $(1, +\infty)$, so $x \to [h_2'(h_2^{-1}(1/x))]^{-1}$ defines an increasing analytic function from $(0, a^{-1})$ to $(0, 1)$. Let

$$g(x) = \int_0^x [h_2'(h_2^{-1}(t^{-1}))]^{-1} dt.$$ 

Then $g$ is also an analytic function on $(0, a^{-1})$, and $0 < g(x) \leq x$. (To see that $g$ is analytic on $(0, a^{-1})$, we take an arbitrary point $z_0 \in (0, a^{-1})$. Then $g(z)$ is written as a sum of a constant and $\int_{z_0}^{x} [h_2'(h_2^{-1}(1/t))]^{-1} dt$. Since the integrand function is analytic, it can be expanded as a power series in a neighborhood of $z_0$, and therefore the integral is a power series in the same neighborhood.) Now, we claim that the function $\varphi$ given by

$$\varphi(x) = g\left(\frac{1}{2} h_2(x)^{-1}\right)$$

satisfies all the requirements described in this lemma.

In fact, $x \to \frac{1}{2} h_2(x)^{-1}$ is an analytic function from $(-\infty, +\infty)$ to $(0, \frac{1}{2a})$. Thus the composition $x \to g\left(\frac{1}{2} h_2(x)^{-1}\right)$ defines an analytic function from $\mathbb{R}$ to $(0, 1)$. To verify the inequalities, since $g(x) \leq x$, we have $\varphi(x) \leq \frac{1}{2} h_2(x)^{-1} \leq h(|x|)^{-1}$. Also,

$$\varphi'(x) = -g\left(\frac{1}{2} h_2(x)^{-1}\right) \cdot \frac{h_2'(x)}{2h_2^2(x)} = -\frac{1}{h_2'(h_2^{-1}(2h_2(x)))} \cdot \frac{h_2'(x)}{2h_2^2(x)}.$$ 

Since both $h_2^{-1}$ and $h_2'$ are increasing, we have $h_2'(h_2^{-1}(2h_2(x))) \geq h_2'(h_2^{-1}(h_2(x))) = h_2'(x)$. So $|\varphi'(x)| \leq h_2(x)^{-2} \leq h(|x|)^{-2}$. The proof is complete. \qed
2.3 The Proof of Theorem 2.1

The implications $6 \Rightarrow 5 \Rightarrow 4 \Rightarrow 2$ and $5 \Rightarrow 3 \Rightarrow 2$ are clear. To show that $2 \Rightarrow 1$ we first observe that, if $\Sigma$ is ACBC, then, to begin with, $\Sigma$ is AC, so $\Sigma_c \oplus \Sigma_u$ must be controllable. To show that $E_u$ must be $\{0\}$ we use the fact that, since $-A_u$ is Hurwitz, there is a $\nu \times \nu$ matrix $P > 0$ such that $PA_u + A_u^TP^{-1}Q > 0$, where $\nu = \dim E_u$. If $C > 0$ and we restrict ourselves to controls bounded in norm by $C$, then it is clear that the derivative of the function $V(x_u) = \langle Px_u, x_u \rangle$ along trajectories of $\Sigma$ is given by

$$
\dot{V}(x_u) = \langle Qx_u, x_u \rangle + 2\langle Px_u, B_u u \rangle \geq \langle Qx_u, x_u \rangle - 2C||P|| ||B_u|| ||x_u||.
$$

So there exists $R > 0$ such that $V(x_u) > 0$ whenever $||x_u|| \geq R$. Therefore, if we choose the initial value of $x_u$ sufficiently far from 0, then the trajectory that starts at such an $x_u$ cannot possibly go to zero.

The implication $7 \Rightarrow 1$ is also easy to prove. Indeed, if (7) holds then in particular $\Sigma$ is stabilizable. Moreover, we may pick, for each $\varepsilon > 0$, a matrix $K_\varepsilon$ such that $||K_\varepsilon|| < \varepsilon$ and $A + BK_\varepsilon$ is stable. Letting $\varepsilon \to 0$, and using the continuous dependence of the eigenvalues of a matrix, we conclude that the spectrum of $A$ is contained in $\{s: \text{Re } s \leq 0\}$.

So, to prove the theorem it suffices to show that $1 \Rightarrow 6$ and $1 \Rightarrow 7$. We will actually show the following fact:

(A) Suppose (1) holds. Then for every $\varepsilon > 0$, there exists a stabilizing feedback $k \in F_{n,m,1}^{w}$ such that

(i) $||k||_{C^1} \leq \varepsilon$,

(ii) there exists an analytic Lyapunov function $V$ of $\Sigma_k$ satisfying $\dot{V}(x) \leq -||k(x)||^2$ for all $x$, where $\dot{V}$ denotes the derivative of $V$ along the trajectories of $\Sigma_k$,

(iii) the linearization $Kx$ of $k(x)$ at the origin also stabilizes $\Sigma$ and the quadratic approximation $V_q$ of $V$ at the origin is a Lyapunov function of $\Sigma_K$. 
To be precise, by “analytic Lyapunov function” for a differential equation \( E : \dot{x} = f(x) \) on \( \mathbb{R}^n \) we mean here an analytic function \( V : \mathbb{R}^n \to \mathbb{R} \) such that \( V(x) > 0 \) if \( x \neq 0 \), \( V(0) = 0 \), \( \lim_{\|x\| \to \infty} V(x) = \infty \), and \( (\nabla V(x), f(x)) \leq 0 \) for all \( x \). The LaSalle Invariance Principle says that, if \( V \) is a Lyapunov function for \( E \), then 0 is GAS equilibrium of \( E \) if and only if the following Invariance Principle Condition (IPC) holds: whenever \( t \to x(t), -\infty < t < +\infty \) is a complete orbit of \( E \) on which \( V \) is a constant, then it follows that \( x(t) \equiv 0 \).

It is clear that the above fact implies (6). In addition, from (iii) of Fact A, we know that the linearization \( Kx \) of \( k(x) \) at 0 also stabilizes \( \Sigma \). Since the norm of the matrix \( K \) is bounded by the \( C^1 \)-norm of \( k \), we see that the above fact also implies (7). So, we only need to prove Fact A.

Before we prove the above fact, we first observe that we can always assume that \( \Sigma_\varepsilon = 0 \). Indeed, a general system satisfying the condition \( \Sigma_\varepsilon = 0 \) can be written in the form

\[
\begin{align*}
\dot{x}_s &= A_s x_s + B_s u, \\
\dot{x}_c &= A_c x_c + B_c u.
\end{align*}
\]

Assume that our theorem is known to be true when \( \Sigma_\varepsilon = 0 \). Then we can find, for every \( \varepsilon > 0 \), an analytic stabilizing feedback \( u = k(x_c) \) such that (i) \( \|k\|_{C^1} \leq \varepsilon \), (ii) there is an analytic Lyapunov function \( V \) for \( (\Sigma_\varepsilon)_K \) which satisfies \( \dot{V}(x_c) \leq -\|k(x_c)\|^2 \), and (iii) the linearization \( Kx_c \) of \( k(x_c) \) also stabilizes \( \Sigma_c \), and the quadratic approximation \( V_q(x_c) \) of \( V(x_c) \) is a Lyapunov function for \( (\Sigma_c)_K \). It is then easy to see that \( u = k(x_c) \) and \( u = Kx_c \) also stabilize \( \Sigma \), and that there is a Lyapunov function \( W \) with the desired properties.

In fact, let \( P \) and \( Q \) be positive definite matrices such that \( PA_s + A_s^t P = -2Q \). Define

\[
W(x_s, x_c) = \langle Px_s, x_s \rangle + \alpha V(x_c),
\]

where \( \alpha \) is chosen so that \( \alpha > 1 + \|B^t P Q^{-1} PB\| \). It is clear that \( W(0) = 0, W(x) > 0 \) if \( x \neq 0 \), and \( W(x) \to \infty \) as \( \|x\| \to \infty \). The derivative \( \dot{W} \) of \( W \) along trajectories of \( \Sigma_k \) is

\[
\dot{W} = -\langle 2Qx_s, x_s \rangle + 2\langle x_s, PBk(x_c) \rangle + \alpha \dot{V}(x_c).
\]
Note that
\[ 2 \langle x_s, PBk(x_c) \rangle = 2 \left( Q^{\frac{1}{2}} z_s, Q^{-\frac{1}{2}} PBk(x_c) \right) \leq ||Q^{\frac{1}{2}} z_s||^2 + ||Q^{-\frac{1}{2}} PBk(x_c)||^2 \]
\[ = \langle Qz_s, x_s \rangle + \langle k(x_c), B^1 PQ^{-1} PBk(x_c) \rangle, \]
and \( \bar{V}(x_c) \leq -||k(x_c)||^2. \) Since \( \alpha > 1 + ||B^1 PQ^{-1} PB||, \) we have
\[ \dot{W}(x) \leq -\langle Qz_s, x_s \rangle - ||k(x_c)||^2. \]
In particular, \( \dot{W} \leq 0, \) so \( W \) is a Lyapunov function for \( \Sigma_k \) which clearly satisfies the bound \( \dot{W} \leq -||k(x_c)||^2. \) To prove the asymptotic stability, we establish the IPC. Let \( t \rightarrow x(t) = (x_s(t), x_c(t)) \) be a complete orbit of \( \Sigma_k \) on which \( W \) is constant. Then \( \dot{W}(x(t)) = 0, \) so \( Qz_s(t), x_s(t) \rangle = 0. \) Since \( Q > 0, \) it follows that \( x_s(t) \equiv 0 \) and we have \( \dot{W}(x(t)) \equiv \alpha \bar{V}(x_c(t)). \) So \( x_c(t) \) is an orbit of \( (\Sigma_c)_{k}. \) To see that the linearization \( Kx_c \) of \( k(x_c) \) stabilizes \( \Sigma, \) we notice that the closed-loop system \( \Sigma_K \) with the feedback \( u = Kx_c \)
\[ \text{is a linear system whose spectrum is the union of that of } A_s \text{ and } A_c + B_c K. \]
Since \( (\Sigma_c)_K \)
is assumed to be asymptotically stable, it follows that the spectrum of \( A_c + B_c K \) is contained in the set \( \{ \lambda : \Re \lambda < 0 \}. \) The spectrum of \( A_s \) is also contained in the set \( \{ \lambda : \Re \lambda < 0 \}, \) so all eigenvalues of the linear system \( \Sigma_K \) have negative real parts. It is also obvious that the quadratic approximation of \( W(x) \) is a Lyapunov function of \( \Sigma_K. \)

Now we give the general proof for our fact with the assumption \( \Sigma_s = 0. \) First we consider the simple case when \( A \) is diagonal over the complex numbers. In that case there is a real non-singular matrix \( M \) such that \( MAM^{-1} \) is skew-symmetric. This means that we can make a linear change of coordinates \( x \rightarrow Mx \) and assume that \( A \) itself is skew-symmetric.

If \( A \) is skew-symmetric, then we can stabilize \( \Sigma \) as follows. Let
\[ V(x) = \frac{1}{2} ||x||^2. \]
We are trying to find a feedback control \( u = k(x) \) such that \( V \) is a Lyapunov function of the closed-loop system \( \Sigma_k. \) For any feedback \( k, \) the derivative \( \dot{V} \) of \( V \) along the trajectories of \( \Sigma_k \) is given by
\[ \dot{V}(x) = \langle Ax, x \rangle + \sum_{i=1}^{m} k_i(x) \langle b_i, x \rangle, \]
where the $b_i$ are the columns of the matrix $B$, and the $k_i$ are the components of $k$. Since $A$ is skew-symmetric, the inner product $\langle Az, z \rangle$ vanishes. So we will have the inequality $\dot{V} \leq 0$ provided that we choose the function $k_i$ in such a way that $k_i(x) \langle b_i, z \rangle \leq 0$ for all $i$'s. An obvious choice is to take

$$k_i(x) = -\varphi(\langle b_i, z \rangle),$$

where $\varphi : \mathbb{R} \to \mathbb{R}$ is any analytic function such that $\varphi(0) = 0$ and $s\varphi(s) > 0$ whenever $s \neq 0$. We claim that the resulting closed-loop system has 0 as a GAS equilibrium.

Indeed, our choice of the $k_i$ suffices to imply the inequality $\dot{V} \leq 0$. To prove the asymptotic stability we need to verify the IPC. Let $t \to x(t)$ be an orbit of $\Sigma_k$ on which $V$ is constant. Since $\dot{V}(x(t)) = -\sum_{i=1}^{m} \langle b_i, x(t) \rangle \varphi(\langle b_i, x(t) \rangle)$, the equality $\dot{V}(x(t)) = 0$ implies $\langle b_i, x(t) \rangle = 0$ for all $t$, and $i = 1, \ldots, m$. The equation $\dot{x}(t) = Az(t) + \sum_{i=1}^{m} k_i(x(t))b_i$ then implies $\dot{x}(t) = Ax(t)$. Hence, repeated differentiation of $\langle b_i, x(t) \rangle = 0$ yields $\langle A^k b_i, x(t) \rangle = 0$ for all integers $k > 0$. The skew-symmetry of $A$ then implies $\langle A^k b_i, x(t) \rangle = 0$ for all $k$. But the vectors $A^k b_i, i = 1, \ldots, m; k = 0, 1, \ldots, n - 1$, span the whole space, because $(A, B)$ is controllable. Therefore $x(t) \equiv 0$, and the IPC follows.

In the above paragraph we choose $\varphi$ arbitrarily with the properties that $\varphi(0) = 0$ and $s\varphi(s) > 0$ if $s \neq 0$. Particularly, if we choose $\varphi(s) = cs$, where $c$ is a positive constant, then the linear feedback $k_i = c \langle b_i, x \rangle, i = 1, \ldots, m$, will stabilize $\Sigma$. More generally, if we choose $\varphi$ such that $\frac{d\varphi}{ds}(0) = c$, then the linearization of $k_i$ is $c \langle b_i, x \rangle$ and therefore stabilizes $\Sigma$. The inequality $\dot{V}(x) \leq -||k(x)||^2$ follows as long as $\varphi$ satisfies $|\varphi(s)| \leq |s|$ for all $s$, which follows automatically if $|\frac{d\varphi(s)}{ds}| \leq 1$.

To conclude the proof for the diagonal case, we have to show that for any $\varepsilon > 0$ we can choose an analytic function $\varphi$ such that $\varphi(0) = 0$, $\frac{d\varphi}{ds}(0) > 0$, $|\frac{d\varphi}{ds}| \leq 1$, and $||k_i||_{C^1} < \varepsilon$. In fact, for any bounded analytic function $\varphi_1$ with the properties $\varphi_1(0) = 0$, $\frac{d\varphi_1}{ds}(0) > 0$ and $|\frac{d\varphi_1}{ds}| \leq 1$, namely $\varphi_1(s) = \tanh(s)$, it follows that $||\varphi_1(\langle b_i, \cdot \rangle)||_{C^1}$ is bounded. (Here, $||\varphi_1(\langle b_i, \cdot \rangle)||_{C^0}$ is bounded because $\varphi_1$ is bounded, and $||\varphi_1(\langle b_i, \cdot \rangle)||_{C^1}$ is bounded because $||\varphi_1(\langle b_i, \cdot \rangle)||_{C^0}$ is bounded and the derivatives of $\varphi_1(\langle b_i, z \rangle)$ are bounded by $|\frac{d\varphi_1}{ds}|$ multiplied by a constant.) Therefore, by taking $\varphi$ to be a sufficiently
small multiple of $\varphi_1$, we have $\|k_i\|_{C^1} < \varepsilon$. The proof for the diagonalable case is now complete.

We now turn to the general case. The proof will be by induction on $n$. To be precise, assume our desired conclusion is not always true. Pick the smallest possible $n$ for which the conclusion fails for some $m$. So there is a system $\Sigma$ with $n$-dimensional state space for which $\Sigma_u = 0$ and $\Sigma_c$ is controllable but our desired conclusion does not hold. From our assumption at the beginning of the proof of our conclusion, we have $\Sigma = \Sigma_c$. Also, the above discussion for diagonalable systems shows that the system matrix of $\Sigma$ cannot be diagonalable. This implies that there exists an eigenvalue $\lambda$ of $A$ for which there is a nonzero eigenvector $v$ and a vector $\tilde{v}$ such that $A\tilde{v} = \lambda \tilde{v} + v$. We pick $\lambda$, $v$, $\tilde{v}$, and keep them fixed in the following discussion.

The eigenvalue $\lambda$ is either 0 or of the form $i\omega$, with $\omega \in \mathbb{R}$, $\omega \neq 0$. Let us consider first the case when $\lambda = 0$. In this case, the vectors $v$ and $\tilde{v}$ can be chosen to be real. Now let $S$, $\tilde{S}_0$ denote, respectively, the real linear spans of $v$ and $\tilde{v}$, and let $\sigma = 1$, $J = 0$. Let $V$, $\tilde{V}_0$ be the sets $\{v\}$, $\{\tilde{v}\}$, respectively. Then

(i) the spaces $S$, $\tilde{S}_0$ are $\sigma$-dimensional,

(ii) $S \cap \tilde{S}_0 = \{0\}$, $AS \subseteq S$, $A\tilde{S}_0 \subseteq \tilde{S}_0 + S$,

(iii) $V$, $\tilde{V}_0$ are bases of $S$, $\tilde{S}_0$,

(iv) $J$ is a $\sigma \times \sigma$ skew-symmetric matrix,

(v) if we write $Ay = \tilde{A}_0y + \tilde{G}_0y$ for $y \in \tilde{S}_0$, with $\tilde{A}_0y \in \tilde{S}_0$, $\tilde{G}_0y \in S$, and let $A_S$ denote the restriction of $A$ to $S$, then $J$ is the matrix of $A_S$ with respect to the basis $V$ and also of $\tilde{A}_0$ with respect to the basis $\tilde{V}_0$, and $\tilde{G}_0$ maps each vector of $\tilde{V}_0$ to the corresponding vector of $V$.

Now, if $\lambda = i\omega$, with $\omega \neq 0$, then we can assume $\omega > 0$. Write $v = v_1 + iv_2$, $\tilde{v} = \tilde{v}_1 + i\tilde{v}_2$. Then $A v_1 = -\omega v_2$, $A v_2 = \omega v_1$, $A \tilde{v}_1 = -\omega \tilde{v}_2 + v_1$, and $A \tilde{v}_2 = \omega \tilde{v}_1 + v_2$. So, if we let $\sigma = 2$, $V = (v_1, v_2)$, $\tilde{V}_0 = (\tilde{v}_1, \tilde{v}_2)$, $S = \text{span } V$, $\tilde{S}_0 = \text{span } \tilde{V}_0$, we see that
(i), . . . , (v) also hold in this case, with

\[ J = \begin{pmatrix} 0 & \omega \\ -\omega & 0 \end{pmatrix} . \]

(Recall that if \( V = (v_1, \ldots, v_n) \), \( W = (w_1, \ldots, w_M) \) are bases of linear spaces \( V, W \), and \( A : V \to W \) is a linear map, then the matrix of \( A \) with respect to \( V, W \) is the matrix \((a_{ij})\) such that \( Av_j = \sum_i a_{ij}w_i \). If \( W = V \), then the matrix of \( A \) with respect to \( V, V \) is the matrix of \( A \) with respect to \( V, V \).)

So, whether \( \lambda \) is real or complex, we have singled out \( \sigma, S, \tilde{S}_0, V, \tilde{V}_0, J \), with \( \sigma = 1 \) or \( 2 \), such that (i), . . . , (v) hold.

The space \( \tilde{S}_0 \) can be extended to a subspace \( S'_0 \) of \( \mathbb{R}^n \) such that \( \tilde{S}_0 \subseteq S'_0 \) and \( \mathbb{R}^n = S'_0 \oplus S \). For \( y \in S'_0 \), we can write \( Ay = A'_0y + G_0y \), with \( A'_0y \in S'_0, G_0y \in S \). Clearly, the restrictions of \( A'_0, G_0 \) to \( \tilde{S}_0 \) are \( \tilde{A}_0, \tilde{G}_0 \).

Consider an arbitrary complementary subspace \( S' \) of \( S \), i.e., an arbitrary subspace \( S' \) such that \( \mathbb{R}^n = S \oplus S' \). Let \( \pi : \mathbb{R}^n \to S'_0 \) be the projection along \( S \), i.e., \( \pi(y) = z \) if \( y = z + x, z \in S'_0, x \in S \). Let \( \tilde{\pi} \) be the restriction of \( \pi \) to \( S' \). Then \( \tilde{\pi} : S' \to S'_0 \) is a linear one-to-one and onto map. For \( y \in S' \), we can write \( Ay = A'y + Gy \) with \( A'y \in S', Gy \in S \).

An easy calculation shows that \( \tilde{\pi} \circ A' = A'_0 \circ \tilde{\pi} \), i.e., the map \( \tilde{\pi} : S' \to S'_0 \) intertwines the operators \( A' : S' \to S' \) and \( A'_0 : S'_0 \to S'_0 \). (Let \( y \in S' \). Write \( y = z + x, z \in S'_0, x \in S \). Then \( Ay = Az + Ax = A'_0z + G_0z + Ax \). But \( Ay = A'y + Gy \). So \( A'y = A'_0z + G_0z + Ax - Gy \). Since \( G_0z + Ax - Gy \in S \), we have \( \pi(G_0z + Ax - Gy) = 0 \). Since \( A'_0z \in S'_0 \), it follows that \( \pi(A'_0z) = A'_0z \). But then \( \tilde{\pi}(A'y) = \pi(A'y) = A'_0z = A'_0\tilde{\pi}(y) \).

Let \( \tilde{S} = \tilde{\pi}^{-1}(\tilde{S}_0) \), and let \( \tilde{V} \) be the basis obtained by mapping \( \tilde{V}_0 \) via \( \tilde{\pi}^{-1} \). Let \( \tilde{A} \) be the restriction of \( A' \) to \( \tilde{S} \). Then \( \tilde{A}\tilde{S} \subseteq \tilde{S} \), and \( \tilde{\pi} \) gives rise to an isomorphism between \( \tilde{S} \) and \( \tilde{S}_0 \) that intertwines the operators \( \tilde{A} \) and \( A'_0 \) and maps the basis \( \tilde{V} \) to \( \tilde{V}_0 \). From this it follows in particular that the matrix of \( \tilde{A} \) with respect to the basis \( \tilde{V}, \tilde{V} \) is \( J \).

Now let \( \tilde{G} \) denote the restriction of \( G \) to \( \tilde{S} \), so \( \tilde{G} : \tilde{S} \to S \). We are interested in computing the matrix \( \tilde{\Gamma} \) of \( \tilde{G} \) with respect to the bases \( \tilde{V}, V \). For this purpose it is
convenient to introduce the linear map \( L : S'_0 \to S \) given by \( Lz = \pi^{-1}z - z \). (If \( z \in S'_0 \) \nolinebreak then \( \pi z = z \), and so \( \pi Lz = z - \pi z = 0 \), so \( Lz \in S \).)

If \( z \in S'_0 \), let \( y = \pi^{-1}z - z + Lz \). Then \( Ay = A'y + Gy \). But

\[
A'y = A'\pi^{-1}z = \pi^{-1}A'_0z = A'_0z + LA'_0z.
\]

So

\[
Ay = \pi^{-1}A'_0z + Gy = A'_0z + LA'_0z + Gy. \tag{2.3.1}
\]

On the other hand,

\[
Ay = A(z + Lz) = Az + ALz = A'_0z + G_0z + ALz. \tag{2.3.2}
\]

Equating these two expressions (2.3.1) and (2.3.2) for \( Ay \), and canceling the common term \( A'_0z \), we get \( LA'_0z + Gy = G_0z + ALz \). Therefore \( G\pi^{-1}z = G_0z + ALz - LA'_0z \).

Now, if we let \( \tilde{L} \) denote the restriction of \( L \) to \( \tilde{S}_0 \), we get in particular

\[
\tilde{G}\pi^{-1}z = \tilde{G}_0z + A_\tilde{S}\tilde{L}z - \tilde{L}\tilde{A}_0z. \tag{2.3.3}
\]

Next recall that \( G_0 \) maps each vector of the basis \( \tilde{V}_0 \) to the corresponding vector of \( V \), so that the matrix of \( G_0 \) with respect to the bases \( \tilde{V}_0, V \) is just the identity. Let \( \tilde{\Lambda} \) denote the matrix with respect to \( \tilde{V}_0, V \). Recall that the matrix of \( A_\tilde{S} \) with respect to \( V \) is \( J \), which is also the matrix of \( \tilde{A}_0 \) with respect to \( \tilde{V}_0 \). So, using 1 to denote the \( \sigma \times \sigma \) identity matrix, from (2.3.3) we find that the matrix of \( G\pi^{-1} \) with respect to \( \tilde{V}_0, V \) is \( 1 + J\tilde{\Lambda} - \tilde{\Lambda}J \). Since \( \pi^{-1} \) maps \( \tilde{V}_0 \) to \( V \), we conclude that

\[
\tilde{\Gamma} = 1 + J\tilde{\Lambda} - \tilde{\Lambda}J.
\]

Summarizing, we have shown:

(* if \( S' \) is an arbitrary linear subspace of \( \mathbb{R}^n \) such that \( \mathbb{R}^n = S \oplus S' \), and we write \( Ay = A'y + Gy \) for \( y \in S' \), with \( A'y \in S' \), \( Gy \in S \), then there exists a \( \sigma \)-dimensional subspace \( \tilde{S} \) of \( S' \) and a basis \( \tilde{V} \) of \( \tilde{S} \), such that if \( \tilde{A}', \tilde{G} \) denote the restrictions of \( A', G \) to \( \tilde{S} \), then the matrix of \( \tilde{A}' \) with respect to \( \tilde{V} \) is \( J \), and the matrix of \( \tilde{G} \) with respect to \( \tilde{V} \), \( \tilde{V} \) is \( 1 + J\tilde{\Lambda} - \tilde{\Lambda}J \), for some \( \sigma \times \sigma \) matrix \( \tilde{\Lambda} \).
So far the choice of $S'$ was arbitrary. We now choose $S'$ in a special way, taking into account the map $B$. We let $\nu$ be the dimension of $BR^m \cap S$ (so that $0 \leq \nu \leq \sigma$). Let $m' = m - \nu$. After a linear change of coordinates in $R^m$, we may assume that the last $\nu$ columns of $B$ are in $S$, and the first $m'$ columns $b_1, \ldots, b_{m'}$ are not. We then pick $S'$ such that $R^n = S \oplus S'$ and $b_i \in S'$ for $i = 1, \ldots, m'$. With this choice of $S'$, if we let $B'$, $B''$ denote the matrices whose columns are, respectively, $b_1, \ldots, b_{m'}$ and $b_{m'+1}, \ldots, b_m$, and write $x = y + z$ for a typical $x \in R^n$, with $y \in S'$, $z \in S$, then our system $\Sigma$ has the form

$$\dot{y} = A'y + B'u', \quad \dot{z} = Jz + Gy + B''u'',$$

where $u' = (u_1, \ldots, u_{m'})$, $u'' = (u_{m'+1}, \ldots, u_m)$. (If $m' = m$, then $u''$ disappears.) We can then make a linear change of coordinates in $R^n$ so that $S', S$ become $R^{n-\sigma} \times \{0\}$ and $\{0\} \times R^\sigma$, respectively. Then we can simply regard $y$, $z$ as column vectors of length $n-\sigma, \sigma$, and think of $A'$, $B'$, $B''$, $G$ as, respectively, an $(n-\sigma) \times (n-\sigma)$, an $(n-\sigma) \times m'$, a $\sigma \times \nu$, and an $(n-\sigma) \times \sigma$ matrix.

The controllability hypothesis implies that $(A', B')$ is controllable. Using this, plus our inductive assumption, we will stabilize (2.3.4) with $u'' \equiv 0$. In view of this, from now on we will ignore the term $B''u''$, and simply relabel $u'$ as $u$ and $m' = m$. So our system now has the form

$$\dot{y} = A'y + B'u, \quad y \in R^{n-\sigma}, \quad u \in R^m, \quad \dot{z} = Jz + Gy, \quad z \in R^\sigma.$$  

(2.3.5)

It is clear that the spectrum of $A'$ is a subset of that of $A$, and that $(A', B')$ is controllable, so $\Sigma' : \dot{y} = A'y + B'u$ satisfies Condition (1) of Theorem 2.1. We now use the inductive hypothesis for $\Sigma'$. We conclude that, for every $\varepsilon > 0$, we can get an analytic feedback $k' : R^{n-\sigma} \to R^m$ such that (i) $||k'||_{C^1} \leq \frac{\varepsilon}{2}$, (ii) there is an analytic Lyapunov function $V$ for $\Sigma_{k'}$ which satisfies $\dot{V}(x) \leq -||k'(x)||^2$ for all $x$, (iii) the linearization $K'x$ of $k'(x)$ stabilizes $\Sigma'$ and the quadratic approximation $V_q$ of $V$ at the origin is a Lyapunov function for $\Sigma_{k'}$. Here, and in what follows, we will use $\dot{h}$, if $h$ is any scalar- or vector-valued function of $y$, to denote the derivative of $h$ along
trajectories of $\Sigma_k$, i.e.,
\[ \dot{h}(y) = Dh(y)(A'y + B'k'(y)), \]
where $Dh$ is the Jacobian matrix of $h$. We will define a new feedback $k(y, z)$ by letting
\[ k(y, z) \overset{\text{def}}{=} k'(y) + \mu(y, z), \]
where $\mu$ is a function to be chosen below. We will then use $\dot{h}_{\text{new}}$ to denote the derivative of a function $h(y, z)$ along trajectories of the closed-loop system $\Sigma_k$ corresponding to the new feedback, so that $\dot{h}_{\text{new}}(x) = Dh(x)(A'z + B'k'(x))$. Recall that $x = (y, z)$, so equivalently, if we use $D_yh$, $D_zh$ to denote the Jacobian matrices of $h$ with respect to $y, z$, we have
\[ \dot{h}_{\text{new}}(y, z) = D_yh(y, z)(A'y + B'k'(y) + B'\mu(y, z)) + D_zh(y, z)(Jz + Gy). \]
In the special case when $h$ only depends on $y$, we have
\[ \dot{h}_{\text{new}}(y, z) = \dot{h}(y) + Dh(y)B'\mu(y, z). \]

We now explain how to choose $\mu$. First, we have to study what happens if we only use the “old” feedback $k'$. Given any initial condition $\bar{x} = (\bar{y}, \bar{z})$, the asymptotic behavior as $t \to +\infty$ of the corresponding trajectory $t \to x(t) = (y(t), z(t))$ can be described explicitly. Indeed, $y(\cdot)$ is a trajectory of $\Sigma_{k'}$, so its limit is zero. Moreover, since $\Sigma_{k'}$ is asymptotically stable, the vector-valued function $y$ decays exponentially as $t \to \infty$. Then $z(t)$ satisfies $\dot{z}(t) = Jz(t) + Gy(t)$, so that $z(t) = e^{tJ}(\bar{z} + \int_0^t e^{-sJ}Gy(s) \, ds)$. Since $y(s)$ decays exponentially as $s \to \infty$, the infinite integral
\[ \xi(\bar{y}) \overset{\text{def}}{=} -\int_0^\infty e^{-sJ}Gy(s) \, ds \]
exists. Then $e^{-tJ}z(t)$ has the limit $\bar{z} - \xi(\bar{y})$ as $t \to \infty$. In particular, since $J$ is skew-symmetric and therefore $e^{tJ}$ is norm-preserving, we conclude that $\lim z(t) = 0$ if and only if $\bar{z} = \xi(\bar{y})$. In other words, the set $M = \{(y, z) : z = \xi(y)\}$ is exactly the set of points $(y, z)$ that are asymptotically driven to $(0, 0)$ by the “old” feedback $k'$. In particular, it is clear that $M$ is invariant under the flow of $\Sigma_{k'}$.

We remark that $M$ is an analytic manifold, since it is the graph of the analytic map $\xi$. To see that $\xi$ is analytic, observe that $\xi(y) = -\int_0^\infty e^{-sJ}G\Phi(s, y) \, ds$, where
\( \Phi \) is the flow of \( \Sigma_{k'}, \) i.e., \( s \rightarrow \Phi(s,y) \) is the solution \( \phi(s) \) of \( \dot{\phi} = A'\phi + B'k'(\phi) \) such that \( \phi(0) = y. \) Note that \( \Sigma_{k'} \) and the linearization of \( \Sigma_{k'} \) are globally asymptotically stable. Also, since \( J \) is skew-symmetric, it follows that \( e^{-sJ} \) is norm preserving and therefore \( ||e^{-sJ}Gz|| \leq ||G|| ||z||. \) So applying Lemma 2.2.3 to \( f(y) = A'y + B'k'(y), \) \( g(y,t) = e^{-tJ}Gy, \) we conclude that \( \xi \) is an analytic function. Let
\[
W(y,z) \overset{\text{def}}{=} \frac{1}{2} ||z - \xi(y)||^2 + V(y). \tag{2.3.6}
\]
From our inductive hypothesis that \( V \) is analytic and the fact that \( \xi \) is analytic, it follows that \( W \) is analytic. We will show below that \( W \) is a Lyapunov function of the closed-loop system for some choice of \( \mu. \)

Note that \( W(0) = 0 \) and \( W(x) > 0 \) if \( x \neq 0. \) Also, \( \lim_{(y,z) \to \infty} W(y,z) = +\infty. \) Let \( \dot{W}_{\text{new}} \) be the derivative of \( W \) along the trajectories corresponding to the “new” feedback \( k(y,z) = k'(y) + \mu(y,z), \) where the function \( \mu \) still remains to be chosen. Then in order to ins ure that \( W \) is a Lyapunov function, we will need the inequality \( \dot{W}_{\text{new}} \leq 0. \) It is clear that
\[
\dot{W}_{\text{new}} = \langle \dot{z}_{\text{new}} - \dot{\xi}_{\text{new}}(y), z - \xi(y) \rangle + \dot{V}(y) .
\]
Now
\[
\dot{z}_{\text{new}} = Jz + Gy,
\]
\[
\dot{\xi}_{\text{new}}(y) = \dot{\xi}(y) + D\xi(y)B'\mu(y,z),
\]
and
\[
\dot{V}(y) = V(y) + \langle \nabla V(y), B'\mu(y,z) \rangle .
\]
Since \( M \) is the graph of \( \xi \) and is invariant under the flow of \( \Sigma_{k'} \), the function \( \dot{\xi} \) is equal to \( J\xi + Gy. \) So
\[
\dot{z}_{\text{new}} - \dot{\xi}_{\text{new}}(y) = J(z - \xi(y)) - D\xi(y)B'\mu(y,z).
\]
Because of the skew-symmetry of \( J \) it follows that \( \langle J(z - \xi(y)), z - \xi(y) \rangle = 0 \) and
\[
\langle \dot{z}_{\text{new}} - \dot{\xi}_{\text{new}}(y), z - \xi(y) \rangle = - \langle D\xi(y)B'\mu(y,z), z - \xi(y) \rangle .
\]
Therefore,
\[
\dot{W}_{\text{new}} = - \langle D\xi(y)B'\mu(y,z), z - \xi(y) \rangle + \dot{V}(y) + \langle \nabla V(y), B'\mu(y,z) \rangle . \tag{2.3.7}
\]
If we let $b_1, \ldots, b_m$ denote the columns of $B'$, and write

$$
\theta_i(y, z) \stackrel{\text{def}}{=} -\langle D\xi(y)b_i, z - \xi(y) \rangle + \langle \nabla V(y), b_i \rangle,
$$

(2.3.8)

we see that $\dot{W}_{\text{new}} = \dot{V} + \sum_{i=1}^{m} \theta_i \mu_i$. So, if we choose

$$
\mu_i(y, z) = -\varphi(y, z) \theta_i(y, z),
$$

(2.3.9)

where $\varphi(y, z)$ is a strict positive function, we will have $\dot{W}_{\text{new}} = \dot{V} - \varphi \sum_{i=1}^{m} \theta^2 \leq 0$. Moreover, if $\varphi(y, z) \leq 1$ for all $(y, z)$, using $\dot{V}(y) \leq -||k'(y)||^2$, we will have

$$
\dot{W}_{\text{new}} \leq -||k'(y)||^2 - ||\mu(y, z)||^2.
$$

(2.3.10)

So we get the bound $\dot{W}_{\text{new}} \leq -||k(y, z)||^2$. If we manage to select $\varphi$ such that $||\varphi \theta||_{C^1} < \frac{\varepsilon}{2}$, then the new feedback $k$ will satisfy $||k||_{C^1} < \varepsilon$. Finally, if $\varphi$ is chosen to be analytic, since $\xi(y)$ and $\nabla V(y)$ are analytic, the new feedback will be analytic. The missing conditions are the stabilities of the new system $\Sigma_k$ and its linearization, and that the quadratic approximation of $W$ is a Lyapunov function for the linearized system.

So we have to show, to conclude our proof, that (i) $\varphi$ can be chosen so that $\varphi$ is analytic, $0 < \varphi(x) \leq 1$ for all $x$, and $||\varphi \theta||_{C^1} \leq \frac{\varepsilon}{2}$, (ii) the resulting feedback $k$ is such that the Lyapunov function $W$ satisfies the IPC, and (iii) the quadratic approximation of $W$ is a Lyapunov function for $\Sigma_k$ and satisfies the IPC, where $K$ is the linearization matrix of $k(x)$ at 0.

First, we define a function $h : \mathbb{R}^+ \to \mathbb{R}^+$ by the formula

$$
h(t) \stackrel{\text{def}}{=} \sup\{M_{i,j}(x) : \|x\| \leq \sqrt{t}, j = 1, 2, \ldots, m; i = 1, 2, \ldots, n\},
$$

(2.3.11)

where

$$
M_{i,j}(x) = \max\{1 + 2|x|, |\theta_j(x)|, |D_i \theta_j(x)|\}
$$

(2.3.12)

and $D_i$ is an operator that takes derivative with respect to the $i$th component of $x$.

It is clear that $h(t)$ is a positive increasing function. From Lemma 2.2.4 we can find an analytic function $\psi : \mathbb{R} \to \mathbb{R}^+$ such that $\psi(t) \leq h(|t|)^{-1}$ and $|\frac{d\psi}{dt}| \leq h(|t|)^{-2}$. Let $\varphi_1(x) = \psi(x^1 x)$. Then $\varphi_1$ is analytic. From the constructions of $\psi$ and $h$, it follows that $0 < \varphi_1(x) \leq h(0)^{-1} \leq 1$ for all $x$. Also, $||\varphi_1 \theta||_{C^1}$ is finite. Indeed, since
\( \varphi_1(x) \theta_j(x) = \psi(x^1 x) \theta_j(x) \leq \frac{\theta_j(x)}{h(x^1 x)}, \) from (2.3.11) and (2.3.12), we see that \( ||\varphi_1 \theta||_{C^0} \) is finite. Write

\[
D_i(\varphi_1 \theta_j) = \frac{d\psi}{dt}(x^1 x) \cdot D_i(x^1 x) \theta_j(x) + \psi(x^1 x) D_i \theta_j(x).
\]  

(2.3.13)

Then from (2.3.11) and (2.3.12) we see that the second term on the right of (2.3.13) is bounded by 1. The first term on the right of (2.3.13) is bounded by \( |\frac{D_i(x^1 x) \theta_j(x)}{h(x^1 x)}| \) because \( |\frac{d\psi}{dt}(t)| \leq h(t)^{-2} \). Since \( |\frac{D_i(x^1 x)}{h(x^1 x)}| \leq 2 \frac{|1|}{h(x^1 x)} \leq 1 \) and \( |\frac{\theta_j(x)}{h(x^1 x)}| \leq 1 \), it follows that the first term on the right of (2.3.13) is also bounded by 1. Therefore, \( ||\varphi_1 \theta||_{C^1} \) is finite. Let \( \varphi \) be a sufficiently small multiple of \( \varphi_1 \). We then get \( ||\varphi \theta||_{C^1} \leq \frac{\pi}{2} \) and Part (i) follows.

We now turn to the proof of Part (ii), namely, the verification of the IPC. Let us pick a complete orbit \( t \rightarrow x(t) = (y(t), z(t)) \) of \( \Sigma_k \), on which \( W \) is constant. From (2.3.10) we have \( \mu(y(t), z(t)) = 0 \), which implies that \( y(\cdot) \) is actually a solution of \( \dot{y} = A'_y y(t) + B'k'(y) \), i.e., a trajectory of \( \Sigma'_k \). Then since \( W_{new} = \dot{V} - \varphi \sum_{i=1}^m \theta_i^2 \leq \dot{V} \leq 0 \), we see that \( \dot{V}(y(t)) \equiv 0 \). So \( y(\cdot) \) is an orbit of \( \Sigma'_k \), on which \( V \) is constant. Since 0 is a GAS equilibrium of \( \Sigma'_k \), it follows that \( y(t) \equiv 0 \). But then \( t \rightarrow z(t) \) is a solution of \( \dot{z} = Jz \). Moreover, from (2.3.8) we have

\[
\theta_i(y(t), z(t)) = -\langle D\xi(y(t)) b_i, z(t) - x(y(t)) \rangle + \langle \nabla V(y(t)), b_i \rangle,
\]

and from (2.3.9) we have \( \theta_i(y(t), z(t)) = 0 \). Since \( \nabla V(y(t)) \equiv 0, y(t) \equiv 0 \), it follows that \( \langle D\xi(0) b_i, z(t) \rangle \equiv 0 \) for all \( i = 1, 2, \ldots, m \). Let the linearization of \( \xi(y) \) at \( y = 0 \) be \( Hy \). Then \( D\xi(0) = H \), so \( \langle H b_i, z(t) \rangle = 0 \) for all \( t \). Since \( \dot{z}(t) = Jz(t) \), we also have \( \langle J H b_i, z(t) \rangle = 0 \). Assume \( H B' \neq 0 \) for the present. Then there exists an \( i \) such that \( H b_i \neq 0 \). In the case \( \sigma = 1 \), it follows immediately that \( z(t) \equiv 0 \) since \( \langle H b_i, z(t) \rangle = 0 \). In the case \( \sigma = 2 \), since \( H b_i \) and \( J H b_i \) form a basis of \( \mathbb{R}^2 \), it follows that the equalities \( \langle H b_i, z(t) \rangle = 0 \) and \( \langle J H b_i, z(t) \rangle = 0 \) imply \( z(t) \equiv 0 \). So all that remains of the IPC is to show that \( H B' \neq 0 \).

Consider an arbitrary orbit \( t \rightarrow y(t) \) of the system \( \Sigma'_k \). Using \( \xi = J\xi + Gy, \hat{\xi} = (D\xi)y, \) and \( \dot{y} = A' y + B' k'(y) \), we conclude that

\[
(JH + G)y = H(A'y + B'K'y).
\]

(2.3.14)
If $HB' = 0$, then (2.3.14) would imply $(HA' - JH)y = Gy$. So $HA' - JH = G$. Let $\hat{H}$ be the restriction of $H$ to the $\sigma$-dimensional subspace $\tilde{S}$. Recall that $\tilde{V}$ and $V$ are the bases of $\tilde{S}$, $S$ described above and $\tilde{\Gamma}$ is the matrix of $G$ with respect to $V$. Let $\mathcal{H}$ be the matrix of $\hat{H}$ with respect to $\tilde{V}$. Since $A'S \subseteq \tilde{S}$, and the matrix of $A'$ with respect to $\tilde{V}$ is $J$, we conclude that $\tilde{\Gamma} = \mathcal{H}J - J\mathcal{H}$. But we know that $\tilde{\Gamma} = 1 + J\tilde{\Lambda} - \tilde{\Lambda}J$. These two equalities together yield a contradiction. Indeed, it is well known that the trace of any commutator is $0$. So the first equality implies that trace $\tilde{\Gamma} = 0$, whereas the second one implies trace $\tilde{\Gamma} = \text{trace } 1 = \sigma \neq 0$. This establishes the inequality $HB' \neq 0$. So Part (ii) is proved.

Finally, to prove Part (iii), we write $W(x) = W_q(x) + W_h(x)$, where $W_q$ is the quadratic approximation of $W$ at $x = 0$ and $W_h$ is its error. Since $V_q$ is positive definite, from (2.3.6) it follows that $W_q$ is also positive definite. In addition, let $Kx$ be the linearization of $k(x)$ at $x = 0$. If we use $W_{q,L}$ to denote the derivative of $W_q$ along the trajectories of $\Sigma_K$ and use $W_{q,H}$ to denote the derivative of $W_q$ along the trajectories of the closed-loop system corresponding to the feedback $u = k(x) - Kx$, then

$$W_{\text{new}}(x) = W_{q,L}(x) + W_{q,H}(x) + W_{h,\text{new}}(x), \quad (2.3.15)$$

which is less than $-||k(x)||^2$ from earlier results. Notice that $W_{q,L}$ consists of some quadratic monomials of $x$, and if $x$ is near zero, then $W_{q,H}$ and $W_{h,\text{new}}$ consist of some monomials of $x$ with degrees higher than two. If we write

$$||k(x)||^2 = ||Kx||^2 + \text{high order terms},$$

then from (2.3.15) we conclude that $W_{q,L}(x) \leq -||Kx||^2$. Therefore $W_q$ is a Lyapunov function for $\Sigma_K$.

To verify that $W_q$ satisfies the IPC for $\Sigma_K$, note that

$$W_{\text{new}} = \dot{V} + \sum_{i=1}^{m} \theta_i \mu_i \quad (2.3.16)$$

(see (2.3.7) and (2.3.8)), where $\theta_i(y, z) = -\langle D\xi(y)b_i, z - \xi(y) \rangle + \langle \nabla V(y), b_i \rangle$ and $\mu_i = -\varphi(x)\theta_i(x)$. Since $W_{q,L}$ consists of all quadratic terms in (2.3.16), it follows that

$$\dot{W}_{q,L} = \dot{V}_{q,L} + \frac{d\varphi}{dt}(0) \sum_{i=1}^{m} (-\langle Hb_i, z - Hy \rangle + \langle \nabla V_q(y), b_i \rangle)^2, \quad (2.3.17)$$
where $\dot{V}_{q,L_{old}}$ denotes the derivative of $V_q$ along the trajectories of $\Sigma_K'$. From the inductive hypothesis, we know that $\dot{V}_{q,L_{old}} \leq 0$. If $t \to (y(t), z(t))$ is a complete orbit of $\Sigma_K$ along which $\dot{W}_{q,L} \equiv 0$, then from (2.3.17) we have

$$\dot{V}_{q,L_{old}} \equiv 0$$

(2.3.18)

and

$$\langle H b_i, z(t) - H y(t) \rangle + \langle \nabla V_q(y), b_i \rangle \equiv 0$$

(2.3.19)

for all $i = 1, 2, \cdots, m$. The inductive hypothesis and (2.3.18) imply that $y(t) \equiv 0$. So from (2.3.19) we have $\langle H b_i, z(t) \rangle \equiv 0$. From the proof of Part (ii) we conclude that $z(t) \equiv 0$. The IPC of $W_q$ then follows. The proof of Theorem 2.1 is complete. \hfill \Box
Chapter 3

The Naive Design

In Chapter 2 we have proved that a linear system $\dot{x} = Ax + Bu$ is bounded stabilizable if and only if the matrix $A$ has no eigenvalues $z$ such that $\Re z > 0$, and all the uncontrollable eigenvalues have a strictly negative real part. The proof of Theorem 2.1 involves a complicated recursive construction. One might think that global stabilization might be achievable by a much simpler approach, e.g., by taking a stabilizer of the form $u(x) = \sigma(h(x))$ where $\sigma$ is —assuming the input is scalar— a “saturation function” such as $\sigma(s) = -\tanh(s)$ or $\sigma(s) = -\text{sign}(s)\min(|s|, 1)$, and $h(x)$ is a linear feedback. It turns out that this is true in some cases, e.g., when the eigenvalues of $A$ with zero real part are simple (see the proof of Theorem 2.1 for diagonal $A$), or when the system under consideration is a double integrator $\dot{x} = y$, $\ddot{y} = u$. On the other hand, it was proved by Fuller in [13] that the simple approach described above cannot possibly work for the $n$-th order integrator, if $n \geq 3$ and $\sigma$ is a saturation function of a special kind.

In this chapter we extend Fuller’s result to more general saturation functions. First we look at some simple examples.

3.1 Some Special Cases

Example 3.1.1 Consider the system $\dot{x} = Ax + Bu$. Without loss of generality, let us suppose that all the eigenvalues of $A$ have zero real part and $(A, B)$ is controllable. Assume that all the eigenvalues of $A$ are simple. Then there exists an invertible matrix $T$ such that $T^{-1}AT$ is skew-symmetric. Therefore the feedback $u = \sigma((T^{-1}B)^{\dagger}T^{-1}x)$ will stabilize the system $\dot{x} = Ax + Bu$, if $\sigma(s)$ is any saturation function with the property that $s\sigma(s) < 0$ whenever $s \neq 0$. 
Indeed, if we choose \( V = ||T^{-1}x||^2 \), then along the trajectories of the closed-loop system we have
\[
\dot{V} = 2\langle T^{-1}Ax, T^{-1}x \rangle + 2\langle T^{-1}Bu, T^{-1}x \rangle.
\]
The skew-symmetry of \( T^{-1}AT \) implies that the first term is zero. So \( \dot{V} \leq 0 \). Using the controllability of \( (A, B) \), we can see that if \( \dot{V} \equiv 0 \) along a trajectory \( t \to x(t) \), then \( x(t) \equiv 0 \). Global asymptotic stability then follows from LaSalle's Invariance Principle. 
\[ \square \]

**Example 3.1.2** Consider the double integrator
\[
\begin{align*}
\dot{x} &= y, \\
\dot{y} &= u.
\end{align*}
\]
Let \( \sigma : \mathbb{R} \to \mathbb{R} \) be any bounded strictly increasing continuous function such that \( \sigma(0) = 0 \). Define \( \Phi(s) = \int_0^s \sigma(t) \, dt \). Then the feedback law \( u = -\sigma(x + y) \) stabilizes the system and \( V(x, y) = \Phi(x) + \Phi(x + y) + y^2 \) is a strict Lyapunov function of the resulting closed-loop system.

In fact, a simple calculation shows that \( \dot{V} = (\sigma(x) - \sigma(x + y))y - [\sigma(x + y)]^2 \). The monotonicity of \( \sigma \) then implies that \( (\sigma(x) - \sigma(x + y))y < 0 \) if \( y \neq 0 \). So \( \dot{V} < 0 \) unless \( (x, y) = (0, 0) \). 
\[ \square \]

**Example 3.1.3** Consider the system
\[
\begin{align*}
\dot{x} &= \alpha y, \\
\dot{y} &= -\alpha x + u, \\
\dot{z} &= \beta w, \\
\dot{w} &= -\beta z + u,
\end{align*}
\]
where \( \alpha, \beta \) are two distinct positive numbers.

Here the matrix \( A \) has four distinct eigenvalues, so we are in a particular case of the situation described in Example 3.1.1. Let \( u = -\sigma(y+w) \) and \( V = x^2 + y^2 + z^2 + w^2 \), where \( \sigma(s) \) is any bounded continuous function with the property that \( s\sigma(s) > 0 \) if \( s \neq 0 \). From the analysis of Example 3.1.1, we know that the feedback law \( u = -\sigma(y + w) \) stabilizes the system and \( V \) is a Lyapunov function. 
\[ \square \]
Example 3.1.4 Consider the system

\[
\begin{align*}
\dot{x} &= y, \\
\dot{y} &= u, \\
\dot{z} &= \alpha w, \\
\dot{w} &= -\alpha z + u,
\end{align*}
\]

where \( \alpha > 0 \) is a constant. Notice that now zero is a double eigenvalue and there is in addition a pair of imaginary eigenvalues.

In this case define \( \sigma \) and \( \Phi \) as in Example 3.1.2. Let \( u = -\sigma(x + y + z) \), \( V = \Phi(x + y + z) + \frac{1}{2}y^2 + \frac{\alpha}{2}z^2 + \frac{\alpha}{2}w^2 \). Then a simple calculation shows that \( \dot{V} = -\sigma^2(x + y + z) \leq 0 \). Clearly, equality holds if and only if \( x + y + z \equiv 0 \). So if \( \dot{V} \equiv 0 \) along a trajectory, then the second equation of the system implies that \( \dot{y} \equiv 0 \). Differentiating the equation \( x + y + z = 0 \) twice, we get \( \dot{y} + \alpha \dot{w} \equiv 0 \), \( -\alpha^2 \dot{z} \equiv 0 \). Thus \( \dot{z} \equiv 0 \). Differentiating the equation \( x + y \equiv 0 \), we then get \( \dot{y} \equiv 0 \). So \( x \equiv 0 \) and \( w \equiv 0 \) as well. Once again global asymptotic stability follows from LaSalle’s Invariance Principle. \( \Box \)

3.2 A Negative Result

As we said earlier, the simple controller used in the examples in Section 3.1 does not work in general. In this section we show that such a control law fails to stabilize multiple integrators with dimension \( \geq 3 \).

Theorem 3.1 For the \( n \)-th order integrator with \( n \geq 3 \), there does not exist a feedback law of the form \( u = k(z) = \sigma(h(x)) \), where \( h : \mathbb{R}^n \to \mathbb{R} \) is a linear function and \( \sigma : \mathbb{R} \to \mathbb{R} \) is a locally Lipschitz function for which both limits \( \lim_{s \to \pm \infty} \sigma(s) \) exist and are nonzero, such that for the resulting closed-loop system the origin is a globally attracting point.

Recall that a point \( p \) is globally attracting for a system \( \Sigma : \dot{x} = f(x) \) if every trajectory \( t \to x(t) \) of \( \Sigma \) satisfies \( x(t) \to p \) as \( t \to +\infty \). In particular, a globally asymptotically stable equilibrium is globally attracting, so our result implies that there does not exist a feedback of the form \( u = \sigma(h(x)) \) which is globally stabilizing in the usual sense.
The restriction \( n \geq 3 \) is essential because, as shown in Example 3.1.2, the double integrator can be stabilized with a saturation of a linear feedback.

The result proved by Fuller in [13] says that if \( \sigma \) is a function such that \( \sigma(s) = \text{sign}(s) \) for \( |s| \geq a, \, a > 0 \), and \( \sigma(s) \) is zero or has the same sign as \( s \) and does not exceed the saturation level for \( |s| < a \), then there is no feedback law of the form \( u = \sigma(h(x)) \) for the multiple integrators with order \( n \geq 3 \), with \( h : \mathbb{R}^n \to \mathbb{R} \) a linear function, such that the resulting closed-loop system is globally asymptotically stable. Theorem 3.1 strengthens the result of [13] by allowing a much larger class of functions \( \sigma \), such as \( \sigma(s) = \tanh(s) \), etc. Also, this result can be used to solve the stabilizability problem for certain systems for which the information provided by the result of [13] does not suffice. For example, consider the system

\[
\begin{align*}
\dot{x} &= y, \\
\dot{y} &= z, \\
\dot{z} &= \sin u. 
\end{align*}
\]

Theorem 3.1 implies that no feedback of the form \( u = \theta(ax + by + cz) \) —where \( \theta \) is a saturation function given by \( \theta(s) = \text{sign}(s)\min\{|s|, A|\} \)— can possibly be stabilizing. (Indeed, the theorem applies directly, using \( \sigma(s) = \sin \theta(s) \), if \( A \) is not an integer multiple of \( 2\pi \). And in the remaining case, when \( A = 2k\pi \) for some integer \( k \), the conclusion is trivial.) However, the result of [13] only yields the conclusion for \( A \leq \pi/2 \).

### 3.3 The proof of Theorem 3.1

We consider an \( n \)-th order integrator

\[
\begin{align*}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= x_3, \\
&\vdots \\
\dot{x}_{n-1} &= x_n, \\
\dot{x}_n &= u, 
\end{align*}
\]

(3.3.1)

where \( n \geq 3 \), and a feedback control of the form \( k(x) = \sigma(h(x)) \), where

\[
h(x) = a_1x_1 + a_2x_2 + \ldots + a_nx_n 
\]

(3.3.2)
is a linear function, and \( \sigma : \mathbb{R} \to \mathbb{R} \) is a locally Lipschitz function such that the limits of \( \sigma(s) \) as \( s \to \pm \infty \) exist and are not equal to zero. We use \( \Sigma \) to denote the closed-loop system obtained from (3.3.1) by plugging in \( u = k(x) \). Our goal is to prove that the origin cannot be globally attracting for \( \Sigma \). The proof will be by contradiction, so from now we assume that \( 0 \) is globally attracting.

We first observe that, if a stabilizing feedback of the above form exists, then necessarily \( a_1 \neq 0 \). (Otherwise, if we take two trajectories with initial conditions all whose components coincide except for the first one, then the corresponding functions \( t \to z_1(t) \) will differ by a nonzero constant, so at least one of the trajectories will not go to \( 0 \) as \( t \to +\infty \).) It then follows that \( \sigma(s) \neq 0 \) for \( s \neq 0 \). (Indeed, assume that \( \sigma(\bar{s}) = 0, \bar{s} \neq 0 \). Then we can find \( \bar{z}_1 \) such that \( a_1\bar{z}_1 = \bar{s} \). The point \( (\bar{z}_1, 0, \ldots, 0) \) is then an equilibrium of \( \Sigma \), contradicting the fact that \( 0 \) is globally attracting.)

Moreover, \( \sigma(s) \) must change sign. Indeed, if \( \sigma(s) \) was \( \geq 0 \) for all \( s \), then it would follow that \( x_n \) is nondecreasing along every trajectory, which is obviously impossible if the origin is globally attracting. The possibility that \( \sigma(s) \leq 0 \) for all \( s \) is ruled out in a similar way. Since we know that \( \sigma(s) \neq 0 \) for \( s \neq 0 \), it follows that the change of sign must occur at \( s = 0 \).

So there are only two possibilities, namely, that \( s\sigma(s) < 0 \) for all \( s \neq 0 \), or that \( s\sigma(s) > 0 \) for all \( s \neq 0 \). In the latter case, the feedback \( k \) can also be expressed as \( k(x) = \bar{\sigma}(\bar{h}(x)) \), where \( \bar{\sigma}(s) = \sigma(-s) \) and \( \bar{h}(x) = -h(x) \), and it is clear that \( s\bar{\sigma}(s) < 0 \) for all \( s \neq 0 \). So we may assume without loss of generality that we are in the first case.

In view of the above considerations, we will assume from now on that

(I) \( s\sigma(s) < 0 \) for \( s \neq 0 \),

(II) \( a_1 \neq 0 \).

(III) the limits \( \lim_{s \to -\infty} \sigma(s) = L, \lim_{s \to +\infty} \sigma(s) = -M \) exist and satisfy \( L > 0, M > 0 \).

It follows in particular that \( \sigma \) is bounded, so we can pick an \( m \) such that

(IV) \( m \geq 1 \) and \( |\sigma(s)| \leq m \) for all \( s \).
We first prove that the coefficient $a_1$ in (3.3.2) is positive. Assume that $a_1 < 0$. Define a polynomial $p(t)$ by

$$
p(t) = -\frac{a_1M}{2n!} t^n - m\left(|a_2| \frac{t^{n-1}}{(n-1)!} + \cdots + |a_n| t\right) + a_1 \bar{x}_1,
$$

where $\bar{x}_1$ is a number to be chosen. Since $M > 0$ and $a_1 < 0$, the leading coefficient of $p(t)$ is positive. Let $A > 0$ be such that $\sigma(s) < -\frac{M}{2}$ for $s > A$. Then it is possible to choose $\bar{x}_1$ such that $\bar{x}_1 < 0$ and $p(t) > 2A$ for all $t \geq 0$. With such a choice of $\bar{x}_1$, we let $\bar{x} = (\bar{x}_1, 0, \cdots, 0)$.

Now consider the trajectory $x(t)$ of $\Sigma$ starting at $x(0) = \bar{x}$. It is clear that $h(x(0)) = a_1 \bar{x}_1 > 2A$. Therefore there exists a maximal interval $[0, \hat{t})$ (with $0 < \hat{t} \leq +\infty$) such that $h(x(t)) > A$ for $t \in [0, \hat{t})$. Then for $t \in [0, \hat{t})$ we have $-m < \sigma(h(x(t))) < -\frac{M}{2}$. So by successive integrations of (3.3.1), we see that, as long as $t \in [0, \hat{t})$:

$$-mt \leq x_n(t) \leq -\frac{M}{2} t,$$

$$-m \frac{t^n}{2} \leq x_{n-1}(t) \leq -\frac{M}{2} \frac{t^n}{2},$$

$$\vdots$$

$$-m \frac{t^{n-1}}{(n-1)!} \leq x_2(t) \leq -\frac{M}{2} \frac{t^{n-1}}{(n-1)!},$$

$$\bar{x}_1 - m \frac{t^n}{n!} \leq x_1(t) \leq \bar{x}_1 - \frac{M}{2} \frac{t^n}{n!}.$$

Notice that the right-hand sides of the above inequalities are nonpositive. If we multiply them by $|a_n|, |a_{n-1}|, \cdots, |a_2|, a_1$, then all the inequalities are preserved except for the last one, which gets reversed because $a_1 < 0$. Adding the resulting inequalities, we get

$$h(x(t)) \geq a_1 x_1(t) + \sum_{i=2}^{n} |a_i| x_i(t) \geq a_1 \bar{x}_1 - a_1 \frac{M}{2} \frac{t^n}{n!} - m \sum_{i=2}^{n} |a_i| \frac{t^{n-i+1}}{(n-i+1)!},$$

whose right-hand side is greater than $2A$ by our choice of $\bar{x}_1$. If $\hat{t}$ was finite, then we could let $t \to \hat{t}$ and conclude that $h(x(\hat{t})) \geq 2A$. But then by continuity we would have $h(x(\hat{t})) > A$ for $t$ in some larger interval $[0, \hat{t} + \alpha)$, $\alpha > 0$, contradicting the choice of $\hat{t}$. So $\hat{t} = +\infty$, and therefore the functions $x_i(t)$, $i = 1, 2, \cdots, n$ are negative and decreasing on the whole interval $(0, +\infty)$. So $x(t)$ cannot tend to zero. This contradiction shows that $a_1$ must be positive.
From now on we shall assume that $a_1 > 0$. We let $\rho = \max\{|a_2|, |a_3|, \ldots, |a_n|\}$.

Let $H$ denote the hyperplane defined by $h(z) = 0$, so that $H$ can be identified with $\mathbb{R}^{n-1}$ via the map $z \rightarrow (x_2, x_3, \ldots, x_n)$. For each $C > 0$ we consider within $H$ two regions $R_-(C), R_+(C)$. We define $R_-(C)$ to be the set of all those points $x \in H$ that satisfy

$$x_n < -C,$$
$$x_{n-1} < \frac{\rho + 2}{a_1} x_n,$$
$$x_{n-2} < \frac{\rho + 2}{a_1} (x_{n-1} + x_n),$$
$$\vdots$$
$$x_2 < \frac{\rho + 2}{a_1} (x_3 + x_4 + \cdots + x_n).$$

Similarly, $R_+(C)$ is the set of those $x \in H$ such that $-x \in R_-(C)$, i.e. that

$$x_n > C,$$
$$x_{n-1} > \frac{\rho + 2}{a_1} x_n,$$
$$x_{n-2} > \frac{\rho + 2}{a_1} (x_{n-1} + x_n),$$
$$\vdots$$
$$x_2 > \frac{\rho + 2}{a_1} (x_3 + x_4 + \cdots + x_n).$$

It is clear that $R_-(C)$ and $R_+(C)$ are not empty. We let $R(C) = R_-(C) \cup R_+(C)$.

We will prove our conclusion by studying the first return map $\Phi$ associated with $H$. Precisely, we will show that, if $C$ is sufficiently large, then the trajectory $\gamma_{x}$ from any $x \in R(C)$ returns to $H$ in a finite time $\tau(x)$ and does so by hitting $H$ at a point $\tilde{x} \in R(C)$, and that $\tilde{x} \in R_+(C)$ if and only if $x \in R_-(C)$. From this it follows easily that a trajectory that starts at some point of $R(C)$ can never go to zero. (Indeed, since the last coordinate of $\gamma_{x}(t)$ must change by at least $2C$ and $|\sigma(s)|$ is bounded by $m$, it follows that $\tau(x) \geq \frac{2C}{m}$. So the times $\tau_j(x)$, defined inductively by $\tau_1(x) = \tau(x)$ and $\tau_{j+1}(x) = \tau(\gamma_{x}(\tau_j(x)))$, go to infinity as $j \rightarrow \infty$ and, clearly, $\gamma_{x}(\tau_j(x))$ does not go to zero.)

We will only study the first return of trajectories from points in $R_-(C)$. (The other case is similar, or can be reduced to the first one by the change of variables $x \rightarrow -x$.)

It will be convenient to work instead with the hyperplane $H^A$ defined by $h(x) + A =
0, where \( A \) is such that
\[
(1 - \delta)L < \sigma(s) < (1 + \delta)L, \quad \text{if } s < -A, \\
-(1 + \delta)M < \sigma(s) < -(1 - \delta)M, \quad \text{if } s > A,
\]
and \( \delta > 0 \) satisfies two inequalities:
\[
\frac{1 - \delta}{1 + \delta} \frac{n^2}{n^2 - 1} > 1, \\
\frac{1 - \delta}{1 + \delta} (n - \frac{1}{2}) > \frac{9}{4}.
\]
Notice that it is possible to choose \( \delta > 0 \) so as to satisfy the last inequality because \( n \geq 3 \). (This is in fact the only point in the whole proof where the condition \( n \geq 3 \) is used.)

We define \( \mathcal{R}_{-}^{A,1}(C) \) to be the set of all those points \( x \in H^A \) that satisfy
\[
x_n < -C, \\
x_{n-1} < \frac{\rho+1}{a_1} x_n, \\
x_{n-2} < \frac{\rho+1}{a_1} (x_{n-1} + x_n), \\
\vdots \\
x_2 < \frac{\rho+1}{a_1} (x_3 + x_4 + \cdots + x_n),
\]
and \( \mathcal{R}_{+}^{A,3}(C) \) to be the set of all those points \( x \in H^A \) that satisfy
\[
x_n > C, \\
x_{n-1} > \frac{\rho+3}{a_1} x_n, \\
x_{n-2} > \frac{\rho+3}{a_1} (x_{n-1} + x_n), \\
\vdots \\
x_2 > \frac{\rho+3}{a_1} (x_3 + x_4 + \cdots + x_n).
\]

As explained before, the constant \( C \) will be required to be “sufficiently large.” The
precise meaning of this is that $C$ should satisfy the following inequalities:

\[
C > 5 ,
\]

\[
C > mA ,
\]

\[
C > a_1 + m\rho ,
\]

\[
C > 3m(\rho + 2) ,
\]

\[
3C > A + 4m\rho ,
\]

\[
C - 1 > \rho(1 + \delta)\max(L, M) ,
\]

\[
3(C - 1) > (2\rho + 3)(1 + \delta)\max(L, M) ,
\]

\[
a_1(C - 1) > 2n\rho(1 + \delta)\max(L, M) ,
\]

\[
a_1(C - 1) > 3n(1 + \rho)(1 + \delta)\max(L, M) .
\]

Our first task is to make sure that

(I) if we start at a point $x(0) = \bar{x} \in \mathcal{R}_-(C)$, then the resulting trajectory $t \to x(t)$ reaches $\mathcal{R}^{A,1}_-(C - 1)$ at some positive time,

and

(II) a trajectory starting at an $x(0) = \bar{x} \in \mathcal{R}^{A,3}_+(C)$ gets to $\mathcal{R}_+(C)$ at some later time.

To prove (I) and (II), notice that the derivative $\dot{h}^*(t)$ of $h^*(t) = h(x(t))$ is given by

\[
\dot{h}^*(t) = a_1\ddot{x}_2(t) + \ldots + a_{n-1}\ddot{x}_n(t) + a_n\sigma(h^*(t)) .
\]

If $x(0) \in \mathcal{R}_-(C)$, then $h^*(0) = 0$ and $\dot{h}^*(0) \leq a_1\ddot{x}_2 + \rho|\ddot{x}_3 + \cdots + \ddot{x}_n| + m\rho$, which is $< 0$ because $a_1\ddot{x}_2 < -(\rho + 2)|\ddot{x}_3 + \cdots + \ddot{x}_n|$ and $|\ddot{x}_n| > m\rho$ (since $C > m\rho$). Then $\dot{h}^*(t) < 0$ for $t$ in some interval of the form $(0, \hat{t})$. Choose $\hat{t}$ to be as large as possible.

(In principle $\hat{t}$ could be infinite.) Clearly, $h^*(t) < 0$ and therefore $\dot{x}_n(t) = \sigma(h^*(t)) > 0$ for $0 < t < \hat{t}$. On the other hand, $\dot{x}_n(t) \leq m$, so that on an interval of length $T$ it is not possible for $x_n(t)$ to increase by more than $mT$. If we let $\hat{t} = 1/m$, we see that on the interval $I = [0, \hat{t}]$ the function $x_n(t)$ satisfies $x_n(t) \leq 1 - C$. In particular (since $C > 1$), $x_n(t) < 0$ on $I$. Since $\dot{x}_{n-1} = x_n$, we see that $x_{n-1}$ is decreasing on $I$, so $x_{n-1}(t) \leq 0$ on $I$, because $x_{n-1}(0) \leq 0$. Continuing in this way, we find that all the functions $x_j(t)$,
for \( j = 1, 2, \ldots, n - 1 \) are decreasing on \( I \), and all the \( x_j(t) \), for \( j = 2, 3, \ldots, n \), are negative on \( I \).

Now consider the trajectory \( x(t) \) on \([0, \tilde{t}]\). Since \( 0 < \sigma(h^*(t)) \leq m \) on \([0, \tilde{t}]\), by repeated integrations we get the following inequalities

\[
0 \leq \gamma_n(t) \leq m t^n, \\
0 \leq \gamma_{n-1}(t) \leq m \frac{t^{n-1}}{(n-1)!},
\]

(3.3.3)

\[
\vdots
\]

\[
0 \leq \gamma_2(t) \leq m \frac{t^{n-1}}{(n-1)!},
\]

\[
0 \leq \gamma_1(t) \leq m \frac{t^n}{n!},
\]

where

\[
\gamma_n(t) = x_n(t) - \bar{x}_n,
\]

\[
\gamma_{n-1}(t) = x_{n-1}(t) - (\bar{x}_{n-1} + \bar{x}_n t),
\]

(3.3.4)

\[
\gamma_2(t) = x_2(t) - (\bar{x}_2 + \bar{x}_3 t + \cdots + \bar{x}_n \frac{t^{n-2}}{(n-2)!}),
\]

\[
\gamma_1(t) = x_1(t) - (\bar{x}_1 + \bar{x}_2 t + \cdots + \bar{x}_n \frac{t^{n-1}}{(n-1)!}).
\]

From (3.3.3), we see that \( x_j(t) \geq \sum_{k=j}^{n} \bar{x}_k \frac{t^{k-j}}{(k-j)!} \). Therefore, for \( i = 2, 3, \ldots, n-1 \), we have

\[
\sum_{j=i+1}^{n} x_j(t) \geq \sum_{k=i+1}^{n} \bar{x}_k \frac{t^{k-j}}{(k-j)!} = \sum_{j=i}^{n-1} (\bar{x}_{j+1} + \bar{x}_{j+2} + \cdots + \bar{x}_n) \frac{t^{j-i}}{(j-i)!}. 
\]

(3.3.5)

Since \( x_j(t) < 0 \) on \([0, \tilde{t}]\) for \( j = 2, 3, \ldots, n \), we have

\[
\dot{h}^*(t) \leq a_1 x_2(t) + \rho |x_3(t) + \cdots + x_n(t)| + m \rho.
\]

From (3.3.5), in particular for \( i = 2 \), we have

\[
|x_2(t) + \cdots + x_n(t)| \leq \sum_{i=2}^{n-1} (\bar{x}_{i+1} + \cdots + \bar{x}_n) \frac{t^{i-2}}{(i-2)!}. 
\]

Also, \( a_1 x_2(t) = \sum_{i=2}^{n} a_1 \bar{x}_i \frac{t^{i-2}}{(i-2)!} + a_1 \gamma_2(t) \). So we have the estimate:

\[
\dot{h}^*(t) \leq \sum_{i=2}^{n} (a_1 \bar{x}_i - \rho (\bar{x}_{i+1} + \cdots + \bar{x}_n)) \frac{t^{i-2}}{(i-2)!} + a_1 m \frac{t^{n-1}}{(n-1)!} + m \rho.
\]
Since $\bar{x} \in \mathcal{R}_-(C)$, it follows that all the coefficients in the summation of the above inequality are negative. In particular, for $t \in I = [0, \bar{t}]$ we have

$$\dot{h}^*(t) \leq a_1 \bar{x}_2 - \rho (\bar{x}_3 + \cdots + \bar{x}_n) + \frac{a_1 m^{2-n}}{(n-1)!} + m \rho,$$

which is bounded above by $2\bar{x}_n + a_1 + m \rho$ (since $a_1 \bar{x}_2 < (\rho + 2)(\bar{x}_3 + \cdots + \bar{x}_n)$ and $m > 1$).

Because $\bar{x}_n < -C$ and $C$ is greater than $a_1 + m \rho$ and $mA$, we see that $h^*(t) < -mA$ for $t \in I$. So $\bar{t} > \bar{t}$, and $h^*(t)$ is decreasing on $I$. Moreover, the bound for $\dot{h}^*$, together with $h^*(0) = 0$, imply that $h^*(\bar{t}) \leq -A$. So there exists a $\bar{t} \in I$ such that $h^*(\bar{t}) = -A$.

To complete the proof of this part, we need to show that $x(\bar{t}) \in \mathcal{R}^{A,1}_-(C - 1)$. Here, the inequality $x_n(\bar{t}) \leq 1 - C$ follows directly from (3.3.3) and the fact that $\bar{t} \leq \bar{t} = \frac{1}{m}$. We have to prove the remaining inequalities:

$$x_i(\bar{t}) \leq \frac{\rho + 1}{a_1} (x_{i+1}(\bar{t}) + \cdots + x_n(\bar{t}))$$

for $i = 2, 3, \cdots, n - 1$.

From (3.3.3) we have

$$x_i(\bar{t}) \leq \sum_{j=i}^n \frac{\bar{x}^j}{(j-i)!} + m \frac{\bar{x}^{n-i+1}}{(n-i+1)!}$$

and from (3.3.5) we have

$$x_{i+1}(\bar{t}) + \cdots + x_n(\bar{t}) \geq \sum_{j=i}^{n-1} \frac{\bar{x}^j}{(j-i)!}.$$

Therefore for $i = 2, 3, \cdots, n$, we see that

$$a_1 x_i(\bar{t}) - (\rho + 1) \sum_{j=i+1}^n x_j(\bar{t}) \leq \sum_{j=i}^n (a_1 \bar{x}_j - (\rho + 1)(\bar{x}_{j+1} + \cdots + \bar{x}_n)) \frac{\bar{x}^j}{(j-i)!} + a_1 m \frac{\bar{x}^{n-i+1}}{(n-i+1)!}.$$

Here all coefficients in the summation on the right-hand side are negative because $\bar{x} \in \mathcal{R}_-(C)$. So in particular the right-hand side is bounded above by

$$a_1 \bar{x}_i - (\rho + 1)(\bar{x}_{i+1} + \cdots + \bar{x}_n) + a_1 m \frac{\bar{x}^{n-i+1}}{(n-i+1)!}.$$

Since $\bar{t} \leq \bar{t} = \frac{1}{m}$, and

$$a_1 \bar{x}_i \leq (\rho + 2)(\bar{x}_{i+1} + \cdots + \bar{x}_n) \leq (\rho + 1)(\bar{x}_{i+1} + \cdots + \bar{x}_n) + \bar{x}_n,$$

we get in fact the upper bound $\bar{x}_n + \frac{a_1}{(n-i+1)!}$. The inequalities $\bar{x}_n < -C$ and $C > a_1$ imply in particular that this is negative. So $x(\bar{t}) \in \mathcal{R}^{A,1}_-(C - 1)$ and (I) is proved.
To prove (II) notice that, if we start a trajectory \( t \rightarrow x(t) \) at a point \( \bar{x} \) in \( R^4_+ \), then at that point \( h^* = -A \), so that \( \bar{x}_n(0) > 0 \). Again, this implies that there is a maximal interval \( J = (0, \hat{t}) \) (with \( \hat{t} \) finite or infinite) on which \( h^* < 0 \), and on that interval \( x_n \) is increasing and therefore positive, and in fact \( \geq C \). So \( x_{n-1} \) is increasing and positive, and then the same is true for \( x_{n-2}, \ldots, x_2 \). Since

\[
\dot{h}^*(t) = a_1 x_2(t) + \ldots + a_{n-1} x_{n-1}(t) + a_n \sigma(h^*(t)),
\]

it follows that

\[
\dot{h}^*(t) \geq a_1 x_2(t) - \rho |x_3(t) + \cdots + x_n(t)| - m \rho.
\]

Notice that Formulas (3.3.3) are still true in this case. So, using the upper bound for \( \gamma_i \), we get

\[
\sum_{j=i+1}^{n} x_j(t) \leq \sum_{j=i}^{n-1} (\bar{x}_{j+1} + \bar{x}_{j+2} + \cdots + \bar{x}_n) \frac{t^{j-i}}{(j-i)!} + m \sum_{j=1}^{n-i} \frac{t^j}{j!}, \quad \text{for } t \in J. \tag{3.3.6}
\]

Also

\[
x_i(t) \geq \sum_{j=i}^{n} \frac{t^{j-i}}{(j-i)!}, \quad \text{for } t \in J. \tag{3.3.7}
\]

Applying (3.3.6) and (3.3.7) for \( i = 2 \), we then see that

\[
\dot{h}^*(t) \geq \sum_{i=2}^{n} (a_1 \bar{x}_i - \rho (\bar{x}_{i+1} + \cdots + \bar{x}_n)) \frac{t^{i-2}}{(i-2)!} - m \rho \sum_{i=1}^{n-2} \frac{t^i}{i!} - m \rho.
\]

Since \( \bar{x} \in R^4_+ \), it follows that \( a_1 \bar{x}_i - \rho (\bar{x}_{i+1} + \cdots + \bar{x}_n) \geq 3 \bar{x}_n > 0 \) for \( i = 2, 3, \ldots, n-1 \). Therefore, on \( J \cap [0,1] \) we have \( \dot{h}^*(t) \geq 3 \bar{x}_n - m \rho (e + 1) > A \) (because \( \bar{x}_n > C > \frac{A+4m \rho}{3} \)), which implies that \( \hat{t} < 1 \) because otherwise \( h^*(1) > h^*(0) + A = 0 \) which is a contradiction. So we see that \( J \subset [0,1], \hat{t} \) is finite, and then \( h^*(\hat{t}) = 0 \).

To see that \( x(\hat{t}) \in R_+ \), we need to show that

\[
x_i(\hat{t}) > \frac{\rho + 2}{a_1} (x_{i+1}(\hat{t}) + \cdots + x_n(\hat{t}))
\]

for \( i = 2, 3, \ldots, n-1 \).

This can be shown by a calculation similar to the one we used to prove (I). In fact, (3.3.6) and (3.3.7) imply that

\[
a_1 x_i(\hat{t}) - (\rho + 2) \sum_{j=i+1}^{n} x_j(\hat{t}) \geq \sum_{j=i}^{n} (a_1 \bar{x}_j - (\rho + 2)(\bar{x}_{j+1} + \cdots + \bar{x}_n)) \frac{t^{j-i}}{(j-i)!} - m (\rho + 2) \sum_{j=1}^{n-i} \frac{t^j}{j!}.
\]
In particular, the right-hand side is bounded below by 

\[ a_1 \bar{x}_i - (\rho + 2)(\bar{x}_{i+1} + \cdots + \bar{x}_n) - m(\rho + 2)e^t. \]

Since \( \bar{x} \in R^{A,3}_+ \), it follows that \( a_1 \bar{x}_i \geq (\rho + 3)(\bar{x}_{i+1} + \cdots + \bar{x}_n) \), and thus \( a_1 \bar{x}_i - (\rho + 2)(\bar{x}_{i+1} + \cdots + \bar{x}_n) \geq \bar{x}_n \), so the lower bound is in fact \( \bar{x}_n - 3m(\rho + 2) > 0 \) (because \( \hat{t} < 1 \) and \( \bar{x}_n > C > 3m(\rho + 2) \)). Therefore \( x(\hat{t}) \in R_+(C) \), and (II) is proved.

In view of (I) and (II) above, our conclusion will follow if we prove:

(III) a trajectory starting at an \( x \in R^{A,1}_-(C - 1) \) gets to \( R^{A,3}_+(C) \) at some later time.

From now on, we write \( K = C - 1 \) and assume that \( t \to x(t) \) is a trajectory such that \( x(0) = \bar{x} \in R^{A,1}_-(K) \). As before, we let \( h^*(t) = h(x(t)) \). Then \( h^*(0) = -A \), and 

\[ \dot{h}^*(0) = \frac{a_1 \bar{x}_2 + \cdots + a_n \bar{x}_n + a_n \sigma(-A)}{a_1 \bar{x}_2 + \cdots + a_n \sigma(-A)} < 0 \] (because \( a_1 \bar{x}_2 < (\rho + 1)(\bar{x}_3 + \cdots + \bar{x}_n) \) and \( \bar{x}_n < -m\rho \), which follows from the facts that \( \bar{x}_n < 1 - C \), \( C > 3m(\rho + 2) \) and \( m > 1 \). So there is a maximal interval \( I = (0, \hat{t}) \) on which \( h^* < -A \). Our goal is to prove that \( \hat{t} \) is finite and \( x(\hat{t}) \in R^{A,3}_+(C) \). Since \( h^*(\hat{t}) \) must be equal to \(-A\) if \( \hat{t} \) is finite, what we need is to show that \( \hat{t} < \infty \) and to prove the inequalities 

\[
\begin{align*}
    x_n(\hat{t}) &> C, \\
x_{n-1}(\hat{t}) &> \frac{c+3}{a_1}x_n(\hat{t}), \\
x_{n-2}(\hat{t}) &> \frac{c+3}{a_1}(x_{n-1}(\hat{t}) + x_n(\hat{t})), \\
&\vdots \\
x_2(\hat{t}) &> \frac{c+3}{a_1}(x_3(\hat{t}) + \cdots + x_n(\hat{t})).
\end{align*}
\]  

(3.3.8)

On \( I \) we have \((1 - \delta)L \leq \dot{x}_n(t) \leq (1 + \delta)L\), so by successive integrations we get the following bounds:

\[
\begin{align*}
    (1 - \delta)Lt &\leq \gamma_n(t) \leq (1 + \delta)Lt, \\
    (1 - \delta)L^2 \frac{t}{2} &\leq \gamma_{n-1}(t) \leq (1 + \delta)L^2 \frac{t}{2}, \\
    &\vdots \\
    (1 - \delta)L \frac{n-1}{(n-1)!} &\leq \gamma_2(t) \leq (1 + \delta)L \frac{n-1}{(n-1)!}, \\
    (1 - \delta)L \frac{n}{n!} &\leq \gamma_1(t) \leq (1 + \delta)L \frac{n}{n!},
\end{align*}
\]  

(3.3.9)
where $\gamma_i(t), i = 1, 2, \ldots, n$ are defined in (3.3.4). Notice that both sides of the above inequalities are positive. If we define

$$f(t) = a_1 \gamma_1(t) + a_2 \gamma_2(t) + \cdots + a_n \gamma_n(t),$$

then

$$f(t) \geq (1 - \delta) La_1 \frac{t^n}{n!} - (1 + \delta) L \rho \left( \frac{t^{n-1}}{(n-1)!} + \frac{t^{n-2}}{(n-2)!} + \cdots + t \right), \quad (3.3.10)$$

and

$$f(t) \leq (1 + \delta) La_1 \frac{t^n}{n!} + (1 + \delta) L \rho \left( \frac{t^{n-1}}{(n-1)!} + \frac{t^{n-2}}{(n-2)!} + \cdots + t \right). \quad (3.3.11)$$

Since $h^*(0) = a_1 \bar{x}_1 + a_2 \bar{x}_2 + \cdots + a_n \bar{x}_n = -A$, it is clear that

$$f(t) = h^*(t) + A - \sum_{k=1}^{n-1} \left( \sum_{i=1}^{n-k} a_i \bar{x}_{i+k} \right) \frac{t^k}{k!},$$

i.e.,

$$h^*(t) + A = f(t) + \sum_{k=1}^{n-1} \left( \sum_{i=1}^{n-k} a_i \bar{x}_{i+k} \right) \frac{t^k}{k!}, \quad (3.3.12)$$

and

$$a_1 \bar{x}_{k+1} + \rho(\bar{x}_{k+2} + \cdots + \bar{x}_n) \leq \sum_{i=1}^{n-k} a_i \bar{x}_{i+k} \leq a_1 \bar{x}_{k+1} - \rho(\bar{x}_{k+2} + \cdots + \bar{x}_n). \quad (3.3.13)$$

From (3.3.10) and (3.3.12), we see that $h^*(t) + A \geq (1 - \delta) La_1 \frac{t^n}{n!} + p(t)$, where $p(t)$ is a polynomial of degree $n - 1$. If $I$ was infinite it would follow that $h^*(t) + A > 0$ for large enough $t$, contradicting the fact that $h^* < -A$ on $I$. So $I$ is finite, and of course $h^*(\bar{t}) = -A$. Now all we need is to establish the inequalities (3.3.8).

First, from (3.3.11), (3.3.12) and (3.3.13), for $t \in I$, we have

$$h^*(t) + A \leq (1 + \delta) La_1 \frac{t^n}{n!} + \sum_{k=1}^{n-1} \left( a_1 \bar{x}_{k+1} - \rho(\bar{x}_{k+2} + \cdots + \bar{x}_n) + (1 + \delta) L \rho \frac{t^k}{k!} \right). \quad (3.3.14)$$

Denote

$$\alpha_k = a_1 \bar{x}_{k+1} - \rho(\bar{x}_{k+2} + \cdots + \bar{x}_n) + (1 + \delta) L \rho, \quad k = 1, 2, \ldots, n - 1.$$}

Since $\bar{x} \in \mathcal{R}_{-}^{A,1}(C)$, we have $a_1 \bar{x}_{k+1} < (\rho + 1)(\bar{x}_{k+2} + \cdots + \bar{x}_n)$ for $k = 1, 2, \ldots, n - 2$ and it then follows that $\alpha_k < \bar{x}_n + (1 + \delta) L \rho < 0$ (because $|\bar{x}_n| > C - 1 > (1 + \delta) L \rho$). Also,
\[ \alpha_{n-1} = a_1 \bar{x}_n + (1 + \delta)L \rho < 0 \text{ (since } |\bar{x}_n| > C - 1 > \frac{(1 + \delta)L \rho}{a_1}). \text{ So all the } \alpha_k \text{ are negative.} \]

Substituting \( t = \hat{t} \) into (3.3.14), we then see that \( (1 + \delta)L a_1 \hat{t}^{n-1} \frac{\hat{x}}{n!} \geq - \sum_{k=1}^{n-1} \alpha_k \hat{t}^{k-1} \frac{k}{k!} \) because \( h^*(\hat{t}) = -A \). Therefore,

\[ L a_1 \frac{\hat{t}^{n-1}}{n!} \geq - \frac{1}{1 + \delta} \sum_{k=1}^{n-1} \alpha_k \hat{t}^{k-1} \frac{k}{k!}. \tag{3.3.15} \]

Since all the \( \alpha_k \) for \( k = 1, 2, \ldots, n - 1 \) are negative, this implies that

\[ L a_1 \frac{\hat{t}^{n-1}}{n!} \geq - \frac{1}{1 + \delta} \alpha_n \hat{t}^{n-2} \frac{(n - 1)!}{(n - 1)!} \tag{3.3.16} \]

which implies that \( \hat{t} \geq \frac{\alpha_n}{(1 + \delta)L a_1} \). Since \( \alpha_n = a_1 \bar{x}_n + (1 + \delta)L \rho \) and \( |\bar{x}_n| > C - 1 > \frac{2n(1 + \delta)L \rho}{a_1} \), we have \( (1 + \delta)L \rho < -\frac{a_1}{2n} \bar{x}_n \), thus \( \alpha_n < \frac{(2n-1)\bar{x}_n}{2n} \). Therefore

\[ \hat{t} \geq \frac{2n-1}{2(1 + \delta)L} |\bar{x}_n| \text{ and } x_n(\hat{t}) \geq \frac{1}{\bar{x}_n} + (1 - \delta)L \hat{t} \geq \left( \frac{1 - \delta}{1 + \delta}(n - \frac{1}{2}) - 1 \right) |\bar{x}_n|, \]

which is greater than \( \frac{3}{4} |\bar{x}_n| \) because \( \frac{1 - \delta}{1 + \delta}(n - \frac{1}{2}) > \frac{3}{4} \). We know that \( |\bar{x}_n| > K = C - 1 \) and \( C > 5 \). Therefore,

\[ x_n(\hat{t}) > \frac{3}{4} |\bar{x}_n| > C. \]

The first inequality in (3.3.8) follows.

Now we establish the other inequalities in (3.3.8), i.e.,

\[ x_i(\hat{t}) \geq \frac{\rho + 3}{a_i} \left( x_{i+1}(\hat{t}) + \cdots + x_n(\hat{t}) \right), \]

for \( i = 2, 3, \ldots, n - 1 \). We need to find a positive lower bound for \( a_1 x_i(\hat{t}) - (\rho + 3)(x_{i+1}(\hat{t}) + \cdots + x_n(\hat{t})) \) for \( i = 2, 3, \ldots, n - 1 \).

Since \( x_i(\hat{t}) = \sum_{j=0}^{n-i} x_j(\hat{t}) = \sum_{j=i+1}^{n} x_j(\hat{t}) \frac{\hat{t}^{j-i}}{(j-i)!} + \gamma_j(\hat{t}) \), we have

\[ \sum_{j=i+1}^{n} x_j(\hat{t}) = \sum_{j=i+1}^{n} \sum_{k=j}^{n} x_k \frac{\hat{t}^{k-j}}{(k-j)!} + \sum_{j=i+1}^{n} \gamma_j(\hat{t}). \tag{3.3.17} \]

The first summation on the right-hand side is equal (see (3.3.5)) to \( \sum_{j=0}^{n-i} (\bar{x}_{j+i+1} + \cdots + \bar{x}_n) \frac{\hat{t}^{j-i}}{(j-i)!} \), i.e. to \( \sum_{j=0}^{n-i} (\bar{x}_{j+i+1} + \cdots + \bar{x}_n) \frac{\hat{t}^{j}}{j!} \). From (3.3.9) it follows that \( \sum_{j=i+1}^{n} \gamma_j(\hat{t}) \) is bounded above by \( \sum_{j=i+1}^{n} (1 + \delta)L \frac{\hat{t}^{j-i}}{(j-i)!} \), which is equal to \( (1 + \delta)L \sum_{j=i+1}^{n-i} \frac{\hat{t}^{j}}{j!} \). Therefore,

\[ \sum_{j=i+1}^{n} x_j(\hat{t}) \leq (\bar{x}_{i+1} + \cdots + \bar{x}_n) + \sum_{j=1}^{n-i} (\bar{x}_{j+i+1} + \cdots + \bar{x}_n + (1 + \delta)L) \frac{\hat{t}^{j}}{j!} + (1 + \delta)L \frac{\hat{t}^{n-i}}{(n-i)!} \tag{3.3.18} \]

Also, from (3.3.9) and (3.3.4), we get

\[ x_i(\hat{t}) \geq \sum_{j=0}^{n-i} \bar{x}_{j+i+1} \frac{\hat{t}^{j}}{j!} + (1 - \delta)L \frac{\hat{t}^{n-i+1}}{(n-i+1)!}. \tag{3.3.19} \]
Therefore, using (3.3.18) and (3.3.19), it follows that

\[
\begin{align*}
& a_1 x_i (\hat{t}) - (\rho + 3) \left( x_{i+1} (\hat{t}) + \cdots + x_n (\hat{t}) \right) \\
& \geq (1 - \delta) L a_1 \frac{\hat{x}^{n-i+1}}{n-i+1} + \left( a_1 \bar{x}_n - (\rho + 3)(1 + \delta)L \right) \frac{\hat{x}^{n-i}}{n-i} \\
& + \left( a_1 \bar{x}_{n-1} - (\rho + 3)(1 + \delta)L \right) \frac{\hat{x}^{n-i-1}}{(n-i-1)!} + \cdots \\
& + \left( a_1 \bar{x}_{i+1} - (\rho + 3)(1 + \delta)L \right) \hat{t} \\
& + a_1 \bar{x}_i - (\rho + 3)(\bar{x}_{i+1} + \cdots + \bar{x}_n) \\
= & \quad q_i (\hat{t}) + r_i (\hat{t}),
\end{align*}
\]

where

\[
\begin{align*}
q_i (\hat{t}) &= a_1 (1-\delta) L \frac{\hat{x}^{n-i+1}}{(n-i+1)!} + \alpha_{n-1} \frac{\hat{x}^{n-i}}{(n-i)!} + \cdots + \alpha_i \hat{t} + \alpha_{i-1} - (2\rho + 3)(1 + \delta)L \frac{\hat{x}^{n-i}}{(n-i)!}, \\
r_i (\hat{t}) &= -3 \bar{x}_n - (2\rho + 3)(1 + \delta)L \frac{\hat{x}^{n-i-1}}{(n-i-1)!} + \cdots \\
& + \left( -3 \sum_{j=i+2}^n \bar{x}_j - (2\rho + 3)(1 + \delta)L \right) \hat{t} \\
& + \left( -3 \sum_{j=i+1}^n \bar{x}_j - (1 + \delta)L \rho \right) \\
& \geq -3 \bar{x}_n + (2\rho + 3)(1 + \delta)L \sum_{k=0}^{n-i-1} \frac{\hat{t}^k}{k!}.
\end{align*}
\]

It is clear that \( r_i (\hat{t}) \geq 0 \) since \( \bar{x}_n < 1 - C \) and \( 3(C - 1) > (2\rho + 3)(1 + \delta)L \). What we need is to show \( q_i (\hat{t}) \geq 0 \). This time we will require an estimation for \( \hat{t} \) more accurate than (3.3.16).

Since all \( \alpha_k \) in the summation of (3.3.15) are negative, if we restrict the summation to the range \( k = i - 1, i, i + 1, \cdots, n - 1 \), the inequality still holds. Dividing both sides of the new inequality by \( \hat{t}^{i-2} \), we then get

\[
L a_1 \frac{\hat{x}^{n-i+1}}{n-i+1} \geq -\frac{1}{1 + \delta} \left( \alpha_{n-1} \frac{\hat{x}^{n-i}}{(n-i)!} + \alpha_{n-2} \frac{\hat{x}^{n-i-1}}{(n-i-2)!} + \cdots + \alpha_{i-1} \frac{1}{(i-1)!} \right).
\]

Multiplying both sides of the above inequality by the constant \( n(n-1) \cdots (n-i+2) \), we find (since \( \Pi_{j=0}^{n-i-1} \frac{\hat{x}^{n-j}}{n-j} \geq \frac{n}{n-1} \) for \( k = 1, 2, \cdots, n-i+1 \)) that

\[
L a_1 \frac{\hat{x}^{n-i+1}}{(n-i+1)!} \geq -\frac{1}{1 + \delta} \frac{n}{n-1} \left( \alpha_{n-1} \frac{\hat{x}^{n-i}}{(n-i)!} + \alpha_{n-2} \frac{\hat{x}^{n-i-1}}{(n-i-1)!} + \cdots + \alpha_{i-1} \right).
\]

Therefore,

\[
q_i (\hat{t}) \geq \left( -\frac{1-\delta}{1 + \delta} \frac{n}{n-1} + 1 \right) \left( \alpha_{n-1} \frac{\hat{x}^{n-i}}{(n-i)!} + \alpha_{n-2} \frac{\hat{x}^{n-i-1}}{(n-i-1)!} + \cdots + \alpha_{i-1} \right) \\
-(2\rho + 3)(1 + \delta)L \frac{\hat{x}^{n-i}}{n-i}.
\]
Since \( \frac{1-\delta}{1 + \delta} \frac{n^2}{n^2 - 1} > 1 \), it follows that \( \frac{1-\delta}{1 + \delta} \frac{n}{n-1} - 1 > \frac{1}{n} \). Thus
\[
q_i(\hat{t}) \geq (\frac{1}{n} \alpha_{n-1} - (2\rho + 3)(1 + \delta)L) \frac{\hat{t}^{n-i}}{(n-i)!}.
\]

Since \( \alpha_{n-1} = a_1 \bar{x}_n + (1 + \delta)L\rho \), we have
\[
q_i(\hat{t}) > (\frac{a_1}{n} \bar{x}_n - 3(1 + \delta)(1 + \rho)L) \frac{\hat{t}^{n-i}}{(n-1)!}.
\]

The assumption that \( |\bar{x}_n| \geq C - 1 > \frac{3n(1+\delta)(1+\rho)L}{a_1} \) then implies that \( q_i(\hat{t}) > 0 \), as desired. \qed
Chapter 4

A Design Using Hidden Layers

In Chapter 2 we have shown that a linear system $\Sigma : \dot{x} = Ax + Bu$ is globally asymptotically stabilizable with a bounded feedback if and only if the following pair of algebraic conditions are satisfied:

(a) all eigenvalues of $A$ have nonpositive real part, and

(b) all eigenvalues of the uncontrollable part of $\Sigma$ have strictly negative real parts

(that is, the pair $(A, B)$ is stabilizable in the ordinary sense).

In some special cases stabilization is possible by simply using a saturated linear feedback law. But as we see in Chapter 3, such saturated linear feedback does not work for general linear systems. In this chapter we present two designs for the general case. One employs linear combinations and compositions of saturation nonlinearities. The other uses linear combinations of saturated linear functions. Both approaches are explicit and constructive.

4.1 Constructions of Feedback

We first define $\mathcal{S}$ to be the class of all functions $\sigma$ from $\mathbb{R}$ to $\mathbb{R}$ such that

- $\sigma$ is locally Lipschitz,

- $s\sigma(s) > 0$ whenever $s \neq 0$,

- $\sigma$ is differentiable at 0 and $\sigma'(0) > 0$,

- $\liminf_{|s| \to \infty} |\sigma(s)| > 0$. 
For any finite sequence $\sigma = (\sigma_1, \cdots, \sigma_k)$ of functions in $S$, we define a set $F_n(\sigma)$ of functions $f$ from $\mathbb{R}^n$ to $\mathbb{R}$ inductively as follows:

- if $k = 0$ (i.e. if $\sigma$ is the empty sequence), then $F_n(\sigma)$ consists of one element, namely, the zero function from $\mathbb{R}^n$ to $\mathbb{R}$;
- $F_n(\sigma_1)$ consists of all the functions $h : \mathbb{R}^n \to \mathbb{R}$ of the form $h(x) = \sigma_1(g(x))$, where $g : \mathbb{R}^n \to \mathbb{R}$ is linear;
- for every $k > 1$, $F_n(\sigma_1, \cdots, \sigma_k)$ is the set of all functions $h : \mathbb{R}^n \to \mathbb{R}$ that are of the form $h(x) = \sigma_k(f(x) + cg(x))$, with $f$ linear, $g \in F_n(\sigma_1, \cdots, \sigma_{k-1})$, and $c \geq 0$.

We also define $G_n(\sigma)$ to be the class of functions $h : \mathbb{R}^n \to \mathbb{R}$ given by

$$h(x) = a_1 \sigma_1(f_1(x)) + a_2 \sigma_2(f_2(x)) + \cdots + a_k \sigma_k(f_k(x)),$$

where $f_1, \cdots, f_k$ are linear functions and $a_1, \cdots, a_k$ are nonnegative constants such that $a_1 + \cdots + a_k \leq 1$.

Next, for an $m$-tuple $l = (l_1, \cdots, l_m)$ of nonnegative integers, define $|l| = l_1 + \cdots + l_m$.

For a finite sequence $\sigma = (\sigma_1, \cdots, \sigma_m, l_i) = (\sigma_1^1, \cdots, \sigma_1^{l_1}, \cdots, \sigma_m^1, \cdots, \sigma_m^{l_m})$ of functions in $S$, we let $F_n^l(\sigma)$ (and respectively $G_n^l(\sigma)$) denote the set of all functions $h : \mathbb{R}^n \to \mathbb{R}^m$ that are of the form $(h_1, \cdots, h_m)$, where $h_i \in F_n(\sigma_1^i, \cdots, \sigma_m^i)$ (and respectively $h_i \in G_n(\sigma_1^i, \cdots, \sigma_m^i)$) for $i = 1, 2, \cdots, m$. (It is clear that $F_n^l(\sigma) = F_n(\sigma)$, $G_n^l(\sigma) = G_n(\sigma)$ if $m = 1$.)

**Definition 4.1.1** Let $\delta > 0$. A function $f : [0, +\infty) \to \mathbb{R}^n$ is eventually bounded by $\delta$ (and write $|f| \leq_{ev} \delta$), if there exists $T > 0$ such that $|f(t)| \leq \delta$ for all $t \geq T$.

**Definition 4.1.2** An $n$-dimensional system $\mathcal{E} : \dot{x} = f(x)$ is SISS (small-input small-state) if for every $\varepsilon > 0$ there is a $\delta > 0$ such that, if $e : [0, +\infty) \to \mathbb{R}^n$ is bounded, measurable, and eventually bounded by $\delta$, then every solution $t \to x(t)$ of $\dot{x} = f(x) + e(t)$ is eventually bounded by $\varepsilon$.

**Definition 4.1.3** For $\Delta > 0, N > 0$, a system $\mathcal{E} : \dot{x} = f(x)$ is SISS$^L(\Delta, N)$ if, whenever $0 < \delta \leq \Delta$, it follows that, if $e : [0, +\infty) \to \mathbb{R}^n$ is bounded, measurable, and
eventually bounded by $\delta$, then every solution of $\dot{x} = f(x) + e(t)$ is eventually bounded by $N\delta$. The system is $S\ddot{I}S\dot{S}_L$ if it is $S\ddot{I}S\dot{S}_L(\Delta, N)$ for some $\Delta > 0, N > 0$.

It is obvious that if a system $\dot{x} = f(x)$ is $S\ddot{I}S\dot{S}_L(\Delta, N)$, then the system $\dot{x} = \lambda f(x)$ is $S\ddot{I}S\dot{S}_L(\lambda \Delta, N)$.

**Definition 4.1.4** For a system $\dot{x} = f(x, u)$, a feedback $u = k(x)$ is stabilizing if 0 is a globally asymptotically stable equilibrium of the closed-loop system $\dot{x} = f(x, k(x))$. If, in addition, this closed-loop system is $S\ddot{I}S\dot{S}_L$, then we will say that $k$ is $S\ddot{I}S\dot{S}_L$-stabilizing.

For a square matrix $A$ we let $s(A)$ denote the number of conjugate pairs of purely imaginary eigenvalues of $A$ (including multiplicity) and we let $z(A)$ denote the multiplicity of zero as an eigenvalue of $A$. We write $\mu(A) = s(A) + z(A)$. Our main result is as follows:

**Theorem 4.1** Let $\Sigma$ be a linear system $\dot{x} = Ax + Bu$ with state space $\mathbb{R}^n$ and input space $\mathbb{R}^m$. Assume that $\Sigma$ is asymptotically null-controllable and does not have an unstable part, i.e., that all the eigenvalues of $A$ have nonpositive real parts and all the eigenvalues of the uncontrollable part of $A$ have strictly negative real parts. Let $\mu = \mu(A)$. Let $\sigma = (\sigma_1, \cdots, \sigma_\mu)$ be an arbitrary sequence of bounded functions belonging to $S$. Then there exists an $m$-tuple $l = (l_1, \cdots, l_m)$ of nonnegative integers such that $|l| = \mu$, for which there are $S\ddot{I}S\dot{S}_L$-stabilizing feedbacks

$$u = -k_\mathcal{F}(x)$$

$$u = -k_\mathcal{G}(x)$$

such that $k_\mathcal{F} \in \mathcal{F}_n^l(\sigma)$, $k_\mathcal{G} \in \mathcal{G}_n^l(\sigma)$. Furthermore, the linearizations of the closed-loop systems are asymptotically stable.

**Remark 4.1.5** Even if we were only interested in stabilization, and did not care for the $S\ddot{I}S\dot{S}_L$ property, our inductive proof of Theorem 4.1 would still require that we prove $S\ddot{I}S\dot{S}_L$ at each step in order to carry out the induction. So the $S\ddot{I}S\dot{S}_L$ property is in
any case a byproduct of our proof, and this is one of the reasons why we have chosen to include it in the conclusion of our theorem. In addition, the $SISS_L$ conclusion is important for at least one other reason, namely that it makes stability a crucial role in our proof for the partially observed case. It is not hard to see that not all stabilizing feedbacks have the $SISS$ property, even if they are linear near the origin. To illustrate this, consider the double integrator: $\dot{x} = y, \dot{y} = u.$

Let $\sigma(s)$ be an odd continuous function such that $s \sigma(s) > 0$ for $s \neq 0$, $\sigma(s) = s$ for $|s| < \frac{1}{2}$ and $\sigma(s) = \frac{1}{s}$ for $s > 1$. Then the feedback $u = -\sigma(x + y)$ stabilizes the double integrator. (This can be established by verifying that

$$V(x, y) = \int_0^{x+y} \sigma(s) \, ds + \frac{1}{2} y^2$$

is a Lyapunov function for the closed-loop system with $u = -\sigma(x + y)$, and applying the LaSalle Invariance Principle.) Let

$$e(t) = \sigma\left(\log(t + 1) + \frac{1}{t+1}\right) - \frac{1}{(t+1)^2}.$$ 

Then clearly $e(t) \to 0$ as $t \to \infty$. But not every trajectory of the system

$$\begin{aligned}
\dot{x} &= y, \\
\dot{y} &= -\sigma(x + y) + e(t)
\end{aligned} \tag{4.1.3}$$

converges to zero. For instance, $x(t) = \log(t + 1)$, $y(t) = \frac{1}{t+1}$ is a solution of (4.1.3), but $x(t) \to \infty$ as $t \to \infty$. \qed

We will say that (4.1.1), (4.1.2) are feedbacks of Type $\mathcal{F}$, $\mathcal{G}$, respectively. In Section 4.4, we will give precise procedures to get these feedbacks. These feedback are also illustrated by diagrams.

### 4.2 Technical lemmas

In this section we present two technical lemmas that will be needed for the proof of Theorem 4.1.
Lemma 4.2.1 Consider an $n$-dimensional linear single-input system

$$\Sigma: \quad \dot{x} = Ax + bu.$$  \hspace{1cm} (4.2.1)

Suppose that $(A, b)$ is a controllable pair and that all the eigenvalues of $A$ have zero real part. Fix a $\nu > 0$.

(i) If $0$ is an eigenvalue of $A$, then there is a linear change of coordinates $Tz = (y_1, \cdots, y_n)' = (\bar{y}', y_n)'$ of $\mathbb{R}^n$ that puts $\Sigma$ in the form

$$\begin{align*}
\dot{\bar{y}} &= A_1 \bar{y} + b_1(y_n + \nu u), \\
y_n &= u,
\end{align*}$$  \hspace{1cm} (4.2.2)

where the pair $(A_1, b_1)$ is controllable and $y_n$ is a scalar variable.

(ii) If $A$ has an eigenvalue of the form $i\omega$, with $\omega > 0$, then there is a linear change of coordinates $Tz = (y_1, \cdots, y_n)' = (\bar{y}', y_{n-1}, y_n)'$ of $\mathbb{R}^n$ that puts $\Sigma$ in the form:

$$\begin{align*}
\dot{\bar{y}} &= A_1 \bar{y} + b_1(y_n + \nu u), \\
\dot{y}_{n-1} &= \omega y_n, \\
\dot{y}_n &= -\omega y_{n-1} + u,
\end{align*}$$  \hspace{1cm} (4.2.3)

where the pair $(A_1, b_1)$ is controllable and $y_{n-1}, y_n$ are scalar variables.

Proof. We first prove (i). If $0$ is an eigenvalue of $A$, then there exists a nonzero $n$-dimensional row vector $v$ such that $vA = 0$. Let $\xi : \mathbb{R}^n \to \mathbb{R}$ be the linear function $x \to vx$. Then, along trajectories of $\Sigma$, $\dot{\xi} = (vb)u$. Controllability of $(A, b)$ implies that $vb \neq 0$. So we may assume that $vb = 1$. Make a linear change of coordinates $Tz = (\bar{z}', \bar{z}_n)'$ so that $z_n \equiv \xi$. Then the system equations are of the form

$$\begin{align*}
\dot{\bar{z}} &= A_1 \bar{z} + \bar{z}_n \bar{b}_1 + \nu \bar{b}_2, \\
\dot{\bar{z}}_n &= u.
\end{align*}$$

It is clear that every eigenvalue of $A_1$ also has zero real part. Now change coordinates again by letting $\bar{y} = \bar{z} + z_n \bar{b}_3$, $y_n = z_n$, where the vector $\bar{b}_3$ will be chosen below. Then the system equations become

$$\begin{align*}
\dot{\bar{y}} &= A_1 \bar{y} + y_n(\bar{b}_1 - A_1 \bar{b}_3) + u(\bar{b}_2 + \bar{b}_3), \\
y_n &= u.
\end{align*}$$
Choose \( \tilde{b}_3 \) to be a solution of \( \tilde{b}_2 + \tilde{b}_3 = \nu(\tilde{b}_1 - A_1 \tilde{b}_3), \) i.e., \( \nu A_1 \tilde{b}_3 + \tilde{b}_3 = \nu \tilde{b}_1 - \tilde{b}_2. \) (This is possible because \( \nu A_1 + I \) is nonsingular.) Let \( \tilde{b}_1 = (\tilde{b}_1 - A_1 \tilde{b}_2). \) The equations now become \( \dot{y} = A_1 \dot{y} + (y_n + \nu u)\dot{b}_1, \) \( \dot{y}_n = u, \) as desired.

We now prove (ii). Let \( \dot{\omega}, \omega > 0, \) be an eigenvalue of \( A. \) Then \( -\omega^2 \) is an eigenvalue of \( A^2. \) So there is a nonzero \( n \)-dimensional row vector \( \tilde{v} \) such that \( \tilde{v}A^2 = -\omega^2 \tilde{v}. \) Let \( \tilde{w} = \omega^{-1} \tilde{v} A. \) Then \( \tilde{w}A = -\omega \tilde{v} \) and \( \tilde{w}A^2 = -\omega^2 \tilde{w}. \) Moreover, \( \tilde{w} \) cannot be a multiple of \( \tilde{v} \) because, if \( \tilde{w} = \lambda \tilde{v}, \) then from the fact that \( \tilde{w} = w^{-1} v A \) we see that \( \lambda \omega \) would be a nonzero real eigenvalue of \( A. \) So the linear span \( S \) of \( \tilde{v} \) and \( \tilde{w} \) is a two-dimensional subspace all whose members \( v \) satisfy \( v A^2 = -\omega^2 v. \) In particular, we can choose \( v \in S \) such that \( vb = 0 \) but \( v \neq 0. \) If we then define \( w \) by \( w = \omega^{-1} v A, \) we have \( wA = -\omega v. \) Moreover, \( wib \) cannot vanish for, if it did, the subspace \( \{ x : vx = wx = 0 \} \) would contain \( b \) and be invariant under \( A, \) contradicting controllability. So, after multiplying both \( v \) and \( w \) by a constant, if necessary, we may assume that \( wib = 1. \) Let \( \xi, \eta \) be the linear functionals \( x \rightarrow vx, x \rightarrow wx. \) Then, along trajectories of \( \Sigma, \dot{\xi} = \omega \eta \) and \( \dot{\eta} = -\omega \xi + u. \) Make a linear change of coordinates \( Tx = (\tilde{z}', z_n-1, z_n)' \) so that \( z_n-1 = \xi, z_n = \eta. \) Then the system equations are of the form

\[
\begin{align*}
\dot{\tilde{z}} &= A_1 \tilde{z} + z_n-1 \tilde{b}_1 + z_n \tilde{b}_2 + u \tilde{b}_5, \\
\dot{z}_n-1 &= \omega z_n, \quad \dot{z}_n = -\omega z_{n-1} + u,
\end{align*}
\]

and every eigenvalue of \( A_1 \) has zero real part. Now change coordinates again by letting

\[
\begin{align*}
\tilde{y} &= \tilde{z} + z_n-1 \tilde{b}_4 + z_n \tilde{b}_5, \\
y_{n-1} &= z_{n-1}, \\
y_n &= z_n,
\end{align*}
\]

where the vectors \( \tilde{b}_4, \tilde{b}_5 \) will be chosen below. Then the system equations become

\[
\begin{align*}
\dot{\tilde{y}} &= A_1 \tilde{y} + y_{n-1}(\tilde{b}_1 - A_1 \tilde{b}_4 - \omega \tilde{b}_5) + y_n(\tilde{b}_2 - A_1 \tilde{b}_5 + \omega \tilde{b}_4) + u(\tilde{b}_3 + \tilde{b}_5), \\
\dot{y}_{n-1} &= \omega y_n, \\
\dot{y}_n &= -\omega y_{n-1} + u, \quad (4.2.4)
\end{align*}
\]

If we could choose \( \tilde{b}_4, \tilde{b}_5 \) such that

\[
\tilde{b}_1 - A_1 \tilde{b}_4 - \omega \tilde{b}_5 = 0 \quad (4.2.5)
\]
and
\[
\tilde{b}_3 + \tilde{b}_5 = \nu (\tilde{b}_2 - A_1 \tilde{b}_5 + \omega \tilde{b}_4), \tag{4.2.6}
\]
then we could let
\[
b_1 = \tilde{b}_2 - A_1 \tilde{b}_5 + \omega \tilde{b}_4 \tag{4.2.7}
\]
and the system equations would become
\[
\begin{align*}
\ddot{y} &= A_1 \ddot{y} + (y_n + \nu u) b_1, \\
\dot{y}_{n-1} &= \omega y_n, \\
\dot{y}_n &= -\omega y_{n-1} + u,
\end{align*}
\]
as desired. To prove the existence of \(\tilde{b}_4\) and \(\tilde{b}_5\), we rewrite (4.2.6) as \((\nu A_1 + I) \tilde{b}_5 = \nu \tilde{b}_2 - \tilde{b}_5 + \nu \omega \tilde{b}_4\), multiply both sides by \(\omega\), and plug in the value of \(\omega \tilde{b}_5\) given by (4.2.5), namely, \(\omega \tilde{b}_5 = \tilde{b}_1 - A_1 \tilde{b}_4\). We end up with the equation
\[
(\nu A_1^2 + A_1 + \nu \omega^2 I) \tilde{b}_4 = \nu A_1 \tilde{b}_1 + \tilde{b}_1 - \nu \omega \tilde{b}_2 + \omega \tilde{b}_3.
\]
Since the eigenvalues of \(A_1\) have zero real part, the matrix \(\nu A_1^2 + A_1 + \nu \omega^2\) is nonsingular, so \(\tilde{b}_4\) exists. \(\square\)

**Lemma 4.2.2** Let \(\omega > 0\). Then for every \(\sigma \in S\) there exist functions \(\varepsilon \to \delta_0(\varepsilon), \delta_1(\varepsilon), \delta_2(\varepsilon)\) from \((0, \infty)\) to \((0, \infty)\) such that

1. whenever \(e_0, e_1, e_2\) are bounded measurable real-valued functions on \([0, \infty)\) such that
\[
\limsup_{t \to +\infty} |e_i(t)| < \delta_i(\varepsilon) \quad \text{for } i = 0, 1, 2, \tag{4.2.8}
\]
then, if \(\gamma = (x(\cdot), y(\cdot)) : [0, \infty) \to \mathbb{R}^2\) is any solution of the system
\[
\begin{align*}
\dot{x} &= \omega y + e_1(t), \\
\dot{y} &= -\omega x - \sigma (y - e_0(t)) + e_2(t), \tag{4.2.9}
\end{align*}
\]
it follows that
\[
\limsup_{t \to +\infty} \|\gamma(t)\| < \varepsilon; \tag{4.2.10}
\]
and
(II) there exist an \( \varepsilon > 0 \) and \( \nu_0, \nu_1, \nu_2 > 0 \) such that \( \delta_i(\varepsilon) = \frac{\varepsilon}{\nu_i} \) for \( i = 0, 1, 2 \) and \( 0 < \varepsilon \leq \overline{\varepsilon} \).

**Proof.** We first observe that it suffices to find \( \delta_0, \delta_1, \delta_2 \) such that

(I) Inequality (4.2.10) holds whenever \( \gamma : [0, \infty) \to \mathbb{R}^2 \) is a solution of (4.2.9) for some triple of functions \( \varepsilon_0, \varepsilon_1, \varepsilon_2 \) such that \( ||\varepsilon_i||_{L^\infty} < \delta_i \) for \( i = 0, 1, 2 \).

Indeed, if (I) holds, and \( \varepsilon_0, \varepsilon_1, \varepsilon_2 \) are bounded measurable and satisfy (4.2.8), then there is a \( T \geq 0 \) such that the restrictions \( \tilde{\varepsilon}_0, \tilde{\varepsilon}_1, \tilde{\varepsilon}_2 \) of \( \varepsilon_0, \varepsilon_1, \varepsilon_2 \) to \( [T, \infty) \) satisfy \( ||\tilde{\varepsilon}_i||_{L^\infty} < \delta_i \).

Applying (I) to the restriction \( \tilde{\gamma} \) of \( \gamma \) to \( [T, \infty) \) yields the desired conclusion.

The hypotheses on \( \sigma \) imply that \( |\sigma| \) is bounded away from 0 on the complement of every interval \(-\theta, \theta\), \( \theta > 0 \). Define

\[
\varphi(\theta) = \inf\{ |\sigma(s)| : |s| \geq \theta \},
\]

\[
\psi(\theta) = \sup\{ |\sigma(s)| : |s| \leq 3\theta \}.
\]

Then \( \psi(\theta) \geq \varphi(\theta) > 0 \). If we let

\[
\kappa = \sigma'(0),
\]

then

\[
\lim_{\theta \to 0^+} \frac{\varphi(\theta)}{\theta} = \kappa
\]

and

\[
\lim_{\theta \to 0^+} \frac{\psi(\theta)}{\theta} = 3\kappa.
\]

For each triple \((\delta_0, \delta_1, \delta_2)\) with \( \delta_i > 0 \), let \( \Gamma(\delta_0, \delta_1, \delta_2) \) denote the set of all \( \gamma : [0, \infty) \to \mathbb{R}^2 \) that are solutions of (4.2.9) for some triple \((\varepsilon_0, \varepsilon_1, \varepsilon_2)\) of functions from \([0, \infty)\) to \( \mathbb{R} \) such that \( ||\varepsilon_i||_{L^\infty} < \delta_i \) for \( i = 0, 1, 2 \).

Now pick \( \alpha > 0, \beta > 0 \), and define the functions

\[
V_{1,\alpha,\beta}(x, y) = (x + \alpha)^2 + (y - \beta)^2,
\]

\[
V_{2,\alpha,\beta}(x, y) = (x + \alpha)^2 + (y + \beta)^2.
\]
Let us compute the derivatives $\dot{V}_{1,\alpha,\beta}$, $\dot{V}_{2,\alpha,\beta}$ of $V_{1,\alpha,\beta}$, $V_{2,\alpha,\beta}$ along a trajectory $\gamma$ that is a solution of (4.2.9) for a triple $e_0, e_1, e_2$. We get, if $\gamma(t) = (x(t), y(t))$:

$$
\dot{V}_{1,\alpha,\beta} = 2(x + \alpha)(\omega y + e_1) + 2(y - \beta)(-\omega x - \sigma(y - e_0) + e_2),
$$

(4.2.18)

$$
\dot{V}_{2,\alpha,\beta} = 2(x + \alpha)(\omega y + e_1) + 2(y + \beta)(-\omega x - \sigma(y - e_0) + e_2).
$$

(4.2.19)

Then

$$
\dot{V}_{1,\alpha,\beta} = 2x(e_1 + \omega \beta) - 2y(\sigma(y - e_0) - e_2 - \alpha \omega) + 2\alpha e_1 + 2\beta(\sigma(y - e_0) - e_2)
$$

(4.2.20)

and

$$
\dot{V}_{2,\alpha,\beta} = 2x(e_1 - \omega \beta) - 2y(\sigma(y - e_0) - e_2 - \alpha \omega) + 2\alpha e_1 - 2\beta(\sigma(y - e_0) - e_2)
$$

(4.2.21)

Now assume that we can pick positive numbers $\alpha, \beta, \theta, \delta_0, \delta_1, \delta_2, A, B, C$ such that

$$
\delta_0 \leq \theta,
$$

(4.2.22)

$$
\delta_1 \leq 2\theta,
$$

(4.2.23)

$$
\delta_1 \leq \omega \beta,
$$

(4.2.24)

$$
A\delta_2 + \alpha \omega \leq \varphi(\theta),
$$

(4.2.25)

$$
\beta \leq 2\theta,
$$

(4.2.26)

$$
\alpha(\delta_1 + \beta \omega) < (2\theta - \beta)B\delta_2,
$$

(4.2.27)

$$
\alpha < C\delta_2,
$$

(4.2.28)

and

$$
A - 1 \geq \max\{B, C\}.
$$

(4.2.29)

Assume that $\gamma \in \Gamma(\delta_0, \delta_1, \delta_2)$, i.e., let $e_0, e_1, e_2$ be such that $\gamma$ is a solution of (4.2.9) and $||e_i||_{L^\infty} < \delta_i$ for $i = 0, 1, 2$. Then the coefficient $e_1 + \omega \beta$ is positive, and $e_1 - \omega \beta$ is negative. So we get the a.e. bounds

$$
\dot{V}_{1,\alpha,\beta} \leq 2\alpha \delta_1 + 2\beta(\sigma(y - e_0) - e_2) - 2y(\sigma(y - e_0) - e_2 - \alpha \omega),
$$

(4.2.30)

$$
\dot{V}_{2,\alpha,\beta} \leq 2\alpha \delta_1 - 2\beta(\sigma(y - e_0) - e_2) - 2y(\sigma(y - e_0) - e_2 - \alpha \omega),
$$

(4.2.31)
valid, respectively, when $x \leq 0$ and $x \geq 0$. If, in addition, $y \geq 2\theta$, then $y - e_0 \geq \theta$, so 
\[ \sigma(y - e_0) > 0. \]
Moreover, since $|y - e_0| \geq \theta$, we have \[ \sigma(y - e_0) \geq \varphi(\theta). \] Then
\[ \sigma(y - e_0) - e_2 - \alpha \omega \geq \varphi(\theta) - \delta_2 - \alpha \omega \geq (A - 1)\delta_2. \quad (4.2.32) \]

In particular $\sigma(y - e_0) - e_2 > 0$, and then the bound (4.2.31) for $\hat{V}_{2,\alpha,\beta}$ becomes
\[ \hat{V}_{2,\alpha,\beta} < 2\alpha\delta_1 - 4\theta(A - 1)\delta_2, \quad (4.2.33) \]
provided, of course, that $x \geq 0$, $y \geq 2\theta$.

To get a bound for $\hat{V}_{1,\alpha,\beta}$, we rewrite (4.2.30) as
\[ \hat{V}_{1,\alpha,\beta} \leq 2\alpha\delta_1 + 2\beta\alpha \omega - 2(y - \beta)(\sigma(y - e_0) - e_2 - \alpha \omega). \quad (4.2.34) \]

Then from (4.2.26) and (4.2.32) we have
\[ \hat{V}_{1,\alpha,\beta} \leq 2\alpha\delta_1 + 2\beta\alpha \omega - (4\theta - 2\beta)(A - 1)\delta_2. \quad (4.2.35) \]

In view of (4.2.28) and (4.2.23), Inequality (4.2.33) implies that $\hat{V}_{2,\alpha,\beta} < 0$ a.e. provided that $x \geq 0$, $y \geq 2\theta$. In view of (4.2.27), it follows from (4.2.35) that $\hat{V}_{1,\alpha,\beta} < 0$ a.e. So we have proved:

(II) if $\alpha, \beta, \theta, \delta_0, \delta_1, \delta_2, A, B, C$ are positive numbers such that (4.2.22)-(4.2.29) hold,
then $V_{1,\alpha,\beta}$ is strictly decreasing along every trajectory $\gamma \in \Gamma(\delta_0, \delta_1, \delta_2)$ which is contained in the region
\[ \mathcal{R}_1(\theta) = \{(x, y) : x \leq 0, y \geq 2\theta\}, \quad (4.2.36) \]
and $V_{2,\alpha,\beta}$ is strictly decreasing along every $\gamma \in \Gamma(\delta_0, \delta_1, \delta_2)$ which is contained in
\[ \mathcal{R}_2(\theta) = \{(x, y) : x \geq 0, y \geq 2\theta\}. \quad (4.2.37) \]

We now fix $\alpha, \beta, \theta, \delta_0, \delta_1, \delta_2, A, B, C$ that satisfy (4.2.22)-(4.2.29), and define, for each $L > 2\alpha$, a closed contour $\mathcal{C}(L)$ as follows (see Figure 4.1). We let $P_0(L)$ be the point $(-L, 2\theta)$. Let $\tilde{C}_1(L)$ be the circle through $P_0(L)$ with center $(-\alpha, \beta)$. Then define $\tilde{C}_1(L)$ to be the arc of $\tilde{C}_1(L)$ obtained by moving clockwise along $\tilde{C}_1(L)$ from $P_0(L)$ to
the point $P_1(L)$ where $\bar{C}_1(L)$ intersects the half-line $\{(x, y) : x = 0, y \geq 2\theta \}$. Then let $\bar{C}_2(L)$ be the circle through $P_1(L)$ with center $(-\alpha, -\beta)$, and define $C_2(L)$ to be the arc of $\bar{C}_2(L)$ obtained by moving clockwise along $\bar{C}_2(L)$ from $P_1(L)$ to the point $P_2(L)$ where $\bar{C}_2(L)$ intersects the half-line $\{(x, y) : x \geq 0, y = 2\theta \}$. Then let $C_3(L)$ be the straight-line segment from $P_2(L)$ to $P_3(L)$, where $P_3(L) = (L, -2\theta)$. The concatenation $C^+(L) = C_1(L) \cup C_2(L) \cup C_3(L)$ is an arc from $P_0(L)$ to $P_3(L)$. By reflecting with respect to the origin, we get another arc $C^-(L) = C_4(L) \cup C_5(L) \cup C_6(L)$ from $P_3(L)$ to $P_0(L)$. The concatenation $C^+(L) \cup C^-(L)$ is our closed contour $\mathcal{C}(L)$. We let $\mathcal{K}(L)$ be the closed region bounded by $\mathcal{C}(L)$, so $\mathcal{K}(L)$ is compact.

We now let $\Gamma_c(\delta_0, \delta_1, \delta_2)$ be the set of those solutions $\gamma$ of (4.2.9) that correspond to continuous functions $e_i : [0, \infty) \to \mathbb{R}$ such that $\|e_i\|_{L^\infty} < \delta_i$ for $i = 0, 1, 2$. If $\gamma \in \Gamma_c(\delta_0, \delta_1, \delta_2)$ then $\dot{\gamma}(t)$ exists for all $t$. If $\gamma(t) \in C_1(L)$ then the derivative at $s = t$ of the function $s \to V_{1,\alpha,\beta}(\gamma(s))$ is $< 0$, so $V_{1,\alpha,\beta}(\gamma(s)) < V_{1,\alpha,\beta}(\gamma(t))$ for all $s > t$ sufficiently close to $t$. This implies that $\sigma(s) \in \text{Int } (\mathcal{K}(L))$ for $s > t$, $s$ near $t$, except possibly when $\gamma(t) = P_0(L)$ or $\gamma(t) = P_1(L)$. A similar conclusion follows if
\( \gamma(t) \in C_2(L) \), using the fact that

\[
\frac{d}{ds} \bigg|_{s=t} \left( V_{i,\alpha,\beta}(\gamma(s)) \right) < 0, \tag{4.2.38}
\]

provided that \( \gamma(t) \neq P_1(L) \) and \( \gamma(t) \neq P_2(L) \).

If \( \gamma(t) = P_1(L) \) then the inequalities \( V_{i,\alpha,\beta}(\gamma(s)) < V_{i,\alpha,\beta}(\gamma(t)) \) hold for \( s > t, s \) near \( t \), both for \( i = 1 \) and \( i = 2 \). This means that \( \gamma(s) \) belongs to \( \text{Int}(\mathcal{K}(L)) \) for \( s > t, s \) near \( t \), whether it lies in \( \mathcal{R}_1(\theta) \) or in \( \mathcal{R}_2(\theta) \). (Since \( \gamma(t) = P_1(L) \), then \( \gamma(s) \) must belong to \( \mathcal{R}_1(\theta) \cup \mathcal{R}_2(\theta) \) for \( s \) near \( t \).) Therefore we have shown:

(III) if \( \gamma \in \Gamma_c(\delta_0,\delta_1,\delta_2) \) and \( \gamma(t) \in \left( C_1(L) \cup C_2(L) \right) - \{P_0(L), P_2(L)\} \), then \( \gamma(s) \in \text{Int}(\mathcal{K}(L)) \) for \( s > t, s \) near \( t \).

It is clear by symmetry that a similar conclusion holds if \( \gamma(t) \in \left( C_4(L) \cup C_5(L) \right) - \{P_3(L), P_5(L)\} \). We now study the missing case, namely, when \( \gamma(t) \in C_3(L) \cup C_6(L) \).

First we need more information about \( P_2(L) \). Write

\[
P_2(L) = (\tilde{L}, 2\theta), \quad P_1(L) = (0, \Lambda + 2\theta). \tag{4.2.39}
\]

Then

\[
(\Lambda - \alpha)^2 + (2\theta - \beta)^2 = (\Lambda + 2\theta - \beta)^2 + \alpha^2 \tag{4.2.40}
\]

and

\[
(\Lambda + 2\theta + \beta)^2 + \alpha^2 = (\tilde{L} + \alpha)^2 + (2\theta + \beta)^2. \tag{4.2.41}
\]

Therefore

\[
\tilde{L}^2 - 2\alpha \tilde{L} = \Lambda^2 + 2(2\theta - \beta)\Lambda \tag{4.2.42}
\]

and

\[
\Lambda^2 + 2(2\theta + \beta)\Lambda = \tilde{L}^2 + 2\alpha \tilde{L} \tag{4.2.43}
\]

From (4.2.42) and (4.2.26) we get \( \Lambda < L \). Then (4.2.42) and (4.2.43) imply:

\[
\tilde{L}^2 + 2\alpha \tilde{L} = \tilde{L}^2 - 2\alpha L + 4\beta \Lambda, \tag{4.2.44}
\]

so that

\[
L^2 - \tilde{L}^2 = 2\alpha(\tilde{L} + L) - 4\beta \Lambda \tag{4.2.45}
\]
and

\[ L - \tilde{L} = 2\alpha - \frac{4\beta \Lambda}{L + \tilde{L}}. \]  

(4.2.46)

If we now impose the further requirement that

\[ \beta \leq \frac{\alpha}{4}, \]  

(4.2.47)

we can conclude that

\[ \frac{4\beta \Lambda}{L + \tilde{L}} < \alpha, \]  

(4.2.48)

and then

\[ L - 2\alpha \leq \tilde{L} \leq L - \alpha. \]  

(4.2.49)

This implies in particular that the \( x \)-coordinate of \( P_2(L) \) is smaller than \( L \), so the segment \( C_3(L) \) has a negative slope, whose absolute value \( \xi \) satisfies

\[ \xi = \frac{4\theta}{L - \tilde{L}}. \]  

(4.2.50)

so that

\[ \frac{2\theta}{\alpha} \leq \xi \leq \frac{4\theta}{\alpha}. \]  

(4.2.51)

The segment \( C_3(L) \) is a level curve of the function \( V_{3,\xi} \), given by

\[ V_{3,\xi}(x, y) = \xi x + y. \]  

(4.2.52)

The derivative of \( V_{3,\xi} \) along trajectories of (4.2.9) is

\[ \dot{V}_{3,\xi} = \xi \omega y + \xi e_1 - \omega x - \sigma(y - e_0) + e_2. \]  

(4.2.53)

If we let \( \mathcal{R}_3(\theta) \) denote the strip \( \{(x, y) : |y| \leq 2\theta\} \), then \( |y - e_0| \leq 3\theta \) on \( \mathcal{R}_3(\theta) \), if \( ||e_0||_{L^\infty} < \delta_0 \). Therefore

\[ \dot{V}_{3,\xi} \leq 2\xi \omega \theta + \xi \delta_1 + \psi(\theta) + \delta_2 - \omega x \]  

(4.2.54)

along any \( \gamma \in \Gamma_c(\delta_0, \delta_1, \delta_2) \), at every time \( t \) such that \( \gamma(t) \in \mathcal{R}_3(\theta) \). If in addition \( \gamma(t) \in C_3(L) \), then \( x \geq \tilde{L} \), and we get

\[ \dot{V}_{3,\xi} \leq 2\xi \omega \theta + \xi \delta_1 + \delta_2 + \psi(\theta) - \omega \tilde{L}. \]  

(4.2.55)
Since \( \hat{L} \geq L - 2\alpha \), and \( \xi \leq \frac{\delta_2}{\alpha} \), we can conclude that

\[
\dot{V}_{3,\xi} \leq \frac{8\theta^2 \omega}{\alpha} + \frac{4\theta \delta_1}{\alpha} + 2\alpha \omega + \delta_2 + \psi(\theta) - \omega L. \tag{4.2.56}
\]

Let \( L_{\text{crit}} \) (the “critical” \( L \)) be defined by

\[
L_{\text{crit}} = \frac{8\theta^2}{\alpha} + \frac{4\theta \delta_1}{\omega} + 2\alpha \omega + \delta_2 + \psi(\theta) + \frac{\psi(\theta)}{\omega}. \tag{4.2.57}
\]

Then, if \( L > L_{\text{crit}} \), we have

\[
\dot{V}_{3,\xi} < 0. \tag{4.2.58}
\]

So we have shown that, if \( L > L_{\text{crit}} \), then

\[
\left. \frac{d}{ds}\right|_{s=t} (V_{3,\xi}(\gamma(s))) < 0 \tag{4.2.59}
\]

if \( \gamma \in \Gamma_c(\delta_0, \delta_1, \delta_2) \), \( \gamma(t) \in C_2(L) \). This implies that \( \gamma(s) \in \text{Int}(\mathcal{K}(L)) \) for \( s > t \), \( s \) near \( t \), except possibly if \( \gamma(t) = P_2(L) \) or \( \gamma(t) = P_3(L) \). If \( \gamma(t) = P_2(L) \) then both functions \( s \to V_{2,\alpha,\beta}(\gamma(s)) \) and \( s \to V_{3,\xi}(\gamma(s)) \) have a negative derivative at \( s = t \), so \( \gamma(s) \) lies to left of \( C_2(L) \) for \( s > t \), \( s \) near \( t \), if \( \gamma(s) \in \mathcal{R}_2(\theta) \), and lies inside the circle \( \tilde{C}_2(L) \) if \( \gamma(s) \in \mathcal{R}_3(\theta) \). In either case, \( \gamma(s) \in \text{Int}(\mathcal{K}(L)) \) for \( s > t \), \( s \) near \( t \). A similar argument works for \( \gamma(t) = P_3(L) \). By symmetry, the same conclusion holds if \( \gamma(t) \in C_6(L) \). So we have proved:

(IV) assume that \( \alpha, \beta, \theta, \delta_0, \delta_1, \delta_2, A, B, C \) are positive numbers such that (4.2.22)-(4.2.29) and (4.2.47) hold. Let \( L_{\text{crit}} \) be defined by (4.2.57). Then, if \( L > L_{\text{crit}} \),

and \( \gamma \in \Gamma_c(\delta_0, \delta_1, \delta_2) \) is such that \( \gamma(t) \in \mathcal{K}(L) \) for some \( t \in [0, \infty) \), it follows that \( \gamma(s) \in \text{Int}(\mathcal{K}(L)) \) for \( s > t \), \( s \) near \( t \).

Using (IV), it is easy to strengthen the conclusion to “\( \gamma(s) \in \text{Int}(\mathcal{K}(L)) \) for all \( s > t \)”.

Indeed, let us first show that \( \gamma(s) \in \mathcal{K}(L) \) for all \( s \geq t \). If this was not true, the set \( I = \{ s : s \geq t, \gamma(s) \notin \mathcal{K}(L) \} \) would be relatively open in [\( t, \infty \)] and nonempty. Let \( \bar{s} = \inf I \). Then \( \bar{s} \) cannot be \( t \), because \( \gamma(s) \in \text{Int}(\mathcal{K}(L)) \) for \( s > t \), \( s \) near \( t \). So \( \bar{s} > t \). But then \( \gamma(\bar{s}) \in \mathcal{K}(L) \), because \( \gamma \) is continuous and \( \mathbb{R}^2 - \mathcal{K}(L) \) is open. But this implies that for some \( \rho > 0 \) we have \( \gamma(s) \in \mathcal{K}(L) \) for \( \bar{s} \leq s \leq \bar{s} + \rho \). So \( \inf I \geq \bar{s} + \rho \), contradicting our choice of \( \bar{s} \). This establishes that \( I \) is empty, i.e. that \( \gamma(s) \in \mathcal{K}(L) \) for
all $s \geq t$. Now let $\rho > 0$ be such that $\gamma(s) \in \text{Int}(\mathcal{K}(L))$ for $t < s < t + \rho$. It is clear that each $\gamma(s)$ such that $s > t$ belongs to a set $\mathcal{K}({\hat L}_s)$ with $L_{\text{crit}} < {\hat L}_s < L$, provided that $s$ is close enough to $t$. Pick an $\bar{s} \in (t, t + \rho)$ for which this is true. Then $\gamma(s) \in \mathcal{K}({\hat L}_{\bar{s}})$ for $s \geq \bar{s}$. Since $\mathcal{K}({\hat L}_{\bar{s}}) \subseteq \text{Int}(\mathcal{K}(L))$, we see that $\gamma(s) \in \text{Int}(\mathcal{K}(L))$ for $s \geq \bar{s}$. So in fact $\gamma(s) \in \text{Int}(\mathcal{K}(L))$ for all $s > t$.

Finally, we can strengthen the above conclusion by dropping the restriction that $\gamma \in \Gamma_c(\delta_0, \delta_1, \delta_2)$, and allowing arbitrary $\gamma \in \Gamma(\delta_0, \delta_1, \delta_2)$. Indeed, if $\gamma \in \Gamma(\delta_0, \delta_1, \delta_2)$, so $\gamma$ is a solution of (4.2.9) corresponding to measurable functions $e_0, e_1, e_2$ such that $\|e_i\|_{L_\infty} < \delta_i$, then we can find for each $e_i$ a sequence $\{e^j_i\}_{j=1}^\infty$ of continuous functions $e^j_i : [0, \infty) \to \mathbb{R}$ such that $\|e^j_i\|_{L_\infty} \leq \|e_i\|_{L_\infty}$ and the $e^j_i$ converge almost everywhere to $e_i$. Suppose that $t \in [0, \infty)$ is such that $\gamma(t) \in \mathcal{K}(L)$. Let $\gamma^i$ be the solutions of (4.2.9) corresponding to $e^0_i, e^1_i, e^2_i$, with initial condition $\gamma^i(0) = \gamma(t)$. Then $\gamma^i(s) \to \gamma(s)$ for each $s$, as $j \to \infty$. Since $\gamma^i(s) \in \mathcal{K}(L)$ for $s \geq t$ (because $\gamma^i \in \Gamma_c(\delta_0, \delta_1, \delta_2)$) and $\mathcal{K}(L)$ is closed, we conclude that $\gamma(s) \in \mathcal{K}(L)$ for $s \geq t$. To prove that $\gamma(s) \in \text{Int}(\mathcal{K}(L))$ for $s > t$, we observe that the conclusion that "$\gamma(s) \in \mathcal{K}(L)$ implies that $\gamma(s) \in \text{Int}(\mathcal{K}(L))$ for $s > \hat{s}$, $s$ near $\hat{s}$" follows, by the same arguments used for $\gamma \in \Gamma(\delta_0, \delta_1, \delta_2)$, provided only that $\gamma \in \Gamma(\delta_0, \delta_1, \delta_2)$ and $\hat{s}$ is such that $\gamma$ is differentiable at $\hat{s}$ and (4.2.9) holds there. In particular, given any $s > t$, we can find an $\hat{s}$ such that $t < \hat{s} < s$, $\gamma$ is differentiable at $\hat{s}$, and (4.2.9) holds at $\hat{s}$. Since $\gamma(\hat{s}) \in \mathcal{K}(L)$, we can conclude that $\gamma(\hat{s}) \in \text{Int}(\mathcal{K}(L))$ for $\hat{s} > \hat{s}$, $\hat{s}$ near $\hat{s}$. Choose $\hat{s} > \hat{s}$ so close to $\hat{s}$ that $\hat{s} < s$ and pick $\hat{L}$ such that $L_{\text{crit}} < \hat{L} < L$ and $\gamma(\hat{s}) \in \mathcal{K}(\hat{L})$. Then $\gamma(s) \in \mathcal{K}(\hat{L}),$ and $\mathcal{K}(\hat{L}) \subseteq \text{Int}(\mathcal{K}(L))$. So $\gamma(s) \in \text{Int}(\mathcal{K}(L))$.

Summarizing, we have proved

(V) Let $\alpha, \beta, \theta, \delta_0, \delta_1, \delta_2, A, B, C$ be positive numbers such that (4.2.22)-(4.2.29) and (4.2.47) hold. Let $L_{\text{crit}}$ be defined by (4.2.57). Then, if $L > L_{\text{crit}}$ and $\gamma \in \Gamma(\delta_0, \delta_1, \delta_2)$ is such that $\gamma(t) \in \mathcal{K}(L)$ for some $t \in [0, \infty)$, it follows that $\gamma(s) \in \text{Int}(\mathcal{K}(L))$ for $t < s$. 

Now let $L_{\text{crit}} < L_1 < L_2$, so that $\mathcal{K}(L_1) \subseteq \text{Int}\mathcal{K}(L_2)$, and divide the closed set $F = \mathcal{K}(L_2) - \text{Int}(\mathcal{K}(L_1))$ into the six closed subsets $F_i$ shown in Figure 4.2, defined by

\[
F_1 = \{(x, y) \in F : x \leq 0, y \geq 2\theta\},
\]
\[
F_2 = \{(x, y) \in F : x \geq 0, y \geq 2\theta\},
\]
\[
F_3 = \{(x, y) \in F : x > 0, |y| \leq 2\theta\},
\]
\[
F_4 = \{(x, y) \in F : x \geq 0, y \leq -2\theta\},
\]
\[
F_5 = \{(x, y) \in F : x \leq 0, y \leq -2\theta\},
\]
\[
F_6 = \{(x, y) \in F : x < 0, y \geq -2\theta\}.
\]

Assume that $\gamma \in \Gamma(\delta_0, \delta_1, \delta_2)$ is such that $\gamma(0) \in \mathcal{K}(L_2)$. We want to show that $\gamma(t) \in \mathcal{K}(L_1)$ for some $t > 0$ (in which case we will have $\gamma(s) \in \mathcal{K}(L_1)$ for all $s \geq t$ as well).

Assume that this is not so, i.e. that $\gamma(t) \in \mathcal{K}(L_2) - \mathcal{K}(L_1)$ for all $t \geq 0$. Suppose there is $t$ such that $\gamma(t) \in F_1$. Then it is not possible for $\gamma(s)$ to belong to $F_1$ for all $s \geq t$. Indeed, if $\gamma(s) \in F_1$ for all $s \geq t$, then the function $s \rightarrow V_{1,\alpha,\beta}(\gamma(s))$ is absolutely
continuous and its derivative is bounded above by \(-\lambda_1\), where
\[
\lambda_1 = (4\theta - 2\beta)(A - 1)\delta_2 - 2\alpha\delta_1 - 2\beta\omega \tag{4.2.60}
\]
(cf. (4.2.35)), so that \(\lambda_1 > 0\) thanks to (4.2.27). If
\[
w_1 = \sup\{V_{1,\alpha,\beta}(p) : p \in F_1\}, \tag{4.2.61}
\]
and \(T_1 = \frac{w_1}{\lambda_1} + 1\), it would follow that \(V_{1,\alpha,\beta}(\gamma(t_1 + t)) < 0\), which is a contradiction. Hence there must be a \(T\) such that \(0 < T < T_1\) and \(\gamma(t + T) \notin F_1\).

Assume now that \(\gamma \in \Gamma_c(\delta_0, \delta_1, \delta_2)\). Let \(\bar{T}\) be the infimum of those \(T \in [0, T_1]\) such that \(\gamma(t + T) \notin F_1\). Then \(\gamma(t + \bar{T}) \in F_1\) but there are points \(s > \bar{T}\) arbitrarily close to \(\bar{T}\) such that \(\gamma(t + s) \notin F_1\). Now let \(S_1\) be the segment joining \(P_0(L_2)\) and \(P_0(L_1)\), and let \(S_2\) be the segment from \(P_1(L_1)\) to \(P_2(L_2)\). Then the boundary \(\partial F_1\) of \(F_1\) is the union of \(S_1, C_1(L_1), S_2\) and \(C_1(L_2)\). If \(\gamma(t + \bar{T}) \in C_1(L_2)\), then we know that \(\gamma(t + s) \in K(L_2)\) for \(s > \bar{T}\). But then \(\gamma(t + s) \in F_1\) for \(s > \bar{T}\), \(s\) near \(\bar{T}\), unless \(\gamma(t + \bar{T}) = P_0(L_2)\) or \(\gamma(t + \bar{T}) = P_1(L_2)\). So \(\gamma(t + \bar{T})\) cannot belong to the arc \(C_1(L_2)\) unless it is one of its endpoints.

Assume that \(\gamma(t + \bar{T}) \in S_1\). The derivative \(\dot{y}\) of the \(y\)-component of \(\gamma\) is given by
\[
\dot{y} = -\omega x - \sigma(y - e_0) + e_2. \tag{4.2.62}
\]
At time \(t + \bar{T}\) we have \(-x \geq L_1\) and \(|y - e_0| \leq 3\theta\), so
\[
\dot{y} \geq \omega L_1 - \psi(\theta) - \delta_2. \tag{4.2.63}
\]
It follows from (4.2.57) that
\[
\omega L_{\text{crit}} > \delta_2 + \psi(\theta). \tag{4.2.64}
\]
Since \(L_1 > L_{\text{crit}}\), we see that \(\dot{y} > 0\) at time \(t + \bar{T}\). Hence \(\gamma(t + s) \in F_1\) for \(s > \bar{T}\), \(s\) near \(\bar{T}\), unless \(\gamma(t + \bar{T})\) is one of the endpoints of \(S_1\). If \(\gamma(t + \bar{T}) = P_0(L_2)\) then both inequalities \(\dot{V}_{1,\alpha,\beta} < 0\) and \(\dot{y} > 0\) hold at time \(t + \bar{T}\), \(s\) near \(\bar{T}\).

So far, we have established that every curve \(\gamma \in \Gamma_c(\delta_0, \delta_1, \delta_2)\) such that \(\gamma(t) \in F_1\) for some \(t\) must hit the set \(C_1(L_1) \cup S_2\) at some time \(i\) such that \(t \leq i \leq t + T_1\). If \(\gamma\) is an arbitrary curve in \(\Gamma(\delta_0, \delta_1, \delta_2)\), we can approximate \(\gamma\) by curves in \(\Gamma_c(\delta_0, \delta_1, \delta_2)\).
Since $C_1(L_1) \cup S_2$ is closed, the conclusion holds for $\gamma \in \Gamma(\delta_0, \delta_1, \delta_2)$ as well. Finally, if $\gamma \in \Gamma(\delta_0, \delta_1, \delta_2)$ never enters $K(L_1)$, then $\gamma(\hat{t})$ must belong to $S_2$, and then $\gamma(\hat{t}) \in F_2$. So we have shown

(VI$_1$) if $\gamma \in \Gamma(\delta_0, \delta_1, \delta_2)$ and $\gamma$ is entirely contained in $K(L_2) - K(L_1)$, then for every $t \in [0, \infty)$ such that $\gamma(t) \in F_1$ there is a $\hat{t}$ such that $t \leq \hat{t} \leq t + T_1$ such that $\gamma(\hat{t}) \in S_2$.

A similar argument can be applied to $F_2$. In this case we use $V_{2,\alpha,\beta}$ instead of $V_{1,\alpha,\beta}$, and rely on the bound $\dot{V}_{2,\alpha,\beta} \leq -\lambda_2$ (cf. (4.2.33)), where

$$\lambda_2 = 4\theta(A - 1)\delta_2 - 2\alpha\delta_1,$$  \hspace{1cm} (4.2.65)

so that $\lambda_2 > 0$ thanks to (4.2.28). To exclude the possibility of exiting through $S_2$, we observe that on $S_2$ we have

$$\dot{x} = \omega y + e_1,$$  \hspace{1cm} (4.2.66)

so that

$$\dot{x} > \omega(A_1 + 2\theta) - \delta_1,$$  \hspace{1cm} (4.2.67)

and then $\dot{x} > 0$ (because $\delta_1 < 2\theta\omega$ thanks to (4.2.24) and (4.2.26)).

Hence, for a suitably chosen $T_2$,

(VI$_2$) the conclusion of (VI$_1$) also holds if we replace $F_1, T_1$ by $F_2, T_2$, and $S_2$ by $S_3$.

Here, of course, $S_3$ is the segment joining $P_2(L_1)$ to $P_2(L_2)$.

Next we show that the same conclusions hold for $F_3$. On $F_3$ we have

$$\dot{y} = -\omega x - \sigma(y - e_0) + e_2$$  \hspace{1cm} (4.2.68)

so that

$$\dot{y} \leq -\omega \tilde{L}_1 + \psi(\theta) + \delta_2.$$  \hspace{1cm} (4.2.69)

Recalling that $\tilde{L}_1 \geq L_1 - 2\alpha$, we have

$$\dot{y} \leq 2\alpha \omega + \psi(\theta) + \delta_2 - \omega L_1.$$  \hspace{1cm} (4.2.70)
Now, (4.2.57) implies that

\[ \omega L_{\text{crit}} > 2\alpha \omega + \psi(\theta) + \delta_2. \quad (4.2.71) \]

So, if we let

\[ \lambda_3 = \omega L_{\text{crit}} - (2\alpha \omega + \psi(\theta) + \delta_2), \quad (4.2.72) \]

we find \( \dot{y} \leq -\lambda_3 \). Therefore every trajectory \( \gamma \in \Gamma(\delta_0, \delta_1, \delta_2) \) that goes through a point of \( F_3 \) must leave \( F_3 \) in time not greater than \( T_3 \), where \( T_3 = \frac{4\delta}{\lambda_3} \). If \( \gamma \) is entirely contained in \( K(L_2) - K(L_1) \), then \( \gamma \) must exit through \( S_4 \), where \( S_4 \) is the segment joining \( P_3(L_2) \) and \( P_3(L_2) \).

If, for \( i = 1, 2, 3, 4, 5, 6 \), we let \( j = i + 1 \) for \( i < 6 \), \( j = 1 \) for \( i = 6 \), and define the segments \( S_6, S_5 \) in an obvious way, we have shown

(VI) there exists a \( T > 0 \) such that, if \( \gamma \) is any trajectory in \( \Gamma(\delta_0, \delta_1, \delta_2) \) such that \( \gamma \)

is entirely contained in \( K(L_2) - K(L_1) \), then, for every \( i \in \{1, 2, 3, 4, 5, 6\} \), \( t \geq 0 \),

if \( \gamma(t) \in F_i \), then there is a \( \tilde{t} \) such that \( t \leq \tilde{t} \leq t + T \) such that \( \gamma(\tilde{t}) \in S_j \).

If \( \gamma \in \Gamma(\delta_0, \delta_1, \delta_2) \) and \( \Gamma \) is contained in \( K(L_2) - K(L_1) \), it follows that \( \gamma(t_0) \) must belong to \( S_1 \) for some \( t_0 \geq 0 \). Then there must exist a \( \hat{t}_0 > t_0 \) such that \( \gamma(t) \in F_1 \) for \( t_0 \leq t \leq \hat{t}_0 \) and \( \gamma(\hat{t}_0) \in S_2 \). And then there must be a \( t_1 > \hat{t}_0 \) such that \( \gamma(t_1) \in S_1 \).

Now, the norm of the right hand side of (4.2.9) is bounded by a fixed constant \( K_1 \) as long as \( (x, y) \in K(L_2) \) and \( |e_i| < \delta_i, i = 0, 1, 2 \). If \( \Delta \) denotes the distance between \( S_1 \) and \( S_2 \), we must have the lower bound

\[ \hat{t}_0 - t_0 \geq \frac{\Delta}{K_1}. \quad (4.2.73) \]

On the other hand, on the interval \([t_0, \hat{t}_0]\) we have \( \dot{V}_{1,\alpha,\beta} \leq -\lambda_1 \). Therefore

\[ V_{1,\alpha,\beta}(\gamma(\hat{t}_0)) \leq V_{1,\alpha,\beta}(\gamma(t_0)) - \frac{\lambda_1 \Delta}{K_1}. \quad (4.2.74) \]

The point \( \gamma(\hat{t}_0) \) belongs to \( C(L) \) for a unique \( L \) (which satisfies \( L_1 \leq L \leq L_2 \)). Then \( \gamma(t) \in K(L) \) for \( t \geq \hat{t}_0 \). In particular, \( \gamma(t_1) \in K(L) \). So \( \gamma(t_1) \in C(L^*) \) for some \( L^* \) such that \( L_1 \leq L^* \leq L \). But then \( \gamma(t_1) \in C_1(L^*) \) and, since \( L^* \leq L \), it follows that \( V_{1,\alpha,\beta}(\gamma(t_1)) \leq V_{1,\alpha,\beta}(\gamma(t_0)) \). So we have proved:
(A) there is a $t_0$ such that $\gamma(t_0) \in S_1$,

(B) for every $t_0$ such that $\gamma(t_0) \in S_1$ there is a $t_1 > t_0$ such that $\gamma(t_1) \in S_1$ and

$$V_{1, \alpha, \beta}(\gamma(t_1)) \leq V_{1, \alpha, \beta}(\gamma(t_0)) - \frac{k\lambda_1\Delta}{K_1}.$$ 

Iterating (B) we get times $t_k$ such that $\gamma(t_k) \in S_1$ and

$$V_{1, \alpha, \beta}(\gamma(t_k)) \leq V_{1, \alpha, \beta}(\gamma(t_0)) - \frac{k\lambda_1\Delta}{K_1}. \tag{4.2.75}$$

This is clearly impossible, since it would imply that $V_{1, \alpha, \beta}$ takes on negative values. So we have reached a contradiction, due to the assumption that there is a $\gamma \in \Gamma(\delta_0, \delta_1, \delta_2)$ that goes through a point in $K(L_2)$ but never enters $K(L_1)$. This proves:

(VII) if $\alpha, \beta, \theta, \delta_0, \delta_1, \delta_2, A, B, C$ are positive numbers such that (4.2.22)-(4.2.29) and (4.2.47) hold, and $L_{\text{crit}}$ is defined by (4.2.57), then, if $L > L_{\text{crit}}$ and $\gamma \in \Gamma(\delta_0, \delta_1, \delta_2)$, there exists a $T \geq 0$ such that $\gamma(t) \in K(L)$ for all $t \geq T$.

We now observe that the region $K(L)$ is contained in the disk

$$D(R) = \{ p \in \mathbb{R}^2 : \|p\| \leq R \} \tag{4.2.76}$$

provided that

$$R \geq L + 2\theta. \tag{4.2.77}$$

Therefore (VII) implies:

(VIII) if $\alpha, \beta, \theta, \delta_0, \delta_1, \delta_2, A, B, C$ are positive numbers such that (4.2.22)-(4.2.29) and (4.2.47) hold, $L_{\text{crit}}$ is defined by (4.2.57), and $R > L_{\text{crit}} + 2\theta$, then every $\gamma \in \Gamma(\delta_0, \delta_1, \delta_2)$ satisfies

$$\limsup_{t \to +\infty} \|\gamma(t)\| < R. \tag{4.2.78}$$

To conclude our proof, it suffices to show that, for every $\varepsilon > 0$, one can always choose $\alpha, \beta, \theta, \delta_0, \delta_1, \delta_2, A, B, C$ so as to satisfy (4.2.22)-(4.2.29) and (4.2.47), and such that $L_{\text{crit}} + 2\theta < \varepsilon$.

Choose

$$\delta_0 = \theta, \; \alpha = K_1 \theta, \; \beta = \frac{\alpha}{4}, \; \delta_1 = \omega \beta, \; \delta_2 = K_2 \theta. \tag{4.2.79}$$
Then (4.2.22), (4.2.24) and (4.2.47) hold trivially. From (4.2.14) and (4.2.15) we know that for any positive number \( r < 1 \) there is a \( \bar{\theta} \) such that
\[
\frac{\varphi(\theta)}{\theta} > r\kappa, \quad \frac{\psi(\theta)}{\theta} < 4\kappa \tag{4.2.80}
\]
for \( 0 < \theta < \bar{\theta} \). From now on, we assume \( 0 < \theta < \bar{\theta} \). Then, (4.2.25) holds if
\[
AK_2 + K_1 \omega \leq r\kappa, \tag{4.2.81}
\]
(4.2.26) holds if
\[
K_1 < 8, \tag{4.2.82}
\]
(4.2.27) holds if
\[
\frac{\omega}{2} K_1^2 < (2 - \frac{K_1}{4})BK_2, \tag{4.2.83}
\]
and (4.2.28) holds if
\[
K_1 < CK_2. \tag{4.2.84}
\]
Let
\[
K_1 \leq \min\{1, \frac{8}{\omega}\}. \tag{4.2.85}
\]
Then (4.2.23) and (4.2.82) holds, and (4.2.83) will follow if we have
\[
2\omega K_1 < 7BK_2. \tag{4.2.86}
\]
So we must find \( K_1, K_2, A, B, C \) so that (4.2.81), (4.2.84), (4.2.85), (4.2.86), and (4.2.29) hold.

Pick
\[
K_2 = \rho \omega K_1. \tag{4.2.87}
\]
Then (4.2.81) is equivalent to \((\rho A + 1)K_1 \omega \leq r\kappa\), i.e. to
\[
K_1 \leq \frac{r\kappa}{(\rho A + 1)\omega}, \tag{4.2.88}
\]
(4.2.86) is equivalent to
\[
\frac{2}{\bar{\theta} \rho} < B, \tag{4.2.89}
\]
(4.2.84) is equivalent to
\[
\frac{1}{\rho \omega} < C. \tag{4.2.90}
\]
To satisfy (4.2.89), (4.2.90), and (4.2.29), we pick

\[ A = \frac{2}{\lambda \rho} + \frac{1}{\rho \omega} + 1, \quad B = C = A - 1. \]  \hspace{1cm} (4.2.91)

Note that we want to choose \( K_1 \) (i.e. \( \alpha \)) as large as possible so that \( L_{\text{crit}} \) defined in (4.2.57) is as small as possible. By choosing \( r \) sufficiently close to 1 and \( \rho \) sufficiently small, we see that the best possible estimate of \( K_1 \) is about \( \frac{7\kappa}{7 + 10\omega} \). For simplicity, we now take \( r = \frac{6}{7} \) and \( \rho = \frac{1}{7} \). Therefore (4.2.89) is written as \( B > 2 \), (4.2.90) is written as \( C > \frac{7}{\omega} \), and we can take \( A = \frac{7}{\omega} + 3 \). With this choice of \( A \), (4.2.88) is equivalent to

\[ K_1 \leq \frac{6\kappa}{7 + 10\omega}. \]  \hspace{1cm} (4.2.92)

So, to satisfy (4.2.22)-(4.2.29), it suffices to require \( K_1 \) satisfy (4.2.85) and (4.2.92). Let

\[ K_1 = \min \left\{ 1, \frac{8\kappa}{\omega}, \frac{6\kappa}{7 + 10\omega} \right\}. \]  \hspace{1cm} (4.2.93)

Then (4.2.85) and (4.2.92) hold. So, if we define \( K_1 \) by (4.2.93), then from (4.2.79), (4.2.87), and (4.2.91) we can find a choice of \( \alpha, \beta, \theta, \delta_0, \delta_1, \delta_2, A, B, C \) that satisfy (4.2.22)-(4.2.29). In particular, we have

\[ \delta_0 = \theta, \quad \delta_1 = \frac{\omega K_1}{4} \theta, \quad \delta_2 = \frac{\omega K_1}{7} \theta. \]  \hspace{1cm} (4.2.94)

Now, we compute \( L_{\text{crit}} \) defined by (4.2.57). From (4.2.79), (4.2.87), and (4.2.80), it follows that \( \frac{8\kappa^2}{\alpha} = \frac{8\kappa}{K_1}, \quad \frac{4\kappa}{\omega} = \theta, \quad 2\alpha = 2K_1 \theta, \quad \frac{8\kappa}{\omega} = \frac{K_1 \theta}{7}, \quad \text{and} \quad \frac{4\kappa}{\omega} \leq \frac{4\kappa}{\omega} \theta \). So \( L_{\text{crit}} \leq K \theta \), where \( K = \frac{8}{K_1} + \frac{15}{7} K_1 + \frac{4\kappa}{\omega} + 1 \). If we choose \( \theta \) such that \( 0 < \theta < \frac{\bar{\theta}}{7} \) and \((K + 2)\theta < \varepsilon\), then, using (VIII), we conclude \( \limsup_{t \to \infty} ||\gamma(t)|| < \varepsilon \) for all \( \gamma \in \Gamma(\delta_0, \delta_1, \delta_2) \). Therefore, if we let \( \nu_0 = K + 2, \quad \nu_1 = \frac{4\kappa}{\omega K_1}, \quad \nu_2 = \frac{7\kappa}{\omega K_1}, \quad \text{and} \quad \varepsilon = \nu_0 \bar{\theta}, \) then our conclusion follows. \( \Box \)

**Remark 4.2.3** The proof of Lemma 4.2.2 actually yields explicit formulas for \( \nu_0, \nu_1, \nu_2 \) and \( \varepsilon \). We can choose \( \varepsilon = \nu_0 \bar{\theta} \), where \( \bar{\theta} \) is any number such that \( \varphi(\theta) > \frac{6}{7} \kappa \theta \) and \( \psi(\theta) < 4\kappa \theta \) for \( 0 < \theta < \frac{\bar{\theta}}{7} \), and \( \nu_0 = (K + 2) = \frac{8}{K_1} + \frac{15}{7} K_1 + \frac{4\kappa}{\omega} + 3 \), \( \nu_1 = \frac{4\kappa}{\omega K_1} \), \( \nu_2 = \frac{7\kappa}{\omega K_1} \), where \( K_1 = \min \left\{ 1, \frac{8}{\omega}, \frac{6\kappa}{7 + 10\omega} \right\} \). In particular, if \( \sigma(s) = \text{sat}(s) \), then \( \kappa = 1 \), and \( K_1 = \frac{6}{7 + 10\omega} \). Therefore, \( \nu_0 = \frac{37}{3} + \frac{40}{3} \omega + \frac{4}{\omega} + \frac{90}{49 + 10\omega} \). When \( \omega = 1 \), from this formula we have \( \nu_0 < 32 \). \( \Box \)
Corollary 4.2.4  For \( n = 1 \) or \( 2 \), let \( J \) be an \( n \times n \) skew-symmetric matrix. Let \( b = 1 \) if \( n = 1 \), and \( b = (0, 1)' \) if \( n = 2 \). Assume that \((J, b)\) is a controllable pair. Then for every \( \sigma \in S \) there exist functions \( \epsilon \to \delta_1(\epsilon), \delta_2(\epsilon) \) from \((0, +\infty)\) to \((0, +\infty)\) such that

(I) whenever \( e_1 : [0, +\infty) \to \mathbb{R} \) and \( e_2 : [0, +\infty) \to \mathbb{R}^n \) are bounded measurable functions such that

\[
\limsup_{t \to +\infty} |e_i(t)| < \delta_i(\epsilon) \quad \text{for } i = 1, 2,
\]

then, if \( \gamma : [0, \infty) \to \mathbb{R}^n \) is any solution of the system

\[
\dot{x} = Jx - \sigma(x_n - e_1(t))b + e_2(t),
\]

where \( x \in \mathbb{R}^n \), it follows that

\[
\limsup_{t \to +\infty} ||\gamma(t)|| < \epsilon;
\]

and

(II) there exist an \( \bar{\epsilon} > 0 \) and \( \nu_1, \nu_2 > 0 \) such that \( \delta_i(\epsilon) = \frac{\epsilon}{\nu_i} \) for \( i = 1, 2 \) and \( 0 < \epsilon \leq \bar{\epsilon} \).

Proof. When \( n = 2 \), the conclusion follows from Lemma 4.2.2. If \( n = 1 \), then (4.2.95) becomes

\[
\dot{x} = -\sigma(x - e_1(t)) + e_2(t).
\]

Let \( \varphi : (0, +\infty) \to (0, +\infty) \) be the function defined by (4.2.11). Assume that \( \sigma'(0) = \kappa \).

Then, for any positive number \( \rho < 1 \), there exists \( \bar{\theta} > 0 \) such that

\[
\varphi(\theta) \geq \rho \kappa \theta \quad \text{for } 0 < \theta \leq \bar{\theta}.
\]

Given \( \epsilon > 0 \), let \( \delta_1, \delta_2 \) satisfy

\[
\delta_1 \leq \min\{\bar{\theta}, \frac{\epsilon}{2}\}, \quad (4.2.96)
\]

\[
\delta_2 < \rho \kappa \delta_1. \quad (4.2.97)
\]

Then, whenever \( |e_i(t)| \leq \delta_i \), we have

\[
|\sigma(x - e_1(t))| \geq \varphi(\delta_1) > \delta_2
\]
valid for $|z| \geq 2\delta_1$. Let $V(x) = |z|^2$. Then, if $|z| \geq 2\delta_1$, it follows that

$$\dot{V}(x) = -2z(\sigma(x - e_1) - e_2) < -4\delta_1(\varphi(\delta_1) - \delta_2) < 0.$$ 

Therefore, $z(t)$ will be eventually $< 2\delta_1 \leq \varepsilon$. Let $\nu_1 = 2, \nu_2 > \frac{2}{\mu}, \varepsilon = 2\tilde{\theta}$. Our conclusion then follows. \hfill \Box

**Corollary 4.2.5** Let $J, b, n, \sigma$ be as in the statement of Corollary 4.2.4. Let $e_1 : [0, \infty) \to \mathbb{R}$ and $e_2 : [0, \infty) \to \mathbb{R}^n$ be bounded measurable functions such that $\lim_{t \to \infty} e_1(t) = 0$. Let $\gamma : [0, \infty) \to \mathbb{R}^n$ be a solution of (4.2.95). Then $\lim_{t \to \infty} \gamma(t) = 0$. \hfill \Box

### 4.3 The Proof of Theorem 4.1

First, we notice that under the conditions of the theorem there exists a linear change of coordinates of the state space that transforms $\Sigma$ into the block form

$$\Sigma : \begin{cases}
\dot{x}_1 & = A_1 x_1 + B_1 u, \; x_1 \in \mathbb{R}^{n_1}, \\
\dot{x}_2 & = A_2 x_2 + B_2 u, \; x_2 \in \mathbb{R}^{n_2},
\end{cases}$$

where (i) $n_1 + n_2 = n$, (ii) all the eigenvalues of $A_1$ have zero real part, (iii) all the eigenvalues of $A_2$ have negative real part, and (iv) $(A_1, B_1)$ is a controllable pair. Suppose that we find a SISS$_L$-stabilizing feedback $u = k(x_1)$ of Type $\mathcal{F}$ or Type $\mathcal{G}$ for the system $\dot{x}_1 = A_1 x_1 + B_1 u$ such that the linearization of the resulting closed-loop system is asymptotically stable. Then it is clear that this same feedback law will work for $\Sigma$ as well. Thus, in order to stabilize $\Sigma$, it is enough to stabilize the "critical subsystem" $\dot{x}_1 = A_1 x_1 + B_1 u$. Without loss of generality, in our proof of the theorem we will suppose that $\Sigma$ is already in this form.

We start with the single-input case, and prove the theorem by induction on the dimension of the system. As discussed earlier, we may assume that all the eigenvalues of $A$ have zero real part and the pair $(A, B)$ is controllable.

For dimension zero, there is nothing to prove. Now assume that we are given a single-input $n$-dimensional system, $n \geq 1$, and suppose that Theorem 4.1 has been
established for all single-input systems of dimension $\leq n - 1$. We consider separately the following two possibilities:

(i) zero is an eigenvalue of $A$,

(ii) zero is not an eigenvalue of $A$.

Recall that $\mu = \mu(A)$. We want to prove that for any finite sequence $\sigma = (\sigma_1, \ldots, \sigma_\mu)$ of bounded functions in $S$, there are $SSS_L$-stabilizing feedbacks $u = -k_F(x)$ and $u = -k_G(x)$ such that $k_F \in F_n(\sigma)$, $k_G \in G_n(\sigma)$, and the linearizations of the resulting closed-loop systems are asymptotically stable.

Assume that $\sigma_\mu'(0) = \kappa > 0$. In Case (i) we apply Part (i) of Lemma 4.2.1 and rewrite our system in the form

\begin{align*}
\ddot{\tilde{y}} &= A_1 \tilde{y} + (y_n + \frac{1}{\kappa} u)b_1, \\
\dot{y}_{n-1} &= \omega y_n,
\end{align*}

where $\tilde{y} = (y_1, \ldots, y_{n-1})'$. (Note that if $n = 1$, only the second equation occurs.) In Case (ii), since $n > 0$, $A$ has an eigenvalue of the form $i\omega$ with $\omega > 0$. So we apply Part (ii) of Lemma 4.2.1 and make a linear transformation that puts $\Sigma$ in the form

\begin{align*}
\ddot{\tilde{y}} &= A_1 \tilde{y} + (y_n + \frac{1}{\kappa} u)b_1, \\
\dot{y}_{n-1} &= \omega y_n, \\
\dot{y}_n &= -\omega y_{n-1} + u,
\end{align*}

where $\tilde{y} = (y_1, y_2, \ldots, y_{n-2})'$. (Naturally, in the special case when $n = 2$, the first equation will be missing.) So, in any case, by using Lemma 4.2.1, we can rewrite our system in the form

\begin{align*}
\ddot{\tilde{y}} &= A_1 \tilde{y} + (y_n + \frac{1}{\kappa} u)b_1, \\
\dot{\tilde{y}} &= J \tilde{y} + ub_0,
\end{align*}

where $J$ is a skew-symmetric matrix, $(J, b_0)$ is a controllable pair, and in Case (i) we have $\tilde{y} = y_n, b_0 = 1$, and in Case (ii) we have $\tilde{y} = (y_{n-1}, y_n)'$, $b_0 = (0, 1)'$.

Let

\begin{equation}
\begin{array}{c}
u = -\sigma_\mu(y_n - \xi v) - \eta v,
\end{array}
\end{equation}

(4.3.4)
where $\xi$ and $\eta$ are constants such that $\xi \eta = 0, \xi + \eta = 1$, and $v$ is to be chosen later. Define $g(s) = \sigma_\mu(s) - \kappa s$. Then

$$y_n + \frac{1}{\kappa} u = (\xi - \frac{1}{\kappa} \eta) v - \frac{1}{\kappa} g(y_n - \xi v).$$

Therefore the first equation of (4.3.3) is written as

$$\dot{\bar{y}} = A_1 \bar{y} + vb_2 - g(y_n - \xi v)b_3,$$  \hspace{1cm} (4.3.5)

where $b_2 = (\xi - \frac{1}{\kappa} \eta)b_1$, $b_3 = \frac{1}{\kappa} b_1$. Notice that $\xi - \frac{1}{\kappa} \eta \neq 0$, and therefore $(A_1, b_2)$ is still a controllable pair. From the inductive hypothesis we know that there exist $\bar{k}_f \in \mathcal{F}_n(\sigma_1, \cdots, \sigma_{\mu-1})$ and $\bar{k}_g \in \mathcal{G}_n(\sigma_1, \cdots, \sigma_{\mu-1})$ (for simplicity we use $\bar{k}$ for both cases) such that the system $\dot{\bar{y}} = A_1 \bar{y} - \bar{k}(\bar{y})b_2$ is $SISL(\bar{\Delta}, \bar{N})$ for some $\bar{\Delta}, \bar{N} > 0$, and the linearization of the closed-loop system is asymptotically stable. (For instance, in Case (ii) there exists a $SISL$-stabilizing feedback $u = -\bar{k}_f$ for $\dot{\bar{y}} = A_1 \bar{y} + b_2 v$ such that $\bar{k}_f \in \mathcal{F}_{n-2}(\sigma_1, \cdots, \sigma_{\mu-1})$. With the natural extension, $\bar{k}_f$ can be also viewed as a function from $\mathbb{R}^n$ to $\mathbb{R}$. Thus $\bar{k} \in \mathcal{F}_n(\sigma_1, \cdots, \sigma_{\mu-1})$.) Let $\lambda > 0$. Then $\dot{\bar{y}} = A_1 \bar{y} - \lambda \bar{k}(\frac{\bar{y}}{\lambda})b_2$ is $SISL(\lambda \bar{\Delta}, \bar{N})$.

Also, since $\bar{k}$ is bounded and locally Lipschitz, and $\bar{k}(0) = 0$, it follows that there exist positive numbers $L, M$ such that

$$|\bar{k}(\bar{y})| \leq L|\bar{y}|$$

and

$$|\bar{k}(\bar{y})| \leq M$$

are valid for all $\bar{y}$.

On the other hand, the second equation of (4.3.3) is written as

$$\dot{\bar{y}} = J \bar{y} - \sigma_\mu(y_n - \xi v)b_0 - \eta vb_0.$$  \hspace{1cm} (4.3.6)

From Corollary 4.2.4, we see that there exist $\Delta_1, \nu_1, \nu_2 > 0$ such that, whenever $\delta \leq \Delta_1$, and $e_1, e_2$ are bounded measurable functions such that $|e_i| \leq e_0 \frac{\delta}{\nu_i}$, $i = 1, 2$, it follows that all the solutions of $\dot{\bar{y}} = J \bar{y} - \sigma_\mu(y_n - e_1) + e_2$ will be eventually bounded by $\delta$. 
Recall that $\sigma_\mu$ is differentiable at 0. So for any $\varepsilon > 0$ there exists $\Delta_2 > 0$ such that

$$|g(s)| \leq \varepsilon |s| \quad \text{for } |s| \leq \Delta_2. \quad (4.3.7)$$

Now, fix a positive number $\alpha < 1$. (To get a feedback of Type $\mathcal{F}$, we choose $\alpha$ close to 1; while to get a feedback of Type $\mathcal{G}$, we choose $\alpha$ close to 0.) Choose $\varepsilon > 0$ so that

$$\varepsilon \left( M(\xi_1 + \eta \xi_2 + \xi) + \Delta_1 \right) |b_0| < \alpha \Delta, \quad (4.3.8)$$

and

$$\varepsilon \tilde{N} L(\xi_1 + \eta \xi_2 + \xi) < \frac{1}{4}. \quad (4.3.9)$$

Then there exists $\Delta_2 > 0$ such that (4.3.7) is satisfied. (If $\sigma_\mu$ is linear near 0, then we can take $\varepsilon = 0$. Therefore (4.3.8) and (4.3.9) are valid and there exists $\Delta_2 > 0$ such that (4.3.7) is still satisfied.)

Now, choose $\delta_0 > 0$ so that

$$\delta_0 \leq \min\{\Delta_1, \alpha \Delta_2\}, \quad (4.3.10)$$

and let $\lambda > 0$ satisfy

$$\lambda \leq 1, \quad (4.3.11)$$

$$\lambda \xi M \leq (1 - \alpha) \Delta_2, \quad (4.3.12)$$

$$\lambda (\xi_1 + \eta \xi_2) M \leq \alpha \delta_0. \quad (4.3.13)$$

Then define

$$\Delta = \min\{\delta_0, \frac{(1 - \alpha) \lambda \delta_0}{\nu^2}, (1 - \alpha) \lambda \Delta\}.$$  

Let $v = -\lambda \tilde{k}(\frac{g}{\lambda})$ in (4.3.5) and (4.3.6). Then it is enough to show that the resulting closed-loop system

$$\dot{\bar{y}} = A_1 \bar{y} - \lambda \tilde{k}(\frac{g}{\lambda}) b_2 - g(y_n + \lambda \xi \tilde{k}(\frac{g}{\lambda})) b_3, \quad (4.3.14)$$

$$\dot{\bar{y}} = J \bar{y} - \sigma_\mu (y_n + \lambda \xi \tilde{k}(\frac{g}{\lambda})) b_0 + \lambda \eta \tilde{k}(\frac{g}{\lambda}) b_0$$

is $S I S S_L(\Delta, N)$ for some $N > 0$. Therefore, the conclusion of Theorem 4.1 for the single-input case is proved.

Indeed, note that the linearization $\Sigma_L$ of (4.3.14) is obviously asymptotically stable because the $\bar{y}$-part in $\Sigma_L$ has all eigenvalues with negative real parts, and the $\bar{y}$-part in
$\Sigma_L$ is just the linearization of $\hat{g} = A_1 \hat{y} - \bar{k}(\bar{y})b_2$ and therefore is asymptotically stable by the inductive hypothesis. As a consequence, the local asymptotic stability of (4.3.14) is guaranteed. To get a feedback of the form of (4.1.1), we take $\xi = 1, \eta = 0$. Then the feedback $u = -\sigma_\mu(y_n + \lambda \bar{k}(\frac{\bar{y}}{\bar{y}}))$ is as desired. To get a feedback of the form of (4.1.2), we take $\xi = 0, \eta = 1$. Then, for any $r > 0$, the closed-loop system of (4.3.3), with

$$u = -r\sigma_\mu\left(\frac{y_n}{r}\right) - r\lambda \bar{k}\left(\frac{\bar{y}}{r}\right), \quad (4.3.15)$$

is $SISS_L(r\Delta, N)$. Choosing $r$ sufficiently small, we can insure that the sum of the coefficients of all saturations in (4.3.15) is bounded by 1. It turns out that the feedback defined by (4.3.15) is as desired. So we only need to show that (4.3.14) is $SISS_L(\Delta, N)$ for some $N > 0$.

Given $\delta \leq \Delta$, let $\bar{e}, \bar{e} : \mathbb{R}_{\geq 0} \to \mathbb{R}$ be bounded measurable functions, eventually bounded by $\delta$, which have the same dimensions of $\bar{y}, \hat{y}$, respectively. We consider the following system:

$$\dot{\hat{y}} = A_1 \hat{y} - \lambda \bar{k}(\frac{\bar{y}}{\lambda})b_2 - g(y_n + \lambda \xi \bar{k}(\frac{\bar{y}}{\lambda}))b_3 + \bar{e}, \quad (4.3.16)$$
$$\dot{\bar{y}} = J \dot{\bar{y}} - \sigma_\mu(y_n + \lambda \xi \bar{k}(\frac{\bar{y}}{\lambda}))b_0 + \lambda \eta \bar{k}(\frac{\bar{y}}{\lambda})b_0 + \bar{\epsilon}.$$

From (4.3.13) we have

$$\lambda \xi |\bar{k}(\frac{\bar{y}}{\lambda})| \leq_{ev} \lambda M \leq \frac{d_0}{\nu_1}, \quad (4.3.17)$$

and (because $\delta \leq \frac{(1-\alpha)\lambda d_0}{\nu_2}$)

$$|\lambda \eta \bar{k}(\frac{\bar{y}}{\lambda})b_0 + \bar{\epsilon}| \leq_{ev} \lambda \eta M + \delta \leq \frac{d_0}{\nu_2}. \quad (4.3.18)$$

So applying Corollary 4.2.4 to the second equation in (4.3.16), we conclude that

$$|\bar{y}| \leq_{ev} \max\{\lambda \xi M \nu_1, \lambda \eta M \nu_2 + \delta \nu_2\} \leq \lambda M (\xi \nu_1 + \eta \nu_2) + \delta \nu_2, \quad (4.3.19)$$

and therefore

$$|y_n + \lambda \xi \bar{k}(\frac{\bar{y}}{\lambda})| \leq_{ev} \lambda M (\xi \nu_1 + \eta \nu_2) + \delta \nu_2 + \lambda \xi M. \quad (4.3.20)$$

Inequality (4.3.13) and $\delta \nu_2 \leq (1 - \alpha) d_0$ then imply that the right-hand side of (4.3.19) is bounded by $d_0$. Therefore, it follows from (4.3.10) that

$$|\hat{y}| \leq_{ev} \min\{\Delta_1, \alpha \Delta_2\}. \quad (4.3.21)$$
So \(|y_n + \lambda \xi \tilde{k}(\frac{\vartheta}{\lambda})| \leq_{ev} \Delta_2\) follows from (4.3.21) and (4.3.12). From (4.3.7) we then get

\[
g(y_n + \lambda \xi \tilde{k}(\frac{\vartheta}{\lambda})) \leq_{ev} \varepsilon |y_n + \lambda \xi \tilde{k}(\frac{\vartheta}{\lambda})|.
\]  

(4.3.22)

In particular, since \(\delta \nu_2 \leq (1 - \alpha)\lambda \delta_0\) and \(\delta_0 \leq \Delta_1\), using (4.3.20), we obtain

\[
|y_n + \lambda \xi \tilde{k}(\frac{\vartheta}{\lambda})| \leq_{ev} \lambda \left( M(\xi \nu_1 + \eta \nu_2 + \xi) + (1 - \alpha)\Delta_1 \right).
\]

Therefore

\[
|g(y_n + \lambda \xi \tilde{k}(\frac{\vartheta}{\lambda}))b_2 + \bar{c}| \leq_{ev} \lambda \varepsilon \left( M(\xi \nu_1 + \eta \nu_2 + \xi) + (1 - \alpha)\Delta_1 \right) ||b_2|| + \delta.
\]  

(4.3.23)

From (4.3.8) we see that the right-hand side of (4.3.23) is bounded by \(\lambda \bar{\Delta}\). Since the system \(\dot{\bar{y}} = A_1 \bar{y} - \lambda \tilde{k}(\frac{\vartheta}{\lambda})b_2\) is \(SISL(\lambda \bar{\Delta}, \bar{N})\), it follows that

\[
|\dot{\bar{y}}| \leq_{ev} \lambda \bar{N} \bar{\Delta}.
\]  

(4.3.24)

Now suppose \(\limsup_{t \to \infty} |\dot{\bar{y}}(t)| = \rho > 0\). Then \(|\dot{\bar{y}}| \leq_{ev} 2\rho\). Since \(|\tilde{k}(\bar{y})| \leq L|\bar{y}|\), we have

\[
|\lambda \xi \tilde{k}(\frac{\vartheta}{\lambda})| \leq_{ev} 2\xi L \rho
\]  

(4.3.25)

and

\[
|\lambda \eta \tilde{k}(\frac{\vartheta}{\lambda})b_0 + \bar{c}| \leq_{ev} 2\eta L \rho + \delta.
\]  

(4.3.26)

Applying Corollary 4.2.4 to the second equation in (4.3.16) in the light of (4.3.25) and (4.3.26), we obtain

\[
|\dot{\bar{y}}| \leq_{ev} \max \{2\xi \nu_1 L \rho, 2\eta \nu_2 L \rho + \nu_2 \delta\} \leq 2(\xi \nu_1 + \eta \nu_2) L \rho + \nu_2 \delta.
\]  

(4.3.27)

(Note that if the right-hand side of (4.3.27) is greater than \(\Delta_1\), then the inequality is trivial since we have already established the fact \(|\dot{\bar{y}}| \leq_{ev} \Delta_1\) in (4.3.21).) From (4.3.22), (4.3.25), and (4.3.27), we have

\[
|g(y_n + \lambda \xi \tilde{k}(\frac{\vartheta}{\lambda}))| \leq_{ev} \varepsilon \left( 2(\xi \nu_1 + \eta \nu_2 + \xi) L \rho + \nu_2 \delta \right).
\]

Now, if

\[
\varepsilon \left( 2(\xi \nu_1 + \eta \nu_2 + \xi) L \rho + \nu_2 \delta \right) + \delta \leq \lambda \bar{\Delta},
\]  

(4.3.28)
then, since the system \( \dot{y} = A_1 \dot{y} - \lambda \dot{k}\left(\frac{\mu}{\lambda}\right)b_2 \) is \( SISS_L(\lambda \bar{\Delta}, \bar{N}) \), it follows that

\[
|\dot{y}| \leq_c \varepsilon \bar{N} \left( 2(\xi \nu_1 + \eta \nu_2 + \xi) L \rho + \nu_2 \delta \right) + \bar{N} \delta. \tag{4.3.29}
\]

If (4.3.28) is not satisfied, then \( \varepsilon \bar{N} \left( 2(\xi \nu_1 + \eta \nu_2 + \xi) L \rho + \nu_2 \delta \right) + \bar{N} \delta > \lambda \bar{N} \bar{\Delta} \). However, from (4.3.24) we see that (4.3.29) still holds. So we have established (4.3.29) in all cases. From (4.3.9) we then get

\[
|\dot{y}| \leq_c \frac{1}{2} \rho + \bar{N}(\varepsilon \nu_2 + 1) \delta. \tag{4.3.30}
\]

Taking the \( \limsup_{\varepsilon \to 0} \) of the left-hand side of (4.3.30), we have

\[
\rho \leq \frac{1}{2} \rho + \bar{N}(\varepsilon \nu_2 + 1) \delta,
\]
i.e., \( \rho \leq 2 \bar{N}(\varepsilon \nu_2 + 1) \delta \). Substituting this into (4.3.27) and (4.3.30), we get

\[
|\dot{y}| \leq_c \left( 4(\xi \nu_1 + \eta \nu_2) L \bar{N}(\varepsilon \nu_2 + 1) + \nu_2 \right) \delta,
\]

and

\[
|\dot{y}| \leq_c 2 \bar{N}(\varepsilon \nu_2 + 1) \delta.
\]

So, if we take

\[
\bar{N} = 2 \bar{N}(\varepsilon \nu_2 + 1)(1 + 2L(\xi \nu_1 + \eta \nu_2)) + \nu_2,
\]

the conclusion then follows.

Next, we deal with the general case of \( m > 1 \) inputs and prove Theorem 4.1 by induction on \( m \).

First, we know from the proof above that the theorem is true if \( m = 1 \). Assume that Theorem 4.1 has been established for all \( k \)-input systems, for all \( k \leq m - 1 \), and consider an \( m \)-input system \( \dot{x} = Ax + Bu \).

Assume without loss of generality that the first column \( b_1 \) of \( B \) is nonzero and consider the Kalman controllability decomposition of the system \( \dot{x} = Ax + b_1 u \) (see [27], Lemma 3.3.3). We conclude that, under a change of coordinates \( y = T^{-1}x \), \( \Sigma_1 \) has the form

\[
\begin{align*}
\dot{y}_1 &= A_1 y_1 + A_2 y_2 + \bar{b}_1 u, \\
\dot{y}_2 &= A_3 y_2,
\end{align*}
\]
where \((A_1, \bar{b}_1)\) is a controllable pair. In these coordinates, \(\Sigma\) has the form

\[
\begin{align*}
\dot{y}_1 &= A_1 y_1 + A_2 y_2 + \bar{b}_1 u_1 + \bar{B}_1 \bar{u}, \\
\dot{y}_2 &= A_3 y_2 + \bar{B}_2 \bar{u},
\end{align*}
\]  

(4.3.31)

where \(\bar{u} = (u_2, \cdots, u_m)'\) and \(\bar{B}_1, \bar{B}_2\) are appropriate matrices. So it suffices to show the conclusion for (4.3.31). Let \(n_1, n_2\) denote the dimensions of \(y_1, y_2\), respectively. Recall that \(\mu = \mu(A)\). Let \(\sigma = (\sigma_1, \cdots, \sigma_\mu)\) be any finite sequence of bounded functions in \(S\). Then, for the single-input controllable system

\[
\dot{y}_1 = A_1 y_1 + \bar{b}_1 u_1,
\]

there is a feedback

\[
u_1 = -k_1(y_1) \quad (4.3.32)\]

such that (i) \(k_1 \in \mathcal{F}_{n_1}(\sigma_1, \cdots, \sigma_\mu)\) (respectively, \(k_1 \in \mathcal{G}_{n_1}(\sigma_1, \cdots, \sigma_\mu)\)), where \(\mu_1 = \mu(A_1)\); (ii) the resulting closed-loop system is \(SILLS(L_1, N_1)\) for some \(\Delta_1, N_1 > 0\); (iii) the linearization of the closed-loop system is asymptotically stable. Since (4.3.31) is controllable, we conclude that the \((m-1)\)-input subsystem \(\dot{y}_2 = A_3 y_2 + \bar{B}_2 \bar{u}\) is controllable as well. By the inductive hypothesis, this subsystem can be stabilized by a feedback

\[
u = -\bar{k}(y_2) = -(k_2(y_2), \cdots, k_m(y_2)) \quad (4.3.33)\]

such that (i) \(\bar{k} \in \mathcal{F}_{n_2}(\sigma_{\mu_1+1}, \cdots, \sigma_\mu)\) (respectively, \(\bar{k} \in \mathcal{G}_{n_2}(\sigma_{\mu_1+1}, \cdots, \sigma_\mu)\)), where \(\bar{I} = (\mu_2, \cdots, \mu_m)\) is an \((m - 1)\)-tuple of nonnegative integers and \(|\bar{I}| = \mu - \mu_1\); (ii) the resulting closed-loop system is \(SILLS(L_2, N_2)\) for some \(\Delta_2, N_2 > 0\); (iii) the linearization of the closed-loop system is asymptotically stable. Now, if the input of the closed-loop system of (4.3.31) with \(u_1, \bar{u}\) given by (4.3.32) and (4.3.33) is bounded measurable and eventually bounded by \(\delta \leq \Delta_2\), then \(|y_2| \leq_{cv} N_2 \delta\). Therefore there is a constant \(C > 0\) such that \(|A_2 y_2(t) - \bar{B}_1 \bar{k}(y_2)| \leq_{cv} C N_2 \delta\). So the feedback given in (4.3.32) and (4.3.33) globally stabilizes (4.3.31), and the resulting closed-loop system is \(SILLS(L, N)\) with \(\Delta = \min\{\Delta_1, \frac{\Delta_2}{2C N_2}, \frac{\Delta_1}{2}\}\), \(N = N_1 + N_2\). So if we let \(I = (\mu_1, \mu_2, \cdots, \mu_m)\) and \(k = (k_1(y_1), k_2(y_2), \cdots, k_m(y_2))\), then \(k \in \mathcal{F}_n(I)\) (respectively, \(k \in \mathcal{G}_n(I)\)) satisfies all the required properties as desired.

\(\Box\)
4.4 An Algorithm

We now present a two-step procedure for computing stabilizing feedbacks of the kind described in the statement of Theorem 4.1. The first step is to transform $\Sigma$ into a special form by means of a linear change of coordinates. The second step is the construction of a stabilizing feedback for the transformed system.

First, we present a lemma which is a direct consequence of the Kalman controllability decomposition.

**Lemma 4.4.1** Let $\Sigma$ be a linear system of the form

$$\dot{x} = Ax + Bu \tag{4.4.1}$$

with state space $\mathbb{R}^n$ and input space $\mathbb{R}^m$. Suppose that all the eigenvalues of $A$ have nonpositive real parts and all the eigenvalues of the uncontrollable part of $\Sigma$ have strictly negative real parts. Then there exists a linear change of coordinates which transforms (4.4.1) into the following form

$$\begin{align*}
\dot{x}_0 &= A_{00}x_0 + A_{01}x_1 + A_{02}x_2 + \cdots + A_{0m}x_m + b_{01}u_1 + b_{02}u_2 + \cdots + b_{0m}u_m, \\
\dot{x}_1 &= A_{11}x_1 + A_{12}x_2 + \cdots + A_{1m}x_m + b_{11}u_1 + b_{12}u_2 + \cdots + b_{1m}u_m, \\
\dot{x}_2 &= A_{22}x_2 + \cdots + A_{2m}x_m + b_{22}u_2 + \cdots + b_{2m}u_m, \\
&\vdots \\
\dot{x}_m &= A_{mm}x_m + b_{mm}u_m. 
\end{align*} \tag{4.4.2}$$

where all the eigenvalues of $A_{00}$ have negative real part, all the eigenvalues of $A_{ii}$ with $i \geq 1$ have zero real part, and all the pairs $(A_{ii}, b_{ii})$, $i \geq 1$, are controllable. (The coordinate $x_i$ may have zero dimension. In that case, there is no equation for $x_i$ in (4.4.2).) \hfill \square

From the proof of Theorem 4.1 we see that, if we find $SISL$-stabilizing feedbacks $u_i = k_i(x_i)$ of Type $\mathcal{F}$ or Type $\mathcal{G}$ for the systems $\Sigma_i : \dot{x}_i = A_{ii}x_i + b_{ii}u_i$ and let $k = (k_1, \cdots, k_m)$ then $k$ is $SISL$-stabilizing for $\Sigma$. So what we need is to stabilize each
\[ \Sigma \] separately, making sure that the \textit{SISS} property holds. To simplify the notation, we consider again an \( n \)-dimensional single-input controllable system of the form

\[ \dot{x} = Ax + bu , \quad (4.4.3) \]

where all the eigenvalues of \( A \) have zero real part. Our goal is to find \textit{SISS}-stabilizing feedbacks of Type \( F \) and Type \( G \) for (4.4.3). For this purpose, we present the next lemma, which is a consequence of Lemma 4.2.1.

**Lemma 4.4.2** Let \( \Sigma : \dot{x} = Ax + bu, \; x \in \mathbb{R}^n \), be a controllable single-input linear system. Suppose that the eigenvalues of \( A \) are:

\[ 0, 0, \cdots, 0, \; \pm \omega_1 i, \pm \omega_2 i, \cdots, \pm \omega_q i. \]

\( p \)

(So that \( p + 2q = n \.) Let \( (\nu_1, \cdots, \nu_{p+q-1}) \) be a finite sequence of positive numbers. Then there is a linear change of coordinates that puts \( \Sigma \) in the following form:

\[
\begin{align*}
\dot{z}_i &= \omega_i y_i \\
\dot{y}_i &= -\omega_i z_i + \sum_{k=i+1}^{q} \left( \prod_{h=i}^{k-2} \nu_h \right) y_k + \sum_{k=i}^{p} \left( \prod_{h=i}^{q+k-2} \nu_h \right) x_k + \left( \prod_{h=i}^{q+p-1} \nu_h \right) u \\
\dot{z}_j &= \sum_{k=j+1}^{p} \left( \prod_{h=q+j}^{q+k-2} \nu_h \right) z_k + \left( \prod_{h=q+j}^{q+p-1} \nu_h \right) u, \quad j = 1, 2, \cdots, p .
\end{align*}
\]

(4.4.4)

Here \( \prod_{h=l}^{m} \nu_h = 1 \) and \( \sum_{h=l}^{m} \nu_h = 0 \) if \( m < l \).

To see that this is a consequence of Lemma 4.2.1, we notice that, if \( p > 0 \), then from Part (i) of Lemma 4.2.1 it follows that there is a linear change of coordinates \((x_1, \cdots, x_n) \rightarrow (y_1, \cdots, y_n) = (\bar{y}', y_n)\) that transforms \( \Sigma \) into the form

\[ \begin{align*}
\dot{\bar{y}} &= A_1 \bar{y} + b_1 (y_n + \nu_{p+q-1} u) , \\
\dot{y}_n &= u .
\end{align*} \]

If \( p > 1 \), then applying Part (i) of Lemma 4.2.1 to the system \( \Sigma_1 : \dot{\bar{y}} = A_1 \bar{y} + b_1 v , \) and using \( \nu_{p+q-2} \) instead of \( \nu \), we find a linear change of coordinates \((y_1, \cdots, y_{n-1}) \rightarrow (z_1, \cdots, z_{n-1}) = (\bar{z}', z_{n-1})\) of \( \mathbb{R}^{n-1} \) that puts \( \Sigma_1 \) in the form

\[
\begin{align*}
\dot{\bar{z}} &= A_2 \bar{z} + b_2 (z_{n-1} + \nu_{p+q-2} v) , \\
\dot{z}_{n-1} &= v ,
\end{align*}
\]
where the pair \((A_2, b_2)\) is controllable. Substituting \(v = y_n + \nu_{p+q-1} u\) and defining \(z_n = y_n\), we see that \(\Sigma\) is transformed into the form

\[
\dot{z} = A_2 z + b_2 (z_{n-1} + \nu_{p+q-2} z_n + \nu_{p+q-2} \nu_{p+q-1} u),
\]

\[
\dot{z}_{n-1} = z_n + \nu_{p+q-1} u,
\]

\[
\dot{z}_n = u.
\]

Continuing in this way, we apply Part (i) or Part (ii) of Lemma 4.2.1 to the first equation of (4.2.2) or (4.2.3), using \(\nu_{p+q-3}, \nu_{p+q-4}\), instead of \(\nu\), until we obtain the representation (4.4.4).

To get a linear transformation that puts \(\Sigma\) into the form (4.4.4), simply let \((\tilde{A}, \tilde{b})\) denote the controllable pair corresponding to (4.4.4). Then \(y = T^{-1} x\), with \(T = R(A, b) R(\tilde{A}, \tilde{b})^{-1}\), is the desired transformation, where \(R(A, b)\) denotes the controllability matrix of \((A, b)\). (See Section 3.3 in [27].)

For simplicity, let \(s = p + q\) and let \(y_{q+i}\) denote \(x_i\). Then (4.4.4) is a system with state \((z_1, \ldots, z_q, y_1, \ldots, y_s)\).

Now let \((\sigma_1, \ldots, \sigma_s)\) be a finite sequence of bounded functions in \(S\). Suppose \(\sigma_i(0) = \kappa_i\). Choose \(\nu_i = \frac{1}{\kappa_{i+1}}\) in (4.4.4). Then, to get a stabilizing feedback of Type \(F\), we choose positive constants \(\lambda_1, \lambda_2, \ldots, \lambda_{s-1}\) so that the following feedback is as desired:

\[
u = -\sigma_s \left( y_s + \lambda_{s-1} \sigma_{s-1} \left( \frac{y_{s-1}}{\lambda_{s-1}} + \lambda_{s-2} \sigma_{s-2} \left( \frac{y_{s-2}}{\lambda_{s-2} \lambda_{s-1}} + \cdots + \lambda_1 \sigma_1 \left( \frac{y_1}{\lambda_1 \lambda_2 \cdots \lambda_{s-1}} \right) \right) \right) \right).
\]

(See Figure 4.3.) If we want to get a stabilizing feedback of Type \(G\), we need to choose positive numbers \(\lambda_1, \lambda_2, \ldots, \lambda_s\) so that the feedback

\[
u = -\lambda_s \sigma_s \left( \frac{y_s}{\lambda_s} \right) - \lambda_{s-1} \lambda_s \sigma_{s-1} \left( \frac{y_{s-1}}{\lambda_{s-1} \lambda_s} \right) - \cdots - \lambda_1 \lambda_2 \cdots \lambda_s \sigma_1 \left( \frac{y_1}{\lambda_1 \lambda_2 \cdots \lambda_s} \right)
\]

is as desired. (See Figure 4.4.)

These constants \(\lambda_1, \lambda_2, \ldots, \lambda_{s-1}\) (and \(\lambda_s\)) are chosen recursively. As in the proof of Theorem 4.1, we first find \((\Delta_1, N_1)\) so that the \((z_1, y_1)\)-part subsystem of (4.4.4) with \(u = -\sigma_1(y_1)\) is \(SISL(\Delta_1, N_1)\), then find \(\lambda_1\) so that the \((z_1, y_1, z_2, y_2)\)-part subsystem of (4.4.4) with \(u = -\sigma_2(y_2) - \lambda_1 \sigma_1 \left( \frac{y_1}{\lambda_1} \right)\) or \(u = -\sigma_2(y_2) - \lambda_1 \sigma_1 \left( \frac{y_2}{\lambda_1} \right)\) is \(SISL(\Delta_2, N_2)\).
for some $\Delta_2, N_2 > 0$, and so on. In the following, we give an algorithm to find $\lambda_1, \lambda_2, \cdots, \lambda_{s-1}$ for (4.4.5). The algorithm for (4.4.6) is similar.

**Step 1:** For $i = 1, 2, \cdots, s$, find $\hat{L}_i \geq 1, M_i > 0$ so that

$$|\sigma_i(t)| \leq \hat{L}_i |t|$$

and

$$|\sigma_i(t)| \leq M_i$$

are valid for all $t$. Note that $\hat{L}_i$ exists because $\sigma_i$ is locally Lipschitz and bounded. Define

$$L_i = \hat{L}_1 \hat{L}_2 \cdots \hat{L}_i.$$ 

Then $L_i, M_i$ will play the roles of $L, M$ in the proof of Theorem 4.1.

**Step 2:** For $i = 1, 2, \cdots, s$, find $\Delta_i, \nu_1^i, \nu_2^i > 0$ such that, whenever $e_1, e_2$ are two bounded measurable functions and $|e_1| \leq_{ev} \frac{\delta}{\nu_1^i}$, $|e_2| \leq_{ev} \frac{\delta}{\nu_2^i}$, with $\delta \leq \Delta_i$, it follows
that the state of $\dot{\phi} = J_i \phi - \sigma_i(y_i + e_1)b_i + e_2$, where $(J_i, b_i)$ is the controllable pair of $(z_i, y_i)$-part in (4.4.4), is eventually bounded by $\delta$.

To get $\Delta_i, \nu^i_1, \nu^i_2$, we apply Remark 4.2.3 and Corollary 4.2.4. First we define

$$\varphi_i(\theta) = \inf\{ |\sigma_i(s)| : |s| \geq \theta \},$$
$$\psi_i(\theta) = \sup\{ |\sigma_i(s)| : |s| \leq 3\theta \}.$$ 

Next, if $i \leq q$, then let $\bar{\theta}_i$ be a positive number such that

$$\varphi_i(\theta) > \frac{6}{7} \kappa_i \theta, \ \psi_i(\theta) < 4\kappa_i \theta, \ \text{for} \ 0 < \theta < \bar{\theta}_i.$$ 

Let

$$K_i = \min\{1, \frac{8}{\omega_i}, \frac{6\kappa_i}{7 + 10\omega_i}\},$$
and choose \( \nu^i_1, \nu^i_2, \Delta^i_1 \) so that
\[
\nu^i_1 \geq \frac{8}{K_i} + \frac{15}{\tau} K_i + \frac{4}{\omega_i} \theta_i + 3,
\]
\[
\nu^i_2 \geq \frac{7}{\omega_i K_i},
\]
\[
\Delta^i_1 \leq \nu^i_1 \theta_i.
\]

Note that \( \nu^i_1, \nu^i_2, \Delta^i_1 \) represent \( \nu_0, \nu_2, \epsilon \) in Remark 4.2.3.

If \( i > q \), then choose any positive number \( \rho_i < 1 \) (if \( \sigma_i \) is linear near 0, then take \( \rho_i = 1 \)) and \( \Delta^i_1 > 0 \) such that
\[
\varphi(\theta) \geq \rho_i \kappa_i \theta \quad \text{for} \quad 0 < \theta \leq \frac{\Delta^i_1}{2},
\]
and let
\[
\nu^i_1 = 2, \quad \nu^i_2 > \frac{2}{\rho_i \kappa_i}.
\]

Here \( \Delta^i_1, \nu^i_1, \nu^i_2 \) represent \( \epsilon, \nu_1, \nu_2 \) at the end of the proof of Corollary 4.2.4.

**Step 3:** Let \( \Delta_0 = \infty, N_0 = 0 \) and fix \( \alpha \in (0, 1) \). Then find \( \epsilon_i, \Delta^i_2, \delta^i_0, \lambda_i, \Delta_i, N_i, i = 1, 2, \ldots, s - 1 \), recursively as shown below.

Let \( \epsilon_i \geq 0 \) satisfy
\[
\epsilon_i(M_i(\nu^i_1 + 1) + \Delta^i_1) \frac{1}{\kappa_i} < \alpha \Delta_{i-1},
\]
\[
\epsilon_i N_{i-1} L_i (\nu^i_1 + 1) < \frac{1}{4},
\]
and find \( \Delta^i_2 > 0 \) so that
\[
|\sigma_i(t) - \kappa_i t| \leq \epsilon_i |t| \quad \text{for} \quad |t| \leq \Delta^i_2.
\]

Then let \( \delta^i_0, \lambda_i, \Delta_i, N_i > 0 \) satisfy
\[
\delta^i_0 \leq \min\{\Delta^i_1, \alpha \Delta^i_2\},
\]
\[
\lambda_i \leq \min\{1, \frac{\alpha \delta^i_0}{\nu^i_2 M_i}, \frac{(1 - \alpha) \Delta^i_2}{M_i}\},
\]
\[
\Delta_i \leq \min\{\delta^i_0, \frac{(1 - \alpha) \lambda_i \delta^i_0}{\nu^i_2}, (1 - \alpha) \lambda_i \Delta_{i-1}\},
\]
\[
N_i \geq 2 N_{i-1} (\epsilon_i \nu^i_2 + 1)(1 + 2 L_i \nu^i_1 + \nu^i_2).
\]

Note that \( \Delta_{i-1}, N_{i-1}, \epsilon_i, \Delta^i_2, \delta^i_0, \lambda_i, \Delta_i, N_i \) represent \( \Delta, \tilde{N}, \epsilon, \Delta_2, \delta_0, \lambda, \Delta, N \) in the proof of Theorem 4.1 for the single-input case.

With the above algorithm, the constants \( \lambda_i, i = 1, 2, \ldots, s - 1 \) which we obtained are such that (4.4.5) globally stabilizes (4.4.4), and the closed-loop system is **SSS**. \( \square \)
Remark 4.4.3 When all $\sigma_i, i = 1, 2, \cdots, s - 1$, are linear near the origin, the algorithm described above is very simple. As an example, we consider the case of $\sigma_i(s) = \text{sat}(s)$, where

$$\text{sat}(s) = \text{sign} \min\{|s|, 1\}. \quad (4.4.7)$$

Then, in the above algorithm, we have $L_i = M_i = \kappa_i = 1$, and $\varphi_i(\theta) = \text{sat}(\theta), \psi_i(\theta) = \text{sat}(3\theta)$. When $i \leq q$, we take $\bar{\theta}_i = \frac{7}{5}$, and

$$K_i = \frac{6}{7 + 10\omega},$$

$$\nu_i \geq \frac{8}{K_i} + \frac{15}{7}K_i + \frac{4}{\omega} + 3, \quad (4.4.8)$$

and $\Delta_i \leq \frac{7}{5}\nu_i$. Since $\nu_i > 3$, we can always choose $\Delta_i = 3$. When $i > q$, we take $\rho_i = 1, \Delta_i = 2$, and

$$\nu_i = 2. \quad (4.4.9)$$

Next, we choose $\varepsilon_i = 0$, and $\Delta_i = 1$ in Step 3 of the above algorithm. Therefore, it is not necessary to calculate $\Delta_i, N_i$. Since $\Delta_i > 1$ and $\Delta_i = 1$, it follows that after fixing $\alpha \in (0, 1)$, we can take $\delta_i = \alpha$. Then $\lambda_i$ is given by

$$\lambda_i \leq \min\left\{\frac{\alpha^2}{\nu_i^2}, 1 - \alpha\right\}. \quad (4.4.10)$$

To summarize, we only need to find $K_i, \nu_i, \lambda_i$ from (4.4.8), (4.4.9) and (4.4.10). The constants $\lambda_i, i = 1, 2, \cdots, s - 1$ which we obtained are as desired. \hfill \square

Example 4.4.4 Consider an oscillator with multiplicity two:

$$\begin{align*}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= -x_1 + x_3, \\
\dot{x}_3 &= x_4, \\
\dot{x}_4 &= -x_3 + u,
\end{align*} \quad (4.4.11)$$

where $u$ is required to satisfy the constraint $|u| \leq \varepsilon$.

In order to get a feedback of the form of

$$-\text{sat} \left( f_1(x) + \lambda \text{sat} \left( \frac{f_2(x)}{\lambda} \right) \right),$$
we need to find a linear transformation that puts (4.4.11) into the form (4.4.4). To do this, we write (4.4.11) as \( \dot{x} = Ax + bu \), and then

\[
A = \begin{pmatrix}
0 & 1 & 0 & 0 \\
-1 & 0 & 1 & 0 \\
0 & 0 & 0 & 1 \\
0 & 0 & -1 & 0
\end{pmatrix}, \quad b = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.
\]

Let

\[
\tilde{A} = \begin{pmatrix}
0 & 1 & 0 & 0 \\
-1 & 0 & 0 & 1 \\
0 & 0 & 0 & 1 \\
0 & 0 & -1 & 0
\end{pmatrix}, \quad \tilde{b} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}.
\]

Then

\[
R(A, b) = \begin{pmatrix}
0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 \\
0 & 1 & 0 & -1 \\
1 & 0 & -1 & 0
\end{pmatrix}, \quad R(\tilde{A}, \tilde{b}) = \begin{pmatrix} 0 & 1 & 1 & -1 \\ 1 & 1 & -1 & -2 \\ 0 & 1 & 0 & -1 \\ 1 & 0 & -1 & 0 \end{pmatrix}.
\]

Define \( y = T^{-1}x \), where \( T = R(A, b)R(\tilde{A}, \tilde{b})^{-1} \). Then \( \dot{y} = \tilde{A}y + \tilde{b}u \), i.e,

\[
\begin{align*}
\dot{y}_1 &= y_2, \\
\dot{y}_2 &= -y_1 + y_4 + u, \\
\dot{y}_3 &= y_4, \\
\dot{y}_4 &= -y_3 + u.
\end{align*}
\]

(4.4.12)

So we need to find \( \lambda > 0 \) so that

\[
u = -\text{sat} \left(y_4 + \lambda \text{sat} \left(\frac{y_2}{\lambda}\right)\right)
\]

(4.4.13)

stabilizes (4.4.12). To get \( \lambda > 0 \), we use the algorithm given by (4.4.8) and (4.4.10).

Note that \( \omega_i = 1 \). So we have \( K_1 = \frac{a_2^2}{17}, \nu_1^1 = 31 \). Let \( \alpha \in (0, 1) \). Then \( \lambda_1 = \min\{\frac{a_2^2}{31}, 1 - \alpha\} \). Choose \( \alpha = \frac{32}{33} \). We obtain \( \lambda_1 = \frac{1}{35} \). So let \( \lambda = \frac{1}{35} \). Then the feedback defined by (4.4.13) globally stabilizes (4.4.12), and the resulting closed-loop system is
\textbf{SISS}. By calculating $T^{-1}$, we obtain that
\[
y = T^{-1}x = \begin{pmatrix}
x_2 + x_3 \\
-x_1 + x_3 + x_4 \\
x_3 \\
x_4
\end{pmatrix}.
\]

Therefore, the feedback
\[
u = -\varepsilon \text{sat} \left( \frac{x_4}{\varepsilon} + \frac{1}{33} \text{sat} \left( \frac{33}{\varepsilon} (-x_1 + x_3 + x_4) \right) \right)
\]
is as desired. \qed

\section{4.5 Multiple Integrators}

In Section 4.4 we not only provided an algorithm but also presented a method for stabilization. Given a system, we can follow this method and get the best possible estimate in each step for the particular system. In this section, we will show a simple procedure to stabilize multiple integrators.

\textbf{Theorem 4.2} Let $0 < \varepsilon \leq \frac{1}{4}$. Then for the $n$-th order integrator
\[
\begin{align*}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= x_3, \\
&\vdots \\
\dot{x}_n &= u,
\end{align*}
\]
(4.5.1)

there exists a feedback law of the form
\[
u = -\sum_{i=1}^{n} \varepsilon^i \text{sat} (h_i(x_1, \ldots, x_n)),
\]
(4.5.2)

where each $h_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \ldots, n$, is a linear function, such that the origin is a globally asymptotically stable state for the resulting closed-loop system.

\textbf{Proof.} We first apply Lemma 4.4.2 to (4.5.1) and conclude that for every $\varepsilon > 0$ there exists a linear change of coordinates $(x_1, \ldots, x_n) \to (y_1, \ldots, y_n)$ which transforms
(4.5.1) into the form

\[
\begin{align*}
\dot{y}_1 &= \varepsilon^{n-1}y_2 + \varepsilon^{n-2}y_3 + \cdots + \varepsilon y_n + u , \\
\dot{y}_2 &= \varepsilon^{n-2}y_3 + \cdots + \varepsilon y_n + u , \\
& \vdots \\
\dot{y}_{n-1} &= \varepsilon y_n + u , \\
\dot{y}_n &= u .
\end{align*}
\] (4.5.3)

(We only need to set \( \nu_i = \frac{1}{\varepsilon} \) in (4.4.4) and then make a dilation

\[
(x_1, x_2, \ldots, x_n) \rightarrow (\varepsilon^{1+2+\cdots+(n-1)}x_1, \varepsilon^{1+2+\cdots+(n-2)}x_2, \ldots, x_n).
\]

We will show that, when \( \varepsilon \leq \frac{1}{4} \), the feedback

\[
u = k(y) = -\varepsilon \text{sat}(y_n) - \varepsilon^2 \text{sat}(y_{n-1}) - \cdots - \varepsilon^n \text{sat}(y_1) \] (4.5.4)

stabilizes (4.5.3), where \( \text{sat}(s) \) is defined in (4.4.7).

To prove this, we observe that, for any trajectory \( t \rightarrow y(t) \) of the resulting closed-loop system of (4.5.3) with the feedback given in (4.5.4), the \( n \)-th coordinate \( y_n(t) \) will enter and stay in the interval \((-\frac{1}{2}, \frac{1}{2})\) after a finite time. This is obvious from the facts that the sign of \( k(y) \) is opposite to that of \( y_n \) if \( |y_n| \geq \frac{1}{2} \), since

\[
|\varepsilon^2 \text{sat}(y_{n-1}) + \cdots + \varepsilon^n \text{sat}(y_1)| \leq \frac{4}{3} \varepsilon^2 \leq \frac{1}{3} \varepsilon < \varepsilon|y_n|,
\]

and that \( \frac{d}{dt}y_n^2(t) \leq -\frac{1}{6} \varepsilon \) if \( |y_n(t)| \geq \frac{1}{2} \). So, after a finite time, \( \text{sat}(y_n) \) will be equal to \( y_n \), and the expression for \( k(y) \) gives

\[
k(y) = -\varepsilon y_n - \varepsilon^2 \text{sat}(y_{n-1}) - \cdots - \varepsilon^n \text{sat}(y_1) .
\] (4.5.5)

Next, we consider the equation for \( \dot{y}_{n-1} \). From (4.5.5), it follows that, after a finite time, the equation for \( \dot{y}_{n-1} \) takes the form

\[
\dot{y}_{n-1} = -\varepsilon^2 \text{sat}(y_{n-1}) - \cdots - \varepsilon^n \text{sat}(y_1) .
\]

With an analysis similar to that done for \( y_n \), we conclude that the coordinate \( y_{n-1}(t) \) will stay in the interval \((-\frac{1}{2}, \frac{1}{2})\) after a finite time, and then \( k(y) \) will be given by the expression

\[
k(y) = -\varepsilon y_n - \varepsilon^2 y_{n-1} - \varepsilon^3 \text{sat}(y_{n-2}) - \cdots - \varepsilon^n \text{sat}(y_1) .
\]
Continuing in this way, we see that after a finite time \( k(y) \) becomes linear in all the coordinates of \( y \), and is given by

\[
k(y) = -\varepsilon y_n - \varepsilon^2 y_{n-1} - \cdots - \varepsilon^n y_1.
\] (4.5.6)

It is clear that the closed-loop system of (4.5.3) with the feedback \( u = k(y) \) given in (4.5.6) is asymptotically stable. So the proof is complete. \( \square \)

In Section 4.4 we described how to get the linear transformation which puts (4.5.1) into (4.5.3). We now develop another method to get the transformation for multiple integrators, which will also allow us to illustrate the stabilization of multiple integrators by means of a concrete systems diagram.

Consider the \( n \)-th order integrator (4.5.1). Let \( \varepsilon > 0 \). Define linear functions \( f_1, f_2, \ldots, f_n \) as follows:

\[
f_1(s_1) = s_1,
\]
\[
f_2(s_1, s_2) = \varepsilon f_1(s_1) + f_1(s_2),
\]
\[
\vdots
\]
\[
f_n(s_1, s_2, \ldots, s_n) = \varepsilon^{n-1} f_{n-1}(s_1, s_2, \ldots, s_{n-1}) + f_{n-1}(s_2, s_3, \ldots, s_n).
\]

Then the change of coordinates \((x_1, \ldots, x_n) \rightarrow (y_1, \ldots, y_n)\) in \( \mathbb{R}^n \) given by

\[
y_n = f_1(x_n),
\]
\[
y_{n-1} = f_2(x_{n-1}, x_n),
\]
\[
\vdots
\]
\[
y_1 = f_n(x_1, x_2, \ldots, x_n),
\]

puts (4.5.1) in the form (4.5.3).

Indeed, from the definition of \( y_k \) we have \( y_k - y_{k+1} = \varepsilon^{n-k} f_{n-k}(x_k, \ldots, x_{n-1}) \) for \( k = 1, 2, \ldots, n - 1 \). Since \( f_{n-k}(s_1, \ldots, s_{n-k}) \) is a linear function, it follows that \( \frac{\partial f_{n-k}}{\partial s_i} \) is the coefficient of \( s_i \) in \( f_{n-k} \). Note that \( \dot{x}_i = x_{i+1} \) for \( i = 1, 2, \ldots, n - 1 \), so we end up with the equality \( \dot{f}_{n-k}(x_k, \ldots, x_{n-1}) = f_{n-k}(x_{k+1}, \ldots, x_n) \). Therefore

\[
\dot{y}_k - \dot{y}_{k+1} = \varepsilon^{n-k} y_{k+1}, \quad k = 1, 2, \ldots, n - 1.
\] (4.5.8)
Adding (4.5.8) for \( k = i, i+1, \ldots, n-1 \), we conclude that

\[
\dot{y}_i - u = \varepsilon^{n-i} y_{i+1} + \varepsilon^{n-i-1} y_{i+2} + \cdots + \varepsilon y_n,
\]

which is the \( i^{th} \) equation in (4.5.3).

The change of coordinates given in (4.5.7) leads to the design of a bounded stabilizing feedback for multiple integrators shown in Figure 4.5.

Figure 4.5: The design of stabilizing feedback for multiple integrators
Chapter 5

Applications

In this chapter we study the output stabilization problem, in which only partial measurements $y = Cz$ are available for control. Under suitable detectability conditions, the standard Luenberger observer construction is shown to carry over to this case. We also show how to extend our result to some other class of systems, such as those obtained by having a saturation nonlinearity in a forward path of integrators. Finally, we work out an example in detail, dealing with the stabilization of an aircraft using elevator rates as controls.

5.1 Output Feedback

Assume we have a system $\dot{x} = Ax + Bu$ with an $\mathbb{R}^p$-valued output $y = Cz$. We assume that $(A, B)$ is stabilizable, $(A, C)$ is detectable, and $A$ has no eigenvalues with strictly positive real part.

We proceed as in the classical case of stabilization by a linear feedback. That is, we consider the composite system

\begin{align}
\dot{x} &= Ax + Bu , \\
y &= Cx , \\
\dot{z} &= Az + Bu + L(y - Cz) , \\
u &= k(z) ,
\end{align}

(5.1.1)

where $k$ is any bounded stabilizing feedback for the system $\dot{x} = Ax + Bu$ and having the SISS property. The error $e = z - x$ then satisfies $\dot{e} = (A - LC)e$. If $(A, C)$ is detectable, then we can choose $L$ so that all the eigenvalues of $A - LC$ are negative. Therefore, $e(t) \to 0$ as $t \to \infty$ for any initial condition $x(0)$ and $z(0)$ of the original
system and its observer. The equation for $x$ then becomes

$$\dot{x} = Ax + Bk(x + e(t)).$$  \hfill (5.1.2)

We know that the trajectories of $\dot{x} = Ax + Bk(x)$ go to zero; the problem is whether the same is true for (5.1.2).

To insure this, we require that the feedback defined in Theorem 4.1 be globally Lipschitz. (Recall that the saturation functions used in Chapter 4 are only required to be locally Lipschitz. If we choose globally Lipschitz saturation functions in our construction, then the stabilizing feedback is guaranteed to be globally Lipschitz.) Therefore, we can rewrite (5.1.2) as

$$\dot{x} = Ax + Bk(x) + \hat{e},$$  \hfill (5.1.3)

where $\hat{e} = B(k(x + e(t)) - k(x))$ is bounded by a constant multiple of $e(t)$. Thus $\hat{e} \to 0$. Then Theorem 4.1 implies that all solutions of (5.1.3) converge to zero, i.e., the full system (5.1.1) is globally asymptotically stable. (Note that local stability is guaranteed.) Summarizing, we have the following result:

**Theorem 5.1** Consider a linear system of the form $\dot{z} = Ax + Bu$, $y = Cx$, such that $A$ has no eigenvalues with positive real part, $(A, B)$ is stabilizable and $(A, C)$ is detectable. Then, for any observer system $\dot{\hat{z}} = Az + Bu + L(y - C\hat{z})$ and for any bounded stabilizing feedback $u \to k(z)$ given by Theorem 4.1, if $k$ is globally Lipschitz, then the composite system $\dot{\hat{z}} = Az + Bk(z)$, $\dot{z} = Az + Bk(z) + L(Cz - C\hat{z})$ has the origin as a globally asymptotically stable equilibrium. \hfill $\square$

We remark that, since the stabilizing feedback given in Chapter 4 causes the resulting closed-loop system to have the SISS property, Theorem 5.1 holds for any linear observer. In [28] the same problem was also considered, but the result in [28] required that all observer poles have real part less than a certain margin. So Theorem 5.1 answers a question posed in [28], namely, whether an arbitrary linear observer can be used for stabilization.
5.2 Cascaded Systems

Consider an \((n + m)\)-dimensional system of the following partitioned form

\[
\begin{align*}
\dot{x} &= Ax + \sum_{i=1}^{m} b_i \sigma_i(z_i), \\
\dot{z} &= u,
\end{align*}
\]

(5.2.1)

where \(x \in \mathbb{R}^n, z = (z_1, \ldots, z_m)' \in \mathbb{R}^m, u = (u_1, \ldots, u_m)' \in \mathbb{R}^m,\) and all \(\sigma_i : \mathbb{R} \rightarrow \mathbb{R}\) are globally Lipschitz functions which are monotonic near zero and satisfy \(\sigma_i(0) = 0.\) Let \(B\) denote \((b_1, \ldots, b_m).\) Assume that the pair \((A, B)\) is stabilizable and that the eigenvalues of \(A\) have nonpositive real part. Then (5.2.1) is stabilizable. Furthermore, if the \(\sigma_i\)'s and their inverses are all smooth functions, then there exists a smooth feedback that stabilizes (5.2.1).

The stabilizability assertion follows trivially. In fact, from the assumptions on the \(\sigma_i\)'s, it follows that there exists some \(\varepsilon > 0\) such that all \(\sigma_i\)'s are monotonic on the interval \((-\varepsilon, \varepsilon).\) From Theorem 4.1 we know that there exists an analytic feedback \(v = \tilde{k}(x) = (\tilde{k}_1(x), \ldots, \tilde{k}_m(x))',\) where \(|\tilde{k}_i(x)| < \varepsilon\) for \(x \in \mathbb{R}^n,\) which stabilizes the system \(\dot{x} = Ax + Bu,\) and such that the closed-loop system has the \(SISS\) property. With the feedback \(v = k(x) = (\sigma_1^{-1}(\tilde{k}_1(x)), \ldots, \sigma_m^{-1}(\tilde{k}_m(x)))',\) the closed-loop system \(\dot{x} = Ax + \sum_{i=1}^{m} b_i \sigma_i(v_i)\) is globally stable. Using the standard “adding an integrator” construction (see, e.g., Lemma 4.8.3 in [27]), we conclude that (5.2.1) is stabilizable.

However, in order to build a stabilizer for (5.2.1) using the “adding an integrator” construction, we would need to first find a Lyapunov function for the system

\[
\dot{x} = Ax + \sum_{i=1}^{m} b_i \sigma_i(k_i(x)),
\]

(5.2.2)

where \(k_i = \sigma_i^{-1}(\tilde{k}_i(x)).\) This is not easy to do. The contribution of this section is to show that one can build a stabilizer without knowing explicitly a Lyapunov function for the system, because we already established the \(SISS\) property for (5.2.2). Next we give details.

Since (5.2.2) has the \(SISS\) property and all \(\sigma_i\)'s are Lipschitz continuous, it follows that the system

\[
\dot{x} = Ax + \sum_{i=1}^{m} b_i \sigma_i(k_i(x) + v_i)
\]

(5.2.3)
is SISS stable with respect to the control \( v = (v_1, \cdots, v_m)' \). Let \( w = z - k(x) \). Then, in the coordinates \((x', w')'\) in the state space, the system (5.2.1) can be written as

\[
\begin{align*}
\dot{x} &= Ax + \sum_{i=1}^{m} b_i \sigma_i(k_i(x) + w_i), \\
\dot{w} &= u - \nabla k(x) \cdot \left( Ax + \sum_{i=1}^{m} b_i \sigma_i(k_i(x) + w_i) \right).
\end{align*}
\]

(5.2.4)

Now let

\[
u = -w + \nabla k(x) \cdot \left( Ax + \sum_{i=1}^{m} b_i \sigma_i(k_i(x) + w_i) \right) = k(x) - z - \nabla k(x) \cdot \left( Ax + \sum_{i=1}^{m} b_i \sigma_i(z_i) \right).
\]

(5.2.5)

Then the resulting closed-loop system for (5.2.1) becomes

\[
\begin{align*}
\dot{x} &= Ax + \sum_{i=1}^{m} b_i \sigma_i(k_i(x) + w_i), \\
\dot{w} &= -w,
\end{align*}
\]

which is stable because \( w(t) \to 0 \) as \( t \to \infty \) and the system (5.2.3) is SISS stable with respect to the control \( v \). So (5.2.5) is a stabilizer for (5.2.1).

As an example to illustrate the procedure of stabilizing a system of the type (5.2.1), we consider the following system shown in Figure 5.1 (see also [33]):

![Cascaded system](image)

Figure 5.1: Cascaded system

\[
\begin{align*}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= x_3, \\
\dot{x}_3 &= \tanh(x_4), \\
\dot{x}_4 &= u.
\end{align*}
\]

(5.2.6)

First, we need to find a smooth feedback which stabilizes the triple integrator

\[
\begin{align*}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= x_3, \\
\dot{x}_3 &= u.
\end{align*}
\]

(5.2.7)
and the feedback has absolute value less than one. To do this, we use the transformation:

\[ y_1 = \frac{1}{8}x_1 + \frac{3}{4}x_2 + x_3, \]
\[ y_2 = \frac{1}{2}x_2 + x_3, \]
\[ y_3 = x_3. \]

Now (5.2.7) takes the form

\[ \dot{y}_1 = \frac{1}{4}y_2 + \frac{3}{4}y_3 + u, \]
\[ \dot{y}_2 = \frac{1}{2}y_3 + u, \]
\[ \dot{y}_3 = u. \]  

(5.2.8)

Let \( \varphi : \mathbb{R} \to \mathbb{R} \) to be a smooth increasing function which has the property that

\[ \varphi(s) = \begin{cases} 
  s, & |s| < \frac{3}{4}, \\
  \text{sign}(s), & |s| > 1.
\end{cases} \]

For every \( \varepsilon > 0 \), we define \( \varphi_\varepsilon(s) = \varepsilon \varphi(\frac{s}{\varepsilon}). \) Then it is not difficult to see that the feedback

\[ u = -\frac{1}{2} \left( \varphi_1(y_3) + \varphi_\frac{1}{2}(y_2) + \varphi_\frac{1}{4}(y_1) \right) \]

stabilizes the system (5.2.8). Therefore

\[ u = k(x_1, x_2, x_3) = -\frac{1}{2} \left( \varphi_1(x_3) + \varphi_\frac{1}{2}(\frac{1}{2}x_2 + x_3) + \varphi_\frac{1}{4}(\frac{1}{8}x_1 + \frac{3}{4}x_2 + x_3) \right) \]

stabilizes the system (5.2.7) and the magnitude of the feedback is less than one. Let \( k(x_1, x_2, x_3) = \tanh^{-1}(k(x_1, x_2, x_3)) \). Then from (5.2.5), we conclude that the feedback

\[ u = k(x_1, x_2, x_3) - x_4 + x_2 \frac{\partial k}{\partial x_1}(x_1, x_2, x_3) + x_3 \frac{\partial k}{\partial x_2}(x_1, x_2, x_3) + \tanh(x_4) \frac{\partial k}{\partial x_3}(x_1, x_2, x_3) \]

stabilizes (5.2.6).

\[ \square \]

5.3 F-8 Longitudinal Flight Control

In this section, we develop in detail an explicit design for an F-8 aircraft. We picked the model in [14], as expanded and corrected in [36], since this has been often considered as a paradigm for many aircraft control problems. Here we use the exact constants and trim conditions—that is, the desired operating points—given in references [3, 36, 37].
Note that the problem considered here is different from the question treated in those references, and only the data is employed. Also, we add a saturation to the model. We pick relatively small values of these saturations to analyze our control design under demanding conditions.

5.3.1 The Model

We rely on [36, 37] (see also [14]) for the following nonlinear model for the F-8 aircraft longitudinal flight dynamics. For convenience, we repeat here all the relevant equations, using the same notations as in the above references:

\[
\begin{align*}
\dot{\mu} &= -\mu q \tan \alpha - g \sin \theta + \frac{L_w}{m} \sin \alpha + \frac{L_t}{m} \sin \alpha_t, \\
\dot{\alpha} &= q + \frac{a}{\mu} \cos \alpha \cos(\alpha - \theta) - \frac{L_w}{i m} \cos \alpha - \frac{L_t}{i m} \cos \alpha \cos(\alpha - \alpha_t), \\
\dot{\theta} &= q, \\
\dot{q} &= (M_\omega + lL_\omega \cos \alpha - l_t L_t \cos \alpha_t - cg)/I_\mu.
\end{align*}
\]

(5.3.1)

We are using the following symbols:

\[
\begin{align*}
\alpha_t &= (1 - a_e)\alpha + \delta_e, \\
L_w &= C_L(\alpha)\bar{q}S, \\
L_t &= C_{L_t}(\alpha_t, \delta_e)\bar{q}S_t, \\
\bar{q} &= \frac{\rho u^2}{2 \cos^2 \alpha}, \\
\gamma &= \theta - \alpha.
\end{align*}
\]

(5.3.2)

Still quoting from [36, 37], the meaning of the above coefficients and variables and their units are:

\(\mu\): forward speed, in ft/sec,

\(\alpha\): wing angle of attack, in rad,

\(\theta, q\): pitch angle, in rad, and pitch rate, in rad/sec,

\(\gamma\): flight path angle, in rad,

\(\alpha_t\): tail angle of attack, in rad,

\(\delta_e\): elevator angle, in rad,
$m$: mass of aircraft, in slugs,

$I_y$: moment of inertia of aircraft about $Y$ axis, in slugs-ft$^2$,

$L_w, L_t$: wing and tail lifts, in lb,

$M_w$: wing moment,

$l$: distance between wing a.c and aircraft c.g., in ft,

$l_t$: distance between tail a.c and aircraft c.g., in ft,

$c$: damping coefficient, in lb-ft-sec,

$C_L, C_{L_t}$: wing and tail lift coefficients,

$\bar{q}$: dynamic pressure, in lb/ft$^2$,

$S, S_t$: wing and tail area, in ft$^2$, and

$\rho$: atmospheric density, in slugs-ft$^3$.

The lift coefficients are complicated nonlinear functions of the angles of attack and elevator angle; for simplicity, and again following [36, 37], we use a cubic approximation:

$$C_L(\alpha) = (C_L^1 \alpha - C_L^2 \alpha^2) ,$$

$$C_{L_t}(\alpha_t, \delta_e) = (C_{L_t}^1 \alpha_t - C_{L_t}^2 \alpha_t^2 + a_c \delta_e) .$$

Define

$$\sigma(s) = \text{sign}(s) \min\{|s|, 0.01\} .$$

The control $u$ is applied to $\delta_e$ by means of an actuator with the following dynamics:

$$\dot{\delta_e} = \sigma(-\delta_e + \delta_{e0} + u),$$

where $\delta_{e0}$ is a desired equilibrium (see below). We have included a saturation nonlinearity which saturates the right hand side at 0.01.
5.3.2 Trim Conditions

As in the references [36, 37], the desired operating point corresponds to an altitude of 30,000 ft, and we also choose, as there, the following values for all parameters:

\[ \rho = 0.00089 \text{ slugs/ft}^3, \]
\[ C_L^1 = 4.0, \]
\[ C_L^2 = 12, \]
\[ a_c = 0.1, \]
\[ a_e = 0.75, \]
\[ S = 375 \text{ ft}^2, \]
\[ S_t = 93.4 \text{ ft}^2, \]  
\[ m = 667.7 \text{ slugs,} \]
\[ I_y = 96800 \text{ slugs \cdot ft}^2, \]
\[ l = 0.189 \text{ ft,} \]
\[ l_t = 16.7 \text{ ft,} \]
\[ M_o = 0 \text{ lb \cdot ft,} \]
\[ c = 38332.8 \text{ lb \cdot ft \cdot sec,} \]
\[ g = 32.2 \text{ ft/sec}^2. \]

The equations (5.3.1), (5.3.2), (5.3.3), and (5.3.5) give rise to a control system. We consider the system with constants given in (5.3.6) and the desired equilibrium at:

\[ \mu_0 = 389.1315833, \alpha_0 = .2400685620, \theta_0 = .2375883269, q_0 = 0, \delta_{e0} = -0.05. \]  

(5.3.7)

The above values are essentially those given in [3] but calculated more precisely using the Maple symbolic manipulation system (the numbers in that reference, namely: \( \mu_0 = 388.7, \alpha_0 = 0.240, \theta_0 = 0.238, q_0 = 0, \delta_{e0} = -0.05 \), give numerical values significantly far from zero when substituted in the equations).
5.3.3 Controller Design

The linearization of the model at the equilibrium (5.3.7) (but not linearizing the saturation, of course) is as follows:

\[
\begin{align*}
\dot{\mu} &= .03895120630(\mu - \mu_0) + 53.20472564(\alpha - \alpha_0) - 31.29545082(\theta - \theta_0) \\
&\quad -95.25527586q + .7601675273(\delta_c - \delta_{c_0}), \\
\dot{\alpha} &= -.0004130994812(\mu - \mu_0) - .2564211668(\alpha - \alpha_0) + .0001993493700(\theta - \theta_0) \\
&\quad + q - .09966311824(\delta_c - \delta_{c_0}), \\
\dot{\theta} &= q, \\
\dot{q} &= -1.061728681(\alpha - \alpha_0) - .3959999999q - 4.71364453(\delta_c - \delta_{c_0}), \\
\dot{\delta}_c &= \sigma(-(\delta_c - \delta_{c_0}) + u). \tag{5.3.8}
\end{align*}
\]

For simplicity in computing the feedback control, we eliminate the small coefficients in (5.3.8) and rewrite the new system as

\[
\begin{align*}
\dot{\mu} &= 53.20472564(\alpha - \alpha_0) - 31.29545082(\theta - \theta_0) - 95.25527586q \\
&\quad + .7601675273(\delta_c - \delta_{c_0}), \\
\dot{\alpha} &= -.2564211668(\alpha - \alpha_0) + q - .09966311824(\delta_c - \delta_{c_0}), \\
\dot{\theta} &= q, \\
\dot{q} &= -1.061728681(\alpha - \alpha_0) - .3959999999q - 4.71364453(\delta_c - \delta_{c_0}), \\
\dot{\delta}_c &= \sigma(-(\delta_c - \delta_{c_0}) + u). \tag{5.3.9}
\end{align*}
\]

From now on, we use symbols to denote the nonzero constants in (5.3.9) and rewrite the system as follows:

\[
\begin{align*}
\dot{\mu} &= a(\alpha - \alpha_0) + b(\theta - \theta_0) + cq + d(\delta_c - \delta_{c_0}), \\
\dot{\alpha} &= e(\alpha - \alpha_0) + q + f(\delta_c - \delta_{c_0}), \\
\dot{\theta} &= q, \\
\dot{q} &= g(\alpha - \alpha_0) + hq + i(\delta_c - \delta_{c_0}), \\
\dot{\delta}_c &= \sigma(-(\delta_c - \delta_{c_0}) + u). \tag{5.3.10}
\end{align*}
\]

Disregarding the saturation, the eigenvalues of (5.3.10) are 0, 0, 0 and

\[
1/2e + 1/2h \pm 1/2(-e^2 + 2eh - h^2 - 4g)^{1/2} \sqrt{-1}.
\]
Following the procedure in [40, 30], we wish to transform (5.3.10) into a system of the following form:

\[
\begin{align*}
\dot{x}_1 &= (1/2e + 1/2h)x_1 + 1/2(-e^2 + 2eh - h^2 - 4g)^{1/2}x_2, \\
\dot{x}_2 &= -1/2(-e^2 + 2eh - h^2 - 4g)^{1/2}x_1 + (1/2e + 1/2h)x_2 + x_4, \\
\dot{x}_3 &= x_4, \quad (5.3.11) \\
\dot{x}_4 &= x_5, \\
\dot{x}_5 &= \sigma(-x_5 + u).
\end{align*}
\]

This can be achieved by changing the variables:

\[
\begin{align*}
x_1 &= \frac{i^2 \Delta}{\Delta} (\alpha - \alpha_0) - \frac{\Delta}{i \gamma z f}(\theta - \theta_0) - \frac{\Delta}{\Delta} q, \\
x_2 &= i\frac{(e i - h i - 2 g f)}{2\Delta} (\alpha - \alpha_0) + \frac{e h}{i(1 - y f)}(\theta - \theta_0) - \frac{f(e i - h i - 2 g f)}{2\Delta} q, \\
x_3 &= -\frac{e h - e g f}{(e i - y f)} (\mu - \mu_0) + \frac{i e h a + i d h g - a e i - i b g d + a g d^2}{e i - y f}(\alpha - \alpha_0) \\
&\quad + \frac{N}{(e i - 2 f i y + e^2 f^2 i y)}(\theta - \theta_0) + \frac{d e v^2 - e f h - e g d + b g f + a g f}{(e i - 2 f i y + e^2 f^2 i y)} q, \\
x_4 &= -\frac{e h - e g f}{e i - y f}(\alpha - \alpha_0) - \frac{e h - e g f}{e i - y f}(\theta - \theta_0) + \frac{e h - e g f}{e i - y f} \gamma, \\
x_5 &= (\delta i - \delta e f_0),
\end{align*}
\]

where \(\lambda = \frac{1}{2}(-e^2 + 2eh - h^2 - 4g)^{1/2}, N = e^2bi + e^2chi - e^2h^2d + efah^2 - ebgf - echg - i g c e - i e h a + 2dehg - fbg h - fah g + f^2c + a g i + i b g - d g^2,\) and \(\Delta = e^2 f i^2 + i^3 e - e f i^2 h - 2 e f^2 g i - g f i^2 + f^2 i b g + f^3 g^2.\)

Finally, we carry out the design following the general outline of the construction in Chapter 4, as applied to a system in the form (5.3.11) and provide a very simple feedback which uses one saturation. (Note that in the closed-loop system there is one other saturation which is part of the original system: the one that already exists in the equation for \(x_5.\))

The feedback we obtain is as follows:

\[
\begin{align*}
u &= -0.9 \sigma(0.2x_3 + 2.2x_4 + 2x_5). \quad (5.3.13)
\end{align*}
\]

These constants 0.9, 0.2, and so forth, are not the ones that would follow from using the very conservative bounds in the proof of Theorem 4.1. We picked much better constants, taking advantage of the knowledge of the equations. (Of course, one may expect that a fair amount of such fine tuning will be necessary in any realistic application.) Using the transformation (5.3.12), we get a control law for the original system (5.3.1).
5.3.4 Stability of the Linear System

We sketch next a proof of the fact that the above design globally stabilizes the linearized system. This is, of course, not enough to guarantee stability when the nonlinear model is used, except locally, but we think it is nonetheless of interest to show the computations. Later, we investigate experimentally the domain of stability for the nonlinear system.

Consider the system

\[
\begin{align*}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= x_3, \\
\dot{x}_3 &= \sigma(-x_3 + u).
\end{align*}
\]  

(5.3.14)

Let \( y_1 = x_1 + x_2, y_2 = x_2 + x_3, y_3 = x_3 \). Then we have

\[
\begin{align*}
\dot{y}_1 &= y_2, \\
\dot{y}_2 &= y_3 + \sigma(-y_3 + u), \\
\dot{y}_3 &= \sigma(-y_3 + u).
\end{align*}
\]  

(5.3.15)

Let

\[
\begin{align*}
u &= -\rho \sigma(ay_1 + by_2).
\end{align*}
\]  

(5.3.16)

Then we have the following result.

**Theorem 5.2** For any \( \varepsilon > 0 \), let \( \sigma(s) = \text{sign}(s) \min\{|s|, \varepsilon\} \). Let \( a, b, \rho \) be positive numbers such that \( b \geq 2a, 2b^2 > 5a, \) and \( 1/2 < \rho < 1 \). Then, the closed-loop system consisting of (5.3.15) together with the control (5.3.16) is globally asymptotically stable.

Choosing \( \varepsilon = 0.01, \rho = 0.9, a = 0.2, b = 2 \) in the theorem, we obtain the feedback (5.3.13). When \( \rho \leq 1/2 \), the conclusion of the theorem can be shown to be true for any \( a, b > 0 \); see the remark following the proof.

**Proof.** Local stability is trivial from linear coordinates. Thus we must prove that every trajectory converges to zero. Let \( (y_1(t), y_2(t), y_3(t)) \) be any closed-loop trajectory. First, from the last equation of (5.3.15), we see that \( y_3 \) will be eventually bounded by \( \rho \varepsilon \). Therefore, for sufficiently large time \( t \),

\[
\begin{align*}
\sigma(-y_3 + u) &= -y_3 - f(y_1, y_2, y_3)\sigma(ay_1 + by_2),
\end{align*}
\]  

(5.3.17)
where \( f \) is a function (which is not required to satisfy any regularity conditions); \( f \) can be taken to satisfy:

- \( 1 - \rho \leq f(y_1, y_2, y_3) \leq \rho \),
- \( f(y_1, y_2, y_3) = \rho \) if \( y_3 \) and \( ay_1 + by_2 \) have different signs.

Without loss of generality, we assume \(|y_3(t)| \leq \rho \varepsilon\) for all \( t \). Therefore (5.3.17) is satisfied, and the first two equations of (5.3.15) become

\[
\begin{align*}
y_1 &= y_2, \\
y_2 &= -f(y_1, y_2, y_3)\sigma(ay_1 + by_2).
\end{align*}
\] (5.3.18)

We will first show that \( y_2 \) becomes eventually bounded.

Let \( H \) denote the line \( ay_1 + by_2 = 0 \). Note that if a trajectory enters the region

\[
\{(y_1, y_2) : ay_1 + by_2 > 0\},
\] (5.3.19)

then it must either (i) cross \( H \), or (ii) satisfy \( y_2(t) \to 0 \) as \( t \to \infty \). Indeed, if \( y_2(t) < 0 \) for some \( t \), then \( H \) will clearly be crossed. So assume that (5.3.19) is entered but (i) and (ii) do not hold and also \( y_2(t) \geq 0 \) for all \( t \). Then, \( \dot{y}_2 < 0 \), so \( y_2(t) \to \alpha > 0 \). Since \( \dot{y}_1 = y_2 > \alpha \), it then follows that \( y_1(t) \to \infty \). As \( t \to \infty \), we have \( \dot{y}_2 < -(1 - \rho)\varepsilon \), which contradicts the fact that \( y_2 \to \alpha \).

The same conclusion is true if a trajectory enters the region \( \{(y_1, y_2) : ay_1 + by_2 < 0\} \).

Without loss of generality, we assume the trajectory crosses \( H \). Let \( (-\frac{b}{a} M, M) \) be any intersection of the trajectory \((y_1(t), y_2(t))\) with \( H \), with \( M > 0 \). Define

\[
\bar{M} = \max\{M, \frac{5\varepsilon}{2b}\}.
\] (5.3.20)

We show that when the trajectory returns to \( H \) (where the \( y_2 \)-coordinate must be negative, since \( \dot{y}_1 = y_2 > 0 \) and \( \dot{y}_2 = 0 \) if \((y_1, y_2)\) is on \( H \) and \( y_2 > 0 \)), \(|y_2(t)|\) has become smaller than \( \frac{5\varepsilon}{2b} \), or smaller than \( M \) by at least a positive constant value. By symmetry, the same conclusion is true if \((y_1(t), y_2(t))\) starts at \((\frac{b}{a} M, -M)\) on \( H \). Since \( y_2 \) is decreasing on \( \{(y_1, y_2) : ay_1 + by_2 > 0\} \) and increasing on \( \{(y_1, y_2) : ay_1 + by_2 < 0\} \),
it follows that $|y_2(t)|$ will be eventually bounded by $\frac{9b}{10M}$. (If the trajectory does not return to $H$, then from the above analysis we know $y_2(t) \to 0$ as $t \to \infty$.)

Indeed, we consider the trajectory starting at $(-\frac{b}{a}M, M)$. If the trajectory first crosses the $y_1$-axis at $\tilde{y}_1 \leq \frac{ob}{10a} \tilde{M}$ ($\tilde{y} > 0$), then, since $y_1(t)$ is decreasing on $\{(y_1, y_2) : y_2 < 0\}$, we conclude that $|y_2(t)| < \frac{o}{10} \tilde{M}$ when $(y_1(t), y_2(t))$ returns to $H$. So our conclusion follows. Now we assume $\tilde{y}_1 > \frac{ob}{10a} \tilde{M}$. Therefore $(y_1(t), y_2(t))$ must cross the line

$$ay_1 + by_2 = \frac{9b}{10} \tilde{M}$$

before reaching the $y_1$-axis. Let $\varphi(t) = ay_1(t) + by_2(t)$. Let $\tilde{t}, \hat{t}$ be the last times so that $\varphi(\tilde{t}) = \frac{ob}{5} \tilde{M}$ and $\varphi(\hat{t}) = \frac{ob}{10} \tilde{M}$, respectively, before $(y_1(t), y_2(t))$ reaches the $y_1$-axis. Then for $t \in [\tilde{t}, \hat{t}]$, we have that $\varphi(t) < ay_2(t) < aM$. Since $\varphi(\hat{t}) - \varphi(\tilde{t}) = \frac{b}{2} \tilde{M}$, it must hold that $\hat{t} - \tilde{t} > \frac{bM}{2aM} \geq 1$ (because $b \geq 2a, \tilde{M} \geq M$). Note also that $\varphi(t) > \varepsilon$ on $[\tilde{t}, \hat{t}]$, by (5.3.20). As $|y_3(t)| \leq \rho \varepsilon$, $\tilde{y}_2(t) < 0$ in this interval (if $y_3 > 0$, this is clear, otherwise $y_2(t)$ and $\varphi(t)$ have opposite signs, and $\sigma(\varphi(t)) = \varepsilon$ implies $\rho \sigma(\varphi(t)) \geq |y_3(t)|$, so sign $(\tilde{y}_2) = \text{sign} (-\rho \sigma(\varphi(t)))$). On $[\tilde{t}, \hat{t}]$, as $\sigma(\varphi(t)) = \varepsilon$, it follows that $\tilde{y}_2(t) \leq -\rho \varepsilon$ whenever $y_3(t) > 0$. Since also $|y_3(t)| \leq \rho \varepsilon$, if $y_3(\hat{t}) > 0$, then (since $\hat{t} - \tilde{t} > 1$), $y_2(\tilde{t}) < 0$. Also, for $t > \tilde{t}$, $y_3(t)$ remains negative as long as $y_2(t) > 0$. So $f(y_1(t), y_2(t), y_3(t)) = \rho$ for such $t$. Therefore, for $t > \tilde{t}$, (5.3.18) becomes

$$\begin{align*}
\dot{y}_1 &= y_2, \\
\dot{y}_2 &= -\rho \varepsilon.
\end{align*}$$

(5.3.21)

Let $\gamma(t) = (\xi(t), \eta(t))$ be the trajectory of (5.3.21) starting at $\gamma(0) = \left( \frac{ob}{10a} \tilde{M}, \tilde{M} \right)$. Then it is clear that $(y_1(t), y_2(t))$ cannot cross the image of $\gamma$ while $y_2(t) > 0$. Since $f(y_1, y_2, y_3) \leq \rho$, it follows by comparing the right hand sides of (5.3.18) and (5.3.21) that there cannot be any crossings even when $y_2(t) < 0$, before crossing $H$.

Now we estimate the first intersection of $\gamma$ with $H$. The curve $\gamma(t)$ can be written exactly:

$$\begin{align*}
\xi(t) &= \frac{ob}{10a} \tilde{M} + \tilde{M} t - \frac{\rho \varepsilon}{2} t^2, \\
\eta(t) &= \tilde{M} - \rho \varepsilon t.
\end{align*}$$
Let $T$ be the first time that $\gamma(T)$ crosses $H$. From

$$a\xi(T) + b\eta(T) = -\frac{a\rho \epsilon}{2} T^2 + (a\bar{M} - b\rho \epsilon)T + \frac{19}{10} b\bar{M} = 0,$$

we get

$$T = \frac{1}{a\rho \epsilon} \left( (a\bar{M} - b\rho \epsilon) + (a\bar{M} + b\rho \epsilon) \sqrt{1 - \frac{ab\rho \epsilon \bar{M}}{5(a\bar{M} + b\rho \epsilon)^2}} \right).$$

Therefore

$$\eta(T) \geq \frac{b}{a} \rho \epsilon - (\bar{M} + \frac{b}{a} \rho \epsilon) \left( 1 - \frac{ab\rho \epsilon \bar{M}}{10(a\bar{M} + b\rho \epsilon)^2} \right) = -\bar{M} + \frac{b\rho \epsilon \bar{M}}{10(a\bar{M} + b\rho \epsilon)}.$$

By (5.3.20), $\bar{M} \geq \frac{5\epsilon}{2b}$, and hence the term $\frac{b\rho \epsilon \bar{M}}{10(a\bar{M} + b\rho \epsilon)}$ is greater than some positive constant independent of $\bar{M}$. So when $(y_1(t), y_2(t))$ returns to $H$, $|y_2(t)|$ is smaller than $\bar{M}$ by at least a positive constant, as we wanted to show. Therefore, we have shown that $|y_2(t)|$ will eventually be bounded by $\frac{5\epsilon}{2b}$.

Define

$$V(y_1, y_2, y_3) = \Phi(ay_1) + \Phi(ay_1 + by_2) + \frac{a}{\rho} y_2^2 + 2by_3^2,$$

where $\Phi(s) = \int_0^s \sigma(t)dt$. It is clear that $V$ is a positive definite and proper function.

After a finite time, the inequalities $|y_2(t)| \leq \frac{5\epsilon}{2b}$ and $|y_3(t)| \leq \rho \epsilon$ will hold, and we have by (5.3.17), along trajectories of (5.3.15), for all large $t$,

$$\dot{V} = ay_2 \left( \sigma(ay_1) - \sigma(ay_1 + by_2) f(y_1, y_2, y_3) / \rho \right) - bf(y_1, y_2, y_3) \sigma^2(ay_1 + by_2)
+ ay_2(1 - f(y_1, y_2, y_3) / \rho)\sigma(ay_1 + by_2) - 4by_3^2 - 4by_3 f(y_1, y_2, y_3) \sigma(ay_1 + by_2).$$

(5.3.22)

Next we show that $V \to 0$ as $t \to \infty$. We examine first the case where $t$ is such that $f(y_1, y_2, y_3) = \rho$. In this case,

$$\dot{V} \leq ay_2 \left( \sigma(ay_1) - \sigma(ay_1 + by_2) \right) - \frac{\rho(1 - \rho)}{2} b \sigma^2(ay_1 + by_2) - \frac{4(1 - \rho)}{1 + \rho} by_3^2.$$

(5.3.23)

Note that this implies that, for each ball $B$ centered at the origin, there is a constant $\kappa = \kappa_B > 0$ such that $\dot{V} < -\kappa$ whenever $y_1, y_2, y_3$ are outside of $B$ and $f(y_1, y_2, y_3) = \rho$.

If instead $f(y_1, y_2, y_3) < \rho$ (i.e., we cannot choose $f(y_1, y_2, y_3) = \rho$), then $y_3(ay_1 + by_3) > 0$. Without loss of generality, we consider the case $y_3(t) > 0$, $ay_1(t) + by_2(t) > 0$. Thus

$$y_3 = \varepsilon - f(y_1, y_2, y_3) \sigma(ay_1 + by_2).$$

(5.3.24)
Also, if \( y_2 > 0 \), then \( \sigma(ay_1) \leq \sigma(ay_1 + by_2) \), and therefore the first term of \( \dot{V} \) in (5.3.22) is smaller than

\[
a|y_2|\sigma(ay_1 + by_2)(1 - f(y_1, y_2, y_3)/\rho).
\]

(5.3.25)

The same inequality is true if \( y_2 < 0 \), since \( \sigma(-s) = -\sigma(s) \). So in the case that \( f(y_1, y_2, y_3) < \rho \), we have

\[
\dot{V} \leq -bf(y_1, y_2, y_3)\sigma^2(ay_1 + by_2) + 2a|y_2|(1 - f(y_1, y_2, y_3)/\rho)\sigma(ay_1 + by_2)
- 4by_3^2 - 4by_3 f(y_1, y_2, y_3)\sigma(ay_1 + by_2).
\]

(5.3.26)

Substituting (5.3.24) into (5.3.26) and using \( f, \sigma \) to denote \( f(y_1, y_2, y_3) \), and \( \sigma(ay_1 + by_2) \), respectively, we have

\[
\dot{V} \leq -bf\sigma^2 + \frac{5a}{\ell} \varepsilon (1 - f/\rho)\sigma - 4b(\varepsilon - f\sigma)^2 - 4bf\sigma(\varepsilon - f\sigma)
= -bf\sigma^2 + \frac{5a}{\ell} \varepsilon (1 - f/\rho)\sigma - 4b\varepsilon^2 + 4b\varepsilon f\sigma
= \frac{5a}{\ell} \varepsilon (\sigma - \varepsilon - (4b - \frac{5a}{\ell})\varepsilon^2 + (4b - \frac{5a}{\ell})f\varepsilon - bf\sigma^2.
\]

The first term is less than or equal to 0. Let \( A, B, C \) denote the coefficients of \( \varepsilon^2, \varepsilon\sigma, \sigma^2 \) in the last three terms. Then (note that \( 2\ell^2 > 5a \))

\[
B^2 - 4AC = (4b - \frac{5a}{\ell})^2 \varepsilon^2 - 4(4b - \frac{5a}{\ell})bf
\leq (4b - \frac{5a}{\ell})f(4bf - \frac{5a}{\ell}f - 4b)
\leq -(4b - \frac{5a}{\ell})\frac{5a}{\ell}f^2 < 0.
\]

Therefore, if \( f(y_1, y_2, y_3) < \rho \), then

\[
\dot{V} < -\kappa < 0,
\]

(5.3.27)

where \( \kappa \) is a constant (independent of \( (y_1, y_2, y_3) \)).

From (5.3.23) and (5.3.27) we conclude that \( \dot{V} \to 0 \) as \( t \to \infty \), and thus for \( i = 1, 2, 3 \), \( y_i(t) \to 0 \) as \( t \to \infty \), as desired. \( \square \)

**Remark 5.3.1** In [30], [29], [40], and [32], only the case \( \rho \leq \frac{1}{2} \) was discussed.

There, we always have \( f(y_1, y_2, y_3) = \rho \), so the first two equations of (5.3.15) will eventually become \( \dot{y}_1 = y_2, \dot{y}_2 = -\rho\sigma(ay_1 + by_2) \). It is then easy to show that

\[
V(y_1, y_2) = \Phi(ay_1) + \Phi(ay_1 + by_2) + \frac{a}{\rho} y_2^2
\]

is a strict Lyapunov function for the two
dimensional system consisting of $y_1$ and $y_2$. In the above theorem we allow $\rho$ to be bigger than $\frac{1}{2}$, and for the particular model studied here, the value $\rho > \frac{1}{2}$ indeed yields far better performance.

5.3.5 Simulation Results

It is guaranteed by our theorems that the control law that we obtained globally stabilizes the linearized model and hence locally the original model. But local stabilization of the nonlinear aircraft model can also be achieved in principle by using linear feedback. Thus we will discuss next, via simulations, the advantages of our control law with respect to such linear feedback, when used for the nonlinear model. Essentially, we obtain a relatively large domain of attraction with respect to the desired equilibrium. Providing explicit bounds on this region of attraction is a difficult problem, and we do not in any way attempt to do so.

We provide now several plots regarding the closed-loop system consisting of (5.3.1-5.3.5) when using the control (5.3.13). We also provide comparisons between our control and the "naive" design, which does not use the saturation in (5.3.13). Finally, we give plots for the behavior that results from leaving the elevator position constant, to illustrate that not applying any control results in far too slowly damped behavior.

The initial values of $\alpha, \theta, q, \delta_c$ in all plots are:

\[
\begin{align*}
\alpha(0) &= \alpha_0 + 0.25 = 0.4900685620, \\
\theta(0) &= \theta_0 + 0.2 = 0.4375883269, \\
q(0) &= 0, \\
\delta_c(0) &= \delta_c0 - 0.04 = 0.01.
\end{align*}
\]

These represent fairly large displacements from equilibrium. The plots will differ depending on the initial value taken for $\mu$.

Figures 5.2, 5.3, and 5.4 are as follows. In each case there are two vertical sets of three plots each. The three left plots are for the design that results when using (5.3.13), while the three right plots are for the design obtained when using instead the "naive"
linear control law
\[ u = -0.9(0.2x_3 + 2.2x_4 + 2x_5), \]
which would stabilize the system in the absence of the saturations in the elevator rate. Observe that (5.3.13) and (5.3.28) coincide for all small deviations \( x \) from the desired equilibrium. The top two plots are for the forward speed \( \mu \), the bottom plots are for \( \dot{\delta}_e \) (so that the effective control being applied is illustrated; notice the saturation at 0.01), and the middle plot displays the variables \( (\alpha, \theta, q, \delta_e) \), plotted according to the linestyles explained in the captions.

In Figure 5.2, the initial value for \( \mu \) is
\[ \mu_0 - 18 = 371.1315833, \]
while in Figure 5.3 it is
\[ \mu_0 - 63 = 326.1315833, \]
and in Figure 5.4 it is
\[ \mu_0 - 100 = 289.1315833. \]

These values were chosen for the following reasons. Simulations show that our design stabilizes the nonlinear system for a wide range of values. Considering in particular the given initial values of \( \alpha, \theta, q, \delta_e \), we wanted to compare values in the range \( \mu \in [-100, 0] \). For this range the "naive" design is stable roughly only in the range \([-62, -18]\) (the domain of attraction is a connected subset of \( \mathbb{R}^5 \)). Thus we considered the extreme cases \(-18\), \(-63\), and \(-100\). In the first of these cases, the "naive" design would stabilize as well, as shown, though the difference in performance is quite striking. In the second and third cases, the "naive" design results in instability.

The original model is marginally stable, and applying no control \( (\dot{\delta}_e \equiv 0) \) results in essentially oscillatory behavior. For purposes of comparison, we include the plots of \( \mu \) and of \( \alpha, \theta, q, \delta_e \) corresponding to the first two cases above, that is, \( \mu_0 - 18 = 371.1315833 \) and \( \mu_0 - 63 = 326.1315833 \) in Figures 5.5-5.6, respectively.
Figure 5.2: $\mu_0 = 371$, saturated and naive design; $\alpha$: ---; $\theta$: --; $g$: · · ·; $\delta_c$: · · ·
Figure 5.3: $\mu_0 = 326$, saturated and naive design; $\alpha$: ---; $\theta$: --.; $g$: ··.; $\delta_x$: ---
Figure 5.4: $\mu_0 = 289$, saturated design; $\alpha$: ——; $\theta$: --; $q$: ···; $\delta_2$: ---
Figure 5.5: $\mu_0 = 371$, no control; $\alpha$: --; $\theta$: --; $q$: ..; $\delta_c$: --

Figure 5.6: $\mu_0 = 326$, no control; $\alpha$: --; $\theta$: --; $q$: ..; $\delta_c$: --
Chapter 6

Discrete-Time Linear Systems

In Chapters 2 to 5 we gave a complete analysis of the global stabilizability of linear continuous-time systems with bounded feedback. In this chapter we discuss the global stabilizability of linear discrete-time systems with bounded feedback. We will present a necessary and sufficient condition for linear discrete-time systems to be globally stabilizable with bounded feedback and present a constructive proof for the sufficient part. Following the philosophy of the proof, one can easily obtain an algorithm to find such bounded stabilizing feedbacks.

6.1 Statement of the Main Results

First, we define the space of feedback functions. Let \( \mathcal{S} \) be the class consisting of functions of the type

\[
\sigma(s) = \delta \text{sat}(s/\delta),
\]

where \( \delta > 0 \), and

\[
\text{sat}(s) = \text{sign}(s) \cdot \min\{|s|, 1\}.
\]

(This is a sub-class of the saturation functions used in Chapter 4.) For an \( m \)-tuple \( l = (l_1, \ldots, l_m) \) of nonnegative integers, and a finite sequence \( \boldsymbol{\sigma} = (\sigma_1, \ldots, \sigma_{|l|}) = (\sigma^1_{l_1}, \ldots, \sigma^{l_1}_{l_1}, \ldots, \sigma^{l_m}_{l_m}) \) of functions in \( \mathcal{S} \), we define sets of functions \( \mathcal{F}_n^l(\boldsymbol{\sigma}) \) and \( \mathcal{G}_n^l(\boldsymbol{\sigma}) \) in a manner entirely analogous to what was done in Chapter 4. Also, we let \( ||\boldsymbol{\sigma}|| \) denote the maximal bound of all \( \sigma_i \). (Since we are using the special class \( \mathcal{S} \), the functions in \( \mathcal{F}_n^l(\boldsymbol{\sigma}) \) and \( \mathcal{G}_n^l(\boldsymbol{\sigma}) \) are symmetric about the origin. Thus there is no difference to write a feedback in a form \( u = k(x) \) or \( u = -k(x) \) with \( k \in \mathcal{F}_n^l(\boldsymbol{\sigma}) \) or \( k \in \mathcal{G}_n^l(\boldsymbol{\sigma}) \).)
Let $\delta > 0$. We say that a function $f : \mathbb{Z}_{\geq 0} \to \mathbb{R}^n$ is eventually bounded by $\delta$ (and write $|f| \leq_{ev} \delta$), if there exists $T > 0$ such that $|f(t)| \leq \delta$ for all $t \geq T$. Given an $n$-dimensional system $\Sigma : x(t + 1) = f(x(t))$, we say that $\Sigma$ is CICS (converging-input converging-state) if, whenever $\{e(t)\}_0^\infty \in l_1$, every solution $t \to x(t)$ of $x(t + 1) = f(x(t)) + e(t)$ converges to zero as $t \to \infty$. For a system $x(t + 1) = f(x(t), u(t))$, we say that a feedback $u = k(x)$ is stabilizing if 0 is a globally asymptotically stable equilibrium of the closed-loop system $x(t + 1) = f(x(t), k(x(t)))$. If, in addition, this closed-loop system is CICS, then we will say that $k$ is CICS-stabilizing.

For a square matrix $A$, let $N(A)$ be the number of eigenvalues $z$ of $A$ such that $|z| = 1$ and $\text{Im} \ z \geq 0$, counting multiplicities. Then we have the following result:

**Theorem 6.1** Let $\Sigma$ be a linear system $x(t + 1) = Ax(t) + Bu(t)$ with state space $\mathbb{R}^n$ and input space $\mathbb{R}^m$. Assume that $\Sigma$ is asymptotically null-controllable and does not have an unstable part, i.e., that all the eigenvalues of $A$ have magnitudes less than or equal to 1, and all the eigenvalues of the uncontrollable part of $A$ have magnitudes strictly less than 1. Let $N = N(A)$. Then, for every $\varepsilon > 0$, there exist a sequence $\sigma = (\sigma_1, \cdots, \sigma_N)$ of functions belonging to $S$ with $||\sigma|| \leq \varepsilon$ and an $m$-tuple $l = (l_1, \cdots, l_m)$ of nonnegative integers such that $|l| = N$, for which there are CICS-stabilizing feedbacks

$$u = k_\mathcal{F}(x) \quad (6.1.1)$$

$$u = k_\mathcal{G}(x) \quad (6.1.2)$$

such that $k_\mathcal{F} \in \mathcal{F}_{\|\sigma\|}^l$, $k_\mathcal{G} \in \mathcal{G}_{\|\sigma\|}^l$.

We will say that (6.1.1), (6.1.2) are feedbacks of Type $\mathcal{F}$, $\mathcal{G}$, respectively.

A linear discrete-time system $\Sigma$ is **bounded feedback stabilizable** (BFS) if there exists a bounded locally Lipschitz feedback $k$ that stabilizes $\Sigma$. A linear discrete-time system $\Sigma$ is **small feedback stabilizable** (SFS) if for every $\varepsilon > 0$ there exists a stabilizing feedback $k$ for $\Sigma$ such that $||k(x)|| \leq \varepsilon$ for all $x$. From Theorem 6.1 we can easily get the following result.

**Theorem 6.2** Let $\Sigma$ be a linear discrete-time system. Then the following conditions are equivalent:
1. \( \Sigma \) is BFS,

2. \( \Sigma \) is SFS,

3. \( \Sigma \) is asymptotically null-controllable and does not have an unstable part.

We remark that, for discrete-time systems, we have a conclusion similar to the one of Theorem 2.1. Here, we only present a short version with a simple proof.

### 6.2 Preliminaries

In this section we present two technical lemmas that will be needed for the proof of Theorem 6.1. The first lemma is similar to Lemma 4.2.1 corresponding to discrete-time systems. The second lemma is similar to Lemma 4.2.2, but we restrict our analysis to the special saturation class defined above. Therefore the proof is much simpler.

**Lemma 6.2.1** Consider an \( n \)-dimensional linear single-input system

\[
\Sigma: \quad x(t + 1) = Ax(t) + bu(t). \tag{6.2.1}
\]

Suppose that \( (A, b) \) is a controllable pair and that all the eigenvalues of \( A \) have magnitude 1.

(i) If \( \lambda = 1 \) or \( \lambda = -1 \) is an eigenvalue of \( A \), then there is a linear change of coordinates

\[
Tx = (y_1, \cdots, y_n)' = (\tilde{y}', y_n)' \quad \text{of} \quad \mathbb{R}^n \quad \text{that transforms } \Sigma \text{ into the form}
\]

\[
\tilde{y}(t + 1) = A_1 \tilde{y}(t) + b_1 (y_n(t) + u(t)) ,
\]

\[
y_n(t + 1) = \lambda (y_n(t) + u(t)) ,
\]

where the pair \( (A_1, b_1) \) is controllable and \( y_n \) is a scalar variable.

(ii) If \( A \) has an eigenvalue of the form \( \alpha + \beta i \), with \( \beta \neq 0 \), then there is a linear change of coordinates \( Tx = (y_1, \cdots, y_n)' = (\tilde{y}', y_{n-1}, y_n)' \) of \( \mathbb{R}^n \) that transforms \( \Sigma \) into the form

\[
\tilde{y}(t + 1) = A_1 \tilde{y}(t) + b_1 (y_n(t) + u(t)) ,
\]

\[
y_{n-1}(t + 1) = \alpha y_{n-1}(t) - \beta (y_n(t) + u(t)) ,
\]

\[
y_n(t + 1) = \beta y_{n-1}(t) + \alpha (y_n(t) + u(t)) ,
\]

where the pair \( (A_1, b_1) \) is controllable and \( y_{n-1}, y_n \) are scalar variables.
Proof. We first prove (i). If \( \lambda = 1 \) or \( \lambda = -1 \) is an eigenvalue of \( A \), then there exists a nonzero \( n \)-dimensional row vector \( v \) such that \( vA = \lambda v \). Let \( \xi : \mathbb{R}^n \to \mathbb{R} \) be the linear function \( x \to vx \). Then, along trajectories of \( \Sigma \), \( \xi(t + 1) = \lambda \xi(t) + (vb)u(t) \). The controllability of \( (A, b) \) implies that \( vb \neq 0 \). So we may assume that \( vb = \lambda \). Make a linear change of coordinates \( Tx = (\bar{x'}, \bar{z}_n) \) so that \( \bar{z}_n \equiv \xi \). Then the system equations are of the form
\[
\bar{z}(t + 1) = A_1 \bar{z}(t) + \bar{z}_n(t) \tilde{b}_1 + u(t) \tilde{b}_2,
\]
\[
\bar{z}_n(t + 1) = \lambda \bar{z}_n(t) + \lambda u(t).
\]
It is clear that every eigenvalue of \( A_1 \) also has magnitude 1. Now change coordinates again by letting \( \bar{y} = \bar{z} + \bar{z}_n \tilde{b}_3 \), \( y_n = \bar{z}_n \), where the vector \( \tilde{b}_3 \) will be chosen below. Then the system equations become
\[
\bar{y}(t + 1) = A_1 \bar{y}(t) + y_n(t)(\tilde{b}_1 + (\lambda I - A_1) \tilde{b}_3) + u(t)(\tilde{b}_2 + \lambda \tilde{b}_3),
\]
\[
y_n(t + 1) = \lambda (y_n(t) + u(t)).
\]
Choose \( \tilde{b}_3 \) to be a solution of \( \bar{b}_2 + \lambda \tilde{b}_3 = (\bar{b}_1 + (\lambda I - A_1) \tilde{b}_3) \), i.e., \( A_1 \tilde{b}_3 = \bar{b}_1 - \bar{b}_2 \). (This is possible because \( A_1 \) is nonsingular.) Let \( b_1 = \bar{b}_1 + (\lambda I - A_1) \tilde{b}_3 \). The equations now become (6.2.2) as desired.

We now prove (ii). Let \( \alpha + \beta i, \beta \neq 0 \), be an eigenvalue of \( A \). Then there exist two independent \( n \)-dimensional row vectors \( v_1, v_2 \) such that \( v_1A = \alpha v_1 - \beta v_2 \), \( v_2A = \beta v_1 + \alpha v_2 \). Let \( \xi, \eta \) be the linear functionals \( x \to v_1x, v_2x \). Then, along trajectories of \( \Sigma \),
\[
\xi(t + 1) = \alpha \xi(t) - \beta \eta(t) + v_1bu(t),
\]
\[
\eta(t + 1) = \beta \xi(t) + \alpha \eta(t) + v_2bu(t).
\]
The controllability of \( (A, b) \) implies that \( v_1b \) and \( v_2b \) cannot both be zero. Without loss of generality, we assume \( v_1b = -\beta \neq 0 \) and let \( \tau = v_2b \). Let
\[
P = \frac{1}{\beta^2 + \tau^2} \begin{pmatrix} \beta^2 + \alpha \tau & \beta (\alpha - \tau) \\ \beta (\tau - \alpha) & \beta^2 + \alpha \tau \end{pmatrix}.
\]
Make a linear change of coordinates $T\mathbf{x} = (\mathbf{z}', z_{n-1}', z_n)'$ so that $(z_{n-1}', z_n)' = P(\xi, \eta)'$. Then the system equations are of the form
\begin{align*}
z(t+1) &= A_1 z(t) + z_{n-1}(t)\hat{b}_1 + z_n(t)\hat{b}_2 + u(t)\hat{b}_3, \\
z_{n-1}(t+1) &= \alpha z_{n-1}(t) - \beta (z_n(t) + u(t)), \\
z_n(t+1) &= \beta z_{n-1}(t) + \alpha (z_n(t) + u(t)),
\end{align*}
and every eigenvalue of $A_1$ has magnitude 1. Now change coordinates again by letting
\[ \tilde{\mathbf{y}} = \mathbf{z} + z_{n-1}\tilde{b}_4 + z_n\tilde{b}_5, \]
where the vectors $\tilde{b}_4$, $\tilde{b}_5$ will be chosen below. Then the last two equations of (6.2.4) are as desired, and the equation of $\tilde{\mathbf{y}}$ becomes
\begin{align*}
\tilde{\mathbf{y}}(t+1) &= A_1 \tilde{\mathbf{y}}(t) + y_{n-1}(t)(\tilde{b}_1 - A_1 \tilde{b}_4 + \alpha \tilde{b}_5) \\
&\quad + y_n(t)(\tilde{b}_2 - A_1 \tilde{b}_5 + \alpha \tilde{b}_4) + u(t)(\tilde{b}_3 - \beta \tilde{b}_4 + \alpha \tilde{b}_5).
\end{align*}
If we could choose $\tilde{b}_4$, $\tilde{b}_5$ such that
\[ \tilde{b}_1 - A_1 \tilde{b}_4 + \alpha \tilde{b}_5 = 0 \]
and
\[ \tilde{b}_3 - \beta \tilde{b}_4 + \alpha \tilde{b}_5 = \tilde{b}_2 - A_1 \tilde{b}_5 - \beta \tilde{b}_4 + \alpha \tilde{b}_5, \]
then we could let
\[ b_1 = \tilde{b}_2 - A_1 \tilde{b}_5 - \beta \tilde{b}_4 + \alpha \tilde{b}_5 \]
and the system equations would become (6.2.3) as desired. To prove the existence of $\tilde{b}_4$ and $\tilde{b}_5$, we rewrite (6.2.7) as $A_1 \tilde{b}_5 = \tilde{b}_2 - \tilde{b}_3$, from which we get $\tilde{b}_5$ because $A_1$ is nonsingular. Then from (6.2.6), we have $(A_1 - \alpha I)\tilde{b}_4 = \tilde{b}_1 + \beta \tilde{b}_5$. Since the eigenvalues of $A_1$ have magnitude 1 and $\alpha \neq \pm 1$, the matrix $A_1 - \alpha I$ is nonsingular, and so $\tilde{b}_4$ exists.

\begin{lemma}
Let $a, b$ be two real constants such that $a^2 + b^2 = 1$, $b \neq 0$. Let $e_j = (e_j(0), e_j(1), e_j(2), \cdots)$, $j = 1, 2$, be two elements of $l_1$. Given $\delta > 0$, let $\varepsilon \in (0, \frac{\delta}{4})$ and $|v(t)| \leq_{cv} \varepsilon$. Then, if $\gamma = (x(\cdot), y(\cdot)) : \mathbb{Z}_{\geq 0} \rightarrow \mathbb{R}^2$ is any solution of the system
\begin{align*}
x(t+1) &= ax(t) - by(t) + bu(t) + e_1(t), \\
y(t+1) &= bx(t) + ay(t) - au(t) + e_2(t),
\end{align*}
\end{lemma}
where
\[ u(t) = \sigma(y(t) + \xi v(t)) + \eta v(t), \quad (6.2.10) \]
and \( \xi + \eta = 1, \xi, \eta \geq 0, \) and \( \sigma(s) = \delta\text{sat}(s/\delta), \) it follows that
\[ \limsup_{t\to+\infty}||\gamma(t)|| < r = \frac{1}{|b|}(7|a| + 4)\varepsilon + 7\varepsilon. \quad (6.2.11) \]

**Proof.** Without loss of generality, we assume \( b > 0. \) Let \( \theta = \arctan\left(\frac{b}{a}\right), \) \( 0 < \theta < \pi, \) if \( a \neq 0, \) and \( \theta = \frac{\pi}{2} \) if \( a = 0. \) Then \( a + bi = e^{i\theta}. \) Let \( z(t) = z(t) + iy(t), c(t) = e_1(t) + ie_2(t). \) Then
\[ z(t + 1) = e^{i\theta}(z(t) - iu(t)) + c(t). \quad (6.2.12) \]

Again, without loss of generality, we assume that \( ||c||_1 < \varepsilon, \) otherwise we can find \( T > 0 \) such that \( \sum_{t \geq T} |c(t)| < \varepsilon, \) and then we only need to consider the solution for \( t \geq T. \) Similarly, we assume \( |v(t)| \leq \varepsilon \) for all \( t. \) So
\[ |z(t + 1)| \leq |z(t) - iu(t)| + |c(t)| \]
\[ = \sqrt{|z(t)|^2 + (y(t) - u(t))^2} + |c(t)| \]
\[ = \sqrt{|z(t)|^2 - u(t)^2(y(t) - u(t)) + |c(t)|} \]
\[ = |z(t)| + w(t) + |c(t)|, \quad (6.2.13) \]

where
\[ w(t) = \frac{-u(t)(2y(t) - u(t))}{|z(t)| + \sqrt{|z(t)|^2 - u(t)(2y(t) - u(t))}}. \quad (6.2.14) \]

If \( |y(t)| \geq 3\varepsilon, \) then from (6.2.10) it follows that
\[ 2\varepsilon \leq |u(t)| \leq \frac{4}{3}|y(t)|, \]
and \( u(t) \) has the same sign as \( y(t). \) So
\[ w(t) \leq -\frac{2\varepsilon \cdot \frac{2}{3}|y(t)|}{2|z(t)|} = -\frac{2\varepsilon^2}{|z(t)|}. \]

From (6.2.13), we have
\[ |z(t + 1)| \leq |z(t)| - \frac{2\varepsilon^2}{|z(t)|} + |c(t)|, \quad \text{if } |y(t)| \geq 3\varepsilon. \quad (6.2.15) \]

When \( |y(t)| < 3\varepsilon, \) since \( |y(t) + \xi v(t)| < 4\varepsilon \leq \delta, \) it follows that
\[ u(t) = y(t) + v(t). \quad (6.2.16) \]
Therefore,
\[ w(t) \leq -\frac{(y(t) + v(t))(y(t) - v(t))}{2|z(t)|} = \frac{v(t)^2 - y(t)^2}{2|z(t)|} \leq \frac{\varepsilon^2}{2|z(t)|}. \]  
(6.2.17)

We conclude that
\[ |z(t + 1)| \leq |z(t)| + \frac{\varepsilon^2}{2|z(t)|} + |e(t)|. \]  
(6.2.18)

In addition,
\[ |y(t + 1)| \geq b|x(t)| - |a|(|y(t)| + |u(t)|) - |e_x(t)| \geq b|x(t)| - (7|a| + 1)\varepsilon. \]

If \( |x(t)| \geq \frac{1}{b}(7|a| + 4)\varepsilon \), then
\[ |y(t + 1)| \geq 3\varepsilon, \]  
(6.2.19)

and also \( |x(t)| \geq 4\varepsilon \), which implies \( |z(t)| \geq 4\varepsilon \). Since \( |e(t)| \leq \varepsilon \), from (6.2.18) it follows that
\[ |z(t + 1)| \leq |z(t)| + \frac{\varepsilon^2}{8\varepsilon} + \varepsilon \leq \frac{41}{32}|z(t)|. \]  
(6.2.20)

On the other hand, since \( |y(t + 1)| \geq 3\varepsilon \), applying (6.2.15) for \( z(t + 2) \), we conclude that
\[ |z(t + 2)| \leq |z(t + 1)| - \frac{2\varepsilon^2}{|z(t + 1)|} + |e(t + 1)|. \]  
(6.2.21)

Using (6.2.18) and (6.2.20) to substitute \( |z(t + 1)| \) in the first and second terms of (6.2.21), we end up with
\[ |z(t + 2)| \leq |z(t)| + \frac{\varepsilon^2}{2|z(t)|} \frac{64\varepsilon^2}{41|z(t)|} + |e(t)| + |e(t + 1)| \]
\[ < |z(t)| - \frac{\varepsilon^2}{|z(t)|} + |e(t)| + |e(t + 1)|. \]

Summarizing, we have proved:

Fact I:

(i) if \( |y(t)| \geq 3\varepsilon \), then
\[ |z(t + 1)| \leq |z(t)| - \frac{2\varepsilon^2}{|z(t)|} + |e(t)|; \]  
(6.2.22)

(ii) if \( |y(t)| < 3\varepsilon \), and \( |x(t)| \geq \frac{1}{b}(7|a| + 4)\varepsilon \), then
\[ |z(t + 2)| \leq |z(t)| - \frac{\varepsilon^2}{|z(t)|} + |e(t)| + |e(t + 1)|. \]  
(6.2.23)
As a consequence of Fact I, we have

Fact II: there exists \( t > 0 \) such that \( z(t) \) is in the region

\[
\mathcal{R} = \{ x + yi : |x| \leq \frac{1}{b} (7|a| + 4) \varepsilon, |y| \leq 3\varepsilon \}.
\]

Indeed, if Fact II were not true, then for any \( t > 0 \) we would have either \(|y(t)| \geq 3\varepsilon\) or \(|z(t)| \geq \frac{1}{b} (7|a| + 4) \varepsilon\). Now we select a sequence \((t_0, t_1, t_2, \ldots)\) of integers in the following way:

- \( t_0 = 0 \),
- for \( j \geq 0 \), if (6.2.22) is true for \( t = t_j \), then \( t_{j+1} = t_j + 1 \); otherwise \( t_{j+1} = t_j + 2 \).

Then we have

\[
|z(t_{j+1})| \leq |z(t_j)| - \frac{\varepsilon^2}{|z(t_j)|} + \sum_{k=t_j}^{t_{j+1}-1} |\epsilon(k)|. \tag{6.2.24}
\]

Summing (6.2.24) for \( j = 0, 1, 2, \ldots, n \), we have

\[
|z(t_{n+1})| \leq |z(0)| - \varepsilon^2 \sum_{k=0}^{n} \frac{1}{|z(t_k)|} + \sum_{k=0}^{t_{n+1}-1} |\epsilon(k)|. \tag{6.2.25}
\]

In particular, we have

\[
|z(t_{n+1})| \leq |z(0)| + \|\epsilon\|_1 = N \tag{6.2.26}
\]

for all \( n \geq 0 \). So from (6.2.25) it follows that

\[
|z(t_{n+1})| \leq |z(0)| - (n + 1) \varepsilon^2 / N + \|\epsilon\|_1. \tag{6.2.27}
\]

Let \( n \to \infty \). Then \( |z(t_{n+1})| \to -\infty \), which is a contradiction. So Fact II is proved.

To complete the proof of the lemma, it is enough to show the next fact.

Fact III: if \( z(T) \in \mathcal{R} \) for some \( T \geq 0 \), then \( |z(t)| \leq r \) for all \( t \geq T \).

Note that if \( z(t) \in \mathcal{R} \), then

\[
|z(t)| \leq \frac{1}{b} (7|a| + 4) \varepsilon + 3\varepsilon. \tag{6.2.28}
\]

If for some \( T_1 \), \( z(T_1) \notin \mathcal{R} \), but \( z(T_1 - 1) \in \mathcal{R} \), then from Fact II it follows that there exists \( T_2 > T_1 \) such that \( z(T_2) \in \mathcal{R} \), and \( z(t) \notin \mathcal{R} \) for \( T_1 \leq t < T_2 \). Now we
select $t_0 = T_1, t_1, t_2, \ldots, t_n = T_2$ as we did above such that (6.2.24) is satisfied for $j = 0, 1, 2, \ldots, n$. Then
\begin{equation}
|z(t_j)| \leq |z(t_0)| + \sum_{k=t_0}^{t_j-1} |e(k)|
\end{equation}
for $1 \leq j \leq n$. Note that $z(t_0) = e^{id}(z(T_1-1)-iu(T_1-1))+e(T_1-1)$, and $z(T_1-1) \in \mathcal{R}$. From (6.2.17) we see that $w(T_1-1) > 0$ becomes possible only if $|y(T_1-1)| < \varepsilon$. When $|y(T_1-1)| < \varepsilon$, we have $|u(T_1-1)| < 2\varepsilon$. So $|z(t_0)| \leq |z(T_1-1)| + 2\varepsilon + |e(T_1-1)|$.

Substituting this into (6.2.29), we end up with
\begin{equation}
|z(t_j)| \leq |z(T_1-1)| + 2\varepsilon + \sum_{k=T_1-1}^{t_j-1} |e(k)|, \quad 0 \leq j \leq n.
\end{equation}

If $t_j - t_{j-1} = 2$, then from (6.2.18) and (6.2.30) we have
\begin{equation}
|z(t_j + 1)| \leq |z(t_j)| + \frac{\varepsilon^2}{2|z(t_j)|} + |e(t_j)| \leq |z(T_1-1)| + \frac{\varepsilon^2}{2|z(t_j)|} + 3\varepsilon.
\end{equation}
Since $|z(t_j)| \geq 3\varepsilon$, we have $|z(t_j + 1)| \leq |z(T_1-1)| + 4\varepsilon$. From (6.2.28) we conclude that
\begin{equation}
|z(t_j + 1)| \leq \frac{1}{b}(7|a| + 4)\varepsilon + 7\varepsilon
\end{equation}
when $t_{j+1} - t_j = 2$. Inequalities (6.2.28), (6.2.30), and (6.2.31) imply that $|z(t)| \leq r$ for $T_1 \leq t < T_2$. So Fact III is established.

\begin{proof}
When $n = 2$, the conclusion follows from Lemma 6.2.2. Indeed, assume $\alpha \pm \beta i$ are the eigenvalues of $J$. Then $\beta \neq 0$. From Lemma 6.2.2, if $\theta \leq \delta/4$, then
\begin{equation}
\limsup_{t \to +\infty} ||\gamma(t)|| < \frac{1}{\theta}(7|a| + 4)\theta + 7\theta.
\end{equation}
\end{proof}

\textbf{Corollary 6.2.3} For $n = 1$ or 2, let $J$ be an $n \times n$ orthonormal matrix. Let $b = 1$ if $n = 1$, and $b = (0,1)^t$ if $n = 2$. Assume that $(J,b)$ is a controllable pair. Then for every $\varepsilon > 0$, $\delta > 0$ there exists $\theta > 0$ such that for any functions $v : \mathbb{Z}_{\geq 0} \to \mathbb{R}$ and $e : \mathbb{Z}_{\geq 0} \to \mathbb{R}^n$, where $v \leq_{cv} \theta, e \in L_1$, if $\gamma : \mathbb{Z}_{\geq 0} \to \mathbb{R}^n$ is any solution of the system
\begin{equation}
z(t+1) = J \left( z(t) - \sigma(x_n(t) - \xi v(t))b + \eta v(t)b \right) + e(t),
\end{equation}
where $\sigma(s) = \delta \text{sat}(s/\delta)$, $\xi + \eta = 1$, $\xi, \eta \geq 0$, it follows that
\begin{equation}
\limsup_{t \to +\infty} ||\gamma(t)|| < \varepsilon.
\end{equation}

\textbf{Proof}. When $n = 2$, the conclusion follows from Lemma 6.2.2. Indeed, assume $\alpha \pm \beta i$ are the eigenvalues of $J$. Then $\beta \neq 0$. From Lemma 6.2.2, if $\theta \leq \delta/4$, then
\begin{equation}
\limsup_{t \to +\infty} ||\gamma(t)|| < \frac{1}{\theta}(7|a| + 4)\theta + 7\theta.
\end{equation}
So if we choose \( \theta = \min\{\frac{\xi}{4}, \frac{\varepsilon}{7 + (7|d| + 4)/|b|}\} \), then the conclusion follows.

Next, we prove the conclusion for \( n = 1 \). If \( n = 1 \), the equation of the system becomes

\[
x(t + 1) = \lambda(x(t) - \sigma(x(t) - \xi v(t)) + \eta v(t)) + e(t),
\]

where \( \lambda = \pm 1 \). For any \( \theta > 0 \), without loss of generality, we assume \( |e|_1 < \theta \). If \( |v(t)| \leq \theta \leq \delta/3 \), then for \( |x(t)| \geq 3\theta \) we have \( |\sigma(x(t) - \xi v(t)) - \eta v(t)| \geq 2\theta \), and \( \sigma(x(t) - \xi v(t)) - \eta v(t) \) has the same sign as \( x(t) \). So if \( |x(t)| \geq 3\theta \), then

\[
|x(t + 1)| \leq |x(t) - \sigma(x(t) - \xi v(t)) + \eta v(t)| + |e(t)| \leq |x(t)| - \theta.
\]

It is easy to see that \( \limsup_{t \to +\infty} |x(t)| \leq 3\theta \). Now, to obtain the conclusion of the corollary, it suffices to take \( \theta = \min\{\delta/3, \varepsilon/3\} \).

\[\square\]

### 6.3 The Proof of Theorem 6.1

First, we notice that under the conditions of the theorem there exists a linear change of coordinates of the state space that transforms \( \Sigma \) into the block form

\[
\Sigma : \begin{cases}
  x_1(t + 1) = A_1x_1(t) + B_1u(t), & x_1(t) \in \mathbb{R}^{n_1}, \\
  x_2(t + 1) = A_2x_2(t) + B_2u(t), & x_2(t) \in \mathbb{R}^{n_2},
\end{cases}
\]

where (i) \( n_1 + n_2 = n \), (ii) all the eigenvalues of \( A_1 \) have magnitude 1, (iii) all the eigenvalues of \( A_2 \) have magnitude less than 1, and (iv) \((A_1, B_1)\) is a controllable pair.

Suppose that we find a CICS-stabilizing feedback \( u = k(x_1) \) of Type \( \mathcal{F} \) or Type \( \mathcal{G} \) for the system

\[
x_1(t + 1) = A_1x_1(t) + B_1u(t)
\]

such that the resulting closed-loop system is asymptotically stable. Then it is clear that this same feedback law will work for \( \Sigma \) as well. Thus, in order to stabilize \( \Sigma \), it is enough to stabilize the "critical subsystem"

\[
x_1(t + 1) = A_1x_1(t) + B_1u(t).
\]

Without loss of generality, in our proof of the theorem we will suppose that \( \Sigma \) is already in this form.

We start with the single-input case, and prove the theorem by induction on the dimension of the system. As discussed earlier, we may assume that all the eigenvalues of \( A \) have magnitude 1 and the pair \((A, B)\) is controllable.
For dimension zero there is nothing to prove. Now assume that we are given a single-input $n$-dimensional system, $n \geq 1$, and suppose that Theorem 6.1 has been established for all single-input systems of dimension less than or equal to $n - 1$. We consider separately the following two possibilities:

(i) 1 or $-1$ is an eigenvalue of $A$,

(ii) neither 1 nor $-1$ is an eigenvalue of $A$.

Recall that $N = N(A)$. We want to prove the existence of $CICS$-stabilizing feedbacks $u = -k_{\mathcal{F}}(x)$ and $u = -k_{\mathcal{G}}(x)$, where $k_{\mathcal{F}} \in \mathcal{F}_n(\sigma)$, $k_{\mathcal{G}} \in \mathcal{G}_n(\sigma)$, for some finite sequence $\sigma = (\sigma_1, \ldots, \sigma_N)$ of functions in $\mathcal{S}$, with $||\sigma|| \leq \varepsilon$.

In Case (i), we apply Part (i) of Lemma 6.2.1 and rewrite our system in the form

\[
\begin{align*}
\bar{y}(t + 1) &= A_1 \bar{y}(t) + (y_n(t) + u(t))b_1, \\
y_{n}(t + 1) &= \lambda(y_n(t) + u(t)),
\end{align*}
\]

where $\bar{y} = (y_1, \ldots, y_{n-1})'$. (Note that if $n = 1$, only the second equation appears.) In Case (ii), since $n > 0$, $A$ has a pair of eigenvalues of the form $\alpha + \beta i$, with $\beta \neq 0$. So we apply Part (ii) of Lemma 6.2.1 and make a linear transformation that puts $\Sigma$ in the form

\[
\begin{align*}
\bar{y}(t + 1) &= A_1 \bar{y}(t) + (y_n(t) + u(t))b_1, \\
y_{n-1}(t + 1) &= \alpha y_{n-1}(t) - \beta(y_n(t) + u(t)), \\
y_{n}(t + 1) &= \beta y_{n-1}(t) + \alpha(y_n(t) + u(t)),
\end{align*}
\]

where $\bar{y} = (y_1, y_2, \ldots, y_{n-2})'$. (In the special case when $n = 2$, the first equation will be missing.) So, by using Lemma 6.2.1, we can rewrite our system in the form

\[
\begin{align*}
\bar{y}(t + 1) &= A_1 \bar{y}(t) + (y_n(t) + u(t))b_1, \\
\bar{y}(t + 1) &= J(\bar{y}(t) + u(t)b_0),
\end{align*}
\]

where $J$ is an orthonormal matrix, $(J, b_0)$ is a controllable pair, and in Case (i) we have $\bar{y} = y_n, b_0 = 1$, and in Case (ii) we have $\bar{y} = (y_{n-1}, y_n)'$, $b_0 = (0, 1)'$. To get a $CICS$-stabilizing feedback, we consider the following system:

\[
\begin{align*}
\bar{y}(t + 1) &= A_1 \bar{y}(t) + (y_n(t) + u(t))b_1 + \bar{c}(t), \\
\bar{y}(t + 1) &= J(\bar{y}(t) + u(t)b_0) + \bar{c}(t),
\end{align*}
\]
where $\bar{c}, \bar{c} \in \ell_1$.

Let

$$u = \sigma_N(-y_n + \xi v) + \eta v = -\sigma_N(y_n - \xi v) + \eta v,$$  \hfill (6.3.5)

where $\xi$ and $\eta$ are constants such that $\xi \eta = 0, \xi + \eta = 1$, $\sigma_N(s) = \varepsilon \text{sat}(s/\varepsilon)$, and $v$ is to be chosen later.

From Corollary 6.2.3 we see that there exists $\theta > 0$, $(\theta < \varepsilon/2)$, such that if $|v(t)| \leq \varepsilon v \theta$, then all trajectories of (6.3.4) satisfy $|\bar{y}| \leq \varepsilon v \varepsilon/2$. Therefore $u(t)$ will eventually become $-y_n(t) + v(t)$, and the first equation of (6.3.4) eventually becomes

$$\bar{y}(t + 1) = A_1 \bar{y}(t) + v(t)b_1 + \bar{c}(t).$$ \hfill (6.3.6)

Note that $(A_1, b_1)$ is controllable and all eigenvalues of $A_1$ have magnitude 1. By the inductive hypothesis, we conclude that there exist

$$v_1 = \bar{k}_\mathcal{F} \in \mathcal{F}_n(\bar{\sigma}) \quad \text{and} \quad v_2 = \bar{k}_\mathcal{G} \in \mathcal{G}_n(\bar{\sigma})$$  \hfill (6.3.7)

for some $\bar{\sigma} = (\sigma_1, \ldots, \sigma_{N-1})$ such that $||\bar{\sigma}|| \leq \theta$, and both $v_1$ and $v_2$ stabilize (6.3.6) for any $\bar{c} \in \ell_1$. Let us use $v$ to denote either of $v_1, v_2$. Then $\bar{y}(t) \to 0$ as $t \to +\infty$. Note that $v$ is a linear function of $\bar{y}$ when $\bar{y}$ is small. So (6.3.4) will eventually become a linear asymptotically stable system with an converging input. We conclude that (6.3.7) and (6.3.5) define a CICS-stabilizing feedback for (6.3.3). To get a feedback of Type $\mathcal{F}$, we set $\xi = 1, \eta = 0$ in (6.3.5) and use $v = v_1$. To get a feedback of Type $\mathcal{G}$, we set $\xi = 0, \eta = 1$ in (6.3.5) and use $v = v_2$. Then, for any $r > 0$, the closed-loop system of (6.3.4) with

$$u = -r \sigma_N(y_n/p) + rv(\bar{y}/p)$$  \hfill (6.3.8)

is CICS. Choosing $r$ sufficiently small, we can bound the sum of the coefficients of all saturations in (6.3.8) by 1. The sequence $\sigma = (\sigma_1, \ldots, \sigma_{N-1}, \sigma_N)$ clearly satisfies $||\sigma|| \leq \varepsilon$. So the proof for the single-input case is completed.

Next, we deal with the general case of $m > 1$ inputs and prove Theorem 6.1 by induction on $m$. 
First, we know from the proof above that the theorem is true if \( m = 1 \). Assume that Theorem 6.1 has been established for all \( k \)-input systems, for all \( k \leq m - 1 \), and let \( \Sigma : x(t + 1) = Ax(t) + Bu(t) \) be an \( m \)-input system.

Assume without loss of generality that the first column \( b_1 \) of \( B \) is nonzero and consider the Kalman controllability decomposition of the system \( \Sigma_1 : x(t + 1) = Ax(t) + b_1 u(t) \) (see [27], Lemma 3.3.3). We conclude that, under a change of coordinates \( y = T^{-1}x \), \( \Sigma_1 \) has the form

\[
\begin{align*}
    y_1(t + 1) &= A_1 y_1(t) + A_2 y_2(t) + \bar{b}_1 u(t), \\
    y_2(t + 1) &= A_3 y_2(t),
\end{align*}
\]

where \((A_1, \bar{b}_1)\) is a controllable pair. In these coordinates \( \Sigma \) has the form

\[
\begin{align*}
    y_1(t + 1) &= A_1 y_1(t) + A_2 y_2(t) + \bar{b}_1 u_1(t) + \bar{B}_1 \bar{u}(t), \\
    y_2(t + 1) &= A_3 y_2(t) + \bar{B}_2 \bar{u}(t),
\end{align*}
\] (6.3.9)

where \( \bar{u} = (u_2, \cdots, u_m)' \) and \( \bar{B}_1, \bar{B}_2 \) are appropriate matrices. So it suffices to show the conclusion for (6.3.9). Let \( n_1, n_2 \) denote the dimensions of \( y_1, y_2 \), respectively. Recall that \( N = N(A) \). For the single-input controllable system

\[
y_1(t + 1) = A_1 y_1(t) + \bar{b}_1 u_1(t),
\]

there is a feedback

\[
u_1 = k_1(y_1)
\] (6.3.10)

such that (i) \( k_1 \in \mathcal{F}_{n_1}(\sigma_1, \cdots, \sigma_{N_1}) \) (respectively, \( k_1 \in \mathcal{G}_{n_1}(\sigma_1, \cdots, \sigma_{N_1}) \)) where \( N_1 = N(A_1) \); (ii) the resulting closed-loop system is CICS; (iii) \( ||\sigma_1|| \leq \varepsilon \), where \( \sigma_1 = (\sigma_1, \cdots, \sigma_{N_1}) \). Since (6.3.9) is controllable, we conclude that the \((m-1)\)-input subsystem

\[
y_2(t + 1) = A_3 y_2(t) + \bar{B}_2 \bar{u}(t)
\]

is controllable as well. By the inductive hypothesis, this subsystem can be stabilized by a feedback

\[
\bar{u} = \bar{k}(y_2) = (k_2(y_2), \cdots, k_m(y_2))
\] (6.3.11)

such that (i) \( \bar{k} \in \mathcal{F}_{n_2}^\bar{I}(\sigma_{N_1+1}, \cdots, \sigma_N) \) (respectively, \( \bar{k} \in \mathcal{G}_{n_2}^\bar{I} \)), where \( \bar{I} = (N_2, \cdots, N_m) \) is an \((m-1)\)-tuple of nonnegative integers and \( |\bar{I}| = N - N_1 \); (ii) the resulting closed-loop system is CICS; (iii) \( ||\sigma_2|| \leq \varepsilon \), where \( \sigma_2 = (\sigma_{N_1+1}, \cdots, \sigma_N) \). Now, if the input of the
closed-loop system of (6.3.9) with \( u_1, \bar{u} \) given by (6.3.10) and (6.3.11) is an \( l_1 \) function \( \mathbb{Z}_{\geq 0} \rightarrow \mathbb{R}^n \), then \( t \rightarrow y_2(t) \) is an \( l_1 \) function because the linearization of (6.3.9) is a stable system. Therefore, \( t \rightarrow A_2y_2(t) + \bar{B}_1\bar{k}(y_2(t)) \) is also an \( l_1 \) function. So the feedback given in (6.3.10) and (6.3.11) globally stabilizes (6.3.9), and the resulting closed-loop system is CICS. So if we let \( l = (N_1, N_2, \ldots, N_m) \) and \( k = (k_1(y_1), k_2(y_2), \ldots, k_m(y_2)) \), then \( k \in \mathcal{F}_n^l(\sigma) \) (respectively, \( k \in \mathcal{G}_n^l(\sigma) \)), \( \sigma = (\sigma_1, \ldots, \sigma_N) \), satisfies all the required properties as desired. \( \square \)

6.4 The Proof of Theorem 6.2

The implication (2) \( \Rightarrow \) (1) is trivial, and the implication (3) \( \Rightarrow \) (2) is given by Theorem 6.1. So we only need to prove the implication (1) \( \Rightarrow \) (3).

Indeed, if \( \Sigma \) is BFS, then, to begin with, \( \Sigma \) is asymptotically null-controllable. So if we write \( \Sigma \) as three parts: \( \Sigma_s \) (stable part), \( \Sigma_u \) (unstable part), and \( \Sigma_c \) (critical part) in a manner analogous to what was done for continuous-time systems in Chapter 2, then \( \Sigma_c \oplus \Sigma_u \) is controllable. Therefore, it suffices to show \( \Sigma_u \) does not exist.

If \( \Sigma_u \) exists, we write \( \Sigma_u \) as \( x^+_u = A_u x_u + B_u u \). Since \( -A_u \) is Hurwitz, there is a matrix \( P > 0 \) of appropriate size such that \( A_u^t P A_u - P = Q > 0 \). Let \( C > 0 \) and assume that there is a stabilizing feedback \( u = k(x) \) which is bounded by \( C \). Define \( V(x_u) = \langle Px_u, x_u \rangle \). Then we have

\[
V(x_u(t + 1)) - V(x_u(t)) = \langle Qx_u(t), x_u(t) \rangle + \langle PB_u u(t), B_u u(t) \rangle \\
+ \langle PA_u x_u(t), B_u u(t) \rangle + \langle PB_u u(t), A_u x_u(t) \rangle \\
\geq \langle Qx_u(t), x_u(t) \rangle - 2C\|P\| \|A_u\| \|B_u\| \|x_u(t)\|.
\]

So there exists \( R > 0 \) such that \( V(x_u(t + 1)) > V(x_u(t)) \) whenever \( \|x_u\| \geq R \). Then it is obvious that the closed-loop system cannot possibly be globally stable. So \( \Sigma \) does not have unstable part. \( \square \)
Chapter 7

Nonlinear Affine Systems

This chapter deals with questions of global stabilizability of nonlinear systems. The text of this chapter appears in [38].

7.1 General Theory

Consider nonlinear affine systems with single-input of the type

$$\Sigma : \dot{x} = f(x) + g(x)u, \quad (7.1.1)$$

where $x \in \mathbb{R}^n$, $f, g \in C^1(\mathbb{R}^n)$. A function $V : \mathbb{R}^n \to \mathbb{R}$ is said to be a control-Lyapunov function (CLF) of $\Sigma$ if $V$ is continuously differentiable, positive definite, proper, and has the property that $L_x V(x) = 0$ implies either $L_f V(x) < 0$ or $x = 0$. As usual, positive definite means $V(x) > 0$ for all $x \neq 0$ and $V(0) = 0$, and properness means that $V(x) \to \infty$ as $|x| \to \infty$ and that the set $\{x : V(x) \leq r\}$ is compact for every $r > 0$. A control-Lyapunov function $V$ is said to satisfy the small control property (SCP) if for any $\varepsilon > 0$ there is a $\delta > 0$ such that for every $x$ with $|x| < \delta, x \neq 0$, there is a $u$ satisfying $|u| < \varepsilon$ and $L_f V(x) + u L_g V(x) < 0$. The following theorem is an extension of the result obtained in [26].

**Theorem 7.1** Let $\Sigma$ be a nonlinear affine system $\dot{x} = f(x) + g(x)u$, where $f, g \in C^1(\mathbb{R}^n)$ and $f(0) = 0$. Let $V$ be a CLF of $\Sigma$. For $r > 2$, let $u_r : \mathbb{R}^n \to \mathbb{R}$ be a function defined by

$$u_r(x) = \begin{cases} 0, & \text{if } L_g V(x) = 0, \\ \frac{-L_f V(x) + \sqrt{L_f V(x)^2 + |L_g V(x)|^r}}{L_g V(x)} & \text{otherwise}. \end{cases} \quad (7.1.2)$$

Let $\dot{V}$ denote the derivative of $V$ along the vector field of the closed-loop system of $\Sigma$ with the feedback $u = u_r(x)$. Then $\dot{V}(x) < 0$ on $\mathbb{R}^n - \{0\}$ and
(i) if \( r \) is an even integer greater than two, then \( u_r \) has the same regularity as \( L_{fV} \) and \( L_{gV} \) on \( \mathbb{R}^n - \{0\} \);

(ii) if \( r > 2 \) is an arbitrary real number, then \( u_r \in C^1(\mathbb{R}^n - \{0\}) \) provided that 
\[ L_{fV}, L_{gV} \in C^1(\mathbb{R}^n - \{0\}) \].

Moreover, in both cases, if \( V \) also satisfies the SCP, then \( u_r \) is continuous at the origin.

**Proof.** The fact that \( \dot{V}(x) < 0 \) on \( \mathbb{R}^n - \{0\} \) and the continuous differentiability (or the regularity when \( r \) is even) of \( u_r \) at the points \( x \) where \( L_{gV}(x) \neq 0 \) follow trivially. The only thing we need to show is the continuous differentiability (or the regularity) of \( u_r \) at the points \( x \neq 0 \) where \( L_{gV}(x) = 0 \).

To do this, we pick an arbitrary point \( x_0 \neq 0 \) such that \( L_{gV}(x_0) = 0 \). Then \( L_{fV}(x_0) < 0 \) because \( V \) is a CLF. Since \( L_{fV} \) is continuous, it follows that there is a neighborhood \( B \) of \( x_0 \) such that \( L_{fV}(x) < 0 \) on \( B \). Therefore, when \( x \in B \), \( u_r(x) \) is expressed as

\[
 u_r(x) = \text{sign}(L_{gV}(x))[L_{gV}(x)]^{r-1} \left( L_{fV}(x) - \sqrt{L_{fV}(x)^2 + L_{gV}(x)^r} \right)^{-1}. \tag{7.1.3}
\]

If \( r > 2 \) and \( L_{fV}, L_{gV} \in C^1(\mathbb{R}^n - \{0\}) \), then the function \( \text{sign}(L_{gV}(\cdot))[L_{gV}(\cdot)]^{r-1} \) is continuously differentiable at \( x_0 \). It follows from Expression (7.1.3) that \( u_r \) is continuously differentiable at \( x_0 \). If \( r \) is an even integer greater than 2, then \( u_r(x) \) is simply expressed as \( L_gV(x)^{r-1} \left( L_{fV}(x) - \sqrt{L_{fV}(x)^2 + L_{gV}(x)^r} \right)^{-1} \) for \( x \in B \). It is clear that \( u_r \) has the same regularity as \( L_{fV} \) and \( L_{gV} \) at \( x_0 \). This shows the conclusions (i) and (ii).

Moreover, if \( V \) satisfies the SCP, then from [26] it follows that \( u_4 \) defined by (7.1.2) is continuous at the origin. Since

\[
|u_r(x) - u_4(x)| = \left| \left( \sqrt{L_{fV}(x) + L_{gV}(x)^r} - \sqrt{L_{fV}(x)^2 + L_{gV}(x)^4} \right)(L_{gV}(x))^{-1} \right|
\leq \left| L_{gV}(x) \right|^{\frac{r}{2} - 1} - L_{gV}(x)
\]

for \( r > 2 \) and \( L_{gV}(x) \neq 0 \), it follows that

\[
\lim_{x \to 0}(u_r(x) - u_4(x)) = 0.
\]
(The inequality follows from the fact $|\sqrt{a+b} - \sqrt{a+c}| \leq |\sqrt{b} - \sqrt{c}|$ for $a, b, c > 0$. The limit follows from $\lim_{x \to 0} L_y V(x) = 0$.) Therefore $u_\epsilon$ is also continuous at the origin. 

The next theorem gives a sufficient condition for a CLF to satisfy the SCP and shows, in some interesting cases, that the control law given by (7.1.2) is continuously differentiable. First, we recall that a one-parameter group of dilations on a real finite dimensional linear space $L$ is a family $\Delta = \{\Delta_\lambda : \lambda > 0\}$ of linear maps from $L$ to $L$, which is of the form $\Delta_\lambda = \exp(A \log \lambda)$ for some diagonalizable matrix $A$ with positive real spectrum. This is equivalent to saying that, for some choice of linear coordinates $(x_1, \ldots, x_n)$, $\Delta$ is given by

$$\Delta_\lambda(x_1, x_2, \ldots, x_n) = \left(\lambda^{i_1} x_1, \lambda^{i_2} x_2, \ldots, \lambda^{i_n} x_n\right) \overset{\text{def}}{=} \lambda^I x \quad (7.1.4)$$

for some multi-index $I = (i_1, i_2, \ldots, i_n)$. We use $|\Delta|$ to denote the largest eigenvalue of $A$, therefore $|\Delta| = \max(i_1, \ldots, i_n)$ if $\Delta$ is given by (7.1.4). A function $\varphi : L \to \mathbb{R}$ is said to be homogeneous of degree $s$ with respect to $\Delta$ if $\varphi(\Delta_\lambda x) = \lambda^s \varphi(x)$ for all $x \in L$ and all $\lambda > 0$.

**Theorem 7.2** Let $\Sigma$ be a nonlinear system $\dot{x} = f(x) + g(x) u$, where $f, g \in C^1(\mathbb{R}^n)$. Let $V$ be a CLF of $\Sigma$ and $\Delta$ be a one-parameter group of dilations on $\mathbb{R}^n$ defined by (7.1.4). Assume that $L_f V$ and $L_g V$ are homogeneous functions of degrees $a$ and $b$, respectively, $(a, b > 0)$ with respect to $\Delta$.

(i) If $a > b$, then $V$ satisfies the SCP.

(ii) If $a - b > |\Delta|$, then for sufficiently large $r$, namely, $\left(\frac{r}{2} - 1\right)b > |\Delta|$, the feedback $u = u_\epsilon(x)$ given by (7.1.2) is differentiable at the origin.

(iii) If $a - b > |\Delta|$ and $L_f V, L_g V \in C^1(\mathbb{R}^n \setminus \{0\})$, then for sufficiently large $r$, namely, $\left(\frac{r}{2} - 1\right)b > |\Delta|$, the feedback $u = u_\epsilon(x)$ given by (7.1.2) belongs to $C^1(\mathbb{R}^n)$.

(iv) If $a - b > |\Delta|$ and $L_f V, L_g V \in C^1(\mathbb{R}^n \setminus \{0\})$, then for $r = 2a/b$ we have $u_\epsilon \in C^1(\mathbb{R}^n)$ and $u_\epsilon$ is homogeneous of degree $a - b$ with respect to $\Delta$. 


We remark that there are different usages when \( r \) is an arbitrary real number or when \( r \) is an integer. If we seek a stabilizing feedback \( u = u_r(x) \) for (7.1.1) such that \( u_r \) is homogeneous with respect to a one-parameter group of dilations, then generally \( r \) is not an integer. So \( u_r \) does not necessarily have the same regularity as \( L_f V \) and \( L_g V \) away from the origin. On the other hand, if we seek a stabilizing feedback \( u = u_r(x) \) which not only is continuously differentiable at the origin but also has the same regularity as \( L_f V \) and \( L_g V \) away from the origin, then \( r \) must be an even integer.

**Proof.** Let \( S \) denote the unit sphere \( \{ x : |x| = 1 \} \) in \( \mathbb{R}^n \) and define

\[
K \overset{\text{def}}{=} \{ x : L_g V(x) = 0 \} \cap S.
\]

Then \( K \) is compact and \( L_f V < 0 \) on \( K \). If \( K = S \), then from the homogeneity of \( L_g V \) it follows that \( L_g V(x) = 0 \), and therefore \( u_r(x) = 0 \) for all \( x \in \mathbb{R}^n \). In this case, all conclusions are totally trivial. So, without loss of generality, we assume \( K \neq S \). Then from the continuity of \( L_f V \) we conclude that there exists a neighborhood \( U \) of \( K \) in \( S \) \((U \neq S)\) such that \( L_f V < 0 \) on \( U \) and

\[
m_1 = \inf \{|L_f V(x)| : x \in U\} > 0.
\]

To prove (i), we define

\[
m = \min \{|L_g V(x)| : x \in S - U\},
\]

and

\[
M = \max \{|L_f V(x)| : x \in S - U\}.
\]

It is clear that \( m > 0 \). Now, given any \( \varepsilon > 0 \), we let \( \lambda > 0 \) satisfy \( \lambda^{b-a} \frac{M}{m} < \varepsilon \). Let \( S_\lambda \) denote the set \( \{ x : \Delta_\lambda x \in S \} \). If \( \Delta_\lambda x \in S - U \), we can choose \( u \) with \( |u| < \varepsilon \) so that

\[
\lambda^{b-a}(L_f V)(\Delta_\lambda x) + u L_g V(\Delta_\lambda x) < 0. \tag{7.1.5}
\]

If \( \Delta_\lambda x \in U \), then simply choosing \( u = 0 \), we see that (7.1.5) is satisfied. Since

\[
L_f V(x) + u L_g V(x) = \left( \lambda^{b-a}(L_f V)(\Delta_\lambda x) + u L_g V(\Delta_\lambda x) \right) \lambda^{-b},
\]
it turns out that for every \( x \in \lambda \) with \( \lambda > \left( \frac{M}{me} \right)^{\frac{1}{n-b}} \) there is a \( u \) with \( |u| < \varepsilon \) such that

\[
L_f V(x) + uL_g V(x) < 0.
\]

Since \( S_{\lambda} = \Delta_{\lambda - 1} S = \Delta_{\lambda - 1} S \), it follows that

\[
\bigcup_{\lambda > \left( \frac{M}{me} \right)^{\frac{1}{n-b}} \lambda} S_{\lambda} = \bigcup_{\lambda > \left( \frac{M}{me} \right)^{\frac{1}{n-b}} \lambda} \Delta_{\lambda} S,
\]

which contains a shell \( \{ x : 0 < |x| < \delta \} \) for some \( \delta > 0 \). For every \( x \) in the shell there is a \( u \) with \( |u| < \varepsilon \) such that \( L_f V(x) + uL_g V(x) < 0 \). So \( V \) satisfies the SCP.

To prove (ii), it is enough to prove \( \lim_{x \to 0} \frac{u_r(x)}{|x|} = 0 \) when \( \left( \frac{x}{2} - 1 \right)b > |\Delta| \), from which we conclude that \( du_r(0) = 0 \) and \( \partial u_r(0) = 0 \) for \( k = 1, 2, \ldots, n \).

Define

\[
M_1 = \max \left\{ \left| \frac{L_f V(x)}{L_g V(x)} \right| : x \in S - U \right\}
\]

and

\[
M_2 = \max \{ |L_g V(x)| : x \in S \}.
\]

Then \( M_i, i = 1, 2, \) are finite. Let \( x \in S \). If \( x \in S - U \), then from (7.1.2) we have

\[
|u_r(\Delta_{\lambda} x)| \leq 2\lambda^{a-b} \left| \frac{L_f V(x)}{L_g V(x)} \right| + \lambda^{\left( \frac{x}{2} - 1 \right)b} |L_g V(x)|^{\frac{\varepsilon}{2} - 1} \\
\leq 2\lambda^{a-b} M_1 + \lambda^{\left( \frac{x}{2} - 1 \right)b} M_2^{\frac{\varepsilon}{2} - 1}.
\]

If \( x \in U - K \), then \( L_f V(\Delta_{\lambda} x) < 0 \). Therefore,

\[
L_f V(\Delta_{\lambda} x) + \sqrt{L_f V(\Delta_{\lambda} x)^2 + |L_g V(\Delta_{\lambda} x)|^2} \leq |L_g V(\Delta_{\lambda} x)|^{\frac{\varepsilon}{2}} = \lambda^{\frac{\varepsilon}{2}} |L_g V(x)|^{\frac{\varepsilon}{2}}.
\]

We conclude that

\[
|u_r(\Delta_{\lambda} x)| \leq \lambda^{\left( \frac{x}{2} - 1 \right)b} |L_g V(x)|^{\frac{\varepsilon}{2} - 1} \leq \lambda^{\left( \frac{x}{2} - 1 \right)b} M_2^{\frac{\varepsilon}{2} - 1}.
\]

If \( x \in K \), then \( g(\Delta_{\lambda} x) = 0 \), and therefore \( u_r(\Delta_{\lambda} x) = 0 \). So, for all \( x \in S \), we have established the following inequality:

\[
| u_r(\Delta_{\lambda} x) | \leq 2\lambda^{a-b} M_1 + \lambda^{\left( \frac{x}{2} - 1 \right)b} M_2^{\frac{\varepsilon}{2} - 1}.
\]

When \( x \in S \) and \( \lambda < 1 \), we have

\[
| \Delta_{\lambda} x | = \sqrt{\left( \lambda^{x_1} x_1 \right)^2 + \left( \lambda^{x_2} x_2 \right)^2 + \ldots + \left( \lambda^{x_n} x_n \right)^2} \geq \lambda |\Delta| \sqrt{x_1^2 + x_2^2 + \ldots + x_n^2} = \lambda |\Delta|.
\]
Therefore, from the above two inequalities we conclude that, for \( z \in S \) and \( \lambda < 1 \),

\[
\frac{|u_r(\Delta \lambda z)|}{|\Delta \lambda z|} \leq 2\lambda^{a-b-|\Delta|}M_1 + \lambda^{(\frac{\xi}{\lambda})b-|\Delta|}M_2. \tag{7.1.6}
\]

Next, for every \( z \neq 0 \) with \( |z| < 1 \) we can always find a \( \lambda > 0 \) such that \( \bar{\lambda} = \Delta \lambda^{-1}z \in S \). Indeed, for every \( z \) with \( 0 < |z| < 1 \) the expression

\[
\rho(\lambda) = \frac{x_1^2}{\lambda^{2i_1}} + \frac{x_2^2}{\lambda^{2i_2}} + \cdots + \frac{x_n^2}{\lambda^{2i_n}}
\]
defines a continuous function: \((0,1] \to \mathbb{R}\). Note that \( \rho(1) = |z|^2 < 1 \) and \( \rho(0+) = +\infty \). So there exists a \( \lambda \in (0,1) \) (actually, a unique \( \lambda \)) such that \( \rho(\lambda) = 1 \). Let \( \bar{\lambda} = \Delta \lambda^{-1}z \).

Then \( |\bar{\lambda}| = 1 \).

To summarize, for \( z \in \mathbb{R}^n \) with \( 0 < |z| < 1 \) there exist \( \bar{\lambda} \in S \) and \( \lambda \in (0,1) \) such that \( z = \Delta \lambda^{-1} \bar{\lambda} \). From (7.1.6) we obtain

\[
\frac{|u_r(z)|}{|z|} = \frac{|u_r(\Delta \lambda \bar{\lambda})|}{|\Delta \lambda \bar{\lambda}|} \leq 2\lambda^{a-b-|\Delta|}M_1 + \lambda^{(\frac{\xi}{\lambda})b-|\Delta|}M_2^{-1}.
\]

When \( (\frac{\xi}{\lambda})b > |\Delta| \), the limit of the right hand side is zero as \( \lambda \to 0 \). So we only need to establish the fact that \( \lambda \to 0 \) implies \( \lambda \to 0 \). This is obvious in view of the fact that \( \lambda^{2i_k} \geq \lambda^{2|\Delta|} \) when \( \lambda < 1 \), and therefore

\[
|\bar{\lambda}| = \frac{x_1^2}{\lambda^{2i_1}} + \frac{x_2^2}{\lambda^{2i_2}} + \cdots + \frac{x_n^2}{\lambda^{2i_n}} \leq \frac{x_1^2 + x_2^2 + \cdots + x_n^2}{\lambda^2|\Delta|} = \frac{|z|^2}{\lambda^2|\Delta|},
\]
i.e., \( \lambda^{|\Delta|} \leq |z| \). The conclusion then follows.

To establish (iii), it suffices to show that all partial derivatives \( \frac{\partial u_r}{\partial x_k} \) are continuous at the origin, because we already have \( u_r \in C^1(\mathbb{R}^n - \{0\}) \) from Theorem 7.1. Since \( \frac{\partial u_r}{\partial x_k}(0) = 0 \), what we need to establish is \( \lim_{x \to 0} \frac{\partial u_r}{\partial x_k}(x) = 0 \), for \( k = 1, 2, \ldots, n \).

We rewrite \( u_r(x) = -\frac{L_f V(x)}{L_g V(x)} \cdot \frac{\sqrt{L_f V(x)^2 + |L_g V(x)|^2}}{L_g V(x)} \) when \( L_g V(x) \neq 0 \) and differentiate the two terms separately. Then, by direct calculation, we get

\[
\frac{\partial u_r}{\partial x_k}(x) = \frac{\partial L_f V(x)}{L_g V(x)} \left( 1 + \frac{L_f V(x)}{\sqrt{L_f V(x)^2 + |L_g V(x)|^2}} \right) - \frac{r|L_g V(x)|^{\frac{r-2}{2}}}{2\sqrt{L_f V(x)^2 + |L_g V(x)|^2}} \frac{\partial L_g V(x)}{\partial x_k}(x)
\]

\[
+ \frac{\partial L_f V(x)}{L_g V(x)^2} \left( L_f V(x) + \sqrt{L_f V(x)^2 + |L_g V(x)|^2} \right) \cdot \frac{\partial L_g V(x)}{\partial x_k}(x). \tag{7.1.7}
\]

We claim that there is a constant \( C = C(r) > 0 \) such that for every \( x \in S \) and \( \lambda \in (0,1) \)

\[
\left| \frac{\partial u_r}{\partial x_k}(\Delta \lambda x) \right| \leq C \left( \lambda^{(\frac{\xi}{\lambda})b-|\Delta|} + \lambda^a \right) \leq C \left( \lambda^{(\frac{\xi}{\lambda})b-|\Delta|} + \lambda^a \right). \tag{7.1.8}
\]
Recall from the beginning of the proof that \( m_1 = \inf \{|L_f V(x)| : x \in U\} > 0 \). Define

\[
M_3 = \max \{|L_f V(x)|, |L_g V(x)| : x \in S\}
\]

and

\[
M_4 = \max \left\{ \left| \frac{\partial L_f V}{\partial x_k}(x) \right|, \left| \frac{\partial L_g V}{\partial x_k}(x) \right| : x \in S, k = 1, 2, ..., n \right\}.
\]

Then \( M_3, M_4 < \infty \).

Let \( x \in S \). There are three possibilities: \( x \in K \), \( x \in U - K \), or \( x \in S - U \). If \( x \in U - K \), then \( L_f V(x) < 0 \). Therefore

\[
L_f V(\Delta \lambda x) + \sqrt{L_f V(\Delta \lambda x)^2 + |L_g V(\Delta \lambda x)|^2} \leq \lambda^{\frac{r}{2}} |L_g V(x)|^{\frac{r}{2}}.
\]

So from (7.1.7) we have

\[
\left| \frac{\partial u}{\partial x_k}(\Delta \lambda x) \right| \leq \left| \frac{\partial L_f V}{\partial x_k}(\Delta \lambda x) \right| \frac{\lambda^{\frac{r}{2}} |L_g V(x)|^{\frac{r}{2}}}{\sqrt{L_f V(\Delta \lambda x)^2 + |L_g V(\Delta \lambda x)|^2}}
\]

\[
+ \left( \frac{r}{2} + 1 \right) \lambda^{\frac{r}{2} - 1} |L_g V(\Delta \lambda x)|^{\frac{r}{2} - 2}
\]

\[
+ \left( \frac{r}{2} + 1 \right) \lambda^{\frac{r}{2} - 1} \left| \frac{\partial L_g V}{\partial x_k}(x) \right| |L_g V(x)|^{\frac{r}{2} - 2}
\]

\[
\leq \left( \frac{M_3^{\frac{r}{2} - 1} M_4}{m_1} + (\frac{r}{2} + 1) M_3^{\frac{r}{2} - 2} M_4 \right) \lambda^{\frac{r}{2} - 1} |L_g V(x)|^{\frac{r}{2} - 2}
\]

Therefore, (7.1.8) is true if we choose

\[
C > M_3^{\frac{r}{2} - 1} M_4/m_1 + (\frac{r}{2} + 1) M_3^{\frac{r}{2} - 2} M_4.
\]

If \( x \in S - U \), then from (7.1.7) we have

\[
\left| \frac{\partial u}{\partial x_k}(\Delta \lambda x) \right| \leq \frac{2\lambda^{a-i_k} |\frac{\partial L_f V}{\partial x_k}(x)| + \lambda^{\frac{r}{2} - 1} |L_g V(x)|^{\frac{r}{2} - 2}}{\lambda^a |L_g V(x)|^{\frac{r}{2} - 2}}
\]

\[
+ \frac{2\lambda^{a+b-i_k} |L_f V(x)| |\frac{\partial L_f V}{\partial x_k}(x)|}{\lambda^{a+b-i_k} |L_g V(x)|^{\frac{r}{2} - 2}} + \lambda^{\frac{r}{2} - 1} |L_g V(x)|^{\frac{r}{2} - 2}
\]

\[
\leq \left( \frac{r}{2} + 1 \right) M_3^{\frac{r}{2} - 2} M_4 \lambda^{\frac{r}{2} - 1} |L_g V(x)|^{\frac{r}{2} - 2} + \left( \frac{2M_4}{m_1} + \frac{2M_3 M_4}{m_1^2} \right) \lambda^{a-b-i_k}.
\]

So (7.1.8) is also true if we choose

\[
C > \left( \frac{r}{2} + 1 \right) M_3^{\frac{r}{2} - 2} M_4,
\]

and

\[
C > \frac{2M_4}{m_1} + \frac{2M_3 M_4}{m_1^2}.
\]
Finally, let \( x_0 \in K \). There are two possibilities. First, \( L_0 V = 0 \) in a neighborhood of \( x_0 \) in \( S \). Second, there exist \( \xi_i \in S - K, i = 1, 2, \cdots \), such that \( \xi_i \to x_0 \). In the first case, \( u_r = 0 \) in a neighborhood of \( x_0 \) in \( \mathbb{R}^n \). Therefore \( \frac{\partial u_r}{\partial x_i}(x_0) = 0 \) for \( k = 1, 2, \cdots, n \), and (7.1.8) holds trivially. In the second case, since (7.1.8) holds for \( x = \xi_i \), letting \( i \to \infty \), we see that (7.1.8) also holds for \( x = x_0 \). Therefore (7.1.8) is established.

When \( (\frac{2}{\gamma} - 1)b - |\Delta| > 0 \), we have \( \lim_{\lambda \to 0^+} \frac{\partial u_r}{\partial x_k}(\Delta \lambda x) = 0 \) for every \( x \in S \). Using the same analysis as in (ii), we get \( \lim_{x \to 0} \frac{\partial u_r}{\partial x_k}(x) = 0 \) for \( k = 1, 2, \cdots n \), as desired.

Finally, we prove (iv). The proof of this part is very simple. In fact, it is clear that, for \( r = \frac{2}{\gamma} \), \( u_r \) is a homogeneous function of degree \( a - b \) with respect to \( \Delta \). So if \( x \in S \) and \( 0 < \lambda < 1 \), then

\[
\left| \frac{u_r(\Delta \lambda x)}{\Delta \lambda (x)} \right| \leq \lambda^{a-b-|\Delta|} |u_r(x)|
\]

and

\[
\left| \frac{\partial u_r}{\partial x_k}(\Delta \lambda x) \right| = \left| \lambda^{a-b-i_i} \frac{\partial u_r}{\partial x_k}(x) \right| \leq \lambda^{a-b-|\Delta|} \left| \frac{\partial u_r}{\partial x_k}(x) \right|.
\]

From the former inequality we conclude that \( \frac{\partial u_r}{\partial x_k}(0) = 0 \) for \( 1 \leq k \leq n \), and from the latter inequality we conclude that \( \lim_{x \to 0} \frac{\partial u_r}{\partial x_k}(x) = 0 \) for \( 1 \leq k \leq n \). (See the analysis at the end of the proof of (ii).) Therefore \( \frac{\partial u_r}{\partial x_k} \) is continuous at the origin. The proof of Theorem 7.2 is now complete.

We remark that in the proof of Theorem 7.2 we used the conditions on \( a, b \) only at the end of each part. Therefore the estimates (7.1.6) and (7.1.8) hold for any pair of positive numbers \( (a, b) \). We will use this fact in the proof of Corollary 7.3.7.

### 7.2 Desingularizing Functions

Theorem 7.1 suggests a way to find a stabilizing feedback for (7.1.1), namely, if we have a control-Lyapunov function for (7.1.1), then we can use the control law given by (7.1.2). Thus, we seek a control-Lyapunov function for (7.1.1). In this section we present a construction of a CLF for certain kinds of composite systems using a result in [20].

Consider the systems

\[ \Sigma_1: \dot{x} = f(x, u) \]
and

$$\Sigma_2 : \dot{x} = f(x, y), \dot{y} = u,$$

where $f$ is a vector field of class $C^1$, $x \in \mathbb{R}^n$, and $u$ is a scalar input. It is well-known that, for a smooth vector field $f$, if $\Sigma_1$ is smoothly stabilizable, so is $\Sigma_2$. (Here and below, “smooth” means $C^\infty$.) In addition, if we find a smooth Lyapunov function of the closed-loop system of $\Sigma_1$ with a smooth stabilizing feedback, we can also construct a smooth feedback that stabilizes $\Sigma_2$. To be precise, if $u = k(x)$ stabilizes $\Sigma_1$ and $V : \mathbb{R}^n \to \mathbb{R}$ is a Lyapunov function of the resulting closed-loop system, then with the control law given by

$$u(x, y) = k_*(x)f(x, y) - y + k(x) - L_0 V(x), \quad (7.2.1)$$

the closed-loop system of $\Sigma_2$ is globally asymptotically stable, where $k_*$ is the differential of $k$ and $g(x, y) = \int_0^1 \frac{\partial f}{\partial y}(x, k(x) + \lambda(y - k(x)))d\lambda$. (See Lemma 4.8.3 in [27], for instance.) If $k, V \in C^1(\mathbb{R}^n)$ but $k$ is not twice differentiable, then the formula (7.2.1) will provide a continuous stabilizing feedback for $\Sigma_2$, which may not necessarily be differentiable. Of course, if $k$ is not differentiable, then (7.2.1) will no longer provide a control law. However, some examples demonstrate the fact that the global stabilization of $\Sigma_2$ might be achievable by a continuously differentiable feedback even if that of $\Sigma_1$ is not. So there may be no need to know whether there exists a differentiable feedback that stabilizes $\Sigma_1$ when we just attempt to stabilize $\Sigma_2$. See also the result in [8].

The work in [20] overcomes the defect deriving from (7.2.1). To apply it, we first recall that a function is a strict Lyapunov function of a closed-loop system $\Sigma_c$ if it is a Lyapunov function of $\Sigma_c$ and its derivatives along any trajectory of $\Sigma_c$ are negative except at the origin. The following lemma is a simple version of a lemma in [20].

**Lemma 7.2.1** Let $\Sigma_1$ be an $n$-dimensional system $\dot{x} = f(x, u)$, where $f : \mathbb{R}^{n+1} \to \mathbb{R}^n$ is a locally Lipschitz function and $f(0) = 0$, and let $\Sigma_2$ be the $(n + 1)$-dimensional system $\dot{x} = f(x, y), \dot{y} = u$. Suppose that there exist a continuous feedback $u = u_0(x)$ and a strict Lyapunov function $V_0 \in C^1(\mathbb{R}^n)$ of the resulting closed-loop system of $\Sigma_1$. Assume there is a scalar function $\varphi \in C^1(\mathbb{R}^{n+1})$ such that
(i) \( \varphi(x, y) = 0 \) is equivalent to \( y = u_0(x) \),

(ii) if we define \( \Phi(x, y) = \int_0^y \varphi(x, s) ds \), then \( \Phi(x, y) \to +\infty \) as \( |y| \to +\infty \),

(iii) \( \Phi(\cdot, u_0(\cdot)) \in C^1(\mathbb{R}^n) \),

(iv) \( \frac{\partial \Phi(x, y)|_{y=u_0(x)}}{\partial x_k} = \frac{\partial \Phi(x, u_0(x))}{\partial x_k} \) for \( k = 1, 2, \ldots, n \).

Then, for every \( \alpha \geq 1 \), the function \( V : \mathbb{R}^{n+1} \to \mathbb{R} \) defined by

\[
V(x, y) = \Phi(x, y) - \Phi(x, u_0(x)) + V_0(x)^\alpha
\]

is a CLF of \( \Sigma_2 \).

A function \( \varphi : \mathbb{R}^{n+1} \to \mathbb{R} \) satisfying the conditions in the lemma is called a desingularizing function of \( \Sigma_2 \). More general functions \( \varphi(x, y) \) were used in [20] and [34]. But the case covered in this lemma is enough for most applications and is easy to apply. Also, the proof of the lemma is significantly simpler than those in [20], [34].

**Proof.** From (iii) it is clear that \( V \in C^1(\mathbb{R}^{n+1}) \). For every \( x \), since \( \frac{\partial V}{\partial y}(x, y) = \varphi(x, y) \), and \( \varphi(x, y) = 0 \) only when \( y = u_0(x) \), it follows easily from (ii) that \( V(x, \cdot) \) attains its minimum value at \( y = u_0(x) \). But \( V(x, u_0(x)) = V_0(x)^\alpha \geq 0 \), and the equality holds only when \( x = 0 \), so \( V \) is positive definite. In addition, from (ii) and the properness of \( V_0 \), we know that \( V \) is proper. So \( V \) is continuously differentiable, proper, and positive definite.

Now consider the derivatives of \( V \) along the vector field of \( \Sigma_2 \). Suppose \( \frac{\partial V}{\partial y}(x, y) = 0 \), i.e., \( y = u_0(x) \). From (iv) we have

\[
(\frac{\partial V}{\partial x_1}(x, y), \ldots, \frac{\partial V}{\partial x_n}(x, y))|_{y=u_0(x)} \cdot f(x, u_0(x)) = \alpha V_0^{\alpha-1}(x) \nabla V_0(x) \cdot f(x, u_0(x)),
\]

which is \( < 0 \) for every \( x \neq 0 \), because \( V_0 \) is a strict Lyapunov function of the system \( \dot{x} = f(x, u_0(x)) \). So \( V \) is a CLF of \( \Sigma_2 \). \( \square \)

From Lemma 7.2.1 and Theorem 7.1 we see that in order to stabilize \( \Sigma_2 \), we only need to find a continuous stabilizing feedback for \( \Sigma_1 \), a strict Lyapunov function in \( C^1(\mathbb{R}^n) \) of the resulting closed-loop system of \( \Sigma_1 \), and a desingularizing function of \( \Sigma_2 \). Sometimes
this makes the stabilization problem much easier than seeking a stabilizing feedback in \( C^1(\mathbb{R}^n) \) for \( \Sigma_1 \). In addition, if we assume that \( \varphi \) has differentiability of a high order, then \( V \) will have differentiability of a high order. In the applications given below, we always choose \( \varphi \in C^2(\mathbb{R}^n) \) so that \( \Phi(\cdot, u_0(\cdot)) \in C^2(\mathbb{R}^n) \), and therefore the derivatives of \( V \) along the vector fields of those systems will be continuously differentiable.

### 7.3 Applications to Planar Systems

In this section we apply the theorems in Section 7.1 to the stabilization problem of planar systems in the form

\[
\begin{align*}
\dot{x} &= f(x, y), \\
y &= u,
\end{align*}
\]  

(7.3.1)

where \( f \in C^1(\mathbb{R}^2) \) and \( f(0, 0) = 0 \).

The study on the stabilizability of (7.3.1) for general real analytic functions \( f \) was considered in [4], [10], [12] and [16]. In [10] it is proved that if \( f \) is an analytic function, then there is a continuous feedback that locally stabilizes (7.3.1) if and only if for every \( \varepsilon > 0 \) there are \((x_1, y_1)\) and \((x_2, y_2)\) with norms less than \( \varepsilon \) such that \( x_1 > 0, x_2 < 0 \) and \( f(x_1, y_1) < 0, f(x_2, y_2) > 0 \). When \( f \) is analytic, sufficient conditions for (7.3.1) to be stabilized by a continuously differentiable feedback were also presented in [12]. In those papers the feedback control laws rely heavily on the algebraic properties of \( f \), and therefore no construction is given to express the feedback laws, even when \( f \) is homogeneous. In this work the interest is in obtaining a constructive control law that stabilizes (7.3.1). We will show a very useful method to stabilize a wide class of these systems. The next theorem is the main result on the stabilizability of (7.3.1).

**Theorem 7.3** Let \( \Sigma \) be a system in the form (7.3.1) with \( f \in C^1(\mathbb{R}^2) \). Assume that there are some indices \( i > 0, j > 0, \) and \( k > \max(i, 2i - j) \) such that the equality

\[
f(\lambda^ix, \lambda^jy) = \lambda^k f(x, y)
\]

(7.3.2)

holds for all \( \lambda > 0 \). Then there is a continuously differentiable feedback such that the origin is a globally asymptotically stable equilibrium of the resulting closed-loop
system if and only if there exist \((x_1, y_1)\) and \((x_2, y_2)\) with \(x_1 > 0, x_2 < 0\) such that \(f(x_1, y_1) < 0, f(x_2, y_2) > 0\).

**Remark 7.3.1** If we only seek a continuous stabilizing feedback, then the assumption \(k > \max(i, 2i-j)\) can be replaced by \(k > \max(0, i-j)\) in Theorem 7.3. This will be seen from the proof of Theorem 7.3. \(\square\)

Note that in the homogeneous case discussed in [12], we have \(i = j = 1\). Therefore, the condition \(k > \max\{i, 2i-j\}\) in the theorem is equivalent to \(k > 1\). When \(k = 1\), the equality (7.3.2) and the continuous differentiability of \(f\) imply that \(f(x, y)\) is linear in \(x\) and \(y\). (It suffices to show both \(f_x\) and \(f_y\) are constants. In fact, for every point \((x, y)\) we have \(f_x(x, y) = f_x(\lambda x, \lambda y)\). By letting \(\lambda \to 0\) in the equality, we see that \(f_x(x, y) = f_x(0, 0)\). So \(f_x\) is a constant. Similarly, \(f_y\) is a constant.) This is summarized in the corollary given below.

**Corollary 7.3.2** Let \(f \in C^1(\mathbb{R}^2)\). If \(f(\lambda x, \lambda y) = \lambda^k f(x, y)\) and \(k \geq 1\), then there is a continuously differentiable feedback that globally stabilizes (7.3.1) if and only if there exist \((x_1, y_1)\) and \((x_2, y_2)\) with \(x_1 > 0, x_2 < 0\) such that \(f(x_1, y_1) < 0, f(x_2, y_2) > 0\). \(\square\)

When \(f(x, y)\) is a polynomial of \(x, y\), if (7.3.2) holds for some indices \(i, j, k\), it will also hold for some choice of \(i, j, k\) which are all integers and where the greatest common divisor \((i, j)\) of \(i\) and \(j\) is equal to 1. Therefore, \(f(x, y)\) is a linear combination of terms \(x^m y^n\) where \(mi + nj = k\). In this case, if \(x^{m_0} y^{n_0}\) is the one of these terms with \(m_0\) the largest possibility, then all other terms must be \(x^{m_0-j} y^{n_0+i}\), i.e.,

\[
f(x, y) = \sum_{l \geq 0, m_0 - j l \geq 0} c_l x^{m_0-j l} y^{n_0+i l},
\]

where all \(c_l\)'s are constants. When \(m_0 \geq 2\), since \(k = m_0 i + n_0 j\) and \(i, j > 0\), it follows immediately that \(k > \max(i, 2i-j)\). Therefore we have the following corollary.

**Corollary 7.3.3** Suppose that \(f(x, y)\) is a polynomial of \(x, y\) with the property (7.3.2) and that \(f(x, 1)\) is a polynomial of a degree greater than or equal to two. Then there is a continuously differentiable feedback such that the origin is a globally asymptotically
stable equilibrium of the corresponding closed-loop system of (7.3.1) if and only if there exist \((x_1, y_1)\) and \((x_2, y_2)\) with \(x_1 > 0, x_2 < 0\) such that \(f(x_1, y_1) < 0, f(x_2, y_2) > 0\). □

In [12] the authors also investigated the global stabilizability of systems in the form

\[
\begin{align*}
\dot{x} &= ax^n + by^m, \\
\dot{y} &= u.
\end{align*}
\]

Clearly, the above corollary implies this case.

**Proof of Theorem 7.3:** The necessary condition is obvious. So we only need to prove the sufficiency. Assume there exist points \((x_1, y_1)\) with \(x_1 > 0\) and \((x_2, y_2)\) with \(x_2 < 0\) such that \(f(x_1, y_1) < 0\) and \(f(x_2, y_2) > 0\). From (7.3-2) we see that there exist \(y'_1\) and \(y'_2\) such that \(f(1, y'_1) < 0\) and \(f(-1, y'_2) > 0\). Without loss of generality, we suppose that \(f(1, y_1) < 0\) and \(f(-1, y_2) > 0\). Let \(u_0(x) = y_1 x^\frac{i}{i} \) if \(x \geq 0\) and \(u_0(x) = y_2 |x|^\frac{i}{i} \) if \(x < 0\). Then for \(x \geq 0\) we have \(f(x, u_0(x)) = f(x, y_1 x^\frac{i}{i}) = x^\frac{i}{i} f(1, y_1) \leq 0\), and the equality holds only when \(x = 0\). Similarly, for \(x < 0\) we have \(f(x, u_0(x)) = |x|^\frac{i}{i} f(-1, y_2) > 0\). Thus \(u = u_0(x)\) is a continuous feedback that stabilizes the one-dimensional system \(\dot{x} = f(x, u)\). In particular, we can choose \(V_0(x) = x^2\) as a strict Lyapunov function of the resulting closed-loop system.

Choose \(d\) to be an odd number such that \(\frac{d}{i} > 2\). Let

\[
\varphi(x, y) = \begin{cases} 
  y^d - \left(y_1 x^\frac{i}{i}\right)^d, & \text{if } x \geq 0, \\
  y^d - \left(y_2 |x|^\frac{i}{i}\right)^d, & \text{if } x < 0.
\end{cases}
\]

Then \(\varphi \in C^2(\mathbb{R}^2)\) and \(\varphi(x, y) = 0\) if and only if \(y = u_0(x)\). Define

\[
\Phi(x, y) = \int_0^y \varphi(x, s)ds.
\]

Then, precisely, we have

\[
\Phi(x, y) = \begin{cases} 
  \frac{1}{d+1} y^{d+1} - \left(y_1 x^\frac{i}{i}\right)^d y, & \text{if } x \geq 0, \\
  \frac{1}{d+1} y^{d+1} - \left(y_2 |x|^\frac{i}{i}\right)^d y, & \text{if } x < 0.
\end{cases}
\]

It is clear that \(\Phi(x, y) \to +\infty\) as \(|y| \to +\infty\). Also, we have an exact expression of \(\Phi(x, u_0(x))\), i.e.,

\[
\Phi(x, u_0(x)) = \begin{cases} 
  -\frac{d}{d+1} \left(y_1 x^\frac{i}{i}\right)^{d+1}, & \text{if } x \geq 0, \\
  -\frac{d}{d+1} \left(y_2 |x|^\frac{i}{i}\right)^{d+1}, & \text{if } x < 0.
\end{cases}
\]
Therefore, \( \Phi(\cdot, u_0(\cdot)) \in C^2(\mathbb{R}) \). Direct calculation shows that
\[
\frac{\partial \Phi}{\partial x}(x, y)|_{y = u_0(x)} = \frac{\partial \Phi(x, u_0(x))}{\partial x}.
\]
So \( \varphi \) satisfies all the conditions of Lemma 7.2.1. Let
\[
V(x, y) = \Phi(x, y) - \Phi(x, u_0(x)) + |z|^d_{i = 1}.
\]
Then \( V \in C^2(\mathbb{R}^2) \) is a CLF of the system (7.3.1) and \( V(\lambda^i x, \lambda^j y) = \lambda^{j(d+1)} V(x, y) \)
for \( \lambda > 0 \). Let \( \Delta \) denote the one-parameter group of dilations on \( \mathbb{R}^2 \) defined by
\( \Delta_{\lambda}(x, y) = (\lambda^i x, \lambda^j y) \). Then \( \frac{\partial V}{\partial x}(x, y) f(x, y) \) and \( \frac{\partial V}{\partial y}(x, y) \) are respectively homogeneous
of degrees \( j(d+1) + k - i \) and \( j(d+1) - j \) with respect to \( \Delta \). Therefore, from Part (iii)
of Theorem 7.2, we conclude that if \( k + j - i > \max(i, j) \), then for \( r \geq 4 \) the feedback
given by (7.1.2) is continuously differentiable, and the resulting closed-loop system has
the origin as a globally asymptotically stable equilibrium. Precisely, we have
\[
u_r(x) = \begin{cases} 
0, & \text{if } \frac{\partial V}{\partial y}(x, y) = 0, \\
-\frac{\frac{\partial V}{\partial x}(x, y) + \sqrt{\left(\frac{\partial V}{\partial x}(x, y)\right)^2 + \left|\frac{\partial V}{\partial y}(x, y)\right|^2}}{\frac{\partial V}{\partial y}(x, y)}, & \text{if } \frac{\partial V}{\partial y}(x, y) \neq 0.
\end{cases}
\tag{7.3.3}
\]
(By choosing \( d \) sufficiently large, we can get the same conclusion for every \( r > 2 \).) To
satisfy the condition \( k + j - i > \max(i, j) \), we need \( k > i \) if \( j \geq i \), and \( k > 2i - j \) if \( i \geq j \).
This is an assumption of the theorem. (To verify Remark 7.3.1, we apply Part (i) of
Theorem 7.2. Therefore only a weaker condition, \( k + j - i > 0 \), is required.) So the
proof is complete.

Now we present some simple examples which show how to stabilize the system (7.3.1)
when there is no continuously differentiable feedback that stabilizes the system \( \dot{x} = f(x, u) \).

**Example 7.3.4** Consider the system
\[
\begin{align*}
\dot{x} &= x(x - 2y^5)(x - y^5), \\
\dot{y} &= u.
\end{align*}
\tag{7.3.4}
\]
It is clear that the function \( f(x, y) = x(x - 2y^5)(x - y^5) \) is homogeneous of degree 15
with respect to the dilations given by \( \Delta_{\lambda}(x, y) = (\lambda^5 x, \lambda y) \). From Corollary 7.3.3 we
know that the system (7.3.4) is stabilizable with continuously differentiable feedback. We now find a stabilizing feedback.

Let $\Sigma_0$ denote the system $\dot{x} = f(x, u)$. Note that $\Sigma_0$ cannot be stabilized by a continuously differentiable feedback. Indeed, to stabilize $\Sigma_0$, we should require $u(x)$ to have the property that $f(x, u(x)) < 0$ for $x > 0$. Precisely, for $x > 0$ we should have $u^5(x) < x < 2u^5(x)$. Therefore, $u(x) > (\frac{2}{3})^\frac{1}{5}$ for $x > 0$. It then follows that $\frac{\partial u}{\partial x}(0)$ does not exist. So, for this example, the formula given by (7.2.1) does not provide a feedback.

To use the method described above, we take $u_0(x) = (\frac{2}{3})^\frac{1}{5}$. Then $u_0$ is continuous and $f(x, u_0(x)) = -\frac{1}{3}x^3$. So the system $\dot{x} = f(x, u_0(x))$ is asymptotically stable. Define

$$\varphi(x, y) = y^{15} - \left(\frac{2}{3}x\right)^3$$

and

$$\Phi(x, y) = \int_0^y \varphi(x, s)ds = \frac{1}{16} y^{16} - \left(\frac{2}{3}x\right)^3 x^3 y.$$ 

Then we see that $\varphi$ satisfies all the conditions of Lemma 7.2.1. Let

$$V(x, y) = \Phi(x, y) - \Phi(x, u_0(x)) + \frac{5}{16} \left(\frac{2}{3}\right)^\frac{15}{2} x^{16} = \frac{1}{16} y^{16} - \left(\frac{2}{3}\right)^3 x^3 y + \frac{20}{16} \left(\frac{2}{3}\right)^\frac{16}{5} x^{\frac{16}{5}}.$$ 

Then $V \in C^2(\mathbb{R}^2)$ is a CLF of the system (7.3.4) and $V$ is homogeneous with respect to the one-parameter group of dilations $\Delta = \{\Delta_\lambda: \Delta_\lambda(x, y) = (\lambda^5x, \lambda y)\}$. By direct calculation, we conclude that the feedback $u = u(x)$, given by $u(x) = 0$ if $y = (\frac{2}{3}x)^{\frac{1}{5}}$ and

$$u(x) = \frac{\frac{32}{27}x^3(x-y^5)(x-2y^5)(\frac{x}{3}y^5 - \frac{2}{3}y) + \sqrt{\left(\frac{32}{27}x^3(x-y^5)(x-2y^5)(\frac{x}{3}y^5 - \frac{2}{3}y)\right)^2 + (y^4 - (\frac{2}{3}x)^{3})^4}}{y^{15}-(\frac{2}{3}x)^{3}}$$

otherwise, is continuously differentiable and the origin is a globally asymptotically stable equilibrium of the resulting closed-loop system.

**Example 7.3.5** Consider the system

$$\begin{align*}
\dot{x} &= (x - 2y^5)(x - y^5), \\
\dot{y} &= u.
\end{align*}$$

(7.3.5)

From Corollary 7.3.3, we know that there exists a continuously differentiable feedback that stabilizes (7.3.5).
To find a stabilizing feedback using the formula (7.3.3), we define \( u_0(x) \) piecewise. Let
\[
    u_0(x) = \begin{cases} 
        \left( \frac{2}{3} x \right)^{\frac{1}{2}}, & \text{if } x \geq 0, \\
        \left( \frac{1}{3} x \right)^{\frac{1}{2}}, & \text{if } x < 0.
    \end{cases}
\]
Then
\[
    f(x, u_0(x)) = \begin{cases} 
        -\frac{1}{3} x^2, & \text{if } x \geq 0, \\
        \frac{2}{3} x^2, & \text{if } x < 0,
    \end{cases}
\]
where \( f(x, y) = (x - 2y^5)(x - y^5) \). So the system \( \dot{x} = f(x, u) \) is stabilized by the feedback \( u = u_0(x) \), and \( V_0(x) = x^2 \) could be a choice of its strict Lyapunov function.
Define
\[
    \varphi(x, y) = \begin{cases} 
        y^{15} - \left( \frac{2}{3} \right)^3 x^3, & \text{if } x \geq 0, \\
        y^{15} - \left( \frac{1}{3} \right)^3 x^3, & \text{if } x < 0.
    \end{cases}
\]
Then \( \varphi \in C^2(\mathbb{R}^2) \) and \( \varphi(x, y) = 0 \) if and only if \( y = u_0(x) \). Let \( \Phi(x, y) = \int_0^y \varphi(x, s)ds \).
Then
\[
    \Phi(x, y) = \begin{cases} 
        \frac{1}{10} y^{16} - \left( \frac{2}{3} \right)^3 x^3 y, & \text{if } x \geq 0, \\
        \frac{1}{10} y^{16} - \left( \frac{1}{3} \right)^3 x^3 y, & \text{if } x < 0.
    \end{cases}
\]
Therefore,
\[
    \Phi(x, u_0(x)) = \begin{cases} 
        -\frac{15}{16} \left( \frac{2}{3} \right)^{\frac{16}{3}} x^{\frac{16}{3}}, & \text{if } x \geq 0, \\
        -\frac{15}{16} \left( \frac{1}{3} \right)^{\frac{16}{3}} x^{\frac{16}{3}}, & \text{if } x < 0
    \end{cases}
\]
is a function in \( C^2(\mathbb{R}^2) \). It is then easy to see that \( \varphi \) satisfies all the conditions of Lemma 7.2.1. Let \( V(x, y) = \Phi(x, y) - \Phi(x, u_0(x)) + x^{\frac{16}{3}} \). Then the feedback given by (7.3.3) is continuously differentiable and stabilizes the system (7.3.5). \( \Box \)

**Remark 7.3.6** Corollary 7.3.3 may fail if \( f(x, y) \) is not homogeneous. To see this, we consider the system
\[
    \begin{align*}
    \dot{x} &= (x - y^5)(x^2 + y^2), \\
    \dot{y} &= u.
    \end{align*}
\]
(7.3.6)
It is pointed out by Dayawansa that there is no continuously differentiable feedback that stabilizes (7.3.6). See also Example 9.9 in [2]. (Generally, we will see that the system
\[
    \begin{align*}
    \dot{x} &= (x - y^5)(x^2 + |y|^k), \\
    \dot{y} &= u
    \end{align*}
\]
(7.3.7)
is stabilizable with continuously differentiable feedback if $k > 4$, and is not stabilizable with continuously differentiable feedback if $k < 4$.)

Note that, if we put $f_1(x, y) = (x - y^2)(x^2 + y^2)$ instead of the function $f(x, y) = (x - y^2)(x^2 + y^2)$ in (7.3.6), then from Corollary 7.3.3 we see that the system will be stabilizable with continuously differentiable feedback, although $f_1$ and $f$ have the same sign at every point. Here $x^2 + y^2$ and $x^2 + y^2$ are both positive definite, but they play different roles because they approach zero with different speeds along the curve $f(x, y) = 0$ as $(x, y) \to (0, 0)$. Generally, we consider the following system

$$
\begin{cases}
\dot{x} &= (x + c \text{ sign } (y)|y|^k)g(x, y), \\
\dot{y} &= u,
\end{cases}
$$

(7.3.8)

where $k > 1$, $c \neq 0$, and $g(x, y) > 0$ except at the origin. (When $g(x, y) = 1$ and $k = 3$, this is Kawski's example. See [16].) We have the following corollary.

**Corollary 7.3.7** Consider the system (7.3.8), where $k > 1$, $c \neq 0$, $g \in C^1(\mathbb{R}^2)$, and $g(x, y) > 0$ for $(x, y) \neq (0, 0)$. Suppose that

$$
g(\lambda^k x, \lambda y) = \lambda^{k_1} h(\lambda, x, y), \quad \text{for } \lambda > 0.
$$

(i) If $k_1 > k - 1$ and $h \in C^0(\mathbb{R}^3)$, then there is a continuously differentiable feedback such that the origin is a globally asymptotically stable equilibrium of the closed-loop system of (7.3.8).

(ii) Conversely, if $k_1 < k - 1$ and $\lim\inf_{\lambda \to 0} h(\lambda, x, y) \neq 0$ when $y \neq 0$, then (7.3.8) cannot be stabilized by a continuously differentiable feedback.

**Proof.** We begin with the proof of the first part. From Remark 7.3.1 we know that there is a continuous feedback $u = \alpha(x, y)$, $\alpha \in C^1(\mathbb{R}^1 - \{0\})$, which stabilizes the system

$$
\dot{x} = x + c \text{ sign } (y)|y|^k,
$$

$$
\dot{y} = u.
$$

Let $\beta(x, y) = \alpha(x, y)g(x, y)$. Then $u = \beta(x, y)$ stabilizes (7.3.8). Particularly, from the proof of Theorem 7.3, we can take $\alpha(x, y)$ to be $u_*(x, y)$ given in (7.3-3), where
\[ f(x, y) = x + c \text{sign}(y)|y|^k \] and \( \frac{\partial^2}{\partial x^2} f, \frac{\partial^2}{\partial y^2} f \) are homogeneous functions of degrees \( d + 1 \) and \( d \), respectively, for some odd number \( d \), with respect to the group of dilations \( \Delta = \{ \Delta(x, y) = (\lambda^k x, \lambda y) \} \). When \( r \) is sufficiently large, from (7.1.6) we obtain
\[
\frac{|\beta|}{|x, y|} \leq C \lambda^{1-k} \text{for } (x, y) \in S \text{ (the unit sphere in } \mathbb{R}^2) \text{ and } 0 < \lambda < 1, \text{ where } C \text{ is a constant. Therefore,}
\]
\[
\frac{|\beta(x^k, y)|}{|x^k, y|} \leq C \lambda^{1-k-\delta} h(\lambda, x, y)
\]
for \((x, y) \in S \) and \( 0 < \lambda < 1 \). Letting \( \lambda \to 0 \), since \( k_1 > k - 1 \), we see that
\[
\lim_{(x, y) \to (0, 0)} \frac{\beta(x, y)}{|x, y|} = 0.
\]
So \( \frac{\partial \beta}{\partial x}(0, 0) = \frac{\partial \beta}{\partial y}(0, 0) = 0 \). In addition, since \( g \in C^1(\mathbb{R}^2) \) and \( \alpha(0, 0) = 0 \), we have
\[
\lim_{(x, y) \to (0, 0)} \frac{\partial \beta}{\partial x}(x, y) = \lim_{(x, y) \to (0, 0)} \left( \frac{\partial \alpha}{\partial x}(x, y) g(x, y) + \alpha(x, y) \frac{\partial g}{\partial x}(x, y) \right)
\]
\[
= \lim_{(x, y) \to (0, 0)} \frac{\partial \alpha}{\partial x}(x, y) g(x, y).
\]
Recall that \( \alpha(x, y) = u_r(x, y) \) and \( r \) is sufficiently large. From (7.1.8) we have
\[
|\frac{\partial \alpha}{\partial x}(x^k, y)| \leq C \lambda^{1-k} \text{ for } (x, y) \in S \text{ and } 0 < \lambda < 1, \text{ where } C \text{ is a constant. Therefore,}
\]
\[
|\frac{\partial \alpha}{\partial x}(x^k, y) g(x^k, y)| \leq C \lambda^{1-k+\delta} h(\lambda, x, y) \text{ for } (x, y) \in S \text{ and } 0 < \lambda < 1. \text{ Since } k_1 > k - 1, \text{ it follows that } \lim_{(x, y) \to (0, 0)} \frac{\partial \alpha}{\partial x}(x, y) g(x, y) = 0, \text{ i.e., } \lim_{(x, y) \to (0, 0)} \frac{\partial \beta}{\partial x}(x, y) = 0.
\]
This proves that \( \frac{\partial \beta}{\partial x} \) is continuous. Similarly, \( \frac{\partial \beta}{\partial y} \) is continuous. So \( \beta \in C^1(\mathbb{R}^2) \).

Next, we prove the second part. Note that, if \( y > 0 \), then \( g(\delta y, y) = y^k h(y, \delta, 1). \) Since \( g(\delta y^k, y) > 0 \) for every \( y > 0 \), it follows that \( h(y, \delta, 1) > 0 \) for all \( y > 0 \). Therefore, from the assumption we have \( \lim_{y \to 0^+} h(y, \delta, 1) > 0 \). Consequently, for every \( \delta \in [-4^k, 4^k] \), there is an open interval \( (0, y_\delta) \) such that \( \inf \{ h(y, \delta, 1) : y \in (0, y_\delta) \} > 0 \). Since \( h(\lambda, x, y) \) is continuous with respect to \( x \), we conclude that there is a neighborhood \( I_\delta \) of \( \delta \) such that \( \inf \{ h(y, x, 1) : y \in (0, y_\delta), x \in I_\delta \} > 0 \). The set of open intervals \( \{ I_\delta \}_{\delta \in [-4^k, 4^k]} \) covers the closed interval \( [-4^k, 4^k] \), and therefore we have a finite number of those intervals covering \( [-4^k, 4^k] \). We then conclude that there is an open interval \( I_1 = (0, \lambda_1) \) such that
\[
m_1 = \inf \{ h(y, x, 1) : y \in I_1, \ x \in [-4^k, 4^k] \} > 0.
\]
So \( h(y, \delta, 1) \geq m_1 \) for \( y \in I_1, \delta \in [-4^k, 4^k] \). Similarly, we have \( |h(y, \delta, -1)| \geq m_2 \) for \( y \in I_2 = (-\lambda_2, 0), \delta \in [-4^k, 4^k] \). Letting \( m = \min \{ m_1, m_2 \} \). It then follows that
\[
g(x, y) \geq m |y|^k, \text{ if } |x| \leq 4^k |y|^k, \ y \in I = (-\lambda_2, \lambda_1).
\]
(7.3.10)
Suppose that $u = u_1(x, y)$ stabilizes (7.3.8) and $u_1 \in C^1(\mathbb{R}^2)$. Without loss of generality, we consider the case $c < 0$, particularly, $c = -1$. Let $f(x, y) = (x - \text{sign}(y)|y|^k)g(x, y)$. Since $f(x, y) > 0$ when $x > |y|^k$, we conclude that $u_1(x, y)$ is positive when $x = y^k$ and $y > 0$. (It is clear that there exists $(x_1, y_1)$ with $x_1 = y_1^k$, $y_1 > 0$, such that $u_1(x_1, y_1) > 0$. If there were a point $(x_2, y_2)$ with $x_2 = y_2^k$, $y_2 > 0$ such that $u_1(x_2, y_2) < 0$, there would be a point $(\hat{x}, \hat{y})$, where $\hat{x}$ is between $x_1$ and $x_2$ with $\hat{x} = \hat{y}^k > 0$, such that $u_1(\hat{x}, \hat{y}) = 0$. Therefore, $(\hat{x}, \hat{y})$ would be an equilibrium, and the closed-loop system cannot be globally stable. Similarly, $u_1(x, y)$ is negative when $x = -|y|^k$ and $y < 0$. We will see that this contradicts the assumption that $u_1 \in C^1(\mathbb{R}^2)$.

First, we claim that for sufficiently small $y_0 \in I$, with $y_0 > 0$, the trajectory of the closed-loop system starting at $(x_0, y_0)$ with $x_0 = y_0^k$ must cross the $y$-axis at $(0, y_1)$ with $y_1 > \frac{y_0}{2}$. Indeed, if the claim were false, there would be a point $(\hat{x}, \hat{y})$ with $0 < \hat{x} < \hat{y}^k$ and $\frac{y_0}{2} < \hat{y} < y_0$ such that $u_1(\hat{x}, \hat{y}) < 0$ and $\left|\frac{u_1(\hat{x}, \hat{y})}{f(\hat{x}, \hat{y})}\right| \geq \frac{m - \frac{1}{2}\mu_0}{y_0} = \frac{m}{2y_0 - 1}$. (The trajectory must go down to the segment between $(0, y_0/2)$ and $(x_0, y_0)$. Therefore, there exists a point at which the tangent slope is greater than the segment slope.) From (7.3.10) it follows that

$$|u_1(\hat{x}, \hat{y})| \geq \frac{1}{2y_0 - 1} (\hat{x} - \hat{y}^k)m \hat{y}^k \geq \frac{m}{2k_1 + 1} y_0^{-k_1 - 1} (\hat{x} - \hat{y}^k).$$

Therefore

$$|u_1(\hat{x}, \hat{y}) - u_1(\hat{y}^k, \hat{y})| \geq \frac{m}{2k_1 + 1} y_0^{-k_1 - 1} (\hat{x} - \hat{y}^k),$$

which contradicts the assumption that $u_1$ is Lipschitz continuous, since $k > k_1 + 1$.

Consider those initial points $(x_0, y_0)$ with $y_0 > 0$ such that the trajectories pass through the points $(0, y_1)$ with $y_1 > \frac{1}{2} y_0$. It is clear that the trajectories must then cross the curve $x = -|y|^k$ at $(x_2, y_2)$ with $y_2 < 0$. Without loss of generality, we assume $|y_2| \leq y_0$; otherwise, with a similar analysis we can consider the initial points $(x_2, y_2)$. Therefore there is a point $(\bar{x}, \bar{y})$ with $-x_0 \leq \bar{x} \leq 0$ and $\frac{1}{2} y_1 \leq \bar{y} \leq y_1$ such that $u_1(\bar{x}, \bar{y}) < 0$ and $\left|\frac{u_1(\bar{x}, \bar{y})}{f(\bar{x}, \bar{y})}\right| \geq \frac{y_1 - \frac{1}{2} y_0}{\bar{x} \bar{y}} = \frac{1}{4y_0^2 - 1}$ (because the trajectory must cross the segment between $(-x_0, y_1/2)$ and $(0, y_1)$). Note that $|\bar{x}| \leq y_0^k \leq (2y_1)^k \leq (4\bar{y})^k$. From
(7.3.10) it then follows that
\[ |u_1(\bar{x}, \bar{y})| \geq \frac{1}{4y_0^{k-1}}|\bar{x} - \bar{y}^k|g(\bar{x}, \bar{y}) \geq \frac{m}{4^{k_1+1}y_0^{k_1-k_1-1}}|\bar{x} - \bar{y}^k|. \]

So
\[ |u_1(\bar{x}, \bar{y}) - u_1(\bar{y}^k, \bar{y})| \geq |u_1(\bar{x}, \bar{y})| \geq \frac{m}{4^{k_1+1}y_0^{k_1-k_1-1}}|\bar{x} - \bar{y}^k|. \]

This contradicts the Lipschitz continuity of \( u_1 \).

From Corollary 7.3.7 it is not difficult to verify the conclusion of Remark 7.3.6. Particularly, we know that the system
\[
\begin{align*}
\dot{x} &= (x - y^5)(x^2 + y^6), \\
\dot{y} &= u
\end{align*}
\]

(7.3.11)
is stabilizable with continuously differentiable feedback. Such a feedback is in the form \( u = \alpha(x, y) \cdot (x^2 + y^6) \), where \( u = \alpha(x, y) \) stabilizes the system
\[
\begin{align*}
\dot{x} &= x - y^5, \\
\dot{y} &= u
\end{align*}
\]

(7.3.12)

However, if we use the procedure in [16] to obtain a stabilizing feedback \( u = \alpha(x, y) \) for (7.3.12), then \( u = \alpha(x, y)(x^2 + y^6) \), which stabilizes (7.3.11), is not continuously differentiable in \( \mathbb{R}^2 \). Indeed, (see Example 6.8 and Example 9.9 in [2]) with the procedure in [16] we have \( \alpha(x, y) = x - y + \frac{5}{6}x^\frac{1}{5} - y^5 \), and \( V(x, y) = x^\frac{5}{6} - xy + \frac{1}{5}y^6 \) is a Lyapunov function of the closed-loop system of (7.3.12) with \( u = \alpha(x, y) \). It is clear that \( u = \alpha(x, y)(x^2 + y^6) \) is not continuously differentiable in \( \mathbb{R}^2 \).

Corollary 7.3.7 applies to very general systems of the type (7.3.8). For instance, if \( g(x, y) = \log(1 + x^2 + y^6) \) and \( k = 5 \), then (7.3.8) can be stabilized by a continuously differentiable feedback because \( g(\lambda x, \lambda y) \) can be expressed as \( \lambda^{k-5}h(\lambda, x, y) \), where \( h(\lambda, x, y) = \lambda^{-5} \log(1 + \lambda^{10}x^2 + \lambda^6y^6) \) if \( \lambda \neq 0 \) and \( h(0, x, y) = 0 \). (It is clear that \( h \) is continuous but not differentiable at \( \lambda = 0 \).) Similarly, if \( g(x, y) = e^{-x_2+y^2} \) for \( (x, y) \neq (0, 0) \) and \( g(0, 0) = 0 \), then for every \( k > 1 \) there exists a continuously differentiable feedback that stabilizes (7.3.8). While for \( g(x, y) = \log^2(1 + x^2 + y^2) \) and \( k = 5 \), there is no continuously differentiable feedback that stabilizes (7.3.8).
We remark that for the second conclusion of Corollary 7.3.7, if \( g \) is analytic at the origin, then we can use the condition that \( h \) is not continuous to replace \( \lim \inf_{\lambda \to 0} h(\lambda, x, y) \neq 0 \). In this case we can see the parallelism of the two conclusions. But, generally, the discontinuity of \( h \) does not imply the limit condition. For instance, let \( g(x, y) = |y|^3 \sin^2 \frac{1}{y} + y^4 + x^2 \) if \( y \neq 0 \) and \( g(x, 0) = x^2 \). Write \( g(\lambda^k x, \lambda y) = \lambda^{k_1} h(\lambda, x, y) \). If \( k_1 < 3 \), then \( h \) is continuous. If \( k_1 \geq 3 \), then \( \lim \inf_{\lambda \to 0} h(\lambda, x, y) = 0 \).

7.4 Applications to Three-dimensional Systems

In this section we consider three-dimensional systems of the type

\[
\begin{aligned}
\dot{x} &= f(x, y), \\
\dot{y} &= g(x, y) + z^2, \\
\dot{z} &= u,
\end{aligned}
\]

(7.4.1)

where both \( f \) and \( g \) are homogeneous functions of degree \( p \) (i.e., \( f(\lambda x, \lambda y) = \lambda^p f(x, y) \)), and \( p, q \) are odd integers. We want to find a continuously differentiable feedback that stabilizes (7.4.1).

Note that for some \( f \) and \( g \), the system

\[
\begin{aligned}
\dot{x} &= f(x, y), \\
\dot{y} &= g(x, y) + u^q
\end{aligned}
\]

(7.4.2)

is not stabilizable with continuously differentiable feedback. For example, let \( f(x, y) = x + y \), \( g(x, y) = 0 \). Since the linearization of (7.4.2) does not vanish at the origin, it follows that, in order to stabilize (7.4.2), the linearization of the closed-loop system of (7.4.2) must have no unstable part. (See Theorem 6.7 in [2], for instance.) So if there is a continuously differentiable feedback \( u = u_1(x, y) \) that stabilizes (7.4.2), then \( u_1^q(x, y) = ax + by + o(|x| + |y|) \) with \( (a, b) \neq (0, 0) \), where \( o(\rho) \) represents a function whose ratio to \( \rho \) goes to zero as \( \rho \to 0 \). But if \( q > 1 \), then \( u_1 \) will not be continuously differentiable at \( (x, y) = (0, 0) \). This is a contradiction. So for general \( f \) and \( g \), we cannot use the formula (7.2.1) to obtain a feedback that stabilizes (7.4.1). The next
Theorem 7.4 Let $\Sigma$ be a system in the form (7.4.1), where $f, g \in C^1(\mathbb{R}^2)$ are two homogeneous functions of degree $p$, and $p, q$ are any two odd integers. Suppose that there exists $(x_1, y_1)$ with $x_1 > 0$ such that $f(x_1, y_1) < 0$.

(i) If $p > 1$, then there is a continuously differentiable feedback such that the origin is a globally asymptotically stable equilibrium of the resulting closed-loop system of (7.4.1).

(ii) If $p = 1$, then there is a continuous feedback such that the origin is a globally asymptotically stable equilibrium of the resulting closed-loop system of (7.4.1).

Proof. First, from the homogeneity of $f$, we know that $f(x_1, y_1) = x_1^pf(1, y_1/x_1)$. Let $b$ denote $y_1/x_1$. Then $f(1, b) < 0$. Therefore, the system $\dot{x} = f(x, bx) = x^p f(1, b)$ has $x = 0$ as a globally asymptotically stable equilibrium. Define (compare with Formula (7.2.1))

$$u_0^a(x, y) = \begin{cases} -g(x, y) + bf(x, y) - (y - bx)^p - x \frac{f(x, y) - f(x, bx)}{y - bx}, & \text{if } y \neq bx, \\ -g(x, bx) + bf(x, bx) - x f_0(x, bx), & \text{if } y = bx, \end{cases}$$

and $V_0(x, y) = x^2 + (y - bx)^2$. Then $(x, y) = (0, 0)$ is a globally asymptotically stable equilibrium of the closed-loop system of (7.4.2) with $u = u_0(x, y)$, and $V_0$ is a strict Lyapunov function. This is obvious by verifying that the derivative of $V_0$ along the vector field of the resulting closed-loop system is $\dot{V}_0 = x^{p+1} f(1, b) - 2(y - bx)^{p+1} < 0$ for $(x, y) \neq (0, 0)$. Note that $u_0^a(x, y)$ defined by (7.4.3) is still homogeneous of degree $p$. Therefore $u_0$ is a homogeneous function of degree $p/q$ and $u = u_0(x, y)$ stabilizes (7.4.2).

Now for an odd integer $d > 1$ we define

$$\varphi(x, y, z) = z^q - u_0^q(z, y).$$

Then $\varphi \in C^2(\mathbb{R}^2)$, and $\varphi(x, y, z) = 0$ if and only if $z = u_0(x, y)$. Let

$$\Phi(x, y, z) = \int_0^z \varphi(x, y, s) \, ds = \frac{1}{q+1} z^{q+1} - u_0^q(z, y) z.$$
Then
\[ \Phi(x, y, u_0(x, y)) = -\frac{qd}{qd + 1} u_0^{qd+1}(x, y). \]

It is clear that \( \varphi \) satisfies all the conditions of Lemma 7.2.1. So \( V : \mathbb{R}^3 \to \mathbb{R} \) defined by
\[ V(x, y, z) = \Phi(x, y, z) - \Phi(x, y, u_0(x, y)) + V_0(x, y)^{\frac{qd+1}{q} + p} \]
is a CLF of (7.4.1), and the feedback \( u = u_r(x, y, z) \) given by (7.1.2) with \( r > 2 \) stabilizes (7.4.1). Let \( \Delta \) denote the one-parameter group of dilations of the form \( \Delta_\lambda(x, y, z) = (\lambda^qx, \lambda^qy, \lambda^qz) \). Then we see that \( V \) is a homogeneous function of degree \( p(qd + 1) \) with respect to \( \Delta \). In addition, both \( f(x, y) \) and \( g(x, y) + z^q \) are homogeneous of degree \( pq \) with respect to \( \Delta \). It follows that \( \frac{\partial V}{\partial x}(x, y, z) \cdot f(x, y) + \frac{\partial V}{\partial y}(x, y, z) \cdot (g(x, y) + z^q) \) and \( \frac{\partial V}{\partial z}(x, y, z) \) are homogeneous of degrees \( p(qd + 1) - q + pq \) and \( p(qd + 1) - p \), respectively.

Comparing with the condition of Theorem 7.2 and using the same notations \( a \) and \( b \), we have \( a = p(qd + 1) - q + pq \) and \( b = p(qd + 1) - p \). Therefore, \( a - b = p(q + p - q) = (p - 1)q + p \).

So if \( p > 1 \), then \( a - b > \max(p, q) \), and the control law given by (7.1.2) for \( r > 2 \) is continuously differentiable. If \( p = 1 \), then \( a - b = p > 0 \), and the control law given by (7.1.2) for \( r > 2 \) is at least continuous at 0.

\[ \square \]

**Remark 7.4.1** For the system (7.4.1) we can also modify a stabilizing feedback control law so that it becomes a homogeneous function with respect to a one-parameter group of dilations and possesses the continuous differentiability as before. This can be obtained when we apply Part (iv) of Theorem 7.2 to the above proof.

\[ \square \]

**Remark 7.4.2** When \( f(x, y) = y^p \), \( g(x, y) = 0 \), the system (7.4.1) becomes
\[
\begin{align*}
\dot{x} &= y^p, \\
\dot{y} &= z^q, \\
\dot{z} &= u,
\end{align*}
\]
where \( p \) and \( q \) are odd integers. We show below that, in a duplicate method, the system can be stabilized by a polynomial feedback.

The stabilization problem for (7.4.4), with \( p = q = 3 \), was proposed in [9] and first solved in [5]. Later, Dayawansa and Martin in [11] proved that for every \( n > 0 \)
and for every odd \( p \) there exists a homogeneous feedback of degree \( p \) which stabilizes
the \( n \)-dimensional system: \( \dot{x}_1 = x_2^p, \cdots , \dot{x}_{n-1} = x_n^p, \dot{x}_n = u \). Theorem 7.4 provides
a continuously differentiable feedback that stabilizes the general system (7.4.4) with
different \( p \) and \( q \) such that \( p > 1 \). We can also stabilize (7.4.4) in a different way
and get an even better result. Particularly, (7.4.4) can be stabilized by a polynomial
feedback.

Indeed, we consider more general systems of the type

\[
\begin{align*}
\dot{x} &= \varphi(y), \\
\dot{y} &= \psi(z), \\
\dot{z} &= u,
\end{align*}
\]  

(7.4.5)

where \( \varphi \) and \( \psi \) are strictly increasing smooth functions with \( \varphi(0) = \psi(0) = 0 \). Let
\( k(x, y) = -(x + y) \). Then \( u = k(x, y) \) stabilizes the system

\[
\begin{align*}
\dot{x} &= \varphi(y), \\
\dot{y} &= \psi(u).
\end{align*}
\]  

(7.4.6)

To see this, we define \( \Phi(x) = \int_0^x \varphi(s)ds \), \( \Psi(x) = -\int_0^x \psi(-s)ds \) and \( V(x, y) = \Psi(x + y) + \Phi(y) \). Then the derivative of \( V \) along any trajectory of the closed-loop
system of (7.4.6), with \( u = k(x, y) \), is \( \dot{V} = -\psi^2(-(x + y)) \leq 0 \), and \( \dot{V} \equiv 0 \) if and
only if \( x + y \equiv 0 \). When \( x + y \equiv 0 \), we have \( \dot{y} \equiv 0 \) and \( \dot{z} \equiv 0 \). Thus \( \varphi(y) \equiv 0 \)
which implies that \( y \equiv 0 \) and \( x \equiv 0 \). By LaSalle’s invariance principle we know that
(7.4.6) is smoothly stabilized by the feedback \( u = k(x, y) \). Thus we obtain the smooth
stabilizability of (7.4.5).

Now, using the formula (7.2.1), we conclude that the following control law stabilizes
(7.4.5):

\[
u(x, y, z) = k_0[x, y]f(x, y, z) - z + k(x, y) - L_0V(x, y),
\]  

(7.4.7)

where \( k_0 = (-1, -1) \) is the differential of \( k \), \( f(x, y, z) = (\varphi(y), \psi(z))' \), and
\( g(x, y, z) = \frac{1}{0} \frac{2f(x, y, k(x, y) + \lambda(z - k(x, y)))}{d\lambda} \). By direct calculation, we obtain

\[
u(x, y, z) = -x - y - z - \varphi(y) - \psi(z) - \frac{(\varphi(y) - \psi(-z - y))(\psi(z) - \psi(-z - y))}{x + y + z},
\]  

(7.4.8)
where \( \frac{\psi(z) - \psi(-x - y)}{x + y + z} \) is considered as \( \psi'(-x - y) \) if \( x + y + z = 0 \).

When \( \varphi \) and \( \psi \) are both polynomials, the feedback given in (7.4.8) is a polynomial. Particularly, when \( \varphi(y) = y^p \) and \( \psi(z) = z^q \), we have

\[
u(x, y, z) = -x - y - z - y^p - z^q - (y^p + (x + y)^q) \sum_{i=0}^{q-1} (-1)^i z^i (x + y)^{q-1-i}.\]

\( \square \)
References


Vita

Yudi Yang

1978 Graduated from the No. 7 Middle School, Jinan, China.
1978-82 Attended Shandong University, Jinan, China.
1982 B.S. in Mathematics, Shandong University, Jinan, China.
1982-85 Graduate work in Mathematics, Institute of Mathematics, Academia Sinica, Beijing, China.
1985 M.S. in Mathematics, Institute of Mathematics, Academia Sinica, Beijing, China.
1985-87 Instructor, Department of Applied Mathematics, Tsinghua University, Beijing, China.
1987-93 Graduate work in Mathematics, Rutgers, The State University of New Jersey, New Brunswick, New Jersey, USA.
1993 Ph.D. in Mathematics.