การศึกษาการวิเคราะห์และออกแบบระบบเชิงเส้นและไม่เป็นเชิงเส้น ในการประยุกต์กับระบบการบิน

The Study of Linear or Nonlinear System Analysis and Design in Application of Aviation

โดย

นาย จตุพร นาคเปี่ยม

นาย พงศ์ระพี แก้วไพรฮะ

อาจารย์ที่ปรึกษา

ดร. ปรเมษฐ์ ประนยานันทน์

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

ภาควิชาวิศวกรรมระบบควบคุม สาขาแมคคาทรอนิกส์

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2547

ปริญญานิพนธ์ปีการศึกษา 2547

ภาควิชาวิศวกรรมระบบควบคุม สาขาแมคคาทรอนิกส์

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง การศึกษาการวิเคราะห์และออกแบบระบบเชิงเส้นและไม่เป็นเชิงเส้นในการประยุกต์กับระบบการบิน

The Study of Linear or Nonlinear System Analysis and Design in Application of Aviation

ผู้จัดทำ

นาย จตุพร	นาคเปี่ยม	รหัสนักศึกษา	44010055
นาย พงศ์ระพี	แก้วไทรฮะ	รหัสนักศึกษา	44010309

....อาจารย์ที่ปรึกษา

(ดร. ปรเมษฐ์ ประนยานันทน์)

การศึกษาการวิเคราะห์และออกแบบระบบเชิงเส้นและไม่เป็นเชิงเส้นในการประยุกต์กับระบบการบิน

นาย จตุพร นาคเปี่ยม
นาย พงศ์ระพี แก้วไทรฮะ

อาจารย์ที่ปรึกษา
ดร. ปรเมษฐ์ ประนยานันทน์

บทคัดย่อ

โดยทั่วไปแล้ว ระบบต่าง ๆ จะมีผลตอบสนองที่ไม่เป็นเชิงเส้น ซึ่งเกิดจากการถูกรบกวนโดยสภาวะแวดล้อมภายนอกและการออกแบบที่ไม่ได้มาตรฐาน ยกตัวอย่างเช่นวงจรไฟฟ้าซึ่งถูกออกแบบให้เป็นระบบเชิงเส้น แต่ในสภาพการทำงานจริงจะมีผลกระทบจากสัญญาณรบกวน, กัมมันตภาพรังสี และสนามแม่เหล็ก ทำให้ผลตอบสนองของวงจรมีความไม่เป็นเชิงเส้นเกิดขึ้น ความไม่เป็นเชิงเส้นนี้อาจทำให้ระบบไม่มีเสถียรภาพ หรือทำงานผิดพลาดได้ การแก้ปัญหาดังกล่าวจำเป็นต้องนำเอาระบบควบคุมไม่เป็นเชิงเส้นมาใช้

โครงงานนี้เริ่มต้นจากการศึกษาทฤษฎีระบบควบคุมเชิงเส้น ซึ่งเป็นความรู้พื้นฐานเกี่ยวกับสมการคุณลักษณะและการพิจารณาเสถียรภาพของระบบ พบว่าการศึกษาทฤษฎีระบบควบคุมจำเป็นต้องอาศัยความรู้พื้นฐานทางคณิตศาสตร์ในเรื่องพีชคณิตเชิงเส้นในการแก้สมการต่าง ๆ ดังนั้น จึงได้มีการรวบรวมเนื้อหาเรื่องพีชคณิตเชิงเส้นบรรจุไว้ในภาคผนวกของหนังสือเล่มนี้ จากความรู้พื้นฐานดังกล่าว จะได้นำไปประยุกต์ใช้ในส่วนของระบบควบคุมไม่เป็นเชิงเส้น ซึ่งมีความซับซ้อนมากกว่า

จุดประสงค์หลักของการจัดทำโครงงานนี้ คือการออกแบบตัวควบคุมไม่เป็นเชิงเส้น เพื่อนำไปใช้กับระบบหรือเครื่องจักรต่าง ๆ ที่มีผลตอบสนองของระบบไม่เป็นเชิงเส้น และมีตัวแปรไม่ทราบค่ารวมอยู่ด้วย ซึ่งระเบียบวิธีการออกแบบตัวควบคุมไม่เป็นเชิงเส้นในที่นี้ถูกแบ่งออกเป็น 2 ส่วน คือ ส่วนควบคุมสถิต และส่วนควบคุมจล ตัวควบคุมสถิตแบบไม่เป็นเชิงเส้นสามารถรับประกันได้ว่า ผลตอบสนองของระบบควบคุมป้อนกลับจะอยู่ภายในช่วงที่จำกัด โดยไม่สนใจว่าตัวแปรไม่ทราบค่าของระบบจะอยู่ภายในช่วงที่จำกัดด้วยหรือไม่ ส่วนตัวควบคุมจลแบบไม่เป็นเชิงเส้นนอกจากจะรับประกันได้ว่าผลตอบสนองของระบบจะอยู่ภายในช่วงที่จำกัดแล้ว ยังส่งผลให้ระบบควบคุมป้อนกลับสามารถติดตามสัญญาณอ้างอิงแบบขึ้นกับเวลาและไม่ขึ้นกับเวลาได้อีกด้วย

ในส่วนท้ายของหนังสือเล่มนี้ ได้แสดงตัวอย่างการนำเอาทฤษฎีระบบควบคุมไม่เป็นเชิงเส้นไปประยุกต์ใช้ในการควบคุมการสั่นของปีกเครื่องบิน ซึ่งแสดงให้เห็นถึงประโยชน์ของระบบควบคุมไม่เป็นเชิงเส้นในการใช้งานจริง

The Study of Linear or Nonlinear System Analysis and Design in Application of Aviation

Jatuporn Nakpiam

Pongrapee Kaewsaiha

Advisor

Dr. Poramate Pranayanuntana

ABSTRACT

In general, all systems have nonlinear modes effected by many strong disturbances and poor design in mechanical parts. For example, an electrical circuit is design to perform linear, but in real operations, it could be disturbed by noise, radiation and electromagnetic field which are nonlinear. These make the system unstable and cause malfunctioning operations. To avoid those problems, nonlinear controllers must be applied.

We first studied in linear system theory which is a basic knowledge about linear equation form of systems and stability theory. It is important that we have a knowledge of linear algebra to deal with these problem equations and we include it as an appendix. Next, we continued with a nonlinear part which is more complicated.

The main topic of our project is to design a nonlinear controller for systems or plants contain nonlinearities and unknown parameters. The design procedures are divided into static and dynamic part. Nonlinear static controllers guarantee a boundedness of the resulting feedback system without any knowledge of a bound on the unknown parameters. Nonlinear dynamic (or adaptive) controllers guarantee not only that the plant state remains bounded, but also that they tend to a desired constant value ("regulation") or asymptotic track a reference signal ("tracking").

At the end of this book, we include an example in applying a nonlinear adaptive controller to control a "wing rock" motion of the aircraft. This shows how useful of a nonlinear control system is in real operations.

THE STUDY OF LINEAR OR NONLINEAR SYSTEM ANALYSIS AND DESIGN IN APPLICATION OF AVIATION

P R O J E C T   R E P O R T

for the Degree of

Bachelor of Engineering (Mechatronics Engineering)

Jatuporn Nakpiam ID. 44010055

Pongrapee Kaewsaiha ID. 44010309

March 2005

THE STUDY OF LINEAR OR NONLINEAR SYSTEM ANALYSIS AND DESIGN IN APPLICATION OF AVIATION

PROJECT REPORT

Submitted in Partial Fulfillment
of the REQUIREMENTS for the

Degree of
Bachelor of Engineering (Mechatronics Engineering)

at the
King Mongkut's Institute of Technology Ladkrabang

by
Jatuporn Nakpiam ID. 44010055
Pongrapee Kaewsaiha ID. 44010309

March 2005

Copy No. 3/4

Approved by the Project Advisor:

P. Pranayanuntana

Dr.Poramate Pranayanuntana

MAR 27, 2005

Date

VITAE

Jatuporn and **Pongrapee** are studying in the 4th year of Mechatronics Engineering, Department of Control Engineering, Faculty of Engineering, King Monkut's Institute of Technology Ladkrabang. They're interested in linear and nonlinear control system theories and designs.

Major : Mechatronics Engineering

J. Nakpiam	[illegible]
Jatuporn Nakpiam	Pongrapee Kaewsaiha
MAR. 27, 2005	MAR. 27, 2005
Date	Date

THE FUTURE BELONGS TO THOSE
WHO BELIEVE IN THE BEAUTY OF THEIR DREAMS

ELEANOR ROOSEVELT,

American first lady (1884-1962)

To our parents

ACKNOWLEDGEMENT

We wish to thank Dr.Poramate Pranayananthana for his guidance, both in the selection of this topic and in its writing. He can also be though of as a big library full of knowledge and experiments from all over the world once in the known history. He is the most active and attentive man we've ever met in our life. It's unbelievable that this project has a steadily advancement than it could be. Without him, this project could only be a trash in a department's box and it will never be real. We're right that we took his first subject when we were in the 3rd-year. This improved our skill in linear control system and recovered the mathematical knowledge from the bottom of our brains. We spend lots of times with him to absorb the essential knowledge and do practically anything to fulfill the requirement of our project.

Unforgettable, we wish to thank his lovely family–his wife, his daughter and his son. They invited us in their small room with a warm welcome that impressed us since the first time we met. Two adorable,curious and naughty kids make a colorful times in that room. I wish to thank Dr.Khongsak Anuntahirunrat for his guidance in using *Visio* program. Also thank all teachers who enhance our knowledge and skills for many years ago.

At last, a sincerely thank is going to D.Knut, the inventor of TeX, that created this useful program. This makes a perfect paper that you've never seen in your life.

ABSTRACT

In general, all systems have nonlinear modes effected by many strong disturbances and poor design in mechanical parts. For example, an electrical circuit is designed to perform linear, but in real operations, it could be disturbed by noise, radiation and electromacnetic field which are nonlinear. These make the system unstable and cause malfunctioning operations. To avoid those problems, nonlinear controllers must be applied.

We first studied in linear system theory which is a basic knowledge about linear equation form of systems and stability theory. It is important that we have a knowledge of linear algebra to deal with these problems and we included it as an appendix. Next, we continued with a nonlinear part which is more complicate.

The main topic of our project is to design a nonlinear controller for systems or plants contain nonlinearities and unknown parameters. The design procedures are divided into static and dynamic part. Nonlinear static controllers guarantee a boundedness of the resulting feedback system without any knowledge of a bound on the unknown parameters. Nonlinear dynamic (or adaptive) controllers guarantee not only that the plant state remains bounded, but also that they tend to a desired constant value ("regulation") or asymptotically tracks a reference signal ("tracking").

At the end of this book, we included an example in applying a nonlinear adaptive controller to control a "wing rock" motion of the aircraft. This shows how useful of a nonlinear control system is in real operations.

Contents

List of Figures

Chapter 1

Introduction

1.1 Overview

In this book, we first started with a "linear system theory" which is a basic knowledge of control systems. It contains a form of linear system in state-space, which is the base form used by all design methods in this book, and design tools for stabilization of linear plants. A linear system theory is related to a linear algebra mentioned in Appendix A., it also need a knowledge of differential equations which are the part of a modern calculus to explain some characteristics of the system. These showed how mathematics are useful and essential for a study in the field of control system.

Next, we continued with a "nonlinear system theory" which is more complicate but useful to explain a system in real operations with an environmental effect and disturbance. Nonlinear stability can be explained by a multivariable circle and Popov criterion which guarantee an *absolute stability* of the system. Lyapunov stability theory is the most popular and used by many design methods of nonlinear controllers in the rest of this book.

We most concerned a nonlinear system as a strict-feedback system contains a "nonlinearity" (φ) and an "unknown constant parameter" (θ) in the feedback part (see Chapter 3). The main topic of this book is to design a nonlinear controller for the system by applying feedback. Controllers are divided into two main part, static and dynamic (adaptive). Nonlinear static controllers guarantee a boundedness of the resulting feedback system without any knowledge of a bound on the unknown parameters. Two design methods provided, we introduced "feedback linearization method" which forces the system to appear linear and "backstepping method" which can avoid cancellation of useful nonlinearity.

Adaptive controllers are dynamic and therefore more complex than the static controllers. The advantage of a nonlinear adaptive controller is it guarantee not only that the plant state remains bounded, but also that it tends to a desired constant value ("regulation") or asymptotically tracks a reference signal ("tracking"). So that a nonlinear adaptive controllers are more useful and applicable in real systems. Almost of this book is about the design of nonlinear adaptive controllers.

1.2 Achievements of adaptive linear control

While in this book we will be preoccupied with nonlinear system, we must not forget that the control of linear plants with unknown parameters was a formidable problem which took almost twenty years to solve. The adaptive control community deserves full credit for providing not only one, but several solutions to this fundamental problem [2]. Each of this solutions was a breakthrough in the development of adaptive control. By the early 1980's, several types of adaptive schemes were proven to provide stable operations and asymptotic tracking. Lyapunov-based design is one of the oldest results of adaptive control. Until recently, however, its limitation has been remove by the recursive design procedures presented in this book, commonly referred to as *backstepping.*

An important feature of traditional adaptive control is its reliance on "certainty equivalence" controllers. This means that a controller is first design as if all the plant parameters were known. The controller parameters are determined as functions of the plants parameters. Given the true values of the plants parameters, the controller parameters are calculated by solving design equations for model-matching, pole-zero, or optimality. When the true plant parameters are unknown, the controller parameters are either estimated directly (direct schemes) or computed by solving the same design equations with plant parameter estimates (indirect schemes). The resulting controller, which is either estimated (direct) or designed for the estimated plant (indirect), is called a *certainty equivalence* controller.

It is not at all obvious that a certainty equivalence controller will work inside and adaptive feedback loop and achieve stabilization and tracing. Even when the plant is stable, bad parameter estimates may yield a destabilizing controller. The situation is more critical when the plant is unstable, because then the controller must achieve stabilization in addition to its tracking task. It is therefore significant that certainty equivalence controllers have been proven to be satisfactory for adaptive control of linear systems.

In spite of major advances in the development of adaptive control schemes for linear system, they have not yet become tools for systematic engineering design. Each adaptive scheme leaves up to the designer the choice of various filters, design coefficients, initialization rules, and so on. It is still unclear how the adaptive system's performance, especially its transient performance, depends on these design choices. Current research activity is aimed at providing the designer with clearer choices and trade-offs between transient performance and robustness. Transient performance improvement is one of the prominent features of the nonlinear control design schemes that we are studying.

1.3 Adaptive control as dynamic nonlinear feedback

If a linear plant contains unknown parameters without any information about their bounds, then in general, it cannot be stabilized by a linear controller. This is true for the simplest scalar plant

$$\dot{x} = u + \theta x, \tag{1.3.1}$$

where u is the control and θ is and unknown constant. If an a priori bound $\bar{\theta}$ were known, $|\theta| \leq \bar{\theta}$, then $u = -2\bar{\theta}$ would be a linear stabilizing controller. If such abound is not known, no linear controller can be designed to guarantee stability of 1.3.1

To examine whether a static nonlinear controller can help, let us try the controller

$$u = -k_1 - k_2 x^3, \tag{1.3.2}$$

where $k_1 > 0$, $k_2 > 0$. The resulting feedback system is

$$\dot{x} = (\theta - k_1)x - k_2 x^3 \tag{1.3.3}$$

For $\theta > k_1$, the equilibrium $x = 0$ is unstable, but the nonlinear term $-k_2 x^3$ prevents $x(t)$ from growing unbounded. It is easy to see that $x(t)$ will converge to one of the two new equilibria $\pm\sqrt{\frac{\theta - k_1}{k_2}}$. Thus, the static nonlinear controller (1.3.2) has achieved boundedness of $x(t)$ without any knowledge of a bound on θ.

Out goal is more ambitious than just boundedness of $x(t)$. We also want to achieve its regulation: $\lim_{t\to\infty} x(t) = 0$. Can this be accomplished by a dynamic nonlinear controller? The answer is affirmative: One such controller is

$$u = -(p+\xi)x, \quad \dot{\xi} = x^2, \tag{1.3.4}$$

where $p > 0$ is a design parameter. The resulting feedback system is of second order:

$$\begin{aligned} \dot{x} &= -(p+\xi) + \theta x \\ \dot{\xi} &= x^2. \end{aligned} \tag{1.3.5}$$

Its stability properties can be checked by examining the derivative of the Lyapunov function

$$V(x,\xi) = \frac{1}{2}x^2 + \frac{1}{2}(\xi-\theta)^2, \tag{1.3.6}$$

which turns out to be nonpositive:

$$\dot{V} = -px^2 - \xi x^2 + \theta x^2 + (\xi-\theta)x^2 = -px^2. \tag{1.3.7}$$

Thus, $V(x(t),\theta(t))$ evaluated along the solution of (1.3.5) is a nonincreasing function of time. This proves that $x(t)$ and $\theta(t)$ remain bounded for all $t \geq 0$. The proof that $\lim_{t\to\infty} x(t) = 0$ is also achieved is given in the next chapter.

How was the dynamics nonlinear controller (1.3.4) conceived? Not as a nonlinear controller, but rather as a parameter adaption scheme! Its dynamics part $\dot{\xi} = x^2$ is, in fact, an update law for ξ as and estimate of θ. Consequently, the estimation error $\xi - \theta$ is penalized in the Lyapunov function(1.3.6).

1.4 Lyapunov-based design

The controller(1.3.4) is and outcome of a systematic Lyapunov design procedure. In this procedure we seek a parameter update update law for the estimate of $\hat{\theta}(t)$,

$$\dot{\hat{\theta}} = \tau(x, \hat{\theta}), \tag{1.4.1}$$

which, along with a control law $u = \alpha(x, \hat{\theta})$, will make the Lyapunov function

$$V(x, \hat{\theta}) = \frac{1}{2}x^2 + \frac{1}{2}(\hat{\theta} - \theta)^2, \tag{1.4.2}$$

a nonincreasing function of time:

$$V(x(t), \hat{\theta}) \leq V(x(t_0), \theta(t_0)) \forall t \geq t_0, \forall t_0 \geq 0. \tag{1.4.3}$$

To this end, we express $\dot{V}$ as a function of u and $\dot{\hat{\theta}}$ and seek $\propto (x, \hat{\theta})$ and $\tau(x, \hat{\theta})$ to guarantee that $\dot{V} \leq -px^2$ with $p > 0$, namely

$$\dot{V} = x(u + \theta x) + (\hat{\theta} - \theta)\dot{\hat{\theta}} \leq -px^2 \tag{1.4.4}$$

Rearranging terms we get

$$xu + \hat{\theta}\dot{\hat{\theta}} + \theta(x^2 - \dot{\hat{\theta}}) \leq -px^2 \tag{1.4.5}$$

Since neither $\alpha(x, \hat{\theta})$ nor $\tau(x, \hat{\theta})$ is allowed to depend on the unknown θ, we must take $\tau(x, \hat{\theta}) = x^2$, that is,

$$\dot{\hat{\theta}} = x^2 \tag{1.4.6}$$

The remaining condition

$$xu + \hat{\theta}x^2 \leq -px^2 \tag{1.4.7}$$

allows us to selected $\alpha(x, \hat{\theta})$ in various ways. The choice which results in the dynamic nonlinear controller(1.3.4) is

$$u = -(p + \hat{\theta})x. \tag{1.4.8}$$

We have thus designed our first Lyapunov-based adaptive scheme shown in Figure 1.1, where s is the complex variable of the Laplace transform. This scheme already exhibits some features of more general schemes to be designed.

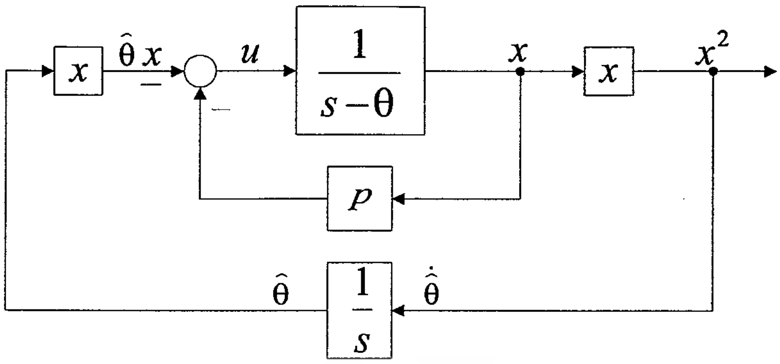

Figure 1.1: Lyapunov-based adaptive scheme for the scalar system $\dot{x} = u + x\theta$.

1.5 Classes of nonlinear systems

The main topic of this book is the design of feedback controllers for nonlinear system with unknown constant parameters. The most important design specification is to achieve asymptotic tracking of a known reference trajectory with the strongest possible form of stability. Another key requirement is that the designed controller should provide effective means for shaping the transient performance and thus allow different performance-robustness trade-offs. the largest classes of nonlinear systems for which the stated design problem is solvable with either state-feedback or output-feedback controllers are not know at this time. the largest classes for which solutions have been obtained are those considered in this book. State feedback solution are given for the so-called class of "parametric pure-feedback system". They are first presented for the subclass of "parametric strict−feedback system", for which the achieved stability and tracking properties are global. By analogy with linear systems, strict-feedback systems are also call "triangular". Output-feedback solutions are restricted to a narrower class of minimum phase systems in which the nonlinearities depend only the output variable. The class of pure-feedback systems with unknown parameters is well represented by the third order system

$$\begin{aligned}\dot{X}_1 &= x_2 + {\varphi_1}^T(x_1, x_2)\theta &\quad (1.5.1)\\ \dot{X}_2 &= x_3 + {\varphi_2}^T(x_1, x_2, x_3)\theta &\quad (1.5.2)\\ \dot{X}_3 &= u + {\varphi_3}^T(x_1, x_2, x_3)\theta, &\quad (1.5.3)\end{aligned}$$

where the $p \times 1$ vector θ is constant and unknown. (In a more general case, the term x_2, x_3, and u can be multiplied by unknown constant parameters provided that the signs of these parameters are known). Apart from the requirement that the dependence of the right–hand side of (1.5.1) on θ be linear,or to be precise, affine, pure-feedback systems are characterized by the structure of the known nonlinearities φ_1, φ_2, andφ_3. The function φ_1 must not depend on x_3, and a further implicit function restriction is imposed on the dependence of φ_1 on x_2, and of φ_2 on x_3. This is restriction is automatically satisfied if φ_1 does not depend on x_2, and φ_2 does not depend on x_3, that is, if we have $\varphi_1(x_1)$ and$\varphi_2(x_1, x_2)$. In this "strict-feedback" case the results are global.

1.6 Adaptive backstepping and tuning function

When the parameter vector θ is known, the pure-feedback restriction essentially amounts to feedback linearizability. However, even when achievable, the goal of feedback linearization is not pursued in this book, because it often leads to cancellation of useful nonlinearities. Backstepping designs are more flexible and do not force the designed system to appear linear and avoid cancellations of useful nonlinearities and often introduce additional nonlinear terms to improve transient performance.

The idea of backstepping is to design a controller for (1.5.1) recursively by considering some of the state variables as "virtual controls" and designing for them intermediate control laws. In (1.5.1), the first virtual control is x_2. It is used to stabilize the first equation as a separate system. Since θ is unknown, this task is solved with an adaptive controller consisting of the control law $\alpha_(x_1)1$ and the update law $\dot{\hat{\theta}} = \tau(x_1)$, as in the Lyapunov-based design in section 1.4. In the next step, the state x_3 is the "virtual control" which is used to stabilize the subsystem consisting of the first two equations of (1.5.1). This is again an adaptive control task, and a new update law $\dot{\hat{\theta}} = \tau(x_1)$ has already been designed in the first step and this does not seem to allow any freedom to proceed further. Chapter 6 and 7 provide two different cuts: *adaptive backstepping* and *tuning functions.*

Adaptive backstepping treats the parameter θ in the second equation of (1.5.1) as a new parameter and assigns to it a new estimate with a new update law. As a result, there are several estimates for the same parameter. This *overparametrization* is avoided in Chapter 7 by considering that in the first step $\dot{\hat{\theta}} = \tau(x_1)$ is not an update law but only a function τ_1. This "tuning function" is used in subsequence recursive steps and the discrepancy $\dot{\hat{\theta}} - \tau(x_1)$ is compensate with additional terms in the controller. Whenever the second derivative $\ddot{\hat{\theta}}$ would appear, it is replaced by the analytic expression for the first derivative of $\tau(x_1)$.

Both adaptive backstepping and tuning functions achieve the goals of stabilization and tracking. The proof of these properties is a direct consequence of the recursive procedure during which a Lyapunov function is constructed for the entire system, including the parameter estimates. For strict-feedback systems, this Lyapunov function provides the proof of global uniform stability and, if $x_1(t)$ is required to follow a trajectory $x_{1,ref}(t)$, also the proof of asymptotic tracking $x_1(t) - x_{1,ref}(t) \to 0$.

The tuning functions approach is an advance form of adaptive backstepping. It has the advantage that the dynamic order of the adaptive controller is minimal. The dimension of the set to which the state and parameter estimates converge is also minimal.

Both adaptive backstepping and tuning functions have crossed the "extended matching" barrier which blocked the traditional Lyapunov-based design. They have achieved this by designing controllers "more intelligent" and "stronger" than certainty equivalence controllers.

Chapter 2

Linear System Theory

2.1 Linear System

A system is called a linear system if for every t_o and any two state-input-output pairs

$$\left.\begin{array}{l} x_i(t_o) \\ u_i(t), \quad t \geq t_o \end{array}\right\} \rightarrow y_i(t), \quad t \geq t_o$$

for $i = 1, 2$ we have

$$\left.\begin{array}{l} x_1(t_o) + x_2(t_o) \\ u_1(t) + u_2(t), \quad t \geq t_o \end{array}\right\} \rightarrow y_1(t) + y_2(t), \quad t \geq t_o \quad (\textit{additivity})$$

and

$$\left.\begin{array}{l} \alpha x_1(t_o) \\ \alpha u_1(t), \quad t \geq t_o \end{array}\right\} \rightarrow \alpha y_1(t), \quad t \geq t_o \quad (\textit{homogeneity})$$

for any real constant α. The first property is called the *additivity* property, the second is called the *homogeneity* property. These two property can be combined as

$$\left.\begin{array}{l} \alpha_1 x_1(t_o) + \alpha_2 x_2(t_o) \\ \alpha_1 u_1(t) + \alpha_2 u_2(t), \quad t \geq t_o \end{array}\right\} \rightarrow \alpha_1 y_1(t) + \alpha_2 y_2(t), \quad t \geq t_o$$

for any real constant α_1 and α_2 are called the *superposition property.* A system is called a nonlinear system if the superposition property does not hold.

If the input $u(t)$ is identically zero for $t \geq t_o$, then the output will be excited exclusively by the initial state $x(t_o)$. this output is called the *zero-input response* and will be denoted by y_{zi} or

$$\left.\begin{array}{ll} x(t_o) & \\ u(t) \equiv 0, & t \geq t_o \end{array}\right\} \rightarrow y_{zi}(t), \quad t \geq t_o$$

If the initial state $x(t_o)$ is zero, then the output will be excited exclusively by the input. This output is called the *zero-state response* and will be denoted by y_{zs} or

$$\left.\begin{array}{ll} x(t_o) = 0 & \\ u(t), & t \geq t_o \end{array}\right\} \rightarrow y_{zs}(t), \quad t \geq t_o$$

The additivity property implies

$$\textit{Output due to} \left\{\begin{array}{ll} x(t_o) & \\ u(t), & t \geq t_o \end{array}\right. = \textit{output due to} \left\{\begin{array}{ll} x(t_o) & \\ u(t) \equiv 0, & t \geq t_o \end{array}\right.$$

$$+ \textit{ output due to} \left\{\begin{array}{ll} x(t_o) = 0 & \\ u(t), & t \geq t_o \end{array}\right.$$

or

$$\textit{Response} = \textit{zero-input response} + \textit{zero-state response}$$

Thus the response of every linear system can be decomposed into the zero-state response and the zero-input response. Furthermore, the two response can be studied separately and their sum yields the complete response. For nonlinear systems, the complete response can be very different from the sum of the zero-input response and zero-state response. Therefore we cannot separate the zero-input and zero-state responses in studying nonlinear systems.

If a system is linear, then the additivity and homogeneity properties apply to zero-state responses. To be more specific,if $x(t_o) = 0$, then the output will be excited exclusively by the input and the state-input-output equation can be simplified as $\{u_i \rightarrow y_i\}$. If the system is linear, then we have $\{u_1 + u_2 \rightarrow y_1 + y_2\}$ and $\{\alpha u_i \rightarrow y_i\}$ for all α and all u_i. A similar remark applies to zero-input responses of any linear system.

2.2 Linear Time-Invariant System

A system is said to be *time invariant* if for every state-input-output pair

$$\left.\begin{array}{l} x(t_o) \\ u(t), \quad t \geq t_o \end{array}\right\} \rightarrow y(t), \quad t \geq t_o$$

and any T, we have

$$\left.\begin{array}{l} x(t_o + T) \\ u(t - T), \quad t \geq t_o + T \end{array}\right\} \rightarrow y(t - T), \quad t \geq t_o + T \quad (time\ shifting)$$

It means that if the initial state is shifted to time $t_o + T$ and the same input waveform is applied from $t_o + T$ instead of from t_o, then the output waveform will be the same except that it starts to appear from time $t_o + T$. In other words, if the initial state and the input are the same, no matter at what time they are applied, the output waveform will always be the same. Therefore, for the time-invariant systems, we can always assume, without loss of generality, that $t_o = 0$. If a system is not time invariant, it is said to be *time varying.*

Time invariance is defined for systems, not for signals. Signals are mostly time varying. If a signal is time invariant such as $u(t) = 1$ for all t, then it is a very simple or a trivial signal. The characteristics of time-invariant systems must be independent of time. For example, an electrical circuit network is time invariant if R, C, and L are constants.

Some physical systems must be modeled as time-varying systems. For example, a burning rocket is a time-varying system, because its mass decreases rapidly with time. Although the performance of an automobile or a TV set may deteriorate over a long period of time, its characteristics do not change appreciable in the first couple of years. Thus a large number of physical systems can be modeled as time-invariant systems over a limited time period.

2.3 State-Space Equation

state-space equation Every linear time-invariant lumped system can be described by a set of equations of the form

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + Du(t) \end{aligned} \tag{2.3.1}$$

For a system with p input, q output, and n state variables, A, B, C,and D are, respectively, $n \times n$, $n \times p$, $q \times n$, and $q \times p$ constant matrices. Applying the Laplace transform to (2.3.1) yields

$$\begin{aligned} s\hat{x}(s) - x(0) &= A\hat{x}(s) + B\hat{u}(s) \\ \hat{y}(t) &= C\hat{x}(s) + D\hat{u}(s) \end{aligned}$$

which implies

$$\hat{x}(s) = (sI - A)^{-1}x(0) + (sI - A)^{-1}B\hat{u}(s) \tag{2.3.2}$$

$$\hat{y}(s) = C(sI - A)^{-1}x(0) + C(sI - A)^{(}-1)B\hat{u}(s) + D\hat{u}(s) \tag{2.3.3}$$

They are algebraic equations. Given $x(0)$ and $\hat{u}(s)$, $\hat{x}(s)$ and $\hat{y}(s)$ can be computed algebraically from (2.3.2) and (2.3.3). Their inverse Laplace transforms yield the time response $x(t)$ and $y(t)$. The equations also reveal the fact that the response of a linear system can be decomposed as the zero-state response and the zero-input response. If the initial state $x(0)$ is zero, then (2.3.3) reduces to

$$\hat{y}(s) = [C(sI-A)^{-1}B + D]\hat{u}(s)$$

Comparing this with $\hat{y}(s) = \hat{G}(s)\hat{u}(s)$ yields

$$\hat{G}(s) = C(sI-A)^{-1}B + D \tag{2.3.4}$$

This related the input-output (or transfer matrix) and state-space description.

2.4 Controllability

Consider the n-dimensional p-input state equation

$$\dot{\vec{x}} = A\vec{x} + B\vec{u} \tag{2.4.1}$$

where A and B are, respectively. $n \times n$ and $n \times p$ real constant matrices. Because the output does not play any role in controllability, we will disregard the output equation in this study.

Definition 2.4.1 *The state equation* (2.4.1) *or the pair* (A, B) *is said to be controllable if for any initial state* $\vec{x}(0) = \vec{x}_o$ *and any final state* $\vec{x}_1$*, there exists an input that transfers* $\vec{x}_0$ *to* $\vec{x}_1$ *in a finite time. Otherwise* (2.4.1) *or* (A, B) *is said to be uncontrollable.* □

Theorem 2.4.2 *The following statements are equivalent.*

1. *The n-dimensional pair* (A, B) *is controllable.*

2. *The* $n \times n$ *matrix*

$$W_c(t) = \int_{\tau=0}^{\tau=t} e^{A\tau}BB^T e^{A^T\tau}d\tau = \int_{\tau=0}^{\tau=t} e^{A(t-\tau)}BB^T e^{A^T(t-\tau)}d\tau \tag{2.4.2}$$

is nonsingular for any $t \geq 0$

3. *The $n \times np$ controllability matrix*

$$\mathcal{C} = [B \;\; AB \;\; A^2B \;\; \ldots \;\; A^{n-1}B] \tag{2.4.3}$$

has rank n (full row rank)

4. *The $n \times (n+p)$ matrix $[A - \lambda I B]$ has full row rank at all $\lambda \in \sigma(A)$*

5. *If, in addition, all eigenvalues of A have negative real parts, then the unique solution of*

$$AW_c + W_cA^T = -BB^T \tag{2.4.4}$$

is positive definite. The solution is called the "Controllability Gramian" *and can be expressed as*

$$W_c = \int_0^\infty e^{A\tau}BB^Te^{A^T\tau}d\tau \tag{2.4.5}$$

□

Example 2.4.3

$$\begin{aligned}\dot{\vec{x}} &= A\vec{x} + B\vec{u} \\ &= \begin{bmatrix} 1 & 1 & -2 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \end{bmatrix}\vec{x} + \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}u \\ \mathcal{C} &= [\vec{b}||A\vec{b}||A^2\vec{b}] = \begin{bmatrix} 0 & 1 & 2 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}\end{aligned}$$

□

2.5 Observability

The concept of observability is dual to that of controllability. Roughly speaking, controllability studies the possibility of steering the state from the input; observability studies the possibility of estimating the state from the output. These two concepts are defined under the assumption that the state equation or,equivalently, all A, B, C, and D are known. Thus the problem of observability is different from the problem of realization or identification, which is to determine or estimate A, B, C, and D from the information collected at the input and output terminals.

consider the n-dimensional p-input q-output state equation

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned} \tag{2.5.1}$$

where A, B, C, and D are, respectively, $n \times n$, $n \times p$, $q \times n$, $q \times p$ constant matrices.

Definition 2.5.1 *The state equation (2.5.1) is said to be observable if for any unknown initial state $x(0)$, there exist a finite $t_1 > 0$ such that the knowledge of the input u and the output $\mathbf{y}$ over $[0, t_1]$ suffices to determine uniquely the initial state $x(0)$. Otherwise, the equation is said to be unobservable.* □

Theorem 2.5.2 *The following statements are equivalent.*

1. *The n-dimension pair (A, C) is observable.*

2. *The $n \times n$ matrix*

$$W_o(t) = \int_0^t e^{A'\tau} C' e^{A\tau} d\tau \tag{2.5.2}$$

is nonsingular for any $t > 0$.

3. *The $nq \times n$ observability matrix*

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \tag{2.5.3}$$

4. *The $(n+q) \times n$ matrix*

$$\begin{bmatrix} A - \lambda I \\ C \end{bmatrix}$$

has full column rank at every eigenvalue, λ, of A.

5. *If, in addition, all eigenvalues of A have negative real parts, then the unique solution of*

$$A'W_o + W_oA = -C'C \tag{2.5.4}$$

is positive definite. The solution is called the "Observability Gramian" *and can be expressed as*

$$W_o = \int_0^{\infty} e^{A'\tau}C'e^{A\tau}d\tau \tag{2.5.5}$$

Let A and C ne $n \times n$ and $q \times n$ constant matrices. We assume that C does not have full row rank, then the output at some output terminal can be expressed as a linear combination of other outputs. Thus the output does not offer any new information regarding the system and the terminal can be eliminated. By deleting the corresponding row, the reduced C will then have full row rank. □

2.6 Introduction to Pole Placement

Most control system can be formulated as shown in Figure 2.1 , in which the plant and the reference signal $r(t)$ are given. The input $u(t)$ of the plant is called the

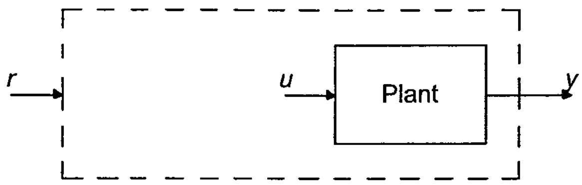

Figure 2.1: Design of control systems.

"actuating signal" or "control signal". The output $y(t)$ of the plant is called the "plant output" or "controlled signal". The problem is to design an overall system so that the plant output will follow as closely as possible to the reference signal $r(t)$. These are two types of control. If the actuating signal $u(t)$ depends only on the reference signal and is independent of the plant output, the control is called an "open-loop control". If the actuating signal depends on both the reference signal and the plant output, the control is called a "close loop" or "feedback control". The open-loop control is, in general, not satisfactory if there are plant parameter variations and/or there are noise and disturbance around the system. A properly designed feedback system, on the other hand, can reduce the effect of parameter variations and suppress noise and disturbance. Therefore feedback control is more widely used in practice.

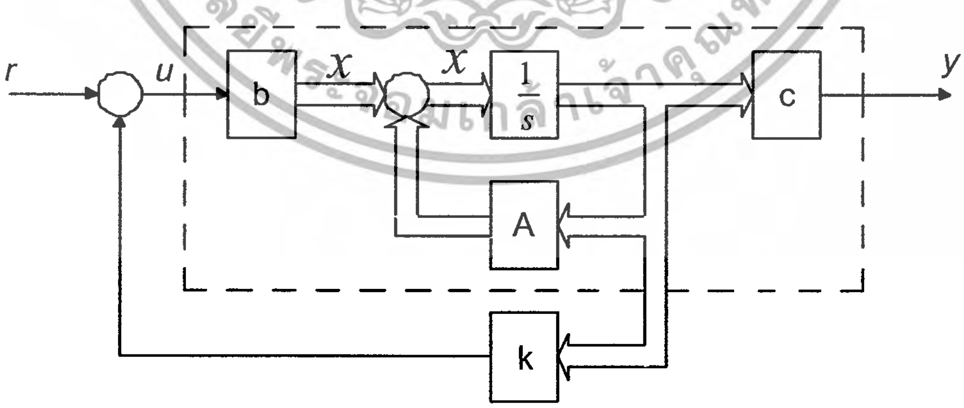

Figure 2.2: State Feedback.

2.7 State Feedback

Consider the n-dimensional single-variable state equation

$$\begin{aligned} \dot{x} &= Ax + bu \\ y &= cx \end{aligned} \tag{2.7.1}$$

where we have assumed $d = 0$ to simplify discussion. In state feedback, the input u is given by

$$u = r - kx = r - [k_1 \ \ k_2 \ \ \dots \ \ k_n]x = r - \sum_{i=1}^{n} k_i x_i \tag{2.7.2}$$

as shown in Figure 2.2. Each feedback gain k_i is a real constant. This is called the *constant gain negative feedback* or, simply, *state feedback*. Substituting (2.7.2) into (2.7.1) yields

$$\begin{aligned} \dot{x} &= (A - bk)x + br \\ y &= cx \end{aligned} \tag{2.7.3}$$

Theorem 2.7.1 *The pair $(A - bk, b)$, for any $1 \times n$ real constant vector k, is controllable if and only if (A, b) is controllable.* □

Example 2.7.2 *Consider the state equation*

$$\dot{x} = \begin{bmatrix} 1 & 2 \\ 3 & 1 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 2 \end{bmatrix} x$$

The state equation can readily be shown to be controllable and observable. Now we introduce the state feedback

$$u = r - \begin{bmatrix} 3 & 1 \end{bmatrix} x$$

Then the state feedback equation becomes

$$\dot{x} = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 2 \end{bmatrix} x$$

The controllability matrix is

$$\mathcal{C} = \begin{bmatrix} 0 & 2 \\ 1 & 0 \end{bmatrix}$$

which is nonsingular. Thus the state feedback equation is controllable. Its observability matrix is

$$\mathcal{O} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$$

which is singular. Thus the state feedback equation is not observable. The reason that the observability property may not be preserved in state feedback will be given later. □

Example 2.7.3 *Consider a plant described by*

$$\dot{x} = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u$$

The A-matrix has characteristic polynomial

$$\Delta(s) = (s-1)^2 - 9 = s^2 - 2s - 8 = (s-4)(s+2)$$

and, consequently, eigenvalues 4 and -2. It is unstable. Let us introduce state feedback $u = r - [k_1 \ \ k_2]x$. Then the state feedback system is described by

$$\dot{x} = \left(\begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} - \begin{bmatrix} k_1 & k_2 \\ 0 & 0 \end{bmatrix} \right) x + \begin{bmatrix} 1 \\ 0 \end{bmatrix} r$$

$$= \begin{bmatrix} 1-k_1 & 3-k_2 \\ 3 & 1 \end{bmatrix} x + \begin{bmatrix} 1 \\ 0 \end{bmatrix} r$$

This new A-matrix has characteristic polynomial

$$\begin{aligned} \Delta_f(s) &= (s-1+k_1)(s-1) - 3(3-k_2) \\ &= s^2 + (k_1-2)s + (3k_2 - k_1 - 8) \end{aligned}$$

It is clear that the roots of $\Delta_f(s)$ or, equivalently, the eigenvalues of the state feedback system can be placed in any positions by selecting appropriate k_1 and k_2. For example, if the two eigenvalues are to be placed at $-1 \pm j2$, then the desired characteristic polynomial is $(s+1+j2)(s+1-j2) = s^2+2s+5$. Equating $k_1 - 2 = 2$ and $3k_2 - k_1 - 8 = 5$ yields $k_1 = 4$ and $k_2 = 17/3$. Thus the state feedback gain $[4 \quad 17/3]$ will shift the eigenvalues from $4, -2$ to $-1 \pm j2$. □

This example shows that state feedback can be used to place eigenvalues in any positions. Moreover the feedback gain can computed by direct substitution. This approach however, will become very involved for three- or higher-dimensional state equations. More seriously, the approach will not reveal how the controllability condition comes into the design. Therefore a more systematic approach is desirable. Before proceeding, we need the following theorem. We state the theorem for $n = 4$; the theorem, however, holds for every positive integer n.

Theorem 2.7.4 *Consider the state equation in (2.7.1) with $n = 4$ and the characteristic polynomial*

$$\Delta(s) = det(sI - A) = s^4 + \alpha_1 s^3 + \alpha_2 d^2 + \alpha_3 s + \alpha_4 \tag{2.7.4}$$

If (2.7.1) is controllable, then it can be transformed by the transformation $\bar{x} = Px$ with

$$Q := P^{-1} = \begin{bmatrix} b & Ab & A^2b & A^3b \end{bmatrix} \begin{bmatrix} 1 & \alpha_1 & \alpha_2 & \alpha_3 \\ 0 & 1 & \alpha_1 & \alpha_2 \\ 0 & 0 & 1 & \alpha_1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

into the controllable canonical form

$$\dot{\bar{x}} = \bar{A}\bar{x} + \bar{b}u = \begin{bmatrix} -\alpha_1 & -\alpha_2 & -\alpha_3 & -\alpha_4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \bar{x} + \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} u$$

$$y = \bar{c}\bar{x} = \begin{bmatrix} \beta_1 & \beta_2 & \beta_3 & \beta_4 \end{bmatrix} \bar{x}$$

Furthermore, the transfer function of (2.7.1) *with* $n = 4$ *equals*

$$\hat{g}(s) = \frac{\beta_1 s^3 + \beta_2 s^2 + \beta_3 s + \beta_4}{s^4 + \alpha_1 s^3 + \alpha_2 s^2 + \alpha_3 s + \alpha_4} \tag{2.7.5}$$

□

Feedback transfer function Consider a plant described by (A, b, c).If (A, b) is controllable, (A, b, c) can be transformed into the controllable form in (2.7.4) and its transfer function can then be read out as, for $n = 4$,

$$\hat{g}(s) = c(sI - A)^{-1}b = \frac{\beta_1 s^3 + \beta_2 s^2 + \beta_3 s + \beta_4}{s^4 + \alpha_1 s^3 + \alpha_2 s^2 + \alpha_3 s + \alpha_4} \tag{2.7.6}$$

After state feedback, the state equation becomes $(A - bk, b, c)$ and is still of the controllable canonical form. Thus the feedback transfer function from r to y is

$$\hat{g}_f(s) = c(sI - A + bk)^{-1}b = \frac{\beta_1 s^3 + \beta_2 s^2 + \beta_3 s + \beta_4}{s^4 + \bar{\alpha}_1 s^3 + \bar{\alpha}_2 s^2 + \bar{\alpha}_3 s + \bar{\alpha}_4} \tag{2.7.7}$$

We see that the numerators of (2.7.6) and (2.7.7) are the same. In other words, state feedback does not affect the zeros of the plant transfer function. This is actually a general property of feedback: *feedback can shift the poles of a plant but has no effect on the zeros.* This can be used to explain by a state feedback may alter the observability property of a state equation. If one or more poles are shifted to coincide with zeros of $\hat{g}(s)$, then the numerator and denominator of $\hat{g}_f(s)$ in (2.7.7) are not coprime. Thus the state equation $(A - bk, c)$ is not observable.

Theorem 2.7.5 *If the n-dimensional state equation in* (2.7.1) *is controllable, then by state feedback* $u = r - kx$, *where* k *is a* $1 \times n$ *real constant vector, the eigenvalues of* $A - bk$ *can arbitrarily be assigned provided that complex conjugate eigenvalues are assigned in pairs.* □

Example 2.7.6 *Consider the state equation*

$$\dot{x} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 5 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \\ 0 \\ -2 \end{bmatrix} u$$

$$y = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} x$$

It is controllable;thus its eigenvalues can be assigned arbitrarily. Because the A-matrix is block triangular, its characteristic polynomial can be obtained by inspection as

$$\Delta(s) = s^2(s^2 - 5) = s^4 + 0 \cdot s^3 - 5s^2 + 0 \cdot s + 0$$

First we compute P that will transform (2.7.6) *into the controllable canonical form. Using* (2.7.4), *we have*

$$P^{-1} = C\bar{C}^{-1} = \begin{bmatrix} 0 & 1 & 0 & 2 \\ 1 & 0 & -3 & 0 \\ 0 & -2 & 0 & -10 \\ -2 & 0 & -10 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & -5 & 0 \\ 0 & 1 & 0 & -5 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 1 & 0 & -3 \\ 1 & 0 & -3 & 0 \\ 0 & -2 & 0 & 0 \\ -2 & 0 & 0 & 0 \end{bmatrix}$$

Its inverse is

$$P = \begin{bmatrix} 0 & 0 & 0 & -\frac{1}{2} \\ 0 & 0 & -\frac{1}{2} & 0 \\ 0 & -\frac{1}{3} & 0 & -\frac{1}{6} \\ -\frac{1}{3} & 0 & -\frac{1}{6} & 0 \end{bmatrix}$$

Let the desired eigenvalues be $-1.5 \pm 0.5j$ and $-1 \pm j$. Then we have

$$\begin{aligned}\Delta_f(s) &= (s+1.5-0.5j)(s+1.5+0.5j)(s+1-j)(s+1+j) \\ &= s^4+5s^3+10.5s^2+11s+5\end{aligned}$$

Thus we have

$$\bar{k} = \left[\begin{array}{cccc} 5-0 & 10.5+5 & 11-0 & 5-0 \end{array}\right]$$

and

$$k = \bar{k}P = \left[\begin{array}{cccc} -\frac{5}{3} & -\frac{11}{3} & -\frac{103}{12} & -\frac{13}{3} \end{array}\right] \tag{2.7.8}$$

This state feedback gain will shift the eigenvalues of the plant from $\{0, \quad 0, \quad \pm j\sqrt{5}\}$ *to* $\{-1.5 \pm 0.5j, \quad -1 \pm j\}$. □

2.8 State Estimators

In order to apply state feedback, we must design a device, called a "state estimator" or "state observer", so that the output of the device will generate an estimation of the state which is under the implicit assumption. However, this assumption may not hold in practice either because the state variables are not accessible for direct connection or because sensors or transducers are not available or very expensive.

2.8.1 Open-Loop State Estimator

Consider the n-dimensional state equation

$$\begin{aligned}\dot{x} &= Ax+Bu \\ y &= Cx\end{aligned} \tag{2.8.1}$$

where A,B and C are given and the input $u(t)$ and the output $y(t)$ are available for us. However, the state x is not available. The problem is to estimate x from u and y with the knowledge of A,B and C. If we know A and B, we can duplicate the original system as

$$\dot{\hat{x}} = A\hat{x} + Bu \tag{2.8.2}$$

as shown in Figure 2.3. Note that the original system could be an electromechanical system and the duplicated system could be an op-amp circuit.The duplication will be called an "open-loop" estimator. Now if (2.8.1) and (2.8.2) have the same initial state, then for any input, we have $\hat{x}(t) = x(t)$ for all $t \geq 0$. Therefore the remaining question is how to find the initial state of (2.8.1) and then set the initial state of (2.8.2) to the state. If (2.8.1) is observable, its initial state $x(0)$ can be computed from u and y over any time interval, say, $[0, t_1]$. We can then compute the state at t_2 and set $\hat{x}(t_2) = x(t_2)$. Then we have $\hat{x}(t) = x(t)$ for all $t \geq t_2$. Thus if (2.8.1) is observable, an open-loop estimator can be used to generate the state vector.

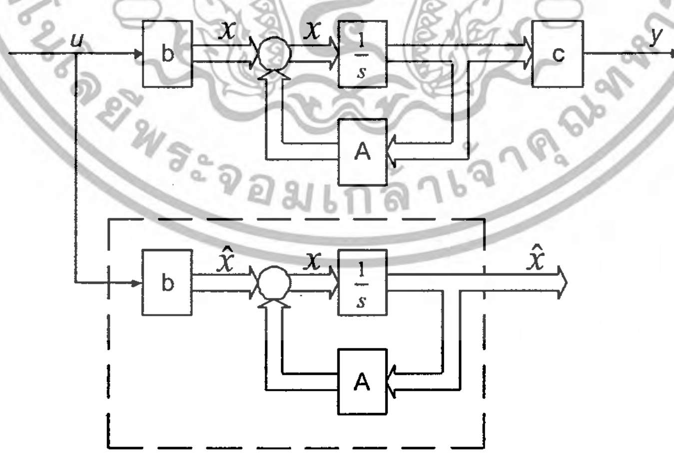

Figure 2.3: Open-loop state estimator.

2.8.2 Closed-Loop State Estimator

There are, however, two disadvantages in using an open-loop estimator. First, the initial state must be computed and set each time we use the estimator. This is very inconvenient. Second, and more seriously, if the matrix A has eigenvalues with positive real parts, then even for a very small difference between $x(t_o)$ and $\hat{x}(t_o)$ for some t_o, which may be caused by disturbance or imperfect estimation of the initial state, the difference between $x(t)$ and $\hat{x}(t)$ will grow with time. Therefore the open-loop estimator is, in general, not satisfactory.

We see from Figure 2.3 that even though the input and output of (2.8.1) are available, we use only the input to drive the open-loop estimator. Now we shall modify the estimator in Figure 2.3 to the one in Figure 2.4, in which the output $y(t) = cx(t)$ of (2.8.1) is compared with $c\hat{x}(t)$. Their difference, passing through an $n \times n$ constant gain vector l, is used as a corresponding term. If the difference is zero, no correction is needed. If the difference is nonzero and if the gain l is properly designed, the difference will drive the estimated state to the actual stste. Such an estimator is called a *closed-loop* or an *asymptotic* estimator

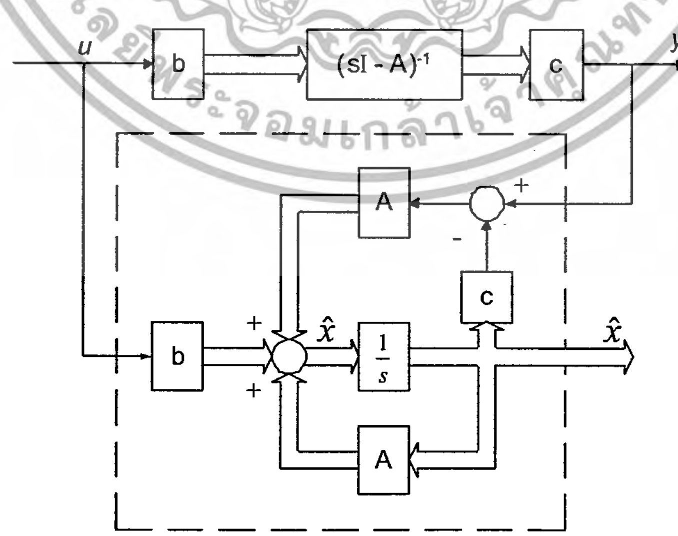

Figure 2.4: Closed-loop state estimator.

or, simply, an estimator.

The open-loop estimator in (2.8.2) is now modified as, following Figure 2.4,

$$\dot{\hat{x}} = A\hat{x} + bu + Iy - c\hat{x}$$

which can be written as

$$\dot{\hat{x}} = (A - Ic)\hat{x} + bu + Iy \tag{2.8.3}$$

and is shown in Figure 2.5. It has two inputs u and y and its output yields an estimated state $\hat{x}$. We define

$$e(t) := x(t) - \hat{x}(t)$$

It is the error between the actual state and the estimated state. Differentiating e and then substituting (2.8.1) and (2.8.3) into it, we obtain

$$\begin{aligned} \dot{e} &= \dot{x} - \dot{\hat{x}} = Ax + bu - (A - Ic)\hat{x} - bu - I(cx) \\ &= (A - Ic)x - (A - Ic)\hat{x} = (A - Ic)(x - \hat{x}) \end{aligned}$$

or

$$\dot{e} = (A - Ic)e \tag{2.8.4}$$

This equation governs the estimation error. If all eigenvalues of $(A - Ic)$ can be assigned arbitrary, then we can control the rate for $e(t)$ to approach zero or, equivalently, for the estimated state to approach the actual state. For example, if all eigenvalues of $(A - Ic)$ have negative real parts smaller than $-\sigma$, then all entries of e will approach zero at rates faster than $e^{-\sigma t}$. Therefore, even if there is a large error between $\hat{x}(t_o)$ and $x(t_o)$ at initial time t_o, the estimated state will approach the actual state rapidly. Thus there is no need to compute the initial state of the original state equation. In conclusion, if all eigenvalues of $(A - Ic)$

are properly assigned, a closed-loop estimator is much more desirable than an open-loop estimator.

As in the state feedback, what constitutes the best eigenvalues is not a simple problem. Probably, they should be placed evenly along a circle inside the sector if an estimator is to be used in state feedback, then the estimator eigenvalues should be faster than the desired eigenvalues of the state feedback. Again, saturation and noise problems will impose constraints on the selection. One way to carry out the selection is by computer simulation.

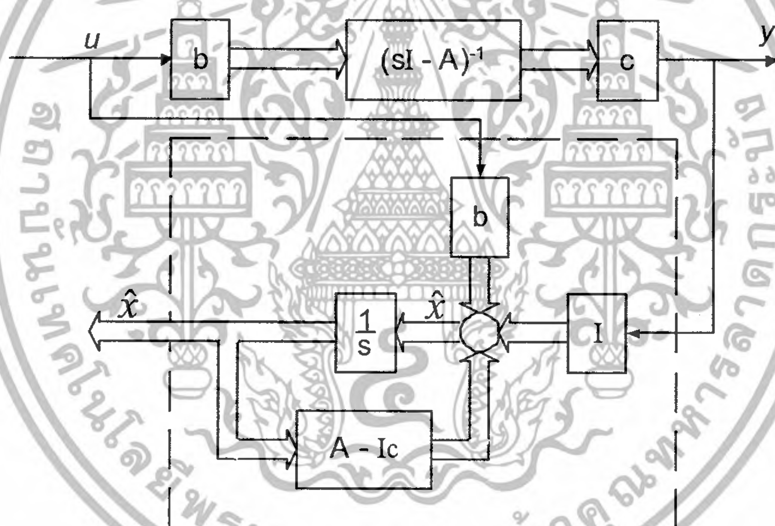

Figure 2.5: Closed-loop state estimator.

Chapter 3

Nonlinear System Theory

Circle Criterion and Popov Criterion

Consider a feedback system

$$\dot{\vec{x}} = A\vec{x} + B\vec{u} \tag{3.0.1}$$

$$\vec{y} = C\vec{x} \tag{3.0.2}$$

$$\vec{u} = -\psi(t, \vec{y}) \tag{3.0.3}$$

as in Figure 3.1 where $\vec{x} \in \mathbb{R}^n$, $\vec{u}, \vec{y} \in \mathbb{R}^p$, (A, B) is controllable, (A, C) is observable, and $\psi : [0, \infty) \times \mathbb{R}^p \to \mathbb{R}^p$ is a memoryless, possibly time-varying nonlinearity which is piecewise continuous in t and locally Lipschitz in $\vec{y}$. The transfer function matrix of the linear system is given by

$$G(s) = C(sI - A)^{-1}B \tag{3.0.4}$$

which is a square strictly proper transfer function. The controllability-observability assumptions ensure that $\{A, B, C\}$ is a minimal realization of $G(s)$.

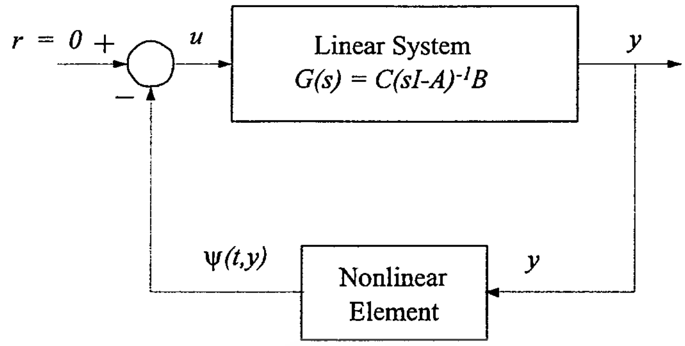

Figure 3.1: Feedback connection of a linear system and a nonlinear element.

Definition 3.0.1 *A memoryless nonlinearity $\psi : [0, \infty) \times \mathbb{R}^p \to \mathbb{R}^p$ is said to satisfy a sector condition if*

$$[\psi(t,\vec{y}) - K_{min}\vec{y}]^T [\psi(t,\vec{y}) - K_{max}\vec{y}] \leq 0, \quad \forall t \geq 0, \forall \vec{y} \in \Gamma \subset \mathbb{R}^p \quad (3.0.5)$$

for some real matrices K_{min} and K_{max}, where $K = K_{max} - K_{min}$ is a positive definite symmetric matrix and the interior of Γ is connected and contains the origin. If $\Gamma = \mathbb{R}^p$, then $\psi(\cdot,\cdot)$ satisfies the sector condition globally, in which case it is said that $\psi(\cdot,\cdot)$ belongs to a sector $[K_{min}, K_{max}]$. If (3.0.5) *holds with strict inequality, then $\psi(\cdot,\cdot)$ is said to belong to a sector (K_{min}, K_{max}).* □

For all nonlinearities satisfying the sector condition (3.0.5), the origin $\vec{x} = \vec{0}$ is an equilibrium point of the system (3.0.1)-(3.0.3). The problem of studying of stability of the origin for a class of nonlinearities that satisfy a given sector condition, such that if we succeed in showing that the origin is asymptotically stable for all nonlinearities in the sector, the system is said to be absolutely stable. This problem was originally formulated by Lure, and is sometimes called *Lure's problem.* Traditionally, absolute stability has been defined for the case when the origin is globally asymptotically stable. To keep up with this tradition, we shall use the phrase "absolute stability" when the sector condition is satisfied globally and the origin is globally asymptotically stable. Otherwise, we shall use the phrase "absolute stability with a finite domain."

Definition 3.0.2 *Consider the system* (3.0.1)-(3.0.3), *where* $\psi(\cdot,\cdot)$ *satisfies a sector condition according to Definition 3.0.1. The system is absolutely stable if the origin is globally uniformly asymptotically stable for any nonlinearity in the given sector. It is absolutely stable with a finite domain if the origin is uniformly asymptotically stable.* □

The concept of a positive real[1] transfer function and a key lemma (Lemma 3.0.5) that relates this concept to the existence of Lyapunov functions play a major role in the frequency-domain conditions of multivariable circle and Popov criterions.

Definition 3.0.3 *A* $p \times p$ *proper rational transfer function matrix* $Z(s)$ *is called positive real if*

- *all elements of* $Z(s)$ *are analytic for* $\mathrm{Re}[s] > 0$.
- *any pure imaginary pole of any element of* $Z(s)$ *is a simple pole and the associated residue matrix of* $Z(s)$ *is positive semidefinite Hermitian, and*
- *for all real* ω *for which* $j\omega$ *is not a pole of any element of* $Z(s)$, *the matrix* $Z(j\omega) + Z^T(-j\omega)$ *is positive semidefinite.*

The transfer function $Z(s)$ *is called strictly positive real*[2] *if* $Z(s-\epsilon)$ *is positive real for some* $\epsilon > 0$. □

When $p = 1$, the frequency domain condition of Definition 3.0.3 reduces to $\mathrm{Re}\, Z(j\omega) \geq 0, \forall\omega \in \mathbb{R}$ which is equivalent to the condition that the Nyquist plot of $Z(j\omega)$ lies in the closed right-half complex plane.

The following lemma gives an equivalent characterization of strictly positive real transfer functions.

[1]See [5] and [7] for the definitions and properties of positive real transfer functions.

[2]The definition of strictly positive real transfer functions is not uniform in the literature; see[7] for various definitions and the relationship between them.

Lemma 3.0.4 *Let $Z(s)$ be a $p \times p$ proper rational transfer function matrix, and suppose* $\det\left[Z(s) + Z^T(-s)\right]$ *is not identically zero. Then, $Z(s)$ is strictly positive real if and only if*

- *$Z(s)$ is* **Hurwitz**; *that is, poles of all elements of $Z(s)$ have negative real parts,*
- *$Z(j\omega) + Z^T(-j\omega) > 0, \quad \forall \omega \in \mathbb{R}$, and*
- *one of the following three conditions is satisfied:*
  1. *$Z(\infty) + Z^T(\infty) > 0$;*
  2. *$Z(\infty) + Z^T(\infty) = 0$ and $\lim_{\omega \to \infty} \omega^2 \left[Z(j\omega) + Z^T(-j\omega)\right] > 0$;*
  3. *$Z(\infty) + Z^T(\infty) \geq 0$ (but not zero nor singular) and there exist positive constants σ_o and ω_o such that*
$$\omega^2 \sigma_{min} \left[Z(j\omega) + Z^T(-j\omega)\right] \geq \sigma_o, \quad \forall |\omega| \geq \omega_o.$$

□

Strictly positive real transfer functions will arise in our analysis as a consequence of the following lemma, known as the Kalman-Yakubovich-Popov lemma or the positive real lemma.

Lemma 3.0.5 *Let $Z(s) = \mathcal{C}(sI - \mathcal{A})^{-1}\mathcal{B} + \mathcal{D}$ be a $p \times p$ transfer function matrix, where $\mathcal{A}$ is Hurwitz, $(\mathcal{A}, \mathcal{B})$ is controllable, and $(\mathcal{A}, \mathcal{C})$ is observable. Then, $Z(s)$ is strictly positive real if and only if there exist a positive definite symmetric matrix P, matrices W and L, and a positive constant ϵ such that*

$$P\mathcal{A} + \mathcal{A}^T P = -L^T L - \epsilon P \tag{3.0.6}$$

$$P\mathcal{B} = \mathcal{C}^T - L^T W \tag{3.0.7}$$

$$W^T W = \mathcal{D} + \mathcal{D}^T \tag{3.0.8}$$

□

3.1 Multivariable Circle Criterion

Lemma 3.1.1 *Consider the system* (3.0.1)-(3.0.3), *where* A *is Hurwitz,* (A, B) *is controllable,* (A, C) *is observable, and* $\psi(\cdot, \cdot)$ *satisfies the sector condition*

$$\psi(t, \vec{y})^T [\psi(t, \vec{y}) - K\vec{y}] \ \leq \ 0, \quad \forall t \geq 0, \forall \vec{y} \in \Gamma \subset \mathbb{R}^p \tag{3.1.1}$$

globally ($\Gamma = \mathbb{R}^p$), *where* K *is a positive definite symmetric matrix. Then, the system is absolutely stable if* $Z(s) = I + KG(s)$ *is strictly positive real. If* (3.1.1) *is satisfied only on a set* $\Gamma \subset \mathbb{R}^p$ *and* $\Gamma \neq \mathbb{R}^p$, *then the same condition on* $Z(s)$ *ensure that the system is absolutely stable with a finite domain.*

Proof *Consider a Lyapunov function candidate*

$$V(\vec{x}) = \vec{x}^T P \vec{x}$$

where P *is a positive definite symmetric matrix to be chosen. The derivative of* $V(\vec{x})$ *along the trajectories of the system* (3.0.1)-(3.0.3) *is given by*

$$\dot{V}(t, \vec{x}) \ = \ \vec{x}^T (PA + A^T P)\vec{x} - 2\vec{x}^T PB\psi(t, \vec{y}) \tag{3.1.2}$$

Since

$$\psi(t, \vec{y})^T [\psi(t, \vec{y}) - K\vec{y}] \leq 0,$$

multiplied by -2 *yields*

$$-2\psi(t, \vec{y})^T [\psi(t, \vec{y}) - K\vec{y}] \ \geq \ 0, \tag{3.1.3}$$

then addition of $-2\psi^T[\psi - K\vec{y}]$ *to the right-hand side of* (3.1.2) *gives an upper bound on* $\dot{V}(t, \vec{x})$. *Therefore,*

$$\begin{aligned} \dot{V}(t, \vec{x}) \ &\leq \ \vec{x}^T (PA + A^T P)\vec{x} - 2\vec{x}^T PB\psi(t, \vec{y}) - 2\psi(t, \vec{y})^T [\psi(t, \vec{y}) - K\vec{y}] \quad (3.1.4) \\ &= \ \vec{x}^T (PA + A^T P)\vec{x} + 2\vec{x}^T (C^T K - PB)\psi(t, \vec{y}) - 2\psi(t, \vec{y})^T \psi(t, \vec{y}). \end{aligned}$$

Suppose there are matrices $P = P^T > 0$ and L and W and a constant $\epsilon > 0$ such that

$$PA + A^T P = -L^T L - \epsilon P \quad (3.1.5)$$

$$PB = C^T K - L^T W \quad (3.1.6)$$

$$W^T W = I + I^T \quad (3.1.7)$$

Then,

$$\begin{aligned} \dot{V}(t,\vec{x}) &\leq -\epsilon\vec{x}^T P\vec{x} - \vec{x}^T L^T L\vec{x} + 2\sqrt{2}\vec{x}^T L^T \psi(t,\vec{y}) - 2\psi(t,\vec{y})^T \psi(t,\vec{y}) \quad (3.1.8) \\ &= -\epsilon\vec{x}^T P\vec{x} - \left[L\vec{x} - \sqrt{2}\psi(t,\vec{y})\right]^T \left[L\vec{x} - \sqrt{2}\psi(t,\vec{y})\right] \\ &\leq -\epsilon\vec{x}^T P\vec{x}. \end{aligned}$$

Thus, we can show that $\dot{V}(t,\vec{x})$ is negative definite provided we can find P, L, W and ϵ that satisfy (3.1.5)-(3.1.7). *Using Lemma 3.0.5, we can see that this is the case if and only if*

$$Z(s) = I + KC(sI - A)^{-1}B \quad (3.1.9)$$

is strictly positive real. Notice that the pair (A, KC) is observable since (A, C) is observable and K is nonsingular. □

The restriction on A to be Hurwitz can be removed by an idea known as *loop transformation* ([6], pp. 50 - 52) or *pole shifting.* The result is as the following theorem, which includes Lemma 3.1.1 as a special case.

Theorem 3.1.2 (Multivariable Circle Criterion) *Consider the system* (3.0.1)-(3.0.3), *where (A, B) is controllable, (A, C) is observable, and $\psi(\cdot,\cdot)$ satisfies the sector condition* (3.0.5) *globally $\Gamma = \mathbb{R}^p$. Then, the system is absolutely stable if*

$$G_T(s) = G(s)\left[I + K_{min}G(s)\right]^{-1} \quad (3.1.10)$$

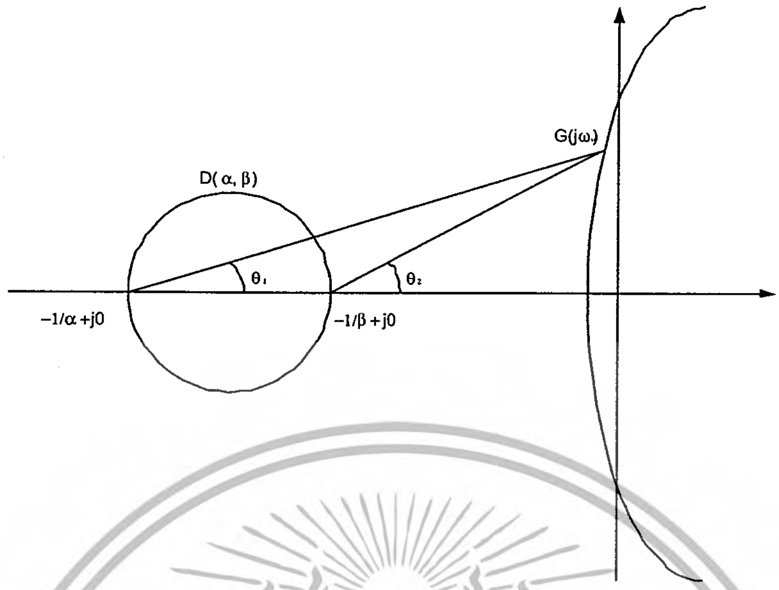

Figure 3.2: Graphical Representation of the circle criterion.

is Hurwitz and

$$\begin{aligned} Z_T(s) &= I + KG_T(s) \\ &= I + (K_{max} - K_{min})\, G(s)\, [I + K_{min}G(s)]^{-1} \\ &= [I + K_{max}G(s)]\,[I + K_{min}G(s)]^{-1} \end{aligned} \tag{3.1.11}$$

is strictly positive real. If (3.0.5) *is satisfied only on a set* $\Gamma \subset \mathbb{R}^p$, *then the conditions given on* $G_T(s)$ *and* $Z_T(s)$ *ensure that the system is absolutely stable with a finite domain.* □

We refer to this theorem as the *multivariable circle criterion*, the reason for using this name is clear for the scalar case $p = 1$.

Theorem 3.1.3 (Circle Criterion) *Consider the system* (3.0.1)-(3.0.3), *where* $\{A, B, C\}$ *is a minimal realization of* $G(s)$ *and* $\psi(\cdot,\cdot)$ *satisfies the sector condition*

$$[\psi(t,y) - \alpha y]\,[\psi(t,y) - \beta y] \;\leq\; 0, \quad \forall t \geq 0,\, \forall y \in \Gamma = [a,b] \subset \mathbb{R}^1 \tag{3.1.12}$$

globally ($\Gamma = [a,b] = (-\infty,\infty) = \mathbb{R}^1$). *Then the system is absolutely stable if one of the following conditions is satisfied, as appropriate:*

1. *If* $0 < \alpha < \beta$ *the Nyquist plot of* $G(j\omega)$ *does not enter the disk* $D(\alpha, \beta)$ *and encircles it* m *times in the counterclockwise direction, where* m *is the number of poles of* $G(s)$ *with positive real parts.*

2. *If* $0 = \alpha < \beta$, $G(s)$ *is Hurwitz and the Nyquist plot of* $G(j\omega)$ *lies to the right of the vertical line defined by* $\mathrm{Re}[s] = -1/\beta$.

3. *If* $\alpha < 0 < \beta$, $G(s)$ *is Hurwitz and the Nyquist plot of* $G(j\omega)$ *lies in the interior of the disk* $D(\alpha, \beta)$.

If the sector condition (3.1.12) *is satisfied only on an interval* $\Gamma = [a, b] \subset \mathbb{R}^1$ *containing the origin* $x = 0$ $(\Gamma \neq \mathbb{R}^1)$, *then the foregoing conditions ensure that the system is absolutely stable with a finite domain.* □

3.2 Multivariable Popov Criterion

Theorem 3.2.1 (Multivariable Popov Criterion) *Consider the system* (3.0.1)-(3.0.3) *where* A *is Hurwitz,* (A, B) *is controllable,* (A, C) *is observable, and* $\psi(\cdot)$ *is a time-invariant nonlinearity that satisfies the sector condition*

$$\psi(\vec{y})^T [\psi(\vec{y}) - K\vec{y}] \leq 0, \quad \forall \vec{y} \in \Gamma \subset \mathbb{R}^p \tag{3.2.1}$$

globally $(\Gamma = \mathbb{R}^p)$ *with a positive definite symmetric matrix* K. *Suppose that* $K\psi(\vec{y})$ *is the gradient of a scalar function and*

$$\int_{\vec{0}}^{\vec{y}} \psi(\vec{s})^T K \, d\vec{s} \;\; \geq \;\; 0, \quad \forall \vec{y} \in \Gamma \subset \mathbb{R}^p \tag{3.2.2}$$

is satisfied globally $(\Gamma = \mathbb{R}^p)$. *Then, the system is absolutely stable if there is* $q \geq 0$, *with* $-1/q$ *not an eigenvalue of* A *such that*

$$Z(s) \;\; = \;\; I + (1 + qs)KG(s) \tag{3.2.3}$$

is strictly positive real. If (3.2.1) *and* (3.2.2) *are satisfied only on a set* $\Gamma \subset \mathbb{R}^p$ *then the same condition on* $Z(s)$ *ensures that the system is absolutely stable with a finite domain.*

Proof *Consider a Lyapunov function candidate of the Lure-type*

$$V(\vec{x}) \;=\; \vec{x}^T P\vec{x} + 2q\int_{\vec{0}}^{\vec{y}} \psi(\vec{s})K\,d\vec{s} \tag{3.2.4}$$

where $q \geq 0$ *is to be chosen. The derivative of* $V(\vec{x})$ *along the trajectories of the system* (3.0.1)-(3.0.3) *is given by*

$$\begin{aligned}\dot{V}(\vec{x}) &= \dot{\vec{x}}^T P\vec{x} + \vec{x}^T P\dot{\vec{x}} + 2q\frac{d}{dt}\left(\int_{\vec{0}}^{\vec{y}} \psi(\vec{s})^T K\,d\vec{s}\right) \qquad (3.2.5)\\ &= (A\vec{x}+B\vec{u})^T P\vec{x} + \vec{x}^T P(A\vec{x}+B\vec{u}) + 2q\frac{\partial}{\partial\vec{y}}\left(\int_{\vec{0}}^{\vec{y}} \psi(\vec{s})^T K\,d\vec{s}\right)\frac{d}{dt}\vec{y}\\ &= \vec{x}^T(PA+A^TP)\vec{x} - 2\vec{x}^T PB\psi(\vec{y}) + 2q\psi(\vec{y})^T KC[A\vec{x} - B\psi(\vec{y})].\end{aligned}$$

Similar to (3.1.3), *multiplying* -2 *to* (3.2.2) *yields*

$$-2\psi(\vec{y})^T\,[\psi(\vec{y}) - K\vec{y}] \;\geq\; 0, \tag{3.2.6}$$

and addition of $-2\psi^T[\psi - K\vec{y}]$ *to the right-hand side of* (3.2.5) *gives an upper bound on* $\dot{V}(\vec{x})$. *Therefore,*

$$\begin{aligned}\dot{V}(\vec{x}) &\leq \vec{x}^T(PA+A^TP)\vec{x} - 2\vec{x}^T PB\psi(\vec{y}) + 2q\psi(\vec{y})^T KC[A\vec{x} - B\psi(\vec{y})] \qquad (3.2.7)\\ &\quad -2\psi(\vec{y})^T\,[\psi(\vec{y}) - K\vec{y}]\\ &= \vec{x}^T(PA+A^TP)\vec{x} - 2\vec{x}^T(PB - qA^TC^TK - C^TK)\psi(\vec{y})\\ &\quad -\psi(\vec{y})^T(2I + qKCB + qB^TC^TK)\psi(\vec{y})\end{aligned}$$

where we have used the fact that the transpose of a scalar is itself. Choose q such that

$$2I + qKCB + qB^TC^TK \geq 0$$

which can be always done.[3] *Suppose there are matrices* $P = P^T > 0$, L *and* W

[3] By choosing q small enough, for example.

and a constant $\epsilon > 0$ *such that*

$$PA + A^T P = -L^T L - \epsilon P \tag{3.2.8}$$

$$PB = C^T K + qA^T C^T K - L^T W \tag{3.2.9}$$

$$W^T W = 2I + qKCB + qB^T C^T K. \tag{3.2.10}$$

Then,

$$\dot{V}(\vec{x}) \leq -\epsilon \vec{x}^T P \vec{x} - [L\vec{x} - W\psi(\vec{y})]^T [L\vec{x} - W\psi(\vec{y})] \leq -\epsilon \vec{x}^T P \vec{x} \tag{3.2.11}$$

which is negative definite. The question of existence of P, L, W *and* ϵ *satisfying* (3.2.8)-(3.2.10) *is obviously answered by Lemma 3.0.5. Consider the transfer function*

$$\begin{aligned} Z(s) &= I + qKCB + (KC + qKCA)(sI - A)^{-1} B \\ &= \mathcal{D} + \mathcal{C}(sI - \mathcal{A})^{-1} \mathcal{B} \end{aligned} \tag{3.2.12}$$

where

$$\mathcal{A} = A,\ \mathcal{B} = B,\ \mathcal{C} = KC + qKCA,\ \mathcal{D} = I + qKCB.$$

Note that

$$Z(\infty) + Z^T(\infty) = W^T W.$$

It is evident that if $-1/q$ *is not an eigenvalue of* A *or* $(1 + q\lambda_i) \neq 0$ *for all* $\lambda_i \in \sigma(A)$, *a set of all eigenvalues of the matrix* A. *Then, the pair* $(A, KC + qKCA)$ *is observable since* (A, C) *is observable. Thus, all the conditions of Lemma 3.0.5 are satisfied and we conclude that there are* P, L, W, *and* ϵ *satisfying* (3.2.8)-(3.2.10) *if and only if* $Z(s)$ *is strictly positive real. The expression of* $Z(s)$ *as in* (3.2.3) *is* *obtained from* (3.2.12), *with a little more manipulation as follows:*

$$\begin{aligned}
Z(s) &= I + qKCB + KC(sI-A)^{-1}B + qKCA(sI-A)^{-1}B \\
&= I + qKC(sI-A)(sI-A)^{-1}B + KC(sI-A)^{-1}B + qKCA(sI-A)^{-1}B \\
&= I + qKC(sI-A+A)(sI-A)^{-1}B + KC(sI-A)^{-1}B \\
&= I + (1+qs)KC(sI-A)^{-1}B \\
&= I + (1+qs)KG(s).
\end{aligned}$$

□

As we have done in the circle criterion, the restriction on A to be Hurwitz can be removed by performing loop transformation ([6], pp. 50 - 52). Note that with $q = 0$, Theorem 3.2.1 reduces to Lemma 3.1.1. This show that the conditions of Theorem 3.2.1 are weaker than those of Lemma 3.1.1. With $q > 0$, absolute stability can be established under less stringent conditions. In the scalar case $p = 1$, we can test the strict positive realness of $Z(s)$ graphically. Choose q such that $Z(\infty) > 0$. Then, by Lemma 3.0.4, $Z(s)$ is strictly positive real if and only if

$$\mathrm{Re}\,[1 + (1 + jq\omega)kG(j\omega)] \;>\; 0, \quad \forall\omega \in \mathbb{R}$$

which is equivalent to

$$\frac{1}{k} + \mathrm{Re}\,[G(j\omega)] - q\omega\,\mathrm{Im}\,[G(j\omega)] \;>\; 0, \quad \forall\omega \in \mathbb{R} \tag{3.2.13}$$

where $G(j\omega) = \mathrm{Re}\,[G(j\omega)] + j\,\mathrm{Im}\,[G(j\omega)]$. If we plot $\mathrm{Re}\,[G(j\omega)]$ versus $\omega\,\mathrm{Im}\,[G(j\omega)]$ with ω as a parameter, then condition (3.2.13) is satisfied if the plot lies to the right of the line that intercepts the point $-1/k + j0$ with a slope $1/q$; see Figure 3.3. Such a plot is known as a Popov plot.

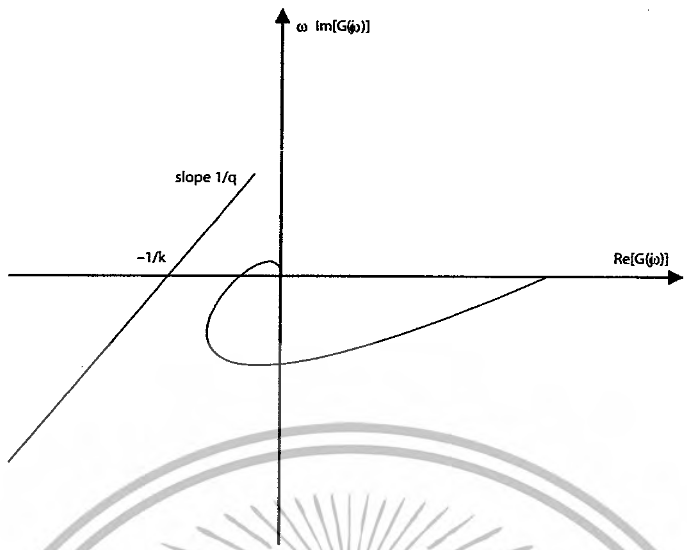

Figure 3.3: Popov plot.

3.3 Introduction to Nonlinear Stability

Nonlinear systems frequently have more than one equilibrium point (an equilibrium point is a point where the system can stay forever without moving, as we shall formalize later). This can be seen by the following simple example.

Example 3.3.1 *Consider the first order system*

$$\dot{x} = -x + x^2 \tag{3.3.1}$$

with initial condition $x(0) = x_o$. *Its linearization is*

$$\dot{x} = -x \tag{3.3.2}$$

The solution of this linear equation is $x(t) = x_o e^{-t}$. *It is plotted in Figure ?? for various initial conditions. The linearized system clearly has a unique equilibrium point at* $x = 0$. *By contrast, integrating equation* $dx/(-x + x^2) = dt$, *the actual response of the nonlinear dynamics* (3.3.1) *can be found to be*

$$x(t) = \frac{x_o e^{-t}}{1 - x_o + x_o e^{-t}}$$

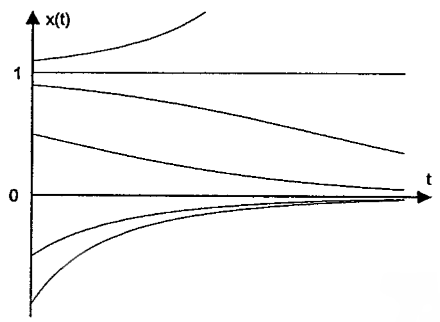

Figure 3.4: Response of the nonlinear system

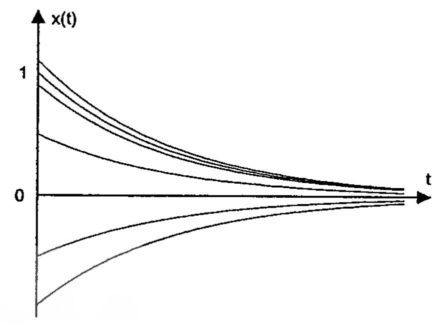

Figure 3.5: Response of the linearized system

This response is plotted in Figure 3.4 for various initial conditions, while the response of its linearized system is plotted in Figure 3.5. The system has two equilibrium points, $x = 0$ and $x = -1$, and its qualitative behavior strongly depends on its initial conditions.

The issue of motion stability can also be discussed with the aid of the above example. For the linearized system, stability is seen by nothing that for any initial condition, the motion always converges to the equilibrium point $x = 0$. However, consider now the actual nonlinear system. While motions starting with $x_o < 1$ will indeed converge to the equilibrium point $x = 0$, those starting with $x_o > 1$ will go to infinity (actually in finite time, a phenomenon known as finite escape time). This means that the stability of nonlinear systems may depend on initial conditions.

In the presence of a bounded external input, stability may also be dependent on the input value. This dependence is highlighted by the so-called bilinear system

$$\dot{x} = xu$$

If input u is chosen to be -1, then the state x converge to 0. If $u = 1$, then $|x|$ tends to infinity. □

3.4 Fundamentals of Lyapunov Theory

3.4.1 Nonlinear Systems and Equilibrium Points

Nonlinear Systems

A nonlinear dynamic system can usually be represented by a set of nonlinear differential equations in the form

$$\dot{\vec{x}} = \vec{f}(\vec{x}, t) \tag{3.4.1}$$

where $\vec{f}$ is an $n \times 1$ nonlinear vector function, and $\vec{x}$ is the $n \times 1$ state vector. A particular value of the state vector is also called a point because it corresponds to a point in the state-space. The number of states n is called the *order* of the system. A solution $\vec{x}(t)$ of the equation (3.4.1) usually corresponds to a curve in state space as t varies from zero to infinity, as already seen in phase plane analysis for the case $n = 2$. This curve is generally referred to as a *state trajectory* or a *system trajectory*.

It is important to note that although equation (3.4.1) does not explicitly contain the control input as a variable, it is directly applicable to feedback control systems. The reason is that equation (3.4.1) can represent the *closed-loop* dynamics of a feedback control system, with the control input being a function of state $\vec{x}$ and time t, and therefore disappearing in the closed-loop dynamics. Specifically, if the plant dynamics are

$$\dot{\vec{x}} = \vec{f}(\vec{x}, \vec{u}, t)$$

and some control law has been selected

$$\vec{u} = \vec{g}(\vec{x}, t)$$

then the closed-loop dynamics is

$$\dot{\vec{x}} = \vec{f}(\vec{x}, \vec{g}(\vec{x}, t), t)$$

which can be rewritten in the form (3.4.1). Of course, equation (3.4.1) can also represent dynamic systems where no control signals are involved, such as a freely swinging pendulum.

A special class of nonlinear systems is that of *linear* systems. The dynamics of linear systems are of the form

$$\dot{\vec{x}} = A(t)\vec{x}$$

where $A(t)$ is an $n \times n$ matrix.

AUTONOMOUS AND NON-AUTONOMOUS SYSTEMS

Definition 3.4.1 *The nonlinear system (3.4.1) is said to be* **autonomous** *if $\vec{f}$ does not depend explicitly on time, i.e., if the system's state equation can be written*

$$\dot{\vec{x}} = \vec{f}(\vec{x}). \tag{3.4.2}$$

Otherwise, the system is called **non-autonomous**. □

A time-invariant plant with dynamics

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

may lead to a non-autonomous closed-loop system if a controller dependent on time t is chosen, i.e., if $\vec{u} = \vec{g}(\vec{x}, t)$.

EQUILIBRIUM POINTS

It is possible for a system trajectory to correspond to only a single point. Such a point is called an equilibrium point. As we shall see later, many stability problems are naturally formulated with respect to equilibrium points.

Definition 3.4.2 *A state $\vec{x}^e$ is an* **equilibrium state** *(or* **equilibrium point***) of the system if once $\vec{x}(t)$ is equal to $\vec{x}^e$, it remains equal to $\vec{x}^e$ for all future time.* □

Mathematically, this means that the constant vector $\vec{x}^e$ satisfies

$$\vec{0} = \vec{f}(\vec{x}^e). \tag{3.4.3}$$

A linear time-invariant system

$$\dot{\vec{x}} = A\vec{x} \tag{3.4.4}$$

has a single equilibrium point (the origin $\vec{0}$) if A is nonsingular. If A is singular, it has infinitely many equilibrium points, which are contained in the null-space of the matrix A, i.e., the subspace defined by $A\vec{x} = \vec{0}$. This implies that the equilibrium points are not isolated, as reflected by the example $\ddot{x} + \dot{x} = 0$, for which all points on the x axis of the phase plane are equilibrium points.

A nonlinear system can have several (or infinitely many) isolated equilibrium points. The following example involves a familiar physical system.

Example 3.4.3 *Consider the pendulum of the below Figure, whose dynamics are given by the following nonlinear autonomous equation*

$$MR^2\ddot{\theta} + k\dot{\theta} + MgR\sin\theta = 0 \tag{3.4.5}$$

where k is a friction coefficient, M the mass, R the pendulum length, and g the gravity constant. Letting $x_1 = \theta, x_2 = \dot{\theta}$, the corresponding state-space equation is

$$\dot{x}_1 = x_2 \tag{3.4.6}$$

$$\dot{x}_2 = -\frac{k}{MR^2}x_2 - \frac{g}{R}\sin x_1 \tag{3.4.7}$$

Therefore, the equilibrium points are given by

$$x_2 = 0, \sin x_1 = 0$$

which leads to the points $(0[2\pi], 0)$ and $(\pi[2\pi], 0)$. Physically, these points correspond to the pendulum resting exactly at the vertical up and down positions. □

In linear system analysis and design, for notational and analytical simplicity, we often transform the linear system equation in such a way that the equilibrium point of interest is $\vec{x}^e$. Then, by introducing a new variable

$$\vec{y} = \vec{x} - \vec{x}^e$$

and substituting $\vec{x} = \vec{y} + \vec{x}^e$ into the equation (3.4.2), a new set of equations on the variable $\vec{y}$ are obtained

$$\dot{\vec{y}} = \vec{f}(\vec{y} + \vec{x}^e) \tag{3.4.8}$$

One can easily verify that there is a one-to-one correspondence between the solution of (3.4.2) and those of (3.4.8), and that in addition, $\vec{y} = \vec{0}$, the solution corresponding to $\vec{x} = \vec{x}^e$, is an equilibrium point of (3.4.8). Therefore, instead of studying the behavior of the equation (3.4.2) in the neighborhood of $\vec{x}^e$, one can equivalently study the behavior of the equations (3.4.8) in the neighborhood of the origin.

NOMINAL MOTION

In some practical problems, we are not concerned with stability around an equilibrium point, but rather with the stability of a *motion*, i.e., whether a system will remain close to its original motion trajectory if slightly perturbed away from it, as exemplified by the aircraft trajectory control problem. We can show that this kind of motion stability problem can be transformed into an equivalent stability problem around an equilibrium point, although the equivalent system is now non-autonomous.

3.4.2 Concepts of Stability

In the beginning of this chapter, we introduced the intuitive notion of stability as a kind of well-behavedness around a desired operating point. However, since nonlinear systems may have much more complex and exotic behavior than linear systems, the mere notion of stability is not enough to describe the essential features of their motion. A number of more refined stability concepts, such as asymptotic stability, exponential stability and global asymptotic stability, are needed. In this section, we define these stability concepts formally, for autonomous systems, and explain their practical meanings.

A few simplifying notions are defined at this point. Let B_R denote the spherical region (or ball) defined by $\|\vec{x}\| < R$ in state-space, and S_R the sphere itself, defined by $\|\vec{x}\| = R$.

STABILITY AND INSTABILITY

Let us see the basic concepts of stability and instability.

Definition 3.4.4 *The equilibrium state $\vec{x} = \vec{0}$ is said to be* **stable** *if, for any $R > 0$, there exists $r > 0$, such that if $\|\vec{x}(0)\| < r$, then $\|\vec{x}(t)\| < R$ for all $t \geq 0$. Otherwise, the equilibrium point is* **unstable**. □

Essentially, stability (also called **stability in the sense of Lyapunov**, or **Lyapunov stability**) means that the system trajectory can be kept arbitrarily close to the origin by starting sufficiently close to it. More formally, the definition states that the origin is stable, if, given that we do not want the state trajectory $\vec{x}(t)$ to get out of a ball of arbitrarily specified radius B_R, a value $r(R)$ can be found such that starting the state from within the ball B_r at time 0 guarantees that the state will stay within the ball B_R thereafter. The geometrical implication of stability is indicated by curve 2 in Figure 3.7 ([54]).

Using mathematical symbols Definition 3.4.4 can be written

$$\forall R > 0, \exists r > 0, \|\vec{x}(0)\| < r \Rightarrow \forall t \geq 0, \|\vec{x}(t)\| < R$$

or, equivalently

$$\forall R > 0, \exists r > 0, \vec{x}(0) \in B_r \Rightarrow \forall t \geq 0, \vec{x}(t) \in B_R$$

Conversely, an equilibrium point is unstable if there exists at least **one** ball B_R, such that for **every** $r > 0$, no matter how small, it is always possible for the system trajectory to start somewhere within the ball B_r and eventually leave the ball B_R (Figure 3.7 [54]). Unstable nodes or saddle points in second-order systems are examples of unstable equilibria. Instability of an equilibrium point is typically undesirable, because it often leads the system into limit cycles or results in damage to the involved mechanical or electrical components.

It is important to point out the qualitative difference between instability and the intuitive notion of "blowing up" (all trajectories close to origin move further and further away to infinity). In linear systems, instability is equivalent to blowing up, because unstable poles always lead to exponential growth of the system states. However, for nonlinear systems, blowing up in only one way of instability. The following example illustrates this point.

Example 3.4.5 (Instability of the Van der Pol Oscillator) *The Van der Pol oscillator is described by*

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1 + (1 - x_1^2)x_2 \end{aligned}$$

The system has an equilibrium point at the origin.

System trajectories starting from any non-zero initial states all asymptotically approach a limit cycle. Furthermore, the ball B_1 ($R = 1$) is easily shown to be within the phase-plane region enclosed by the limit cycle. Therefore, system trajectories starting from an arbitrarily small ball B_r will eventually get out of

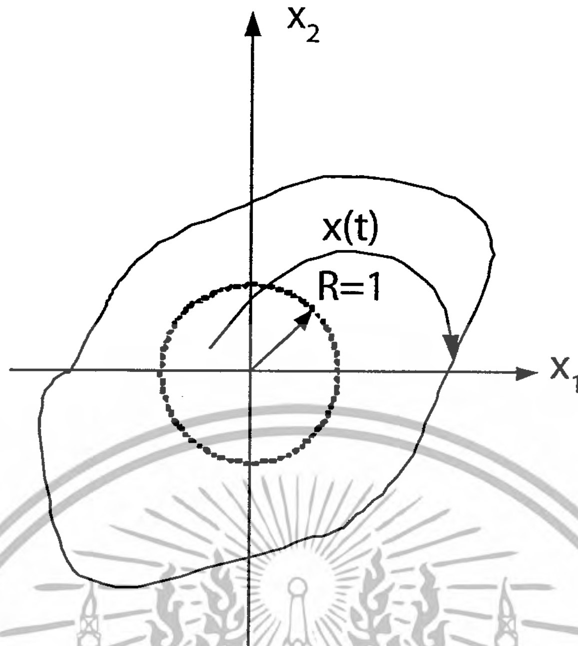

Figure 3.6: Unstable origin of the Van der Pol Oscillator

the ball B_1 to approach the limit cycle Figure 3.9. This implies instability of the origin.

Thus, even though the state of the system does remain around the equilibrium point in the certain sense, it cannot stay arbitrarily close to it. This is the fundamental distinction between stability and instability. □

ASYMPTOTIC STABILITY AND EXPONENTIAL STABILITY

In many engineering applications, Lyapunov stability is not enough. For example when a satellite's attitude is disturbed from it nominal position, we not only want the satellite to maintain its attitude in a range determined by the magnitude of the disturbance, i.e., Lyapunov stability, but also require that the attitude gradually go back to its original value. This type of engineering requirement is captured by the concept of asymptotic stability.

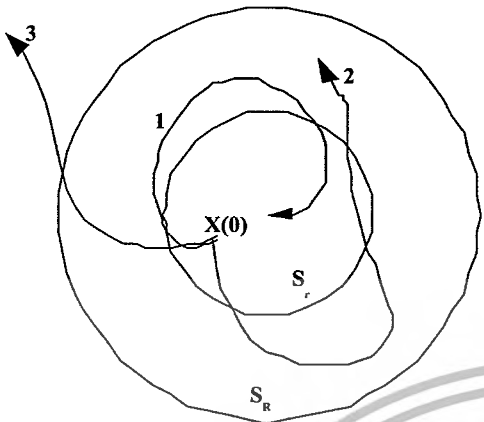

Figure 3.7: Concepts of stability

Definition 3.4.6 *An equilibrium point* $\vec{0}$ *is* **asymptotically stable** *if it is stable, and if in addition there exists some* $r > 0$ *such that* $\|\vec{x}(0)\| < r$ *implies that* $\vec{x}(t) \to \vec{0}$ *as* $t \to \infty$. □

Asymptotic stability means that the equilibrium is stable, and that in addition, states started close to $\vec{0}$ actually converge to $\vec{0}$ as time t goes to infinity.

Figure 3.10 shows that the system trajectories starting from within the ball B_r converge to the origin. The ball B_r is called a **domain of attraction** of the equilibrium point (while **the** domain of attraction of the equilibrium point refers to the largest such region, i.e., to the set of all points such that trajectories initiated at these points eventually converge to the origin). An equilibrium point which is Lyapunov stable but not asymptotically stable is called **marginally stable**.

One may question the need for the explicit stability requirement in the definition above, in view of the second condition of state convergence to the origin. However, it is easy to build counter-example that show that state convergence does not necessarily imply stability. For instance, a simple system studied by Vinograd has trajectories of the form shown in Figure 3.8.

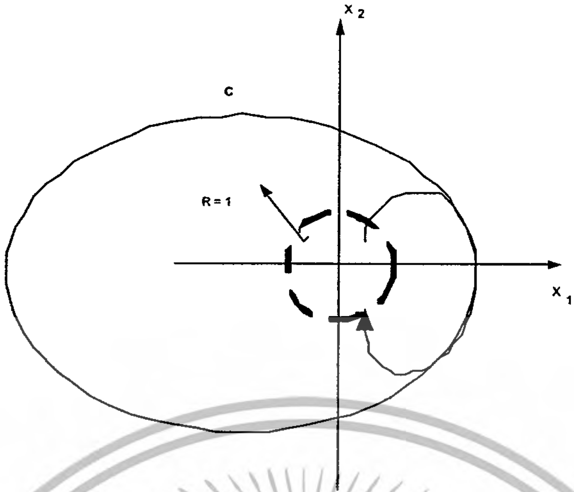

Figure 3.8: State convergence does not imply stability

All the trajectories starting from non-zero initial points within the unit disk first reach the curve C before converging to the origin. Thus, the origin is **unstable** in the sense of Lyapunov, despite the state convergence. Calling such a system unstable is quite reasonable, since a curve such as C may be outside the region where the model is valid - for instance, the subsonic and super sonic dynamics of a high-performance aircraft are radically different, while, with the problem under study using subsonic dynamic models, C could be in the supersonic range.

In many engineering applications, it is still not sufficient to know that a system will converge to the equilibrium point after infinite time. There is a need to estimate **how fast** the system trajectory approaches $\vec{0}$. The concept of **exponential stability** can be used for this purpose.

Definition 3.4.7 (Exponential Stability) *An equilibrium point* $\vec{0}$ *is* **exponentially stable** *if there exist two strictly positive numbers* α *and* λ *such that*

$$\forall t > 0,\ \|\vec{x}(t)\| \leq \alpha \|\vec{x}(0)\| e^{-\lambda t} \tag{3.4.9}$$

in some ball B_r *around the origin.* □

In words, (3.4.9) means that the state vector of an exponentially stable system converges to the origin faster than an exponential function. The positive number λ is often called the **rate** of exponential convergence. For instance, the system

$$\dot{x} = -(1 + \sin^2 x)x$$

is exponentially convergent to $x = 0$ with a rate $\lambda = 1$. Indeed, its solution is

$$x(t) = x(0)e^{-\int_0^t \left[1+\sin^2(x(\tau))\right] d\tau}$$

and therefore

$$|x(t)| \leq |x(0)|\, e^{-t}$$

Note that exponential stability implies asymptotic stability. But asymptotic stability does not guarantee exponential stability, as can be seen from the system

$$\dot{x} = -x^2, \qquad x(0) = 1 \tag{3.4.10}$$

whose solution is $x = \dfrac{1}{1+t}$, a function slower than any exponential function $e^{-\lambda t}$ (with $\lambda > 0$).

The definition of exponential convergence provides an explicit bound on the state at any time, as seen in (3.4.9). By writing the positive constant α as $\alpha = e^{\lambda \tau_0}$, it is easy to see that, after a time of $\tau_0 + \dfrac{1}{\lambda}$, the magnitude of the state vector decreases to less than 35% ($\approx e^{-1}$) of its original value (at $t = 0$), similarly to the notion of time-constant in a linear system. After $\tau_0 + \dfrac{3}{\lambda}$, the state magnitude $\|\vec{x}(t)\|$ will be less than 5% ($\approx e^{-3}$) of $\|\vec{x}(0)\|$.

LOCAL AND GLOBAL STABILITY

The above definitions are formulated to characterize the **local** behavior of systems, i.e., how the state evolves after starting near the equilibrium point. Local properties tell little about how the system will behave when the initial state is some distance away from the equilibrium as seen for the nonlinear system in Example 3.4.3. Global concepts are required for this purpose.

Definition 3.4.8 *If asymptotic (or exponential) stability holds for any initial states, the equilibrium point is said to be asymptotically (or exponentially) stable* **in the large**, *it is also called* **globally** *asymptotically (or exponentially) stable.*

□

For instance, in Example 3.3.1 the linearized system is globally asymptotically stable, but the original system is not. The simple system in (3.4.10) is also global asymptotically stable, as can be seen from its solutions.

Linear time-invariant systems are either asymptotically stable, of marginally stable, or unstable, as can be seen from the modal decomposition of linear system solutions; linear asymptotic stability is always global and exponential, and linear instability always implies exponential blow-up. **This explains why the refined notion of stability introduced here were not previously encountered in the study of linear systems.** They are explicitly needed only for nonlinear systems.

3.4.3 Linearization and Local Stability

Lyapunov's linearization method is concerned with the **local** stability of a nonlinear system. It is a formalization of the intuition that a nonlinear system should behave similarly to its linearized approximation for small range motions. Because all physical systems are inherently nonlinear, Lyapunov's linearization method serves as the fundamental **justification of using linear control techniques** in practice, i.e., shows that stable design by linear control guarantees the stability of the original physical system locally.

Consider the autonomous system in (3.4.2), and assume that $\vec{f}(\vec{x})$ is continuously differentiable. Then the system dynamics can be written as

$$\dot{\vec{x}} = \left(\frac{\partial \vec{f}}{\partial \vec{x}}\right)_{\vec{x}=\vec{0}} \vec{x} + \vec{f}_{h.o.t}(\vec{x}) \tag{3.4.11}$$

where $\vec{f}_{h.o.t}$ stands for higher-order terms in $\vec{x}$. Note that the above Taylor expansion starts directly with the first-order term, due to the fact that $\vec{f}(\vec{0}) = \vec{0}$, since $\vec{0}$ is an equilibrium point. Let us use the constant matrix A to denote the Jacobian matrix of $\vec{f}$ with respect to $\vec{x}$ at $\vec{x} = \vec{0}$ (an $n \times n$ matrix of elements $\partial f_i / \partial x_j$)

$$A = \left(\frac{\partial \vec{f}}{\partial \vec{x}}\right)_{\vec{x}=\vec{0}}.$$

Then, the system

$$\dot{\vec{x}} = A\vec{x} \tag{3.4.12}$$

is called the **linearization** (or **linear approximation**) of the original nonlinear system at the equilibrium point $\vec{0}$.

Note that, similarly, starting with an autonomous nonlinear system with a control input $\vec{u}$

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

such that $\vec{f}(\vec{0}, \vec{0}) = \vec{0}$, we can write

$$\dot{\vec{x}} = \left(\frac{\partial \vec{f}}{\partial \vec{x}}\right)_{\vec{x}=\vec{0},\vec{u}=\vec{0}} \vec{x} + \left(\frac{\partial \vec{f}}{\partial \vec{u}}\right)_{\vec{x}=\vec{0},\vec{u}=\vec{0}} \vec{u} + \vec{f}_{h.o.t}(\vec{x}, \vec{u})$$

where $\vec{f}_{h.o.t}$ stands for higher-order terms in $\vec{x}$ and $\vec{u}$. Letting A denote the Jacobian matrix of $\vec{f}$ with respect to $\vec{x}$ at $(\vec{x} = \vec{0}, \vec{u} = \vec{0})$, and B denote the Jacobian matrix of $\vec{f}$ with respect to $\vec{u}$ at the same point (an $n \times m$ matrix of elements $\partial f_i / \partial u_j$, where m is the number of inputs)

$$A = \left(\frac{\partial \vec{f}}{\partial \vec{x}}\right)_{\vec{x}=\vec{0},\vec{u}=\vec{0}} \qquad B = \left(\frac{\partial \vec{f}}{\partial \vec{u}}\right)_{\vec{x}=\vec{0},\vec{u}=\vec{0}}$$

the system

$$\dot{\vec{x}} = A\vec{x} + B\vec{u}$$

is the linearization (or linear approximation) of the originally nonlinear system at $(\vec{x} = \vec{0}, \vec{u} = \vec{0})$.

Furthermore, the choice of a control law of the form $\vec{u} = \vec{u}(\vec{x})$ (with $\vec{u}(\vec{0}) = \vec{0}$) transforms the original system into an autonomous closed-loop system, having $\vec{x} = \vec{0}$ as an equilibrium point. Linearly approximating the control law as

$$\vec{u} = \left(\frac{d\vec{u}}{d\vec{x}}\right)_{\vec{x}=\vec{0}} \vec{x} = G\vec{x}$$

the closed-loop dynamics can be linearly approximated as

$$\dot{\vec{x}} = \vec{f}(\vec{x}, \vec{u}(\vec{x})) = (A + BG)\vec{x}$$

Of course, the same linear approximation can be obtained by directly considering the autonomous closed-loop system

$$\dot{\vec{x}} = \vec{f}(\vec{x}, \vec{u}(\vec{x})) = \vec{f}_1(\vec{x})$$

and linearizing the function $\vec{f}_1$ with respect to $\vec{x}$, at its equilibrium point $\vec{x} = \vec{0}$.

In practice, finding a system's linearization is often most easily done simply by neglecting any term of order higher than 1 in the dynamics, as we now illustrate.

Example 3.4.9 *Consider the system*

$$\dot{x}_1 = x_2^2 + x_1 \cos x_2$$

$$\dot{x}_2 = x_2 + (x_1 + 1)x_1 + x_1 \sin x_2$$

Its linearized approximation about $\vec{x} = \vec{0}$ is

$$\dot{x}_1 \approx 0 + x_1 \cdot 1 = x_1$$

$$\dot{x}_2 \approx x_2 + 0 + x_1 + x_1 x_2 \approx x_2 + x_1$$

The linearized system can thus be written

$$\dot{\vec{x}} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \vec{x}.$$

A similar procedure can be applied for a controlled system. Consider the system

$$\ddot{x} + 4\dot{x}^5 + (x^2+1)u = 0$$

The system can be linearly approximated about $\vec{x} = \vec{0}$ as

$$\ddot{x} + 0 + (0+1)u \approx 0$$

i.e., the linearized system can be written

$$\ddot{x} = -u.$$

Assume that the control law for the original nonlinear system has been selected to be

$$u = \sin x + x^3 + \dot{x}\cos^2 x$$

then the linearized closed-loop dynamics is

$$\ddot{x} + \dot{x} + x = 0$$

□

The following result makes precise the relationship between the stability of the linear system (3.4.12) and that of the original nonlinear system (3.4.2).

Theorem 3.4.10 (Lyapunov's Linearization Method)

- *If the linearized system is strictly stable (i.e., if all eigenvalues of A are strictly in the left-half complex plane), then the equilibrium point is asymptotically stable (for the actual nonlinear system).*

- *If the linearized system is unstable (i.e., if at least one eigenvalue of A is strictly in the right-half complex plane), then the equilibrium point is unstable (for the nonlinear system).*

- *If the linearized system is marginally stable (i.e., all eigenvalues of A are in the left-half complex plane, but at least one of them is on the $j\omega$ axis), then one cannot conclude anything from the linear approximation (the equilibrium point may be stable, asymptotically stable, or unstable for the nonlinear system).* □

A summary of the theorem is that it is true by **continuity**. If the linearized system is strictly stable, or strictly unstable, then since the approximation is valid "not too far" from the equilibrium, the nonlinear system itself is locally stable, or locally unstable. However, if the linearized system is marginally stable, the higher-order terms in (3.4.11) can have a decisive effect on whether the nonlinear system is stable or unstable. As we shall see in the next section, simple nonlinear systems may be globally asymptotically stable while their linear approximations are only marginally stable: one simply cannot infer any stability property of a nonlinear system from its marginally stable linear approximation.

Example 3.4.11 *As expected, it can be shown easily that the equilibrium states ($x_1 = \pi[2\pi], \dot{x}_1 = 0$) of the pendulum of Example 3.4.3 are unstable. Indeed, first note that, in a neighborhood of $x_1 = \pi$*

$$\sin x_1 = \sin\pi + (\pi - x_1) + h.o.t. = (\pi - x_1) + h.o.t..$$

Thus, the matrix A corresponding to the system's linearization in that neighborhood is

$$A = \begin{bmatrix} 0 & 1 \\ a_2 & -a_1 \end{bmatrix}.$$

Since $a_1 > 0$ and $a_2 > 0$, the linear approximation is unstable, and therefore so is the nonlinear system at that equilibrium point. □

Example 3.4.12 *Consider the first order system*

$$\dot{x} = ax + bx^5.$$

The origin 0 *is one of the two equilibrium points of this system. The linearization of this system around the origin is*

$$\dot{x} = ax$$

The application of Lyapunov's linearization method indicates the following stability properties of the nonlinear system

- $a < 0$: *asymptotically stable;*
- $a > 0$: *unstable;*
- $a = 0$: *cannot tell from linearization.*

In the third case, the nonlinear system becomes

$$\dot{x} = bx^5$$

The linearization method fails while, as we shall see, the direct method to be described can easily solve this problem. □

Lyapunov's linearization theorem shows that linear control design is a matter of consistency: one must design a controller such that the system remain in its "linear range". It also stresses major limitations of linear design: how large is the linear range? What is the **extent** of stability (how large is r in Definition 3.4.4? These questions motivate a deeper approach to the nonlinear control problem, Lyapunov's direct method.

3.5 Lyapunov's Direct Method

The basic philosophy of Lyapunov's direct method is the mathematical extension of a fundamental physical observation: if the total **energy** of a mechanical (or electrical) system is continuously dissipated, then the system, **whether linear or nonlinear** must eventually settle down to an equilibrium point. Thus, we may conclude the stability of a system by examining the variation of a single **scalar** function.

Specifically, let us consider the nonlinear mass-damper-spring system in Figure 3.9 ([54]), whose dynamic equation is

$$m\ddot{x} + b\dot{x}\,|\dot{x}| + k_0 x + k_1 x^3 = 0 \tag{3.5.1}$$

with $b\dot{x}\,|\dot{x}|$ representing nonlinear dissipation or damping, and $(k_0 x + k_1 x^3)$ representing a nonlinear spring term. Assume that the mass is pulled away from the natural length of the spring by a large distance, and then released. Will the resulting motion be stable? It is very difficult to answer this question using the definitions of stability, because the general solution of this nonlinear equation is unavailable. The linearization method cannot be used either because the motion starts outside the linear range (and in any case the system's linear approximation is only marginally stable). However, examination of the system energy can tell us a lot about the motion pattern. The total mechanical energy of the system is the sum of its kinetic energy and its potential energy

$$V(\vec{x}) = \frac{1}{2}m\dot{x}^2 + \int_0^x (k_0 y + k_1 y^3)\,dy = \frac{1}{2}m\dot{x}^2 + \frac{1}{2}k_0 x^2 + \frac{1}{4}k_1 x^4 \tag{3.5.2}$$

Comparing the definitions of stability and mechanical energy, one can easily see some relations between the mechanical energy and the stability concepts

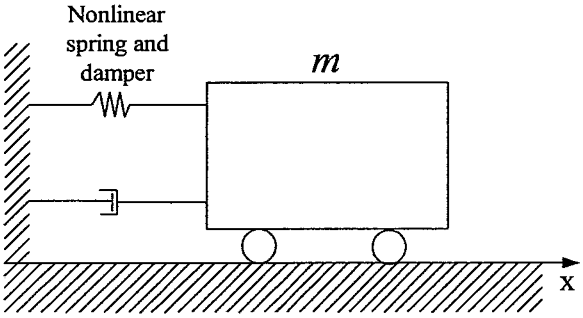

Figure 3.9: A nonlinear mass-damper-spring-system.

described earlier:

- zero energy corresponds to the equilibrium point ($x = 0, \dot{x} = 0$).
- asymptotic stability implies the convergence of mechanical energy to zero.
- instability is related to the growth of mechanical energy.

These relations indicate that the value of a scalar quantity, the mechanical energy, indirectly reflects the magnitude of the state vector; and furthermore, that the stability properties of the system can be characterized by the variation of the mechanical energy of the system.

The rate of energy variation during the system's motion is obtained easily by differentiating the first equality in (3.5.2) and using (3.5.1)

$$\dot{V}(\vec{x}) = m\dot{x}\ddot{x} + (k_0x + k_1x^3)\dot{x} = \dot{x}(-b\dot{x}\,|\dot{x}|) = -b\,|\dot{x}|^3 \tag{3.5.3}$$

Equation (3.5.3) implies that the energy of the system, starting from some initial value is continuously dissipated by the damper until the mass settles down, i.e., until $\dot{x} = 0$. Physically, it is easy to see that the mass must finally settle down at the natural length of the spring, because it is subjected to a non-zero spring force at any position other than the natural length.

The direct method of Lyapunov is based on a generalization of the concepts in the above mass-spring-damper system to more complex systems. **Faced with a set of nonlinear differential equations, the basic procedure of Lyapunov's direct method is to generate a scalar "energy-like" function for the dynamic system, and examine the time variation of that scalar function. In this way, conclusions may be drawn on the stability of the set of differential equations without using the difficult stability definitions or requiring explicit knowledge of solutions.**

3.5.1 Positive Definite Function and Lyapunov Functions

The energy function in (3.5.2) has two properties. The first is a property of the function itself: it is strictly positive unless both state variables x and $\dot{x}$ are zero. The second property is a property associated with the dynamics (3.5.1): the function is monotonically decreasing when the variables x and $\dot{x}$ vary according to (3.5.1). **In the Lyapunov's direct method, the first property is formalized by the notion of positive definite functions, and the second is formalized by the so-called Lyapunov functions.** Let us discuss positive definite functions first.

Definition 3.5.1 *A scalar continuous function $V(\vec{x})$ is said to be* **locally positive definite** *if $V(\vec{0}) = 0$ and, in a ball B_{R_0}*

$$\vec{x} \neq \vec{0} \Rightarrow V(\vec{x}) > 0.$$

If $V(\vec{0}) = 0$ and the above property holds over the whole state space, then $V(\vec{x})$ is said to be **globally positive definite.** □

For instance, the function

$$V(\vec{x}) = \frac{1}{2}MR^2x_2^2 + MR(1 - \cos x_1)$$

which is the mechanical energy of the pendulum of Example 3.4.3, is locally positive definite. The mechanical energy (3.5.2) of the nonlinear mass-damper-spring system is globally positive definite. Note that, for that system, the kinetic energy $\frac{1}{2}m\dot{x}^2$ is **not** positive definite by itself, because it can equal zero for non-zero value of x.

The above definition implies that the function V has a unique minimum at the origin $\vec{0}$. Actually, given any function having a **unique** minimum in a certain ball, we can construct a locally positive definite function simply by adding a constant to that function. For example, the function $V(\vec{x}) = x_1^2 + x_2^2 - 1$ is a lower bounded function with a unique minimum at the origin, and the addition of the constant 1 to it makes it a positive definite function. Of course, the function shifted by a constant has the same time-derivative as the original function.

Let us describe the geometrical meaning of locally positive definite functions. Consider a positive definite function $V(\vec{x})$ of two state variable x_1 and x_2. Plotted in a 3-dimensional space, $V(\vec{x})$ typically corresponds to a surface looking like an upward cup (Figure 3.10 [54]). The lowest point of the cup is located at the origin.

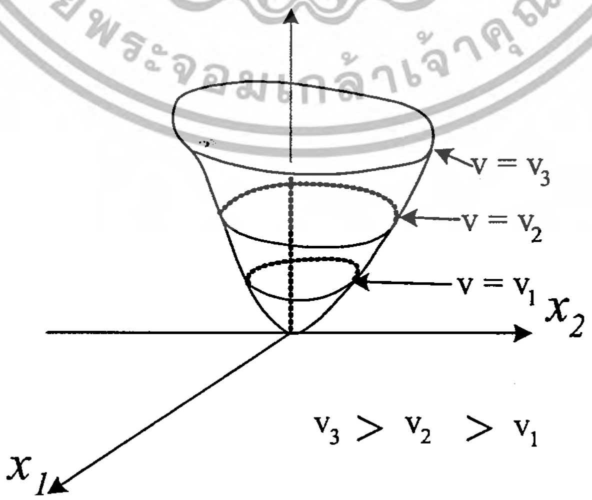

Figure 3.10: Typical shape of a positive definite function $V(x_1, x_2)$.

A few related concepts can be defined similarly, in a local or global sense, i.e., a function $V(\vec{x})$ is **negative definite** if $-V(\vec{x})$ is positive definite; $V(\vec{x})$ is **positive semi-definite** if $V(\vec{0}) = 0$ and $V(\vec{x}) \geq 0$ for $\vec{x} \neq \vec{0}$; $V(\vec{x})$ is **negative semi-definite** if $-V(\vec{x})$ is positive semi-definite. The prefix "semi" is used to reflect the possibility of V being equal to zero for $\vec{x} \neq \vec{0}$. These concepts can be given geometrical meanings similar to the ones given for positive definite functions.

With $\vec{x}$ denoting the state of the system (3.4.2), a scalar function $V(\vec{x})$ actually represents an implicit function of time t. Assuming that $V(\vec{x})$ is differentiable, its derivative with respect to time can be found by the chain rule,

$$\dot{V} = \frac{dV(\vec{x})}{dt} = \frac{\partial V}{\partial \vec{x}}\dot{\vec{x}} = \frac{\partial V}{\partial \vec{x}}\vec{f}(\vec{x})$$

We see that, because $\vec{x}$ is required to satisfy the autonomous state equations (3.4.2), $\dot{V}$ only depends on $\vec{x}$. It is often referred to as "the derivative of V along the system trajectory". For the system (3.5.1), $\dot{V}(\vec{x})$ is computed in (3.5.3) and found to be negative. Functions such as V in that example are given a special name because of their importance in Lyapunov's direct method.

Definition 3.5.2 *If, in a ball B_{R_0}, the function $V(\vec{x})$ is positive definite and has continuous partial derivatives, and if its time derivative along any state trajectory of system* (3.4.2) *is negative semi-definite, i.e.,*

$$\dot{V}(\vec{x}) \leq 0$$

Then $V(\vec{x})$ is said to be a **Lyapunov function** *for the system* (3.4.2). □

3.5.2 Equilibrium Point Theorems

The relations between Lyapunov functions and the stability of systems are made precise in a number of theorems in Lyapunov's direct method. Such theorems usually have local and global versions. The local versions are concerned with stability properties in the neighborhood of equilibrium point and usually involve a locally positive definite function.

LYAPUNOV THEOREM FOR LOCAL STABILITY

Theorem 3.5.3 (Local Stability) *If, in a ball B_{R_0}, there exists a scalar function $V(\vec{x})$ with continuous first partial derivatives such that*

- *$V(\vec{x})$ is positive definite (locally in B_{R_0})*
- *$\dot{V}(\vec{x})$ is negative semi-definite (locally in B_{R_0})*

then the equilibrium point $\vec{0}$ is stable. If, actually, the derivative $\dot{V}(\vec{x})$ is locally negative definite in B_{R_0}, then the stability is asymptotic. □

In applying the above theorem for analysis of a nonlinear system, one goes through the two steps of choosing a positive definite function, and then determining its derivative along the path of the nonlinear systems. The following example illustrates this procedure.

Example 3.5.4 (Local Stability) *A simple pendulum with viscous damping is described by*

$$\ddot{\theta} + \dot{\theta} + \sin\theta = 0.$$

Consider the following scalar function

$$V(\vec{x}) = (1 - \cos\theta) + \frac{\dot{\theta}^2}{2}.$$

One easily verifies that this function is locally positive definite. As a matter of fact, this function represents the total energy of the pendulum, composed of the sum of the potential energy and the kinetic energy. Its time-derivative is easily found to be

$$\dot{v}(\vec{x}) = \dot{\theta}\sin\theta + \dot{\theta}\ddot{\theta} = -\dot{\theta}^2 \leq 0.$$

Therefore, by invoking the above theorem, one concludes that the origin is a stable equilibrium point. In fact, using physical insight, one easily sees the reason why $\dot{V}(\vec{x}) \leq 0$*, namely that the damping term absorbs energy. Actually,* $\dot{V}$ *is precisely the power dissipated in the pendulum. However, with this Lyapunov function, one cannot draw conclusions on the asymptotic stability of the system, because* $\dot{V}(\vec{x})$ *is only negative semi-definite.* □

The following example illustrates the asymptotic stability result.

Example 3.5.5 (Asymptotic stability) *Let us study the stability of the non-linear system defined by*

$$\dot{x}_1 = x_1(x_1^2 + x_2^2 - 1) - x_2$$
$$\dot{x}_2 = x_1 + x_2(x_1^2 + x_2^2 - 1)$$

around its equilibrium point at the origin. Given the positive definite function

$$V(x_1, x_2) = x_1^2 + x_2^2$$

its derivative along any system trajectory is

$$\dot{V} = 2(x_1^2 + x_2^2)(x_1^2 + x_2^2 - 1).$$

Thus, $\dot{V}$ *is locally negative definite in the 2-dimensional ball* B_1*, i.e., in the region defined by* $x_1^2 + x_2^2 < 1$*. Therefore, the above theorem indicates that the origin is asymptotically stable.* □

LYAPUNOV THEOREM FOR GLOBAL STABILITY

The above theorem applies to the local analysis of stability. In order to assert **global asymptotic stability** of a system, one might naturally expect that the ball B_{R_0} in the above local theorem has to be expanded to be the whole state-space. This is indeed necessary, but it is not enough. An additional condition on the function V has to be satisfied: $V(\vec{x})$ must be **radially unbounded**, by which we mean that $V(\vec{x}) \to \infty$ as $\|\vec{x}\| \to \infty$ (in other words, as $\vec{x}$ tends to infinity in **any** direction). We then obtain the following powerful result:

Theorem 3.5.6 (Global Stability) *Assume that there exists a scalar function V of the state $\vec{x}$, with continuous first order derivatives such that*

- *$V(\vec{x})$ is positive definite*
- *$\dot{V}(\vec{x})$ is negative definite*
- *$V(\vec{x}) \to \infty$ as $\|\vec{x}\| \to \infty$*

then the equilibrium at the origin is globally asymptotically stable. □

The reason for the radial unboundedness condition is to assure that the contour curves (or contour surfaces in the case of higher order systems) $V(\vec{x}) = V_\alpha$ correspond to closed curves. If the curves are not closed, it is possible for the state trajectories to drift away from the equilibrium point, even though the state keeps going through contours corresponding to smaller and smaller $V_\alpha' s$. For example, for the positive definite function $V = \dfrac{x_1^2}{1+x_1^2} + x_2^2$, the curves $V(\vec{x}) = V_\alpha$ for $V_\alpha > 1$ are open curves.

Example 3.5.7 (A class of first-order systems) *Consider the nonlinear system*

$$\dot{x} + c(x) = 0$$

where c is any continuous function of the same sign as its scalar argument x, i.e.,

$$xc(x) > 0 \quad for \quad x \neq 0.$$

Intuitively, this condition indicates that $-c(x)$ "pushes" the system back towards its rest position $x = 0$, but is otherwise arbitrary. Since c is continuous, it also implies that $c(0) = 0$

Consider as the Lyapunov function candidate the square of the distance to the origin

$$V = x^2.$$

The function V is radially unbounded, since it tends to infinity as $|x| \to \infty$. Its derivative is

$$\dot{V} = 2x\dot{x} = -2xc(x).$$

Thus $\dot{V} < 0$ as long as $x \neq 0$, so that $x = 0$ is globally asymptotically stable equilibrium point.

For instance, the system

$$\dot{x} = \sin^2 x - x$$

is globally asymptotically convergent to $x = 0$, since for $x \neq 0$, $\sin^2 x \leq |\sin x| < |x|$. Similarly, the system

$$\dot{x} = -x^3$$

is globally asymptotically convergent to $x = 0$. Notice that while this system linear approximation ($\dot{x} \approx 0$) is inconclusive, even about local stability, the actual nonlinear system enjoys a strong stability property (global asymptotic stability). □

Example 3.5.8 *Consider the system*

$$\dot{x}_1 = x_2 - x_1(x_1^2 + x_2^2)$$

$$\dot{x}_2 = -x_1 - x_2(x_1^2 + x_2^2)$$

The origin of the state-space is an equilibrium point for this system. Let V be the positive definite function

$$V(\vec{x}) = x_1^2 + x_2^2.$$

The derivative of V along any system trajectory is

$$\dot{V}(\vec{x}) = 2x_1\dot{x}_1 + 2x_2\dot{x}_2 = -2(x_1^2 + x_2^2)^2$$

which is negative definite. Therefore, the origin is a globally asymptotically stable equilibrium point. Note that the globalness of this stability result also implies that the origin is the **only** *equilibrium point of the system.* □

REMARKS

Many Lyapunov functions may exist for the same system. For instance, if V is a Lyapunov function for a given system, so is

$$V_1 = \rho V^\alpha$$

where ρ is any strictly positive constant and α is any scalar (not necessarily an integer) larger than 1. Indeed, the positive-definiteness of V implies that of V_1, the positive definiteness (or positive semi-definiteness) of $-\dot{V}$ implies that of $-\dot{V}_1$, and (the radial unboundedness of V (if applicable) implies that of V_1.

More importantly, for a given system, specific choices of Lyapunov functions may yield more precise results than others. (See P.67 [54])

Along the same lines, it is important to realize that the theorems in Lyapunov analysis are all **sufficiency** theorems. If for a particular choice of Lyapunov function candidate V, the conditions on $\dot{V}$ are not met, one cannot draw any conclusions on the stability or instability of the system – the only conclusion one should draw is that a different Lyapunov function candidate should be tried.

3.5.3 Lyapunov Analysis of Linear Time-Invariant Systems

Definition 3.5.9 *A square matrix M is symmetric if $M = M^T$. A square matrix M is skew-symmetric if $M = -M^T$.* □

Definition 3.5.10 *A square $n \times n$ matrix M is* **positive definite** *(P.D.) if*

$$\vec{x} \neq \vec{0} \Rightarrow \vec{x}^T M \vec{x} > 0.$$

In other words, a matrix M is positive definite if the quadratic function $\vec{x}^T M \vec{x}$ is a positive definite function. This definition implies that to every positive definite matrix is associated a positive definite function. Obviously, the converse is not true. □

LYAPUNOV FUNCTIONS FOR LINEAR TIME-INVARIANT SYSTEMS

Given a linear system of the form $\dot{\vec{x}} = A\vec{x}$, let us consider a quadratic Lyapunov function candidate

$$V = \vec{x}^T P \vec{x}$$

where P is a given symmetric positive definite matrix. Differentiating the positive definite function V along the system trajectory yields another quadratic form

$$\dot{V} = \dot{\vec{x}}^T P \vec{x} + \vec{x}^T P \dot{\vec{x}} = -\vec{x}^T Q \vec{x} \tag{3.5.4}$$

where

$$A^T P + PA = -Q. \tag{3.5.5}$$

The question, thus, is to determine whether the symmetric matrix Q defined by the so-called **Lyapunov equation** (3.5.10) above, is itself P.D. . If this is the case, then V satisfies the conditions of the basic theorem of last section, and the origin is globally asymptotically stable. However, this "natural" approach may lead to inconclusive result, i.e., Q may be not positive definite even for stable systems.

Example 3.5.11

$$A = \begin{bmatrix} 0 & 4 \\ -8 & -12 \end{bmatrix}.$$

If we take $P = I$, then

$$-Q = A^T P + PA = \begin{bmatrix} 0 & -4 \\ -4 & -24 \end{bmatrix}.$$

The matrix Q is not positive definite. Therefore, no conclusion can be drawn from the Lyapunov function on whether the system is stable or not. □

A more useful way of studying a given linear system using scalar quadratic functions is, instead, to derive a positive matrix P from a given positive definite matrix Q, i.e.,

- choose a positive definite matrix Q
- solve for P from the Lyapunov equation (3.5.5)
- check whether P is P.D.

If P is P.D., then $\frac{1}{2}\vec{x}^T P \vec{x}$ is a Lyapunov function for the linear system and global asymptotical stability is guaranteed. Unlike the previous approach of going from a given P to a matrix Q, this technique of going from a given Q to a matrix P always leads to conclusive results for stable linear systems, as seen from the following theorem.

Theorem 3.5.12 *A necessary and sufficient condition for a LTI system $\dot{\vec{x}} = A\vec{x}$ to be strictly stable is that, for any symmetric P.D. matrix Q, the unique matrix P solution of the Lyapunov equation* (3.5.5) *be symmetric positive definite.* □

The above theorem shows that **any** positive definite matrix Q can be used to determine the stability of a linear system. A simple choice of Q is the identity matrix.

Example 3.5.13 *Consider again the second-order system of Example 3.5.11. Let us take $Q = I$ and denote P by*

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix}$$

where, due to symmetry of $P, p_{21} = p_{12}$. Then the Lyapunov equation is

$$\begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} \begin{bmatrix} 0 & 4 \\ -8 & -12 \end{bmatrix} + \begin{bmatrix} 0 & -8 \\ 4 & -12 \end{bmatrix} \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

whose solution is

$$p_{11} = 5, p_{12} = p_{22} = 1.$$

The corresponding matrix

$$P = \begin{bmatrix} 5 & 1 \\ 1 & 1 \end{bmatrix}$$

is positive definite, and therefore the linear system is globally asymptotically stable. Note that we have solved for P directly. □

Even thought the choice $Q = I$ is motivated by computational simplicity, it has a surprising property: the resulting Lyapunov analysis allows us to get the best estimate of the state convergence rate, as we shall see in later courses.

Chapter 4

Nonlinear Static Controller Design

4.1 Introduction

Let's start with a simple scalar plant.

$$\dot{x} = u + \theta x \tag{4.1.1}$$

where u is the control and θ is an unknown constant. If the parameter vector θ

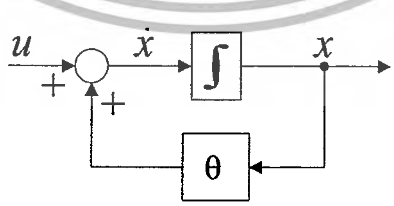

Figure 4.1: Linear scalar plant of (4.1.1)

is known–with the knowledge about their boundedness $|\theta| \leq \bar{\theta}$–then, it's possible to design a nonlinear controller in static part be applying system's feedback. One we can choose is

$$u = -k_1 x - k_2 x^3 \tag{4.1.2}$$

where $k_1 > 0$, $k_2 > 0$. The resulting feedback system is

$$\dot{x} = (\theta - k_1)x - k_2 x^3 \tag{4.1.3}$$

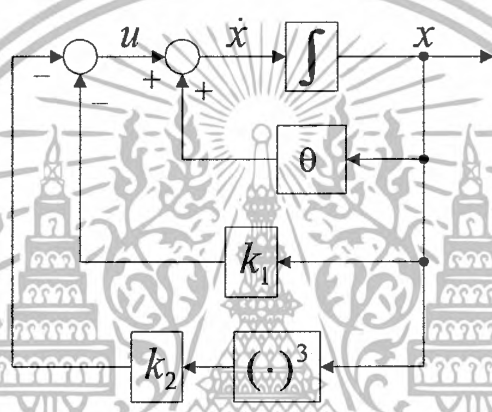

Figure 4.2: The resulting feedback system of (4.1.3)

For $\theta > k_1$, the equilibrium $x = 0$ is unstable, but the nonlinear term $-k_2 x^3$ prevents $x(t)$ from growing unbounded. It is easy to see that $x(t)$ will converge to one of the two equilibria $\pm\sqrt{\frac{\theta - k_1}{k_2}}$. Thus, the static nonlinear controller (4.1.2) has achieved boundedness of $x(t)$ without any knowledge of a bound on θ.

To achieve the desired static controller ,we have two methods available which are *feedback linearization* method and *backstepping* method.

4.2 Feedback Linearization

4.2.1 Introduction

Consider a class of nonlinear systems of the form

$$\begin{aligned}\dot{x} &= f(x) + G(x)u \\ y &= h(x)\end{aligned}$$

and pose the question of whether there exist a state feedback control

$$u = \alpha(x) + \beta(x)v$$

and a change of variables

$$z = T(x)$$

that transform the nonlinear system into an equivalent linear system. We start with *input-state linearization*, where the full state equation is linearized. Next, we introduce the notion of *input-output linearization*, where the emphasis is on linearizing the input-output map from u to y even if the state equation is only partially linearized. State feedback control of feedback linearizable systems is discussed where we deal with stabilization, tracking and regulation via integral control problems. Finally, we present the differential geometric approach to feedback linearization which allows us to characterize the class of feedback linearizable systems by geometric conditions.

4.2.2 Input-State Linearization

Inspection of the state equation.

$$\begin{aligned}\dot{x}_1 &= x_2 \\ \dot{x}_2 &= -a[\sin(x_1 + \delta) - \sin\delta] - bx_2 + cu\end{aligned}$$

shows that we can choose u as

$$u = \frac{a}{c}[\sin(x_1 + \delta) - \sin\delta] + \frac{v}{c}$$

to cancel the nonlinear term $a[\sin(x_1 + \delta) - \sin\delta]$. This cancellation results in the linear system

$$\begin{aligned}\dot{x}_1 &= x_2\\ \dot{x}_2 &= -bx_2 + v\end{aligned}$$

Thus, the stabilization problem for the nonlinear systems has been reduced to a stabilization problem for a controllable linear system. We can proceed to design a stabilizing linear state feedback control

$$v = k_1x_1 + k_2x_2$$

to locate the eigenvalues of the closed-loop system

$$\begin{aligned}\dot{x}_1 &= x_2\\ \dot{x}_2 &= k_1x_1 + (k_2 - b)x_2\end{aligned}$$

in the open left-half plane. The overall state feedback control law is given by

$$u = (\frac{a}{c})[\sin(x_1 + \delta) - \sin\delta] + \frac{1}{c}(k_1x_1 + k_2x_2)$$

The general idea of nonlinearity cancellation is that we should not expected to be able to cancel nonlinearities in every nonlinear system. There must be a certain structural property of the system that allows us to perform such cancellation. It is not hard to see that to cancel a nonlinear term $\alpha(x)$ by subtraction, the control u and the nonlinearity $\alpha(x)$ must always appear together as a sum $u + \alpha(x)$. To cancel a nonlinear term $\gamma(x)$ by division, the control u and the nonlinearity $\gamma(x)$ must always appear as a product $\gamma(x)u$. If the matrix $\gamma(x)$ is nonsingular in the domain of interest, then it can be cancelled by $u = \beta(x)v$, where $\beta(x) = \gamma^{-1}(x)$ is the inverse of the matrix $\gamma(x)$. Therefore, the ability to use feedback to convert a nonlinear state equation into a controllable linear state equation by cancelling nonlinearities requires the nonlinear state equation to have structure

$$\dot{x} = Ax + B\beta^{-1}(x)[u - \alpha(x)] \tag{4.2.1}$$

where A is $n \times n$, B is $n \times p$, the pair (A, B) is controllable, and the functions $\alpha : \mathbb{R}^n \to \mathbb{R}^p$ and $\beta : \mathbb{R}^n \to \mathbb{R}^{p \times p}$ are defined in a domain $D_x \subset \mathbb{R}^n$ that contains the origin. The matrix $\beta(x)$ is assumed to be nonsingular for every $x \in D_x$. Notice that β^{-1} here denotes the inverse of the matrix $\beta(x)$ for every x,and not the inverse map of the function $\beta(x)$. If the state equation takes the form (4.2.1), then we can linearize it via the state feedback

$$u = \alpha(x) + \beta(x)v \tag{4.2.2}$$

to obtain the linear state equation

$$\dot{x} = Ax + Bv \tag{4.2.3}$$

For stabilization, we design $v = Kx$ such that $A + BK$ is Hurwitz. The overall nonlinear stabilizing state feedback control is

$$u = \alpha(x) + \beta(x)Kx \tag{4.2.4}$$

Even if the state equation does not have the structure (4.2.1) for one choice of state variable, it might do so for another choice. Consider, for example, the system

$$\begin{aligned} \dot{x}_1 &= a\sin(x_2) \\ \dot{x}_2 &= -{x_1}^2 + u \end{aligned}$$

We cannot simply choose u to cancel the nonlinear term $a \sin x_2$. However, if we first change the variables by the transformation

$$\begin{aligned} z_1 &= x_1 \\ z_2 &= a\sin x_2 = \dot{x}_1 \end{aligned}$$

then z_1 and z_2 satisfy

$$\begin{aligned}\dot{z}_1 &= z_2\\ \dot{z}_2 &= a\cos x_2(-{x_1}^2+u)\end{aligned}$$

and the nonlinearity can be cancelled by the control

$$u = {x_1}^2 + \frac{1}{a\cos(x_2)}v$$

which is well defined for $\frac{-\pi}{2} < x_2 < \frac{\pi}{2}$. The state equation in the new coordinate (z_1, z_2); that is,

$$\begin{aligned}x_1 &= z_1\\ x_2 &= \sin^{-1}(\frac{z_2}{a})\end{aligned}$$

which is well defined for $-a < z_2 < a$. The transformed state equation is given by

$$\begin{aligned}\dot{z}_1 &= z_2\\ \dot{z}_2 &= a\cos(\sin^{-1}(\frac{z_2}{a}))(-{z_1}^2+u)\end{aligned}$$

When a change of variable $z = T(x)$ is used to transform the state equation from the x-coordinates to the z-coordinates, the map T must be invertible, that is, it must have an inverse map $T^{-1}(\cdot)$ such that $x = T^{-1}(z)$ for all $z \in T(D_x)$ where D_x is the domain of T. Moreover, since the derivatives of z and x should be continuous, we require that both $T(\cdot)$ and $T^{-1}(\cdot)$ be continuously differentiable map with a continuously differentiable inverse is known as a *diffeomorphism*. Now we have all the elements we need to define input-state linearizable systems.

Definition 4.2.1 *A nonlinear system*

$$\dot{x} = f(x) + G(x)u \tag{4.2.5}$$

where $f : D_x \to \mathbb{R}^n$ and $G : D_x \to \mathbb{R}^{n\times p}$ are sufficiency smooth on a domain $D_x \subset \mathbb{R}^n$, is said to be input-state linearizable if there exists a diffeomorphism

$T : D_x \to \mathbb{R}^n$ such that $D_x = T(D_x)$ contains the origin and the change of variables $z = T(x)$ transforms the system (4.2.5) *into the form*

$$\dot{z} = Az + B\beta^{-1}(x)[u - \alpha(x)] \tag{4.2.6}$$

with (A, B) controllable and $\beta(x)$ nonsingular for all $x \in D_x$. □

Setting

$$\alpha_o(z) = \alpha(T^{-1}(z)) \text{ and } \beta_o(z) = \beta(T^{-1}(z))$$

we can write equation (4.2.6) as

$$\dot{z} = Ax + B\beta_o{}^{-1}(z)[u - \alpha_o(z)] \tag{4.2.7}$$

which takes the form (4.2.1). It is more convenient, however, to express α and β in the x-coordinates since the state the state feedback control is implemented in these coordinate.

Suppose we are given an input-state linearizable system (4.2.5). Let $z = T(x)$ be a change of variables that brings the system into the form (4.2.6). We have

$$\dot{z} = \frac{\partial T}{\partial x}\dot{x} = \frac{\partial T}{\partial x}[f(x) + G(x)u] \tag{4.2.8}$$

On the other hand, from (4.2.6),

$$\dot{z} = Az + B\beta^{-1}(x)[u - \alpha(x)] \tag{4.2.9}$$

From (4.2.8) and (4.2.9), we see that the equality

$$\frac{\partial T}{\partial x}[f(x) + G(x)u] = AT(x) + B\beta^{-1}(x)[u - \alpha(x)]$$

must hold for all x and u in the domain of interest. By first taking $u = 0$, we split the equation into two:

$$\frac{\partial T}{\partial x}f(x) = AT(x) - B\beta^{-1}(x)\alpha(x) \tag{4.2.10}$$

$$\frac{\partial T}{\partial x}G(x) = B\beta^{-1}(x) \tag{4.2.11}$$

Therefore, we conclude that any function $T(\cdot)$ that transform (4.2.5) into the form (4.2.6) must satisfy the partial differential equation (4.2.10)-(4.2.11). Alternatively, if there is a map $T(\cdot)$ that satisfy (4.2.10) - (4.2.11) for some α ,β,A and B with the desired properties, then it can be easily seen that the change of variable $z = T(x)$ transforms (4.2.5) into (4.2.6). Hence, the existence of T,α,β,A,and B that satisfy the partial differential equations (4.2.10) - (4.2.11) is a necessary and sufficient condition for the system (4.2.5) to be input-state linearizable.

When a nonlinear system is input-state linearizable, the map $z = T(x)$ that transforms the system into the form (4.2.6) is not unique. Probably the easiest way to see this point is to notice that if we apply the linear state transformation $\zeta = Mz$, with a nonsingular M, to (4.2.6) then the state equation in the ζ-coordinates will be

$$\dot{\zeta} = MAM^{-1}\zeta + MB\beta^{-1}(x)[u - \alpha(x)]$$

which is still of the form (4.2.6), but with different A and B matrices. Therefore, the composition of the transformations $z = T(x)$ and $\zeta = Mz$ gives a new transformation that transform the system into the special structure of (4.2.6). The nonuniqueness of T can be exploited to simplify the partial differential equations (4.2.10)-(4.2.11). To illustrate the idea without complications, we will restrict our discussions to single-input systems ($p = 1$) and write the single-column input matrix G as g. In this case, for any controllable pair (A, B) we can find a nonsingular matrix M that transforms (A, B) into a controllable canonical form; that is, $MAM^{-1} = A_c + B_c\lambda^T$ and $MB = B_c$, where

$$A_c = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & 0 & 1 \\ 0 & \dots & \dots & 0 & 0 \end{bmatrix}_{n\times n} \quad and \quad B_c = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}_{n\times 1}$$

The term $B_c\lambda^T\zeta = B_c\lambda^T MT(x)$ can be include in the term $B_c\beta^{-1}(x)\alpha(x)$. Therefore, without loss of generality, we can assume that the matrices A and B in (4.2.10)-(4.2.11) are the canonical form matrices A_c and B_c. Let

$$T(x) = \begin{bmatrix} T_1(x) \\ T_2(x) \\ \vdots \\ T_{n-1}(x) \\ T_n(x) \end{bmatrix}$$

It can be easily verified that

$$A_cT(x) - B_c\beta^{-1}(x)\alpha(x) = \begin{bmatrix} T_2(x) \\ T_3(x) \\ \vdots \\ T_{n-1}(x) \\ -\alpha(x)/\beta(x) \end{bmatrix} \quad \textit{and} \quad B_c\beta^{-1}(x) = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1/\beta(x) \end{bmatrix}$$

where α and β are scalar functions. Using these expressions in (4.2.10)-(4.2.11) simplifies the partial differential equations. Equation (4.2.10) simplifies to

$$\begin{aligned} \frac{\partial T_1}{\partial x}f(x) &= T_2(x) \\ \frac{\partial T_2}{\partial x}f(x) &= T_3(x) \\ &\vdots \\ \frac{\partial T_{n-1}}{\partial x}f(x) &= T_n(x) \\ \frac{T_n}{\partial x}f(x) &= -\alpha(x)/\beta(x) \end{aligned}$$

The first $n-1$ equations show that the components T_2 to T_n of T are determined functions of the first component T_1. The last equation defines α/β in terms of T_1

Equation (4.2.11) simplifies to

$$\begin{aligned}
\frac{\partial T_1}{\partial x} g(x) &= 0 \\
\frac{\partial T_2}{\partial x} g(x) &= 0 \\
&\vdots \\
\frac{\partial T_{n-1}}{\partial x} g(x) &= 0 \\
\frac{T_n}{\partial x} g(x) &= -1/\beta(x) \neq 0
\end{aligned}$$

We need to search for a function $T_1(x)$ that satisfies

$$\frac{\partial T_1}{\partial x} g(x) = 0, \quad i = 1, 2, \ldots, n-1; \quad \frac{\partial T_n}{\partial x} g(x) \neq 0 \tag{4.2.12}$$

where

$$T_{i+1}(x) = \frac{\partial T_i}{\partial x} f(x), \quad i = 1, 2, \ldots, n-1$$

If there is a function $T_{i+1}(x)$ that satisfies (4.2.12), then β and α are given by

$$\beta(x) = \frac{1}{(\partial T_n/\partial x) g(x)}; \quad \alpha(x) = -\frac{(\partial T_n/\partial x) f(x)}{(\partial T_n/\partial x) g(x)} \tag{4.2.13}$$

Thus we have reduced the problem to solving (4.2.12) for T_1. We will see that the existence of a function T_1 that satisfies (4.2.12) can be characterized by a necessary and sufficient condition in terms of the functions f,g, and their partial derivatives. Furthermore, if $T_1(x)$ satisfies (4.2.12) in a domain D, then for each $x_o \in D$ there is a neighborhood N of x_o such that the restriction of $T(x)$ to (N) is a diffeomorphism.

Example 4.2.2 *As a first example, we reconsider the system*

$$\dot{x} = \begin{bmatrix} a\sin(x_2) \\ -x_1^2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u = f(x) + gu$$

and see how we can arrive at the transformation we used earlier by solving (4.2.12)*. The open-loop system has an equilibrium point at $x = 0$. We want to find $T_1(x)$ that* *satisfies the conditions*

$$\frac{\partial T_1}{\partial x}g = 0; \frac{\partial T_2}{\partial x}g \neq 0$$

with $T_1(0) = 0$, *where*

$$T_2 = \frac{\partial T_1}{\partial x}f(x)$$

From the condition $[\partial T_1/\partial x]g = 0$,*we have*

$$\frac{\partial T_1}{\partial x}g = \frac{\partial T_1}{\partial x_2} = 0$$

So,T_1 *must be independent of* x_2.*Therefore,*

$$T_2(x) = \frac{\partial T_1}{\partial x}a\sin(x_2)$$

The condition

$$\frac{\partial T_2}{\partial x}g = \frac{\partial T_2}{\partial x_2} = \frac{\partial T_1}{\partial x_1}a\cos(x_2) \neq 0$$

is satisfied in the domain where $\cos(x_2) \neq 0$ *by any choice of* $T_1 = T_1(x_1)$ *such that* $\frac{\partial T_1}{\partial x_1} \neq 0$. *Taking* $T_1(x_1) = x_1$ *results in the transformation we used earlier. Other choices of* T_1 *could have been made. For example,* $T_1(x_1) = x_1 + x_1^3$ *would give another change of variables that transforms the system into the form* (4.2.6).

□

Example 4.2.3 *A synchronous generator connected to an infinite bus may be represented by a third-order model*

$$\dot{x} = f(x) + gu$$

with

$$f(x) = \begin{bmatrix} x_2 \\ -a[(1+x_3)\sin(x_1+\delta) - \sin(\delta)] - bx_2 \\ -cx_3 + d[\cos(x_1+\delta) - \cos(\delta)] \end{bmatrix}; \quad g = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

where $a, b, c, d,$ and δ are positive constants. The open-loop system has equilibrium at $x = 0$. We want to find $T_1 = T_1(x)$ that satisfies the conditions

$$\frac{\partial T_1}{\partial x}g = 0; \quad \frac{\partial T_2}{\partial x}g = 0; \quad \frac{\partial T_3}{\partial x}g \neq 0$$

with $T_1(0) = 0$, where

$$T_2(x) = \frac{\partial T_1}{\partial x}f(x); \quad T_3(x) = \frac{\partial T_2}{\partial x}f(x)$$

from the condition $[\partial T_1/\partial x]g = 0$,we have

$$\frac{\partial T_1}{\partial x}g = \frac{\partial T_1}{\partial x_3} = 0$$

We choose T_1 independent of x_3. Therefore,

$$T_2(x) = \frac{\partial T_1}{\partial x_1}x_2 - \frac{\partial T_1}{\partial x_2}a[(1 + x_3)\sin(x_1 + \delta) - \sin(\delta)] + bx_2$$

from the condition $[\partial T_2/\partial x]g = 0$,we have

$$\frac{\partial T_2}{\partial x}g = \frac{\partial T_2}{\partial x_3} = -a\sin(x_1 + \delta)\frac{\partial T_1}{\partial x_2} = 0$$

We choose T_1 independent of x_2. Therefore,T_2 simplifies to

$$T_2(x) = \frac{\partial T_1}{\partial x_1}x_2$$

and

$$T_3(x) = \frac{\partial T_2}{\partial x_1}x_2 - \frac{\partial T_1}{\partial x_1}a[(1 + x_3)\sin(x_1 + \delta) - \sin(\delta)] + bx_2$$

Hence,

$$\frac{\partial T_3}{\partial x}g = \frac{\partial T_3}{\partial x_3} = -a\sin(x_1 + \delta)\frac{\partial T_1}{\partial x_1}$$

and the condition $[\partial T_3/\partial x]g \neq 0$ is satisfied in the domain $0 < x_1 + \delta < \pi$ with any choice of $T_1 = T_1(x_1)$ such that $(\partial T_1/\partial x_1) \neq 0$ on this domain. The simple choice $T_1 = x_1$ satisfies this requirement as well as the condition $T_1(0) = 0$. Thus, *we choose* $oT_1 = x_1$ *and find from the previous expressions that the change of variables* $z = T(x)$ *is given by*

$$\begin{aligned} z_1 &= T_1(x) = x_1 \\ z_2 &= T_2(x) = x_1 \\ z_1 &= T_3(x) = a[(1+x_3)\sin(x_1+\delta) - \sin(\delta)] + bx_2 \end{aligned}$$

The inverse of this transformation $x = T^-1(z)$ *is defined for all* $0 < z_1 + \delta < \pi$ *and given by*

$$\begin{aligned} x_1 &= z_1 \\ x_2 &= z_2 \\ x_3 &= -1 - \frac{z_3 + bz_2 - a\sin(\delta)}{a\sin(z_1+\delta)} \end{aligned}$$

The functions β *and* α *are given by*

$$\begin{aligned} \beta(x) &= \frac{1}{(\partial T_3/\partial x)g} = \frac{1}{(\partial T_3/\partial x_3)} = \frac{1}{a\sin(x_1+\delta)} \\ \alpha(x) &= -\frac{(\partial T_3/\partial x)f(x)}{(\partial T_3/\partial x)g(x)} \\ &= \frac{-a(1+x_3)\cos(x_1+\delta)f_1(x) - bf_2(x) - a\sin(x_1+dt)f_3(x)}{a\sin(x_1+dt)} \end{aligned}$$

The state equation in the z-coordinates is

$$\begin{aligned} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= z_3 \\ \dot{z}_3 &= -a\sin(x_1+\delta)[u+\alpha(x)] \end{aligned}$$

which can be linearized by the linearizing feedback control

$$u = \alpha(x) - \frac{1}{a\sin(x_1+\delta)}v$$

Notice that the state equation in the z-coordinates is valid only in the domain $0 < z_1 + \delta < \pi$, *which is the domain over which the change of variables* $z = T_1(x)$ *is a well-defined diffeomorphism.* □

4.2.3 Input-Output Linearization

When certain output variable are interest, as in tracking control problems, the state model is described by state and output equations. Linearizing the state equation, as we have done in the previous section, does not necessarily linearize the output equation. For example, if the system

$$\begin{aligned} \dot{x}_1 &= a \sin x_2 \\ \dot{x}_2 &= -{x_1}^2 + u \end{aligned}$$

has an output $y = x_2$, then the change of variables and state feedback control

$$z_1 = x_1, \quad z_2 = a \sin x_2, \quad and \quad u = {x_1}^2 + \frac{1}{a \cos x_2} v$$

yield

$$\begin{aligned} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= v \\ y &= \sin^{-1} \frac{z_2}{a} \end{aligned}$$

While the state equation is linear, solving a tracking control problem for y is still complicated by the nonlinearity of the output equation. Inspection of both the state and output equations in the x-coordinates shows that, if we use the state feedback control $u = {x_1}^2 + v$, we can linearize the input-output map from u to y, which will be described by the linear model

$$\begin{aligned} \dot{x}_2 &= v \\ y &= x_2 \end{aligned}$$

We can now proceed to solve the tracking control problem using linear control theory. This discussion shows that sometimes it is more beneficial to linearize the input-output map even at the expense of leaving part of the nonlinear state equation. This is the input-output linearization problem which we will address in this section. One catch about input-output linearization is that the linearized input-output map may not account for all the dynamics of all the dynamics of the system. In the foregoing example, the full system is described by

$$
\begin{aligned}
\dot{x}_1 &= a \sin x_2 \\
\dot{x}_2 &= v \\
y &= x_2
\end{aligned}
$$

Note that the state variable x_1 is not connected to the output y. In other words, the linearizing feedback control has mode x_1 unobservable from y. When we design tracking control, we should make sure that the variable x_1 is well behaved; that is, stable or bounded in some sense. A native control design that uses only the linear input-output map may result in an ever-growing signal $x_1(t)$. For example, suppose we design a linear control to stabilize the output y at a constant value y_R. The corresponding solution of the $x_1(t)$ equation is $x_1(t) = x_1(0) + ta \sin y_R$. So, for any $y_R \neq 0$, the variable $x_1(t)$ ill grow unbounded. This internal stability issue will be address in the context of input-output linearization.

Consider the single-input-single-output:(SISO) system

$$
\begin{aligned}
\dot{x} &= f(x) + g(x)u \\
y &= h(x)
\end{aligned}
$$

where $f, g, andh$ are sufficiently smooth in a domain $D \subset \mathbb{R}^n$. The simplest case of input-output linearization arises when the system is both input-state linearizable. Starting with an input-state linearizable system, let $T_1(x)$ be a solution of (4.2.12). Suppose the output function $h(x)$ happens to be equal to $T_1(x)$. For example, in the robotic manipulator example 12.3, we choose $T_1(x) = x_1$; this could indeed be the output of interest, since in this problem we are usually interested in controlling the angle x_1. If $h(x) = T_1(x)$, then the change of variables $z = T(x)$ and state feedback control $u = \alpha(x) + \beta(x)v$ yield the system

$$\begin{aligned} \dot{z} &= A_c z + B_c v \\ y &= C_c z \end{aligned}$$

where (A_c, B_c, C_c) is a canonical form representation of a chain of n integrators; that is, A_c and B_c take the form (4.2.2) and

$$C_c = [1 \;\; 0 \;\; \dots \;\; 0 \;\; 0]_{1\times n} \tag{4.2.14}$$

In this system, both the state and output equations are linear. For a given output function $h(x)$, we can find out whether or not $h(x)$ satisfies the condition (4.2.12) by direct verification; we do not need to solve partial differential equations. The condition (4.2.12) can be interpreted as a restriction on the way the derivatives of y depend on u. To see this point, set $\psi_1(x) = h(x)$. The derivative $\dot{y}$ is given by

$$\dot{y} = \frac{\partial \psi_1}{\partial x}[f(x) + g(x)u]$$

If $[\partial\psi_1/\partial x]g(x) = 0$, then

$$\dot{y} = \frac{\partial \psi_1}{\partial x} f(x) \triangleq \psi_2(x)$$

If we continue to calculate the second derivative of y, denoted by $y^{(2)}$, we obtain

$$y^{(2)} = \frac{\partial \psi_2}{\partial x}[f(x) + g(x)u]$$

Once again, if $[\partial\psi_2/\partial x]g(x) = 0$ then

$$y^{(2)} = \frac{\partial \psi_2}{\partial x} f(x) \triangleq \psi_3(x)$$

Repeating this process, we see that if $h(x) = \psi_1(x)$ satisfies (4.2.12); that is,

$$\frac{\partial \psi_i}{\partial x} = 0, \quad i = 1, 2, \dots, n-1; \quad \frac{\partial \psi_n}{\partial x} g(x) \neq 0$$

where

$$\psi_{i+1}(x) = \frac{\partial \psi_i}{\partial x} f(x), \quad i = 1, 2, \dots, n-1$$

then u does not appear in the equations of $y, \dot{y}, \ldots, y^{(n-1)}$ and appears in the equation of $y^{(n)}$ with a nonzero coefficient

$$y^{(n)} = \frac{\partial \psi_n}{\partial x} f(x) + \frac{\partial \psi_n}{\partial x} g(x) u$$

This equation shows clearly that the system is input-output linearizable since the state control

$$u = \frac{1}{\frac{\partial \psi_n}{\partial x} g(x)} [-\frac{\partial \psi_n}{\partial x} f(x) + v]$$

reduces the input-output map to

$$y^{(n)} = v$$

which is a chain of n integrators. If u appears in the equation of one of the derivative $\dot{y}, \ldots, y^{(n-1)}$ and the coefficient of u (when it appear) is nonzero, then we can again linearize the input-output map. In particular, if $h = \psi_1(x)$ satisfies

$$\frac{\psi_i}{\partial x} g(x) = 0, \quad i = 1, 2, \ldots, r-1; \quad \frac{\partial \psi_r}{\partial x} g(x) \neq 0$$

for some $1 \leq r < n$, then the equation of $y^{(r)}$ is given by

$$y^{(r)} = \frac{\partial \psi_r}{\partial x} f(x) + \frac{\partial \psi_r}{\partial x} g(x) u$$

and the state feedback control

$$u = \frac{1}{\frac{\partial \psi_r}{\partial x} g(x)} [-\frac{\partial \psi_r}{\partial x} f(x) + v]$$

linearizes the input-output map to the chain of r integrators

$$y^{(r)} = v$$

In this case, the integrator r is called the *"relative degree"* of the system, according to the following definition.

Definition 4.2.4 *A nonlinear system*

$$\dot{x} = f(x) + g(x)u \tag{4.2.15}$$

$$y = h(x) \tag{4.2.16}$$

where $f : D \to \mathbb{R}^n$, $g : D \to \mathbb{R}^n$, *and* $h : D \to \mathbb{R}$ *are sufficiently smooth on a domain* $D \subset \mathbb{R}^n$, *is said to have relative degree* $r, 1 \leq r \leq n$, *in a region* $D_o \subset D$ *if*

$$\frac{\partial \psi_i}{\partial x} g(x) = 0, \quad i = 1, 2, \ldots, r-1; \quad \frac{\partial \psi_r}{\partial x} g(x) \neq 0 \tag{4.2.17}$$

for all $x \in D_o$ *where*

$$\psi_1(x) = h(x) \ \ and \ \ \psi_{i+1}(x), \quad i = 1, 2, \ldots, r-1 \tag{4.2.18}$$

If the system (4.2.15)-(4.2.16) has relative degree r, *then it is input-output linearizable. If it has relative degree* n, *then it both input-state and input-output linearizable.* □

Example 4.2.5 *Consider the controlled Van der Pol equation*

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1 + \epsilon(1 - x_1^2)x_2 + u, \quad \epsilon > 0 \end{aligned}$$

with output $y = x_1$. *Calculating the derivatives of the output, we obtain*

$$\begin{aligned} \dot{y} &= \dot{x}_1 = x_2 \\ \ddot{y} &= \dot{x}_2 = -x_1 + \epsilon(1 - x_1^2)x_2 + u \end{aligned}$$

Hence, The system has relative degree 2 in $\mathbb{R}^2$. *Therefore, it is both input-state and input-output linearizable. If the output is* $y = x_2$, *then*

$$\dot{y} = -x_1 + \epsilon(1 - x_1^2)x_2 + u$$

and the system has relative degree 1 in $\mathbb{R}^2$. *If the output is* $y = x_1 + x_2^2$, *then*

$$\dot{y} = x_2 + x_2[-x_1 + \epsilon(1 - x_1^2)x_2 + u]$$

and the system has relative degree 1 in $D_o = \{x \in \mathbb{R}^2 | x_2 \neq 0\}$. □

4.3 Backstepping Technique

Although the system can be linearized by applying feedback linearization method, it often leads to cancellation of useful nonlinearities. Backstepping designs are more flexible and do not force the designed system to appear linear. They can avoid cancellations of useful nonlinearities and often introduce additional nonlinear terms to improve their transient performance.

Example 4.3.1 *We are now trying to construct clf's for a second-order system*

$$\dot{x} = x\xi \quad (4.3.1\text{-a})$$
$$\dot{\xi} = u \quad (4.3.1\text{-b})$$

which is uncontrollable at $x = 0$. *We need to select a clf* V_a *for the system* (4.2.19) *which guarantees global stability of the system, so we need* V_a *to be negative definite (or* $\dot{V}_a < 0$*). We choose one possible form,*

$$\dot{V}_a(x, z) = -x^4 - z^2, \qquad z = \xi - \xi_{des} \quad (4.3.2)$$

and choose a Control Lyapunov Function (clf)

$$V_a(x, z) = \frac{1}{2}x^2 + \frac{1}{2}z^2 \quad (4.3.3)$$

which is positive definite (P.D.) and radially unbounded (R.U.). We apply $\dot{x}$ *and*

$\dot{z}$ in the derivative $\dot{V}_a$

$$\begin{aligned} \dot{V}_a(x,z) &= x\dot{x} + z\dot{z} \\ &= x(x\xi) + z(\dot{\xi} - \dot{\xi}_{des}) \\ &= x^2\xi + z(\dot{\xi} - \dot{\xi}_{des}) \end{aligned} \tag{4.3.4}$$

Compare with the desired value $\dot{V}_a = -x^4 - z^2$ we get

$$\alpha = \xi_{des} = -x^2 \tag{4.3.5}$$

$$\begin{aligned} \dot{\alpha} &= \dot{\xi}_{des} = -2x\dot{x} \\ &= -2x^2\xi \end{aligned} \tag{4.3.6}$$

Apply α and $\dot{\alpha}$ in (4.3.4)

$$\begin{aligned} \dot{V}_a(x,z) &= x^2\xi + z(\dot{\xi} - \dot{\alpha}) \\ &= x^2(z+\alpha) + z(u + 2x^2\xi) \\ &= x^2z - x^4 + z(u + 2x^2\xi) \\ &= -x^4 + z(u + 2x^2\xi + x^2) \end{aligned} \tag{4.3.7}$$

Again, compare with the desired $\dot{V}_a = -x^4 - z^2$, we get

$$\begin{aligned} u + 2x^2\xi + x^2 &= -z \quad \textit{or} \\ u &= -2x^2 - (2x^2+1)\xi \end{aligned} \tag{4.3.8}$$

be the corresponding control u which matches the desired condition $\dot{V}_a = -x^4 - z^2$, yield the resulting system in the $(x\xi)$ coordinates

$$\dot{x} = x\xi \tag{4.3.9-a}$$

$$\dot{\xi} = -2x^2 - (2x^2+1)\xi \tag{4.3.9-b}$$

and its equilibrium $(0,0)$ is Globally Asymptotically Stable (GAS.) □

The x^2 term in feedback loop still being nonlinear, but this condition help in decreasing a setting time of the closed-loop system as shown in Figure 4.3.

However, we have to be sure that the system will be not overheated, damaged, or saturated by this rapidly growing of a control input u.

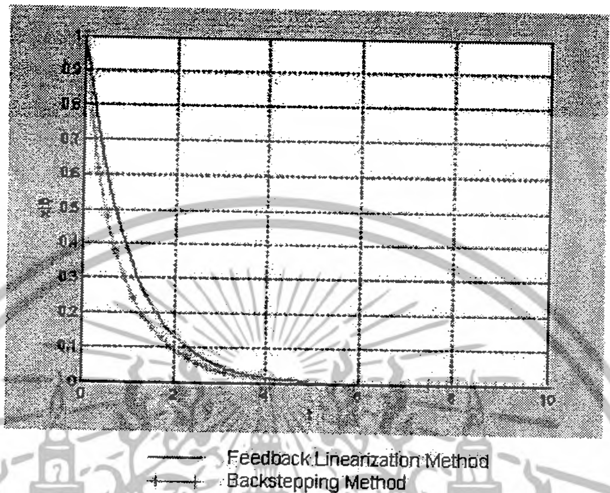

Figure 4.3: Decreasing in setting time caused by nonlinearity

Consider a nonlinear control system

$$\dot{\vec{x}} \;=\; f(\vec{x},t). \tag{4.3.10}$$

Theorem 4.3.2 (LaSalle-Yoshizawa) *Let $\vec{x} = \vec{0}$ be an equilibrium point of* (4.3.10) *and suppose f is locally Lipschitz in $\vec{x}$ uniformly in t. Let $V : \mathbb{R}^n \to \mathbb{R}_+$ be a continuously differentiable, positive definite and radially unbounded function $V(\vec{x})$ such that*

$$\dot{V} \;=\; \frac{\partial V}{\partial \vec{x}}(\vec{x})f(\vec{x},t) \leq \; -W(\vec{x}) \leq 0, \qquad \forall t \geq 0, \forall \vec{x} \in \mathbb{R}^n, \tag{4.3.11}$$

where W is a continuous function. Then, all solutions of (4.3.10) *are globally uniformly bounded and satisfy*

$$\lim_{t\to\infty} W(\vec{x}(t)) \;=\; 0 \tag{4.3.12}$$

In addition, if $W(\vec{x})$ is positive definite, then the equilibrium $\vec{x} = \vec{0}$ is globally uniformly asymptotically stable (GUAS). □

For the regulation task, the designed system is usually time-invariant,

$$\dot{\vec{x}} = f(\vec{x}), \tag{4.3.13}$$

in which case we are interested in its *invariant sets.* A set M is called an invariant set of (4.3.13) if any solution $\vec{x}(t)$ that belongs to M at some time instant t_1 must belong to M for all future and past time:

$$\vec{x}(t_1) \in M \Rightarrow \vec{x}(t) \in M, \quad \forall t \in \mathbb{R} \tag{4.3.14}$$

A set Ω is *positively invariant* if this is true for all future time only:

$$\vec{x}(t_1) \in \Omega \Rightarrow \vec{x}(t) \in \Omega, \quad \forall t \geq t_1. \tag{4.3.15}$$

Can we guarantee convergence to a desired invariant set? A rewarding answer to this question is provided by LaSalle's Invariant Theorem and its asymptotic stability corollary:

Theorem 4.3.3 (LaSalle) *Let Ω be a positively invariant set of* (4.3.13). *Let $V : \Omega \to \mathbb{R}_+$ be a continuously differentiable function $V(\vec{x})$ such that $\dot{V}(\vec{x}) \leq 0, \forall \vec{x} \in \Omega$. Let $E = \left\{ \vec{x} \in \Omega \mid \dot{V}(\vec{x}) = 0 \right\}$, and let M be the largest invariant set contained in E. Then every bounded solution $\vec{x}(t)$ starting in Ω converges to M as $t \to \infty$.* □

Corollary 4.3.4 (Asymptotic Stability) *Let $\vec{x} = \vec{0}$ be the only equilibrium of* (4.3.13). *Let $V : \mathbb{R}^n \to \mathbb{R}_+$ be a continuously differentiable, positive definite, radially unbounded function $V(\vec{x})$ such that $\dot{V}(\vec{x}) \leq 0, \forall \vec{x} \in \mathbb{R}^n$. Let $E = \left\{ \vec{x} \in \mathbb{R}^n \mid \dot{V}(\vec{x}) = 0 \right\}$, and suppose that no solution other than $\vec{x}(t) \equiv \vec{0}$ can stay forever in E. Then the origin is globally asymptotically stable (GAS).* □

These invariance results motivate us to closely examine the invariant subsets of E. The convergence properties of the designed system are stronger if the dimension of M is lower. In the most favorable case of asymptotic stability, the largest invariant subset M of E is just the origin $\vec{x} = \vec{0}$. Our aim in backstepping technic thus is to render the dimension of M as low as possible.

Assumption 4.3.5 *Consider the system*

$$\dot{\vec{x}} = f(\vec{x}) + g(\vec{x})u, \quad f(\vec{0}) = \vec{0}, \tag{4.3.16}$$

where $\vec{x} \in \mathbb{R}^n$ is the state and $u \in \mathbb{R}$ is the control input. There exist a continuously differentiable feedback control law

$$u = \alpha(\vec{x}), \quad \alpha(\vec{0}) = 0. \tag{4.3.17}$$

and a smooth, positive definite, radially unbounded function $V : \mathbb{R}^n \to \mathbb{R}$ such that

$$\frac{\partial V}{\partial \vec{x}}(\vec{x})\left[f(\vec{x}) + g(\vec{x})\alpha(\vec{x})\right] \leq -W(\vec{x}) \leq 0, \qquad \forall \vec{x} \in \mathbb{R}^n, \tag{4.3.18}$$

where $W : \mathbb{R}^n \to \mathbb{R}$ is positive semidefinite. □

Under this assumption, the control (4.3.17), applied to the system (4.3.16), guarantees global boundedness of $\vec{x}$, and via the LaSalle-Yoshizawa Theorem (Theorem 4.3.2), the regulation of $W(\vec{x})$:

$$\lim_{t\to\infty} W(\vec{x}(t)) = 0. \tag{4.3.19}$$

A stronger convergence result is obtained using LaSalle's Theorem (Theorem 4.3.3) with $\Omega = \mathbb{R}^n$: $\vec{x}(t)$ converges to the largest invariant set M contained in the set $E = \{\vec{x} \in \mathbb{R}^n \mid W(\vec{x}) = 0\}$. Clearly, if $W(\vec{x})$ is positive definite, the control (4.3.17) renders $\vec{x} = \vec{0}$ the GAS equilibrium of (4.3.16).

Lemma 4.3.6 (Integrator Backstepping) *Let the system* (4.3.16) *be augmented by an integrator:*

$$\dot{\vec{x}} = f(\vec{x}) + g(\vec{x})\xi \tag{4.3.20-a}$$

$$\dot{\xi} = u, \tag{4.3.20-b}$$

and suppose that (4.3.20-a) *satisfies Assumption 4.3.5 with* $\xi \in \mathbb{R}$ *as its control.*

(i) If $W(\vec{x})$ *is positive definite, then*

$$V_a(\vec{x}, \xi) = V(\vec{x}) + \frac{1}{2}[\xi - \alpha(\vec{x})]^2 \tag{4.3.21}$$

is a CLF for the full system (4.3.20)*, that is, there exists a feedback control* $u = \alpha(\vec{x}, \xi)$ *which renders* $\vec{x} = \vec{0}, \xi = \alpha(\vec{0}) = 0$ *the GAS equilibrium of* (4.3.20)*. One such control is*

$$u = -c(\xi - \alpha(\vec{x})) + \dot{\alpha}(\vec{x}) - \frac{\partial V}{\partial \vec{x}}(\vec{x})g(\vec{x}), \quad c > 0. \tag{4.3.22}$$

(ii) If $W(\vec{x})$ *is positive semidefinite, then there exists a feedback control which renders* $\dot{V}_a \leq -W_a(\vec{x}, \xi) \leq 0$ *such that* $W_a(\vec{x}, \xi) > 0$ *whenever* $W(\vec{x}) > 0$ *or* $\xi \neq \alpha(\vec{x})$*. This guarantees global boundedness and convergence of* $\begin{bmatrix} \vec{x}(t) \\ \xi(t) \end{bmatrix}$ *to the largest invariant set* M_a *contained in the set*
$E_a = \left\{ \begin{bmatrix} \vec{x} \\ \xi \end{bmatrix} \in \mathbb{R}^{n+1} \mid W(\vec{x}) = 0, \xi = \alpha(\vec{x}) \right\}$

While the choice of control (4.3.22) is simple, this control may not be desirable because it involves cancellation of nonlinearities, some of which may be useful.

Chapter 5

Nonlinear Adaptive Controller Design

5.1 Introduction

All methods in the previous chapter are the design procedures employ *static* feedback which guarantee that in the presence of uncertain bounded nonlinearities, the closed-loop state remains bounded. Now, we focus on the design procedures in the **Dynamic** part where the uncertainties are more specific. They consist of unknown constant parameters which appear linearly in the system equations. In the presence of such parametric uncertainties, we will be able to achieve both boundedness of the closed-loop states and convergence of the tracking error to zero.

What we need in designing a dynamic feedback is a parameter estimation. The dynamic part of the controller is designed as a *parameter update law* with which the static part is continuously *adapted* to new parameter estimates, hence its name: *Adaptive control law.* An adaptive controllers guarantees not only that the plant state x remains bounded, but also that it tends to a desired constant value ("regulation") or asymptotically tracks a reference signal ("tracking").

5.2 Adaptation as dynamic feedback

Consider the simplest nonlinear system

$$\dot{x} = u + \theta\varphi(x) \tag{5.2.1}$$

where θ is the unknown constant parameter.

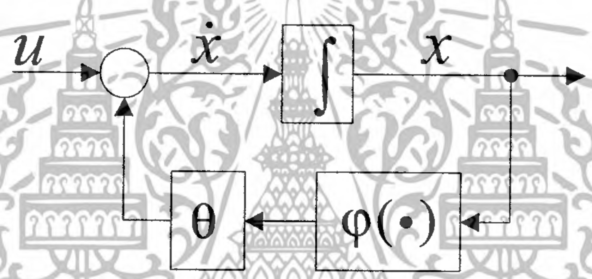

Figure 5.1: Block diagram of the system (5.2.1)

Even if we do not know a bound for θ, we can apply a nonlinear static feedback, as explained in preceding section, to guarantees global boundedness of $x(t)$. Now, we try to achieve a system regulation by applying a dynamic feedback controller.

If θ were known, the control

$$u = -\theta\varphi(x) - c_1 x, \quad c_1 > 0 \tag{5.2.2}$$

would render the derivative of $V_0(x) = \frac{1}{2}x^2$ negative definite: $\dot{V}_0 = -c_1 x^2$. Of course the control law (5.2.2) can not be implemented, since θ is unknown. Instead, one can employ its *certainty-equivalence* form in which θ is replaced by an estimate $\hat{\theta}$:

$$u = -\hat{\theta}\varphi x - c_1 x \tag{5.2.3}$$

Substituting (5.2.3) into (5.2.1), we obtain

$$\dot{x} = -c_1 x + \tilde{\theta}\varphi(x) \tag{5.2.4}$$

where $\tilde{\theta}$ is the *parameter error*:

$$\tilde{\theta} = \theta - \hat{\theta} \tag{5.2.5}$$

The derivative of $V_0(x) = \frac{1}{2}x^2$ becomes

$$\dot{V}_0 = -c_1 x^2 + \tilde{\theta} x \varphi(x) \tag{5.2.6}$$

Since the second term is indefinite and contains the unknown parameter error $\tilde{\theta}$, we can not conclude anything about the stability of (5.2.4). We make the controller dynamic with an update law for $\hat{\theta}$. To design this update law, we augment V_0 with a quadratic term in the parameter error $\tilde{\theta}$,

$$V_1(x, \tilde{\theta}) = \frac{1}{2}x^2 + \frac{1}{2\gamma}\tilde{\theta}^2 \tag{5.2.7}$$

where $\gamma > 0$ is the *adaptation gain*. The derivative of this function is

$$\begin{aligned} \dot{V}_1 &= x\dot{x} + \frac{1}{\gamma}\tilde{\theta}\dot{\tilde{\theta}} \\ &= -c_1 x^2 + \tilde{\theta} x \varphi(x) + \frac{1}{\gamma}\tilde{\theta}\dot{\tilde{\theta}} \\ &= -c_1 x^2 + \tilde{\theta}\left[x\varphi(x) + \frac{1}{\gamma}\dot{\tilde{\theta}}\right] \end{aligned} \tag{5.2.8}$$

The second term is still indefinite and contains $\tilde{\theta}$ as a factor. However, the situation is much better than in (5.2.6), because we now have the dynamics of $\dot{\tilde{\theta}} = -\dot{\hat{\theta}}$ at our disposal. With the appropriate choice of $\dot{\hat{\theta}}$ we can cancel the indefinite term. Thus, we choose the update law

$$\dot{\hat{\theta}} = -\dot{\tilde{\theta}} = \gamma x \varphi(x) \tag{5.2.9}$$

which yields

$$\dot{V}_1 = -c_1 x^2 \leq 0 \tag{5.2.10}$$

The resulting adaptive system consists of (5.2.1) with the control (5.2.3) and the update law (5.2.9), and is shown in Figure 5.2. In Figure 5.3, this system is redrawn in its closed-loop form consisting of (5.2.4) and (5.2.9), namely

$$\begin{aligned} \dot{x} &= -c_1 x + \tilde{\theta}\varphi(x) \\ \dot{\tilde{\theta}} &= -\gamma x \varphi(x) \end{aligned} \tag{5.2.11}$$

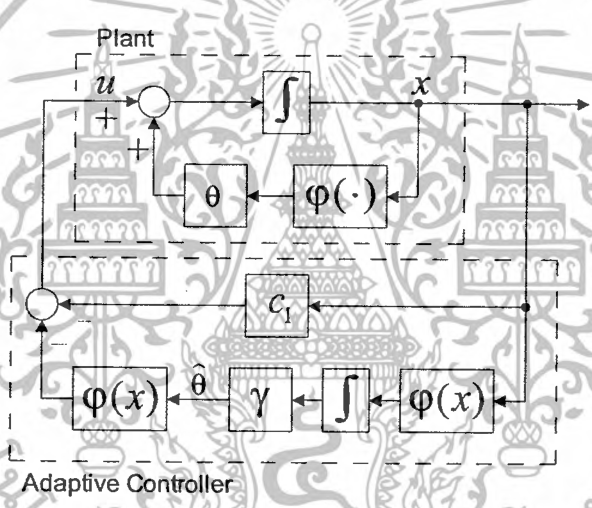

Figure 5.2: The closed-loop adaptive system (5.2.11)

Because $\dot{V}_1 \leq 0$, the equilibrium $x = 0$, $\tilde{\theta} = 0$ of (5.2.11) is globally stable. In addition, the desired regulation property $\lim_{t\to\infty} x(t) = 0$ follows from the LaSalle-Yoshizawa theorem. The adaptive nonlinear controller which guarantees these properties is given by (5.2.4) and (5.2.9):

$$\begin{aligned} u &= -c_1 x - \hat{\theta}\varphi(x) \\ \dot{\hat{\theta}} &= \gamma x \varphi(x) \end{aligned} \tag{5.2.12}$$

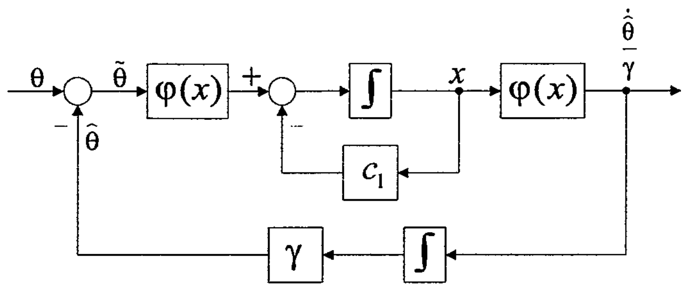

Figure 5.3: An equivalent representation of (5.2.11)

One may think that the above adaptive design is so straightforward because (5.2.1) is a first-order system. In fact, this is due to the *matching condition*: The terms containing unknown parameters in (5.2.1) are in the span of the control, that is, they can be directly cancelled by u when θ is known. To illustrate this point, let us consider the following second-order system, where again the uncertain term is "matched" by the control u:

$$\begin{aligned} \dot{x}_1 &= x_2 + \varphi_1(x_1) \\ \dot{x}_2 &= \theta\varphi_2(x) + u \end{aligned} \tag{5.2.13}$$

If θ were known, we would be able to use a static design procedures as mentioned in preceding sections. First view x_2 as the virtual control, design the stabilizing function

$$\alpha_1(x_1) = -c_1x_1 - \varphi_1(x_1) \tag{5.2.14}$$

and then form the Lyapunov function

$$V_c(x) = \frac{1}{2}x_1^2 + \frac{1}{2}(x_2 - \alpha_1(x_1))^2 \tag{5.2.15}$$

whose derivative is rendered negative definite

$$\dot{V}_c(x) = -c_1x_1^2 - c_2(x_2 - \alpha_1)^2 \tag{5.2.16}$$

by the control

$$u = -c_2(x_2 - \alpha_1) - x_1 + \frac{\partial \alpha_1}{\partial x_1}(x_2 + \varphi_1) - \theta\varphi_2(x) \tag{5.2.17}$$

Since θ is unknown, we again replace it with its estimate $\hat{\theta}$ in (5.2.17) to obtain the adaptive control law:

$$u = -c_2(x_2 - \alpha_1) - x_1 + \frac{\partial \alpha_1}{\partial x_1}(x_2 + \varphi_1) - \hat{\theta}\varphi_2(x) \tag{5.2.18}$$

This results in the error system ($z_1 = x_1, z_2 = x_2 - \alpha_1$):

$$\frac{d}{dt}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} -c_1 & 1 \\ -1 & -c_2 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \varphi_2(x) \end{bmatrix}\tilde{\theta} \tag{5.2.19}$$

Then we augment (5.2.15) with a quadratic term in the parameter error $\tilde{\theta}$ to obtain the Lyapunov function:

$$V_1(z, \tilde{\theta}) = V_c + \frac{1}{2\gamma}\tilde{\theta}^2 = \frac{1}{2}z_1^2 + \frac{1}{2}z_2^2 + \frac{1}{2\gamma}\tilde{\theta}^2 \tag{5.2.20}$$

Its derivative is

$$\dot{V}_1 = -c_1 z_1^2 - c_2 z_2^2 + \tilde{\theta}\left[z_2\varphi_2 - \frac{1}{\gamma}\dot{\hat{\theta}}\right] \tag{5.2.21}$$

The choice of update law

$$\dot{\hat{\theta}} = \gamma\varphi_2 z_2 \tag{5.2.22}$$

eliminates the $\tilde{\theta}$-term in (5.2.21) and renders the derivative of the Lyapunov function (5.2.20) nonpositive:

$$\dot{V}_1 = -c_1 z_1^2 - c_2 z_2^2 \leq 0 \tag{5.2.23}$$

This implies that the $z = 0, \tilde{\theta} = 0$ equilibrium point of the closed-loop adaptive system consisting of (5.2.19) and (5.2.22) (see block diagram in Figure 5.4)

$$\begin{aligned} \frac{d}{dt}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} &= \begin{bmatrix} -c_1 & 1 \\ -1 & -c_2 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \varphi_2(x) \end{bmatrix}\tilde{\theta} \\ \dot{\tilde{\theta}} &= -\gamma\begin{bmatrix} 0 & \varphi_2 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \end{aligned} \tag{5.2.24}$$

is globally stable and, in addition, $x(t) \to 0$ as $t \to \infty$.

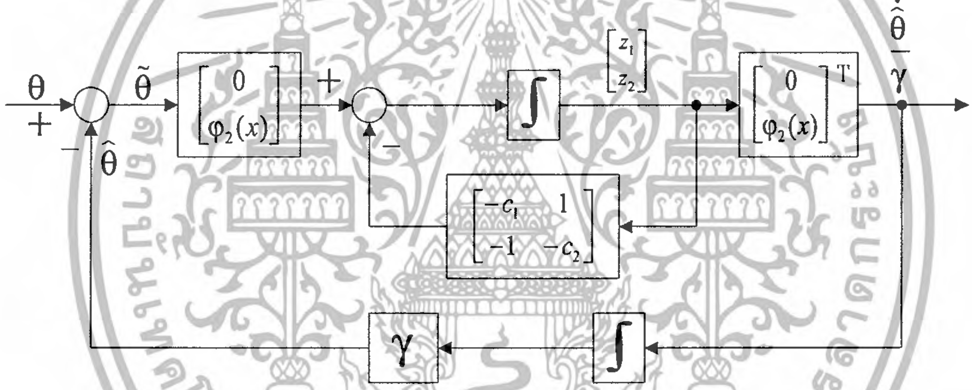

Figure 5.4: The closed-loop adaptive system (5.2.24)

Chapter 6

Adaptive Backstepping Designs

6.1 Adaptive integrator backstepping

The adaptive design in the above examples was simple because of the matching: The parametric uncertainty was in the span of the control. We now move to the more general case of *extended matching*, where the parametric uncertainty enters the system one integrator before the control does:

$$\dot{x}_1 = x_2 + \theta\varphi(x_1) \tag{6.1.1-a}$$

$$\dot{x}_2 = u \tag{6.1.1-b}$$

We use this example to introduce *adaptive backstepping.*

Again, if θ were known, we would design the stabilizing function for x_2 as doing with the static part

$$\alpha_1(x_1, \theta) = -c_1 x_1 - \theta\varphi(x_1) \tag{6.1.2}$$

With the Lyapunov function

$$V_c(x, \theta) = \frac{1}{2}x_1^2 + \frac{1}{2}(x_2 - \alpha_1(x_1, \theta))^2 \tag{6.1.3}$$

whose derivative is rendered negative definite

$$\dot{V}_c(x,\theta) = -c_1x_1^2 - c_2(x_2 - \alpha_1(x_1,\theta))^2 \tag{6.1.4}$$

by the control

$$u = -c_2(x_2 - \alpha_1(x_1,\theta)) - x_1 + \frac{\partial\alpha_1}{\partial x_1}(x_2 + \theta\varphi) \tag{6.1.5}$$

Since θ is unknown and appears one equation before the control does, this old method is no longer available because the dependence of $\alpha_1(x_1) = -c_1x_1 - \theta\varphi(x_1)$ on the unknown parameter makes it impossible to continue the procedure. However, we can utilize the idea of integrator backstepping.

Step 1. If x_2 were the control, an adaptive controller for (6.1.1-a) would be given by (5.2.12):

$$\begin{aligned} \alpha_1(x_1,\vartheta_1) &= -c_1z_1 - \vartheta_1\varphi(x_1) && (6.1.6)\\ \dot{\vartheta}_1 &= \gamma z_1\varphi(x_1) && (6.1.7)\end{aligned}$$

In the above equations we have replaced the parameter estimate $\hat{\theta}$ with the estimate ϑ_1, which denotes the estimate generated in this design step. As we will see, there will be another estimate generated in the next step. With (6.1.6) and the new error variable $z_2 = x_2 - \alpha_1$, the $\dot{z}_1$-equation becomes

$$\dot{z}_1 = -c_1z_1 + z_2 + (\theta - \vartheta_1)\varphi \tag{6.1.8}$$

The derivative of the Lyapunov function

$$V_1(x,\vartheta_1) = \frac{1}{2}z_1^2 + \frac{1}{2\gamma}(\theta - \vartheta_1)^2 \tag{6.1.9}$$

along the solutions of (6.1.8) is

$$\begin{aligned} \dot{V}_1 &= z_1\dot{z}_1 - \frac{1}{\gamma}(\theta - \vartheta_1)\dot{\vartheta}_1 && (6.1.10)\\ &= z_1z_2 - c_1z_1^2 + (\theta - \vartheta_1)\left[\varphi_1z_1 - \frac{1}{\gamma}\dot{\vartheta}_1\right] \\ &= z_1z_2 - c_1z_1^2 \end{aligned}$$

Step 2. The derivative of z_2 is now expressed as

$$\begin{aligned} \dot{z}_2 &= \dot{x}_2 - \dot{\alpha}_1 \\ &= u - \frac{\partial \alpha_1}{\partial x_1}\dot{x}_1 - \frac{\partial \alpha_1}{\partial \vartheta_1}\dot{\vartheta}_1 \end{aligned}$$

Substituting (6.1.1-a) and the update law (6.1.7) results in

$$\begin{aligned} \dot{z}_2 &= \tfrac{\partial \alpha_1}{\partial x_1}(x_2 + \theta\varphi) - \tfrac{\partial \alpha_1}{\partial \vartheta_1}\gamma\varphi z_1 \\ &= u - \tfrac{\partial \alpha_1}{\partial x_1}x_2 - \tfrac{\partial \alpha_1}{\partial \vartheta_1}\gamma\varphi z_1 - \theta\tfrac{\partial \alpha_1}{\partial x_1}\varphi \end{aligned} \tag{6.1.11}$$

At this point we need to select a Lyapunov function and design u to render its derivative nonpositive. Our first attempt is the augmented Lyapunov function

$$V_2(z_1, z_2, \vartheta_1) = V_1(z_1, \vartheta_1) + \frac{1}{2}z_2^2$$

whose derivative, using (6.1.10) and (6.1.11), is

$$\begin{aligned} \dot{V}_2 &= \dot{V}_1 + z_2\dot{z}_2 \\ &= -c_1 z_1^2 + z_2\left[z_1 + u - \frac{\partial \alpha_1}{\partial x_1}x_2 - \frac{\partial \alpha_1}{\partial \vartheta_1}\gamma\varphi z_1 - \theta\frac{\partial \alpha_1}{\partial x_1}\varphi\right] \end{aligned}$$

The control u should now be able to cancel the indefinite terms in $\dot{V}_2$. To deal with the terms containing the unknown parameter θ, we will try to employ the existing estimate ϑ_1:

$$u = -z_1 - c_2 z_2 + \frac{\partial \alpha_1}{\partial x_1}x_2 + \frac{\partial \alpha_1}{\partial \vartheta_1}\gamma\varphi z_1 + \theta\frac{\partial \alpha_1}{\partial x_1}\varphi_1$$

From the resulting derivative

$$\dot{V}_2 = -c_1 z_1^2 - c_2 z_2^2 - (\theta - \vartheta_1)\frac{\partial \alpha_1}{\partial x_1}\varphi_1 z_2$$

we see that we have no design freedom left to cancel the $(\theta - \vartheta_1)$-term. To overcome this difficulty, we replace ϑ_1 in the expression for u with a *new* estimate ϑ_2:

$$u = -z_1 - c_2 z_2 + \frac{\partial \alpha_1}{\partial x_1}x_2 + \frac{\partial \alpha_1}{\partial \vartheta_1}\gamma\varphi z_1 + \theta\frac{\partial \alpha_1}{\partial x_1}\varphi \tag{6.1.12}$$

With the choice (6.1.12), the z_2-equation becomes

$$\dot{z}_2 = -c_2 z_2 - z_1 - (\theta - \vartheta_2)\frac{\partial \alpha_1}{\partial x_1}\varphi \tag{6.1.13}$$

The presence of the new parameter estimate ϑ_2 suggests the following augmentation of the Lyapunov function:

$$\begin{aligned} V_2(z_1, z_2, \vartheta_1, \vartheta_2) &= V_1 + \frac{1}{2}z_2^2 + \frac{1}{2\gamma}(\theta - \vartheta_2)^2 \\ &= \frac{1}{2}(z_1^2 + z_2^2) + \frac{1}{2\gamma}\left[(\theta - \vartheta_1)^2 + (\theta - \vartheta_2)^2\right] \end{aligned} \tag{6.1.14}$$

The derivative of V_2 is

$$\begin{aligned} \dot{V}_2 &= \dot{V}_1 + z_2\dot{z}_2 - \frac{1}{\gamma}(\theta - \vartheta_2)\dot{\vartheta}_2 \\ &= z_1 z_2 - c_1 z_1^2 + z_2\left[-c_2 z_2 - z_1 - (\theta - \vartheta_2)\frac{\partial \alpha_1}{\partial x_1}\varphi\right] - \frac{1}{\gamma}(\theta - \vartheta_2)\dot{\vartheta}_2 \\ &= -c_1 z_1^2 - c_2 z_2^2 - (\theta - \vartheta_2)\left(\frac{\partial \alpha_1}{\partial x_1}\varphi z_2 + \frac{1}{\gamma}\dot{\vartheta}_2\right) \end{aligned} \tag{6.1.15}$$

Now the $(\theta - \vartheta_2)$-term can be eliminated with the update law

$$\dot{\vartheta}_2 = -\gamma\frac{\partial \alpha_1}{\partial x_1}\varphi z_2 \tag{6.1.16}$$

which yields

$$\dot{V}_2 = -c_1 z_1^2 - c_2 z_2^2 \tag{6.1.17}$$

The equations (6.1.13) and (6.1.16) along with (6.1.8) and (6.1.7) form the error system representation of the resulting closed-loop adaptive system:

$$\begin{aligned} \dot{z}_1 &= -c_1 z_1 + z_2 + (\theta - \varphi_1)\varphi \\ \dot{z}_2 &= -c_2 z_2 - z_1 - (\theta - \vartheta_2)\tfrac{\partial \alpha_1}{\partial x_1}\varphi \\ \dot{\vartheta}_1 &= \gamma\varphi z_1 \\ \dot{\vartheta}_2 &= -\gamma\tfrac{\partial \alpha_1}{\partial x_1}\varphi z_2 \end{aligned} \tag{6.1.18}$$

The matrix form of this system,

$$\begin{aligned} \frac{d}{dt}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} &= \begin{bmatrix} -c_1 & 1 \\ -1 & -c_2 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} \varphi & 0 \\ 0 & -\frac{\partial \alpha_1}{\partial x_1}\varphi \end{bmatrix}\begin{bmatrix} \theta - \vartheta_1 \\ \theta - \vartheta_2 \end{bmatrix} \\ \frac{d}{dt}\begin{bmatrix} \vartheta_1 \\ \vartheta_2 \end{bmatrix} &= \gamma\begin{bmatrix} \varphi & 0 \\ 0 & -\frac{\partial \alpha_1}{\partial x_1}\varphi \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \end{aligned} \tag{6.1.19}$$

makes its properties more visible:

- The constant system matrix has negative terms along its diagonal, while its off-diagonal terms are skew-symmetric, and
- The matrix that multiplies the parameter errors in the $\dot{z}$-equation is used in the update laws for the parameter estimates.

The stability properties of (6.1.19) follow from ((6.1.14) and (6.1.17): The LaSalle-Yoshizawa theorem establishes that $z_1, z_2, \vartheta_1, \vartheta_2$ are bounded, and $z \to 0$ as $t \to \infty$. Since $z_1 = x_1$, x_1 is also bounded and converges to zero. The bounedness of x_2 then follows from the boundedness of α_1 (defined in (6.1.6)) and the fact that $x_2 = z_2 + \alpha_1$. Using (6.1.12) we conclude that the control u is also bounded. Finally, we note that the regulation of z and x_1 does not imply the regulation of x_2: From $z_2 = x_2 - \alpha_1$and (6.1.6) we see that $x_2 + \vartheta_1\varphi(0)$ will converge to zero. Thus, x_2 is not guaranteed to converge to zero unless $\varphi(0) = 0$. However, x_2 will converge to a constant value:

$$\lim_{t\to\infty} = -\theta\varphi(0) \triangleq x_2^e \tag{6.1.20}$$

This can be seen from (6.1.1-a): Since x_1 and $\dot{x}_1$ converge to zero, so does $x_2 + \theta\varphi(0)$.

With the above example we have illustrated the idea of adaptive backstepping. To formulate it as a design tool analogous to an integrator backstepping design, we start with the assumption that an adaptive controller is known for an initial system.

Assumption 6.1.1 *Consider the system*

$$\dot{x} = f(x) + F(x)\theta + g(x)u \tag{6.1.21}$$

where $x \in \mathbb{R}^n$ is the state, $\theta \in \mathbb{R}^p$ is a vector of unknown constant parameters, *and $u \in \mathbb{R}$ is the control input. There exist an adaptive controller*

$$\begin{aligned} u &= \alpha(x, \vartheta) \\ \dot{\vartheta} &= T(x, \vartheta) \end{aligned} \tag{6.1.22}$$

with parameter estimate $\vartheta \in \mathbb{R}^q$, and a smooth function $V(x, \vartheta) : \mathbb{R}^{n+q} \to \mathbb{R}$ which is positive definite and radially unbounded in the variables $(x, \vartheta - \theta)$ such that for all $(x, \vartheta) \in \mathbb{R}^{n+q}$:

$$\begin{aligned} \frac{\partial V}{\partial x}(x, \vartheta)[f(x) + F(x)\theta g(x)\alpha(x, \vartheta)] + \frac{\partial V}{\partial \vartheta}(x, \vartheta)T(x, \vartheta) \\ \leq -W(x, \vartheta) \leq 0 \end{aligned} \tag{6.1.23}$$

where $W : \mathbb{R}^{n+q} \to \mathbb{R}$ is positive semidefinite. □

Under this assumption, the control (6.1.22), applied to the system (6.1.21), guarantees global boundedness of $x(t)$, $\vartheta(t)$ and, by the LaSalle-Yoshizawa theorem (Theorem 2.1), regulation of $W(x(t), \vartheta(t))$. Adaptive backstepping allows us to achieve the same properties for the augmented system.

Lemma 6.1.2 *[Adaptive Backstepping] Let the system (6.1.21) be augmented by an integrator,*

$$\begin{aligned} \dot{x} &= f(x) + F(x)\theta + g(x)\xi \\ \dot{\xi} &= u \end{aligned} \tag{6.1.24}$$

where $\xi \in \mathbb{R}$. Consider for this system the dynamic feedback controller

$$\begin{aligned} u = & -c(\xi - \alpha(x, \vartheta)) + \frac{\partial \alpha}{\partial x}(x, \vartheta)[f(x) + F(x)\bar{\vartheta} + g(x)\xi] + \frac{\partial \alpha}{\partial \vartheta}T(x, \vartheta) \\ & -\frac{\partial V}{\partial x}(x, \vartheta)g(x), \quad c > 0 \end{aligned} \tag{6.1.25}$$

$$\dot{\vartheta} = T(x, \vartheta) \tag{6.1.26}$$

$$\dot{\bar{\vartheta}} = -\Gamma\left[\frac{\partial \alpha}{\partial x}(x, \vartheta)F(x)\right]^T (\xi - \alpha(x, \vartheta)) \tag{6.1.27}$$

where $\bar{\vartheta}$ is a new estimate of θ, $\Gamma = \Gamma^T > 0$ is the adaptation gain matrix. Under Assumption 6.1.1, this adaptive controller guarantees global boundedness of $x(t)$,

$\xi(t)$, $\vartheta(t)$, $\bar{\vartheta}(t)$ and regulation of $W(x(t),\vartheta(t))$ and $\xi(t)-\alpha(x(t),\vartheta(t))$. These properties can be established with the Lyapunov function

$$V_a(x,\xi,\vartheta,\bar{\vartheta}) = V(x,\vartheta) + \frac{1}{2}[\xi-\alpha(x,\vartheta)]^2 + \frac{1}{2}(\theta-\bar{\vartheta})^T\Gamma^{-1}(\theta-\bar{\vartheta}) \quad (6.1.28)$$

Proof *With the error variable $z=\xi-\alpha(x,\vartheta)$,* (6.1.24) *is rewritten as*

$$\begin{aligned} \dot{x} &= f(x)+F(x)\theta+g(x)[\alpha(x,\vartheta)+z] && (6.1.29)\\ \dot{z} &= u-\frac{\partial\alpha}{\partial x}[f(x)+F(x)\theta+g(x)(\alpha(x,\vartheta)+z)]-\frac{\partial\alpha}{\partial\vartheta}T(x,\vartheta) && (6.1.30) \end{aligned}$$

Note that in (6.1.30) *the derivative of ϑ was replaced by the update law* (6.1.26). *Introducing a new parameter estimate $\bar{\vartheta}$, we augment the Lyapunov function:*

$$V_a(x,\xi,\vartheta,\bar{\vartheta}) = V(x,\vartheta) + \frac{1}{2}z^2 + \frac{1}{2}(\theta-\bar{\vartheta})^T\Gamma^{-1}(\theta-\bar{\vartheta}) \quad (6.1.2)$$

Using (6.1.23), *it is easy to show that the derivative of* (6.1.2) *satisfies*

$$\begin{aligned} \dot{V}_a &= \frac{\partial V}{\partial x}(f+F\theta+g\alpha+gz)+\frac{\partial V}{\partial\vartheta}T\\ &\quad +z\left[u-\frac{\partial\alpha}{\partial x}(f+F\theta+g(\alpha+z))-\frac{\partial\alpha}{\partial\vartheta}T\right]-\dot{\bar{\vartheta}}^T\Gamma^{-1}(\theta-\bar{\vartheta})\\ &= \frac{\partial V}{\partial x}(f+F\theta+g\alpha)+\frac{\partial V}{\partial\vartheta}T\\ &\quad +z\left[u-\frac{\partial\alpha}{\partial x}(f+F\theta+g(\alpha+z))-\frac{\partial\alpha}{\partial\vartheta}T+\frac{\partial V}{\partial x}g\right]-\dot{\bar{\vartheta}}^T\Gamma^{-1}(\theta-\bar{\vartheta})\\ &\le -W(x,\vartheta)+z\left[u-\frac{\partial\alpha}{\partial x}(f+F\vartheta+g(\alpha+z))-\frac{\partial\alpha}{\partial\vartheta}T+\frac{\partial V}{\partial x}g\right]\\ &\quad -\left[\frac{\partial\alpha}{\partial x}Fz+\dot{\bar{\vartheta}}^T\Gamma^{-1}\right](\theta-\bar{\vartheta}) \qquad (6.1.3) \end{aligned}$$

The $(\theta-\bar{\vartheta})$-term is now eliminated with the update law

$$\dot{\bar{\vartheta}} = -\Gamma\left(\frac{\partial\alpha}{\partial x}F\right)^T z \quad (6.1.4)$$

and the control (6.1.25) *is chosen to make the bracketed term multiplying z in* (6.1.3) *equal to $-cz$:*

$$u = -cz + \frac{\partial\alpha}{\partial x}(f+F\bar{\vartheta}+g(\alpha+z)) + \frac{\partial\alpha}{\partial\vartheta}T - \frac{\partial V}{\partial x}g \quad (6.1.5)$$

This results in the desired nonpositivity of $\dot{V}_a$*:*

$$\dot{V}_a \leq -W(x,\vartheta) - cz^2 \leq 0 \tag{6.1.6}$$

From (6.1.28) *and* (6.1.6) *we conclude that* $V(x,\vartheta)$, $\bar{\vartheta}$ *and* z *are bounded. By Assumption 6.1.1, this means that* $x(t)$ *and* $\vartheta(t)$ *are bounded. Hence,* $\xi = z + \alpha(x,\vartheta)$ *and* u *are bounded. By LaSalle-Yoshizawa Theorem, the boundedness of all the signals combined with* (6.1.6) *proves the regulation of* $W(x(t)), \vartheta(t)$ *and* $z(t)$. □

6.2 Recursive design procedures

Through repeated application of adaptive backstepping, the design procedure is now generalized to nonlinear systems which can be transformed into the *parametric strict-feedback form*

$$\begin{aligned}
\dot{x}_1 &= x_2 + \varphi_1^T(x_1)\theta \\
\dot{x}_2 &= x_3 + \varphi_2^T(x_1,x_2)\theta \\
&\;\;\vdots \\
\dot{x}_{n-1} &= x_n + \varphi_{n-1}^T(x_1,\ldots,x_{n-1})\theta \\
\dot{x}_n &= \beta(x)u + \varphi_n^T(x)\theta
\end{aligned} \tag{6.2.1}$$

where $\beta(x) \neq 0$ for all $x \in \mathbb{R}^n$ The reason for the name "parametric strict-feedback" can be deduced from the block diagram in Figure 6.1, where, except for the integrators, there are only feedback paths. This implies that the nonlinearities depend only on variables which are "fed back."

For systems in the form (6.2.1), the number of design steps required is equal to the degree n of the system. At each step, an error variable z_i, a stabilizing function α_i, and a parameter estimate ϑ_i are generated. As a result, if a system

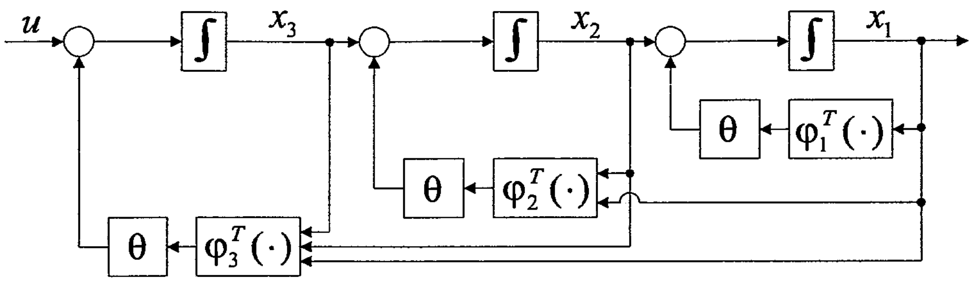

Figure 6.1: Block diagram of a third-order parametric strict-feedback system with $\beta(x) = 1$.

contains p unknown parameters, the overparametrized adaptive controller may employ as many as pn parameter estimates. A schematic representation of this design procedure is given in Figure 6.2.

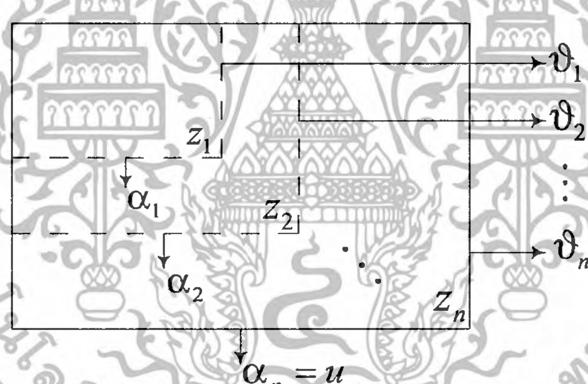

Figure 6.2: The design procedure for overparametrized schemes.

6.3 Extended Matching Design

In the designing of adaptive nonlinear controllers using recursive design procedures, each step or each level of state analysis genetates one error variable $\dot{z}_i$ and one update law $\dot{\vartheta}_1$ that cause the increase in the number of parameter estimates, called an overparametrization. This could be an undesirable feature, since it rapidly increased the dynamic order of the resulting adaptive controller. When the uncertain parameters are only one integrator away from the control, called an "extended matching" condition, we are now present a method which can avoid an overparametrization. The idea of applying an extended matching design is to avoid choosing an update law in step 1, but postpone this procedure to step 2 which cause a $\dot{\hat{\theta}}$ to be appear. Then we choose an appropriate parameter update law $\dot{\hat{\theta}}$ which can eliminate the error parameter term ($\tilde{\theta}$) as shown in the design procedures.

Again, we consider the nonlinear system (6.1.1-a)

$$\begin{aligned} \dot{x}_1 &= x_2 + \theta\varphi(x_1) \\ \dot{x}_2 &= u \end{aligned}$$

and modify its two-step design.

Step 1. With $z_1 = x_1$ and x_2 viewed as the virtual control in the $\dot{z}_1$-equation, we define the first stabilizing function α_1 as in (6.1.6)

$$\alpha_1 = -c_1 z_1 - \hat{\theta}\varphi \tag{6.3.1}$$

Comparing (6.3.1) with (6.1.6), we see that the parameter estimate ϑ_1 has been replaced by the parameter estimate ϑ_1. The difference in notation indicates that in this design procedure only one estimate $\hat{\theta}$ of the unknown parameter will be used.

The first Lyapunov function is now chosen as

$$V_1(z_1, \hat{\theta}) = \frac{1}{2}z_1^2 + \frac{1}{2\gamma}\tilde{\theta}^2 \tag{6.3.2}$$

where $\tilde{\theta} = \theta - \hat{\theta}$ is the parameter error, and $\gamma > 0$ is the adaptation gain. With $z_2 = x_2 - \alpha_1$, the derivative of V_1 is

$$\dot{V}_1 = z_1 z_2 - c_1 z_1^2 + \tilde{\theta}\left(\varphi z_1 - \frac{1}{\gamma}\dot{\hat{\theta}}\right) \tag{6.3.3}$$

We postpone the choice of update law for $\hat{\theta}$ until the next step. The first error subsystem becomes

$$\dot{z}_1 = -c_1 z_1 + z_2 + \tilde{\theta}\varphi \tag{6.3.4}$$

Step 2. The derivative of $z_2 = x_2 - \alpha_1$ is

$$\begin{aligned} \dot{z}_2 &= u - \frac{\partial \alpha_1}{\partial x_1}(x_2 + \theta\varphi) - \frac{\partial \alpha_1}{\partial \hat{\theta}}\dot{\hat{\theta}} \\ &= u - \frac{\partial \alpha_1}{\partial x_1}x_2 - \hat{\theta}\frac{\partial \alpha_1}{\partial x_1}\varphi - \tilde{\theta}\frac{\partial \alpha_1}{\partial x_1}\varphi - \frac{\partial \alpha_1}{\partial \hat{\theta}}\dot{\hat{\theta}} \end{aligned} \tag{6.3.5}$$

To design the control u, we consider the augmented Lyapunov function

$$V_2 = V_1 + \frac{1}{2}z_2^2 = \frac{1}{2}z_1^2 + \frac{1}{2}z_2^2 + \frac{1}{2\gamma}\tilde{\theta}^2 \tag{6.3.6}$$

The only difference between (6.3.6) and (6.1.14) is the absence of the new parameter error $(\theta - \vartheta_2)$ in (6.3.6). In view of (6.3.3) and (6.3.5), the derivative of V_2 is

$$\begin{aligned} \dot{V}_2 &= z_1 z_2 - c_1 z_1^2 + \tilde{\theta}\left(\varphi z_1 - \frac{1}{\gamma}\dot{\hat{\theta}}\right) \\ &\quad + z_2\left[u - \frac{\partial \alpha_1}{\partial x_1}x_2 - \hat{\theta}\frac{\partial \alpha_1}{\partial x_1}\varphi - \tilde{\theta}\frac{\partial \alpha_1}{\partial x_1}\varphi - \frac{\partial \alpha_1}{\partial \hat{\theta}}\dot{\hat{\theta}}\right] \\ &= -c_1 z_1^2 + \tilde{\theta}\left[\varphi z_1 - z_2\frac{\partial \alpha_1}{\partial x_1}\varphi - \frac{1}{\gamma}\dot{\hat{\theta}}\right] \\ &\quad + z_2\left[z_1 + u - \frac{\partial \alpha_1}{\partial x_1}x_2 - \hat{\theta}\frac{\partial \alpha_1}{\partial x_1}\varphi - \frac{\partial \alpha_1}{\partial \hat{\theta}}\dot{\hat{\theta}}\right] \end{aligned} \tag{6.3.7}$$

In the last equation, all the terms containing $\tilde{\theta}$ have been grouped together. To eliminate them, the update law is chosen as

$$\dot{\hat{\theta}} = \gamma\left(\varphi z_1 - \frac{\partial \alpha_1}{\partial x_1}\varphi z_2\right) \tag{6.3.8}$$

Then, the last bracketed term in (6.3.7) will be rendered equal to $-c_2 z_2^2$ with the control

$$u = -z_1 - c_2 z_2 + \frac{\partial \alpha_1}{\partial x_1}x_2 + \hat{\theta}\frac{\partial \alpha_1}{\partial x_1}\varphi + \frac{\partial \alpha_1}{\partial \hat{\theta}}\dot{\hat{\theta}} \tag{6.3.9}$$

where for $\dot{\hat{\theta}}$ we use the analytical expression of the update law (6.3.8). Substituting the expressions (6.3.8) and (6.3.9) into (6.3.7) we obtain

$$\dot{V}_2 = -c_1 z_1^2 - c_2 z_2^2 \leq 0 \tag{6.3.10}$$

and the error system becomes (see block diagram in Figure 6.3)

$$\begin{aligned} \frac{d}{dt}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} &= \begin{bmatrix} -c_1 & 1 \\ -1 & -c_2 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} \varphi \\ -\frac{\partial \alpha_1}{\partial x_1}\varphi \end{bmatrix}\tilde{\theta} \\ \dot{\hat{\theta}} &= -\gamma \begin{bmatrix} \varphi & -\frac{\partial \alpha_1}{\partial x_1}\varphi \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \end{aligned} \tag{6.3.11}$$

We can say that the extended matching design is unique and only applicable with level-one systems, because the result $\dot{\hat{\theta}}$ perfectly matches the desired update law. In the case of level two, said that the control is three steps away, postponing the choosing of parameter update law to step 3 cause a $\ddot{\hat{\theta}}$ to be appear which is not available to acquire such that derivative from the system. The next chapter introduces *tuning function design* based on the idea of recursive design procedures. This method can avoid an overparametrization of the system which has more than one levels of uncertainties. It also provides characteristics of regulation and tracking which are the advantage of applying adaptive nonlinear controllers.

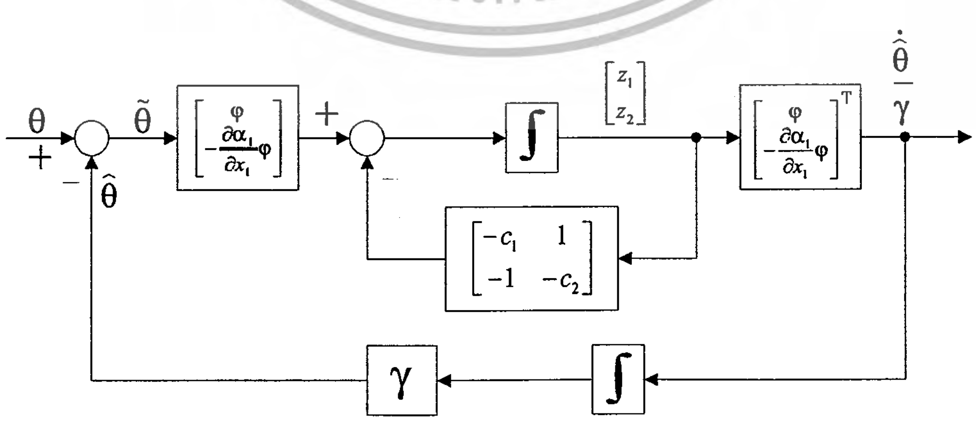

Figure 6.3: The closed-loop adaptive system (6.3.11).

Chapter 7

Tuning Functions Design

7.1 Introduction

The adaptive backstepping designs for a plant with unknown parameters is a starting point for more elaborate adaptive designs which lead to new properties of the designed controller and the resulting feedback system. One of the improvement to be achieved with the tuning functions design in this chapter is the reduction of the dynamic order of the adaptive controller to its minimum: The number of parameter estimates is equal to the number of unknown parameters. This minimum-order design is advantageous not only for implementation, but also because it guarantees the strongest achievable stability and convergence properties.

In the tuning functions procedure the parameter update law is designed recursively. At each consecutive step, we design a tuning function as a potential update law. In contrast to adaptive backstepping, these intermediate update laws are not implemented. Instead, the controller uses them to compensate for the effect of parameter estimation transients. Only the final tuning function is used as the parameter update law.

7.2 Adaptive Control Lyapunov Functions

The basic idea of the Lyapunov approach to adaptive control is to design a control law and a parameter update law to guarantee that the derivative of a suitable Lyapunov function is nonpositive. We are therefore sent to search for a tripple: Lyapunov function, control law, and update law. For a class of nonlinear systems called parametric-strict-feedback systems we will be able to make this search systematic.

To begin with, let us investigate the possibility of adaptive design for the system

$$\dot{x} = f(x) + F(x)\theta + g(x)u, \quad x \in \mathbb{R}^n, \quad u \in \mathbb{R} \tag{7.2.1}$$

where $\theta \in \mathbb{R}^p$ is a vector of unknown constant parameters, and $f(x), F(x)$ and $g(x)$ are smooth. For simplicity let $f(0) = 0$, $F(0) = 0$, so that $x = 0$ is an equilibrium of the uncontrolled plant.

7.2.1 Departure from certainty equivalence

Much of the traditional adaptive control employs some form of "certainty equivalence" thinking. Following this path one first performs a design for the case when the exact value of θ is known. Suppose that this nontrivial task is completed and that its result is a feedback control $u = \alpha_c(x, \theta)$ which stabilizes the equilibrium $x = 0$ with respect to a known Lyapunov function $V_c(x, \theta)$. The subscript 'c' stands for "certainty equivalence". We know that $V_c(x, \theta)$ is positive definite and radially unbounded in x for all θ, and that there exists a function $W(x, \theta)$, which is also positive definite in x for all θ, such that

$$\frac{\partial V_c}{\partial x}[f(x) + F(x)\theta + g(x)\alpha_c(x, \theta)] \leq -W(x, \theta) \tag{7.2.2}$$

How can we exploit the knowledge of $\alpha_c(x,\theta)$ and $V_c(x,\theta)$ for adaptive design when θ is not known? The certainty equivalence idea is to replace θ by an estimate $\hat{\theta}(t)$ obtained from a parameter update law

$$\dot{\hat{\theta}} = \Gamma\tau(x,\hat{\theta}) \tag{7.2.3}$$

where the adaptation gain matrix Γ is positive definite. We want to select u and τ to guarantee that the derivative of a Lyapunov function is nonpositive. For the system (7.2.1),(7.2.3), a Lyapunov function candidate is

$$V(x,\hat{\theta}) = V_c(x,\hat{\theta}) + \frac{1}{2}\tilde{\theta}^T\Gamma^{-1}\tilde{\theta} \tag{7.2.4}$$

where the "certainty equivalence" from of V_c is augmented by a term quadratic in the parameter estimation error

$$\tilde{\theta} = \theta - \hat{\theta} \tag{7.2.5}$$

Upon the substitution of $F(x)\theta = F(x)\hat{\theta} + F(x)\tilde{\theta}$, the derivative of $V(x,\hat{\theta})$ along the solutions of (7.2.1), (7.2.3) is

$$\dot{V} = \frac{\partial V_c}{\partial x}[f(x) + F(x)\hat{\theta} + g(x)u] + \frac{\partial V_c}{\partial x}\Gamma\tau + \tilde{\theta}^T\left(\frac{\partial V_c}{\partial x}F(x)\right)^T - \tilde{\theta}^T\tau \tag{7.2.6}$$

To eliminate the indefinite dependence of $\dot{V}$ on the unknown parameter error $\tilde{\theta}$, we select τ to cancel the last two terms in (7.2.6):

$$\tau(x,\hat{\theta}) = \left(\frac{\partial V_c}{\partial x}F(x)\right)^T \tag{7.2.7}$$

With this choice of $/ta$, the expression (7.2.6) is reduced to

$$\dot{V} = \frac{\partial V_c}{\partial x}[f(x) + F(x)\hat{\theta} + g(x)u] + \frac{\partial V_c}{\partial x}\Gamma\left(\frac{\partial V_c}{\partial x}F(x)\right)^T \tag{7.2.8}$$

Our next task is to select a control law $u = \alpha(x,\hat{\theta})$ to make $\dot{V}$ nonpositive. The "certainty equivalence" control $u = \alpha_c(x,\hat{\theta})$ fails to achieve this because then (7.2.2) and (7.2.8) yield

$$\dot{V} \leq -W(x,\hat{\theta}) + \frac{\partial V_c}{\partial \hat{\theta}}\Gamma\left(\frac{\partial V_c}{\partial x}F(x)\right)^T \tag{7.2.9}$$

Clearly, $\dot{V}$ is not nonpositive because a sign-indefinite term is added to $-W(x,\hat{x})$. In search of a better control law $\alpha(x,\hat{\theta})$, we augment $\alpha_c(x,\hat{\theta})$ by $\alpha_T(x,\hat{\theta})$,

$$\alpha(x,\hat{\theta}) = \alpha_c(x,\hat{\theta}) + \alpha_T(x,\hat{\theta}) \tag{7.2.10}$$

The substitution of (7.2.10) into (7.2.8) shows that the desired nonpositivity $\dot{V} \leq -W(x,\hat{\theta})$ will be achieved if α_T can be found to satisfy

$$\frac{\partial V_c}{\partial x} g(x)\alpha_T(x,\hat{\theta}) + \frac{\partial V_c}{\partial \hat{\theta}}\Gamma\left(\frac{\partial V_c}{\partial x}F(x)\right)^T = 0 \tag{7.2.11}$$

This condition for α_T demonstrates the difficulty of adaptive design for a general nonlinear system (7.2.1). It is easy to see that α_T satisfying (7.2.11) is unlikely to exist The scalar quantity $\frac{\partial V_c}{\partial x}g(x)$ may be zero at a set of points. Still, the condition (7.2.11) is of interest because of an important special case, which will be the starting point of our recursive design. The special case is the "extended matching" studied in the previous chapter. In this case, a smooth vector-valued function $\varphi : \mathbb{R}^{n+p} \to \mathbb{R}^p$ is known such that $\frac{\partial V_c}{\partial \hat{\theta}}$ can be factored as follows:

$$\frac{\partial V_c}{\partial \hat{\theta}} = \frac{\partial V_c}{\partial x} g(x)\varphi(x,\hat{\theta})^T \tag{7.2.12}$$

Then, irrespective of the zeros of $\frac{\partial V_c}{\partial x}g(x)$, an α_T which satisfies (7.2.11) is

$$\alpha_T(x,\hat{\theta}) = -\varphi(x,\hat{\theta})^T\Gamma\left(\frac{\partial V_c}{\partial x}F(x)\right)^T = -\varphi(x,\hat{\theta})^T\Gamma\tau(x,\hat{\theta}) \tag{7.2.13}$$

We observe that, in addition to its "certainty equivalence" part α_c, the adaptive control law α contains a part α_T which is proportional to τ, that is, to $\dot{\hat{\theta}}$ (see (7.2.3), (7.2.10), and (7.2.13)). In this way, the adaptive control law takes into account the parameter estimation transients. When the parameter estimate is constant, the control law reduces to the "certainty equivalence" control. Let eu examine an example of a system for which (7.2.12) is satisfied.

Example 7.2.1 *Consider the problem of designing an adaptive controller for the system*

$$\begin{aligned} \dot{x}_1 &= x_2 + \varphi(x_1)^T\theta \\ \dot{x}_2 &= u \end{aligned} \tag{7.2.14}$$

where $\theta = [\theta_1, \theta_2]^T$ *is an unknown constant parameter vector, and the vector-valued function* $\varphi(x_1) = [\varphi_1(x_1), \varphi_2(x_1)]^T$ *is known and smooth. We dealt with this system in the extended-matching design. If the parameter* θ *were known, backstepping would result in the* θ*-dependent change of coordinates*

$$\begin{aligned} z_1 &= x_1 \\ z_2 &= x_2 + \varphi(x_1)^T\theta + c_1x_1 \end{aligned} \tag{7.2.15}$$

and the control law

$$u = \alpha_c(x,\theta) = -z_1 - c_2z_2 - \left(\frac{\partial\varphi^T}{\partial x_1} + c_1\right)\left(x_2 + \varphi(x_1)^T\theta\right) \tag{7.2.16}$$

which $c_1, c_2 > 0$*, which results in the closed-loop system*

$$\dot{z} = Az, \quad A = \begin{bmatrix} -c_1 & 1 \\ -1 & -c_2 \end{bmatrix} \tag{7.2.17}$$

Due to the structure of A*, an appropriate Lyapunov function is*

$$v_c(x,\theta) = \frac{1}{2}z(x,\theta)^T z(x,\theta) \tag{7.2.18}$$

Observing from (7.2.1) *and* (7.2.14) *that*

$$f(x) = \begin{bmatrix} x_2 \\ 0 \end{bmatrix}, \quad F(x) = \begin{bmatrix} \varphi x_1{}^T \\ 0 \end{bmatrix}, \quad g(x) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{7.2.19}$$

and evaluating

$$\frac{\partial V_c}{\partial x} = z^T \begin{bmatrix} 1 & 0 \\ \frac{\partial\varphi^T}{\partial x_1}\theta & 1 \end{bmatrix} \tag{7.2.20}$$

with (7.2.19), (7.2.20), *and* (7.2.16), *it is easy to show that*

$$\frac{\partial V_c}{\partial x}[f(x) + F(x)\theta + g(x)\alpha_c(x,\theta)] = -c_1 z_1^2 - c_2 z_2^2 \tag{7.2.21}$$

Let us now evaluate the partial derivatives appearing in (7.2.11)*:*

$$\frac{\partial V_c}{\partial \theta} = z^T e_2 \varphi^T = z_2 \varphi^T \tag{7.2.22}$$

$$\frac{\partial V_c}{\partial x} g = z^T e_2 = z_2 \tag{7.2.23}$$

where (7.2.23) *is immediate from* (7.2.20) *and* (7.2.19). *A comparison of* (7.2.22) *and* (7.2.23) *reveals that* $\frac{\partial V_c}{\partial \hat\theta} = \frac{\partial V_c}{\partial x} g \varphi^T$, *so that* α_c *is given by* (7.2.13)*:*

$$\alpha_\tau(x,\hat\theta) = -\varphi^T \Gamma \left(\frac{\partial V_c}{\partial x} F(x)\right)^T = -\varphi^T \Gamma \varphi \left[\, 1, \ \ \frac{\partial \varphi^T}{\partial x_1}\hat\theta + c_1 \,\right] z \tag{7.2.24}$$

Taking for simplicity $\Gamma = I$, *the resulting adaptive control law is*

$$\begin{aligned} u \;=\; & \alpha(x,\hat\theta) = -z_1 - c_2 z_2 - \left(c_1 + \frac{\partial \varphi^T}{\partial x_1}\theta\right)\left(x_2 + \varphi(x_1)^T\theta\right) \\ & -\varphi^T \varphi \left[\, 1, \ \ \frac{\partial \varphi^T}{\partial x_1}\hat\theta + c_1 \,\right] z \end{aligned} \tag{7.2.25}$$

and the corresponding parameter update law (7.2.7) *is*

$$\dot{\hat\theta} = \tau(x,\hat\theta) = \left(\frac{\partial V_c}{\partial x} F(x)\right)^T = \varphi \left[\, 1 \ \ , \ \ \frac{\partial \varphi^T}{\partial x_1}\hat\theta + c_1 \,\right] z \tag{7.2.26}$$

Note that in (7.2.25) *and* (7.2.26) *we use* $z(x,\hat\theta)$ *instead of* $z(x,\theta)$. *With the choice of* α *and* τ *given by* (7.2.25) *and* (7.2.26), *the derivative* $\dot V$ *of the Lyapunov function* $V(x,\hat\theta) = \frac{1}{2} z(x,\hat\theta)^T z(x,\hat\theta) + \frac{1}{2}\tilde\theta^T\tilde\theta$ *is guaranteed to be nonpositive:* $\dot V = -c_1 z_1^2 - c_2 z_2^2$. *This assures that both* x *and* $\hat\theta$ *are bounded. A standard argument using the LaSalle-Yoshizawa Theorem proves that also* $x(t) \to 0$. □

In the above example, the desired factorization (7.2.12) of $\frac{\partial V_c}{\partial \hat\theta}$ is a consequence of a particular feature of the system (7.2.14). The unknown parameter appears in the first, while the control appears only in the second equation. It is not hard to see that the same factorization (7.2.12) would be a possible for a higher-order plant, provided that *the unknown parameter is separated from the control input by at most one integrator.* So the factorization (7.2.12) is not a fortuitous event, but a structural property. For systems with this "extended matching" property, the above simple adaptive design is feasible. However, most systems fail to posses the "extended matching" property.

A benchmark example is the third-order system

$$\begin{aligned} \dot{x}_1 &= x_2 + \varphi(x_1)^T\theta \\ \dot{x}_2 &= x_3 \\ \dot{x}_3 &= u \end{aligned} \tag{7.2.27}$$

which has the form of (7.2.14) augmented by an integrator. In this system, θ and u are separated by two integrators and we unable to find α_T which satisfies (7.2.11). We will solve this problem with a recursive design which will circumvent the obstacle posed by the restrictive condition (7.2.11).

7.2.2 Certainty equivalence for a modified system

Condition (7.2.11) was dictated by our choice of the Lyapunov function $V_c(x, \hat{\theta})$ as the "certainty equivalence" form of $V_c(x, \theta)$. The only good thing we know about $V_c(x, \hat{\theta})$ is that it works when the factorization (7.2.12) is possible. Otherwise, we do not know how to remove the indefinite term preventing the nonpositivity of $\dot{V}$ in (7.2.9). Having recognized that a cause of our difficulties is $V_c(x, \theta)$, we now embark on a search for Lyapunov functions more suitable for adaptive control. The key idea is to counteract the effect of $\dot{\hat{\theta}}$ and thus prevent the parameter estimate transients from destroying the nonpositivity of the Lyapunov derivative

We say that the system

$$\dot{x} = f(x) + F(x)\theta + g(x)u \tag{7.2.28}$$

is **globally adaptively stabilizable** if there exist a function $\alpha(x,\hat{\theta})$ smooth on $(\mathbb{R}^n \setminus \{0\}) \times \mathbb{R}^p$ with $\alpha(0,\hat{\theta}) \equiv 0$, a smooth function $\tau(x,\hat{\theta})$, and a positive definite symmetric $p \times p$ matrix Γ, such that the dynamic controller

$$u = \alpha(x,\hat{\theta}) \tag{7.2.29}$$

$$\dot{\hat{\theta}} = \Gamma\tau(x,\hat{\theta}) \tag{7.2.30}$$

guarantees that the solution $(x(t),\hat{\theta}(t))$ is globally bounded, and $x(t) \to 0$ as $t \to \infty$, for all $\theta \in \mathbb{R}^p$.

Our approach is to replace the problem of adaptive stabilization of the original system (7.2.28) by a problem of nonadaptive stabilization of a modified system.

Definition 7.2.2 *A smooth function $V_a : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}_+$, positive definite and radially unbounded in x for each θ, is called an* ***adaptive control Lyapunov function (aclf)*** *for (7.2.28)if there exists a positive definite symmetric matrix $\Gamma \in \mathbb{R}^{p\times p}$ such that for each $\theta \in \mathbb{R}^p$, $V_a(x,\theta)$ is a* ***clf*** *for the modified system*

$$\dot{x} = f(x) + F(x)\left(\theta + \Gamma\left(\frac{\partial V_a}{\partial \theta}\right)^T\right) + g(x)u \tag{7.2.31}$$

that is, V_a satisfies

$$\inf_{u\in\mathbb{R}}\left\{\frac{\partial V_a}{\partial x}\left[f(x) + F(x)\left(\theta + \Gamma\left(\frac{\partial V_a}{\partial \theta}\right)^T\right) + g(x)u\right]\right\} < 0 \tag{7.2.32}$$

□

We now show how to design an adaptive controller (7.2.29) - (7.2.30) when an aclf is known.

Theorem 7.2.3 *The following two statements are equivalent:*

1. *There exists a triple (α, V_a, Γ) such that $\alpha(x,\theta)$ globally asymptotically stabilizes (7.2.32) at $x = 0$ for each $\theta \in \mathbb{R}^p$ with respect to the Lyapunov function $V_a(x,\theta)$.*

2. *There exists an aclf $V_a(x,\theta)$ for (7.2.28).*

Moreover, if an aclf $V_a(x,\theta)$ exists, then (7.2.28) is globally adaptively stabilizable.

Proof (1 ⇒ 2) *Obvious because 1 implies that there exists a continuous function $W : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}_+$, positive definite in x for each θ, such that*

$$\frac{\partial V_a}{\partial x}\left[f(x) + F(x)\left(\theta + \Gamma\left(\frac{\partial V_a}{\partial \theta}\right)^T\right) + g(x)\alpha(x,\theta)\right] \leq -W(x,\theta) \quad (7.2.33)$$

Thus $V_a(x,\theta)$ is a clf for (7.2.31) for each $\theta \in \mathbb{R}^p$, and therefore it is an aclf for (7.2.28).

(2 ⇒ 1) *The proof of this part is based on Sontag's constructive proof [171] of Artstein's theorem [4]. We assume that V_a is an aclf for* (7.2.28)*, that is, a clf for* (7.2.31). *Sontag's formula* (??) *applied to* (7.2.31) *gives a control law smooth on $(\mathbb{R}^n \setminus \{0\}) \times \mathbb{R}^p$:*

$$\alpha(x,\theta) = \begin{cases} -\dfrac{\frac{\partial V_a}{\partial x}\tilde{f} + \sqrt{\left(\frac{\partial V_a}{\partial x}\tilde{f}\right)^2 + \left(\frac{\partial V_a}{\partial x}g\right)^4}}{\frac{\partial V_a}{\partial x}g} & , \quad \frac{\partial V_a}{\partial x}g(x,\theta) \neq 0 \\ 0 & , \quad \frac{\partial V_a}{\partial x}g(x,\theta) = 0 \end{cases} \quad (7.2.34)$$

where

$$\tilde{f}(x,\theta) = f(x) + F(x)\left(\theta + \Gamma\left(\frac{\partial V_a}{\partial \theta}\right)^T\right) \quad (7.2.35)$$

With the choice (7.2.34)*,inequality* (7.2.33) *is satisfied with the continuous function*

$$W(x,\theta) = \sqrt{\left(\frac{\partial V_a}{\partial x}\tilde{f}(x,\theta)\right)^2 + \left(\frac{\partial V_a}{\partial x}g(x,\theta)\right)^4} \quad (7.2.36)$$

which is positive definite in x for each θ, because (7.2.32) implies that $\frac{\partial V_a}{\partial x}\tilde{f}(x,\theta) < 0$ whenever $\frac{\partial V_a}{\partial x}g(x,\theta) = 0$ and $x \neq 0$. We note that the control law $\alpha(x,\theta)$ will be continuous at $x = 0$ if and only if the aclf V_a satisfies the following property, called the small control property *[171]: For each $\theta \in \mathbb{R}^p$ and for any $\varepsilon > 0$ there is a $\delta > 0$ such that, if $x \neq 0$ satisfies $|x| \leq \delta$, then there is some u with $|u| \leq \varepsilon$ such that*

$$\frac{\partial V_a}{\partial x}\left[f(x) + F(x)\left(\theta + \Gamma\left(\frac{\partial V_a}{\partial \theta}\right)^T\right) + g(x)u\right] < 0 \tag{7.2.37}$$

Assuming the existence of an aclf we now show that (7.2.28) is globally adaptively stabilizable. Since ($2 \Rightarrow 1$), there exists a triple (α, V_a, Γ) and a function W such that (7.2.33) is satisfied, that is,

$$\frac{\partial V_a}{\partial x}[f(x) + F(x)\theta + g(x)\alpha(x,\theta)] + \frac{\partial V_a}{\partial \theta}\Gamma\left(\frac{\partial V_a}{\partial x}F(x)\right)^T \leq -W(x,\theta) \tag{7.2.38}$$

Consider the Lyapunov function candidate

$$V(x,\hat{\theta}) = V_a(x,\hat{\theta}) + \frac{1}{2}(\theta - \hat{\theta})^T\Gamma^{-1}(\theta - \hat{\theta}) \tag{7.2.39}$$

With the help of (7.2.38), the derivative of V along the solutions of (7.2.28), (7.2.29), and (7.2.30) is

$$\begin{aligned}
\dot{V} &= \frac{\partial V_a}{\partial x}\left[f + F\theta + g\alpha(x,\hat{\theta})\right] + \frac{\partial V_a}{\partial \hat{\theta}}\Gamma\tau(x,\hat{\theta}) - \tilde{\theta}^T\tau(x,\hat{\theta}) \\
&= \frac{\partial V_a}{\partial x}\left[f + F\theta + g\alpha(x,\hat{\theta})\right] + \frac{\partial V_a}{\partial \hat{\theta}}\Gamma\tau(x,\hat{\theta}) + \frac{\partial V_a}{\partial x}F\tilde{\theta} - \tilde{\theta}^T\tau(x,\hat{\theta}) \\
&\leq -W(x,\hat{\theta}) - \frac{\partial V_a}{\partial \hat{\theta}}\Gamma\left(\frac{\partial V_a}{\partial x}F\right)^T + \frac{\partial V_a}{\partial \hat{\theta}}\Gamma\tau(x,\hat{\theta})
\end{aligned} \tag{7.2.40}$$

Choosing

$$\tau(x,\hat{\theta}) = \left(\frac{\partial V_a}{\partial x}(x,\hat{\theta})F(x)\right)^T \tag{7.2.41}$$

we get

$$\dot{V} \leq -W(x,\hat{\theta}), \quad \forall\theta \in \mathbb{R}^p \tag{7.2.42}$$

Thus, the equilibrium $x = 0$, $\hat{\theta} = \theta$ *of* (7.2.28), (7.2.29), *and* (7.2.30) *is globally stable, and by the LaSalle-Yoshizawa Theorem,* $x(t) \to 0$, *that is,* (7.2.28) *is globally adaptively stabilizable.* □

The adaptive controller constructed in the proof of Theorem 7.2.3 consists of a control law $u = \alpha(x, \theta)$ given by (7.2.34), and an update law $\dot{\hat{\theta}} = \Gamma\tau(x, \hat{\theta})$ with (7.2.41).

It is of interest to interpret this controller as a certainty equivalence controller. The control law $\alpha(x, \theta)$ given by (7.2.34) is stabilizing for the modified system (7.2.31) but may not be stabilizing for the original system (7.2.28). However, as the proof of Theorem 7.2.3 shows, its certainty equivalence form $\alpha(x, \hat{\theta})$ is an adaptive globally stabilizing control law for the original system (7.2.28). Hence, if a certainty equivalence approach is to be applied to a nonlinear system, the system is to be modified to require a control law which anticipates the parameter estimation transients. In the proof of Theorem 7.2.3, this is achieved by incorporating the *tuning function* τ in the control law α. Indeed, the formula (7.2.34) for α depends on τ via

$$\frac{\partial V_a}{\partial x}\tilde{f}(x, \theta) = \frac{\partial V_a}{\partial x} f + \tau(x, \theta)^T \left(\theta + \Gamma \left(\frac{\partial V_a}{\partial \theta} \right)^T \right) \tag{7.2.43}$$

which is obtained by combining (7.2.35) and (7.2.41). Using (7.2.41) to rewrite the inequality (7.2.38) as

$$\frac{\partial V_a}{\partial x}[f(x) + F(x)\theta + g(x)\alpha(x, \theta)] + \frac{\partial V_a}{\partial \theta}\Gamma\tau(x, \theta) \le -W(x, \theta) \tag{7.2.44}$$

it is not difficult to see that the control law (7.2.34) containing (7.2.43) prevents τ from destroying the nonpositivity of the Lyapunov derivative.

Remark 7.2.4 *A relevant question remains unanswered: If there exists an aclf for* (7.2.28)*, is this system globally asymptotically stabilizable for each* θ *(and vice* *versa)? In other words, does the existence of a pair* α, V_a *satisfying* (7.2.33) *for some* $\Gamma > 0$ *imply the existence of a pair* (α^0, V_a^0) *satisfying* (7.2.33) *for* $\Gamma = 0$ *(and vice versa)? Adaptive Lyapunov designs available in the literature [59, 65, 69, 94, 156, 157, 186] are all for systems which are not only globally adaptively stabilizable, but aldo globally asymptotically stabilizable for each* θ.

As is always the case in adaptive control, in the proof of Theorem 7.2.3 we used a Lyapunov function $V(x, \hat{\theta})$ given by (7.2.39), which is quadratic in the parameter error $\theta - \hat{\theta}$. The quadratic form is suggested by the linear dependence of (7.2.28) on θ, and the fact that θ cannot be used for feedback. We will now show that the quadratic form of (7.2.39) is both necessary and sufficient for the existence of an aclf.

We say that system (7.2.28) is **globally adaptively quadratically stabilizable** if it is *globally adaptively stabilizable* and, in addition, these exist a smooth function $V_a(x, \theta)$ positive definite and radially unbounded in x for each θ, and a continuous function $W(x, \theta)$ positive definite in x for each θ, such that for all $(x(0), \hat{\theta}(0)) \in \mathbb{R}^{n+p}$ and all $\theta \in \mathbb{R}^p$, the derivative of (7.2.39) along the solutions of (7.2.28), (7.2.29), (7.2.30) is given by (7.2.42).

Corollary 7.2.5 *The system* (7.2.28) *is globally adaptively quadratically stabilizable if and only if there exists an aclf* $V_a(x, \theta)$.

Proof *The 'if' part is contained in the proof of Theorem 7.2.3 where the Lyapunov function* $V(x, \hat{\theta})$ *is in the form* (7.2.39). *To prove the 'only if' part, we start by assuming global adaptive quadratic stabilizability of* (7.2.28), *and first show that* $\tau(x, \hat{\theta})$ *must be given by* (7.2.41). *The derivative of* V *along the solutions of* (7.2.28), (7.2.29), (7.2.30) *given by* (7.2.40), *is rewritten as*

$$\begin{aligned}\dot{V} = & \ \tfrac{\partial V_a}{\partial x} \left[f + F\hat{\theta} + g\alpha(x,\hat{\theta})\right] + \frac{\partial V_a}{\partial \hat{\theta}}\Gamma\tau(x,\hat{\theta}) - \hat{\theta}^T \left(\left(\frac{\partial V_a}{\partial \theta}\right)^T - \tau\right) \\ + & \quad \theta^T \left(\left(\frac{\partial V_a}{\partial \theta}\right)^T - \tau\right) \end{aligned} \tag{7.2.45}$$

This expression has to be nonpositive to satisfy (7.2.42)*. Since it is affine in* θ*, it can be nonpositive for all* $(x,\hat{\theta}) \in \mathbb{R}^{n+p}$ *and all* $\theta \in \mathbb{R}^p$ *only if the last term is zero, that is, only if* τ *is defined as in* (7.2.41)*. Then, it is straightforward to verify that*

$$\begin{aligned} & \frac{\partial V_a}{\partial x}\left[f(x) + F(x)\left(\hat{\theta} + \Gamma\left(\frac{\partial V_a}{\partial \hat{\theta}}\right)^T\right) + g(x)\alpha(x,\hat{\theta})\right] \\ = & \ \dot{V} + \left(\hat{\theta}^T + \frac{\partial V_a}{\partial \hat{\theta}}\Gamma\right)\left(\tau - \left(\frac{\partial V_a}{\partial x}F\right)^T\right) \\ \leq & \ -W(x,\hat{\theta}) \end{aligned} \tag{7.2.46}$$

for all $x, \hat{\theta} \in \mathbb{R}^{n+p}$*. By* $(1 \Rightarrow 2)$ *in Theorem 7.2.3,* $V_a(x,\theta)$ *is an aclf for* (7.2.28).

□

The above analysis applies also to the case where the unknown parameter enter the control vector field:

$$\dot{x} = f(x) + F(x)\theta + [g(x) + G(x)\theta]u \tag{7.2.47}$$

In this case, the existence of an aclf V_a is equivalent to the existence of a clf for the system

$$\begin{aligned}\dot{x} \ = & \ f(x) + F(x)\left(\theta + \Gamma\left(\frac{\partial V_a}{\partial \theta}\right)^T\right) \\ & + \left[g(x) + G(x)\left(\theta + \Gamma\left(\frac{\partial V_a}{\partial \theta}\right)^T\right)\right]u \end{aligned} \tag{7.2.48}$$

The extension to the multi-input case is also straightforward.

It is of interest to examine the input-output properties of the system resulting from the application of the adaptive control law $\alpha(x, \hat{\theta})$ to the plant (7.2.1):

$$\dot{x} = f(x) + F(x)\hat{\theta} + g(x)\alpha(x, \hat{\theta}) + F(x)\tilde{\theta} \tag{7.2.49}$$

7.2.3 Adaptive backstepping via aclf

With Theorem 7.2.3, the problem of adaptive stabilization is reduced to the problem of finding an aclf. We now address the problem of systematic construction of an aclf. Our aim is a recursive approach because we already know how to find aclf's for systems with the extended matching property, and expect to recursively enlarge this initial class of systems with repeated use of backstepping. So, we assume that an aclf is known for an initial system, and construct a new aclf for the initial system augmented by an integrator.

Lemma 7.2.6 *If the system*

$$\dot{x} = f(x) + F(x)\theta + g(x)u \tag{7.2.50}$$

is globally adaptively quadratically stabilizable with $\alpha \in \mathbb{C}^1$, then the augmented system

$$\begin{aligned} \dot{x} &= f(x) + F(x)\theta + g(x)\xi \\ \dot{\xi} &= u \end{aligned} \tag{7.2.51}$$

is also globally adaptively quadratically stabilizable.

Proof *Since system* (7.2.50) *is globally adaptively stabilizable, then by Corollary 7.2.5 there exists an aclf $V_a(x, \theta)$, and by Theorem 7.2.3, it satisfies* (7.2.33) *with a control law $u = \alpha(x, \theta)$. We will now show that*

$$V_1(x, \xi, \theta) = V_a(x, \theta) + \frac{1}{2}(\xi - \alpha(x, \theta))^2 \tag{7.2.52}$$

is an aclf for the augmented system (7.2.51) *by showing that it satisfies*

$$\frac{\partial V_1}{\partial(x,\xi)}\begin{bmatrix} f+F\left(\theta+\Gamma\left(\frac{\partial V_1}{\partial\theta}\right)^T\right)+g\xi \\ \alpha_1(x,\xi,\theta)\end{bmatrix} \leq -W-(\xi-\alpha)^2 \tag{7.2.53}$$

with the control law

$$\begin{aligned} u=\alpha_1(x,\xi,\theta)= &-\frac{\partial V_a}{\partial x}g-(\xi-\alpha)+\frac{\partial\alpha}{\partial x}(f+F\theta+g\xi) \\ &+\frac{\partial\alpha}{\partial\theta}\Gamma\left(\frac{\partial V_1}{\partial x}F\right)^T+\frac{\partial V_a}{\partial\theta}\Gamma\left(\frac{\partial\alpha}{\partial x}F\right)^T \end{aligned} \tag{7.2.54}$$

Let us start by introducing for brevity $z=\xi-\alpha(x,\theta)$. *With* (7.2.52) *we compute*

$$\frac{\partial V_1}{\partial(x,\xi)}\begin{bmatrix} f+F\theta+g\xi \\ \alpha_1(x,\xi,\theta)\end{bmatrix}=\frac{\partial V_1}{\partial x}(f+F\theta+g\xi)+\frac{\partial V_1}{\partial\xi}\alpha_1(x,\xi,\theta) \tag{7.2.55}$$

$$\begin{aligned} &=\left(\frac{\partial V_a}{\partial x}-z\frac{\partial\alpha}{\partial x}\right)(f+F\theta+g\xi)+z\alpha_1 \\ &=\frac{\partial V_a}{\partial x}(f+F\theta+g\alpha)+\frac{\partial V_a}{\partial x}gz-z\frac{\partial\alpha}{\partial x}(f+F\theta+g\xi)+z\alpha_1 \\ &=\frac{\partial V_a}{\partial x}(f+F\theta+g\alpha)+z\left(\alpha_1+\frac{\partial V_a}{\partial x}g-\frac{\partial\alpha}{\partial x}(f+F\theta+g\xi)\right) \end{aligned}$$

On the other hand, in view of (7.2.52), *we have*

$$\frac{\partial V_1}{\partial(x,\xi)}\begin{bmatrix} F\Gamma\left(\frac{\partial V_1}{\partial\theta}\right)^T \\ 0\end{bmatrix}=\frac{\partial V_1}{\partial x}F\Gamma\left(\frac{\partial V_1}{\partial\theta}\right)^T \tag{7.2.56}$$

$$\begin{aligned} &=\left(\frac{\partial V_a}{\partial x}-z\frac{\partial\alpha}{\partial x}\right)F\Gamma\left(\frac{\partial V_a}{\partial\theta}\right)^T \\ &=\frac{\partial V_a}{\partial x}F\Gamma\left(\frac{\partial V_a}{\partial\theta}\right)^T-z\left(\frac{\partial\alpha}{\partial\theta}\Gamma\left(\frac{\partial V_1}{\partial x}F\right)^T+\frac{\partial V_a}{\partial\theta}\Gamma\left(\frac{\partial\alpha}{\partial x}F\right)^T\right) \end{aligned}$$

Adding (7.2.55) *and* (7.2.56) *with* (7.2.33) *and* (7.2.54), *we get*

$$\begin{aligned} \frac{\partial V_1}{\partial(x,\xi)}\begin{bmatrix} f+F\left(\theta+\Gamma\left(\frac{\partial V_1}{\partial\theta}\right)^T\right)+g\xi \\ \alpha_1(x,\xi,\theta)\end{bmatrix}&=\frac{\partial V_a}{\partial x}(f+F\theta+g\alpha)+\frac{\partial V_a}{\partial x}\Gamma\left(\frac{\partial V_a}{\partial\theta}\right)^T \\ +z\left(\alpha_1+\frac{\partial V_a}{\partial x}g-\frac{\partial\alpha}{\partial x}(f+F\theta+g\xi)-\frac{\partial\alpha}{\partial\theta}\Gamma\left(\frac{\partial V_1}{\partial x}F\right)^T-\frac{\partial V_a}{\partial\theta}\Gamma\left(\frac{\partial\alpha}{\partial x}F\right)^T\right)& \\ \leq -W(x,\theta)-z^2 & \end{aligned} \tag{7.2.57}$$

This proves by Theorem 7.2.3 that $V_1(x,\xi,\theta)$ is an aclf for system (7.2.51), and by Corollary 7.2.5, this system is globally adaptively quadratically stabilizable. □

The new tuning function for system (7.2.51) is determined by the new aclf V_1 and given by

$$\begin{aligned}\tau_1(x,\xi,\theta) &= \left(\frac{\partial V_1}{\partial(x,\xi)}\begin{bmatrix} F \\ 0\end{bmatrix}\right)^T = \left(\frac{\partial V_1}{\partial x}F\right)^T = \left[\left(\frac{\partial V_a}{\partial x} - (\xi-\alpha)\frac{\partial \alpha}{\partial x}\right)F\right]^T \\ &= \tau(x,\theta) - \left(\frac{\partial \alpha}{\partial x}F\right)^T(\xi-\alpha) \qquad (7.2.58)\end{aligned}$$

We note that the new tuning function τ_1 is obtained by augmenting the initial tuning function τ with the term $-\left(\frac{\partial\alpha}{\partial x}F\right)^T(\xi-\alpha)$ which accounts for the fact that the aclf V_a is augmented by $\frac{1}{2}(\xi-\alpha(x,\theta))^2$

The form of the control law $\alpha_1(x,\xi,\theta)$ in (7.2.54) is of particular interest. It consists of two parts, $\alpha_1 = \alpha_{1,c} + \alpha_{1,\tau}$. The first part,

$$\alpha_{1,c}(x,\xi,\theta) = 3\frac{\partial V_a}{\partial x}g - (\xi-\alpha) + \frac{\partial\alpha}{\partial x}(f + F\theta + g\xi) \qquad (7.2.59)$$

would become the "certainty equivalence" control law for the augmented system (7.2.51) if we were to set $\Gamma = 0$. The second part consists of two terms.

$$\alpha_{1,\tau}(x,\xi,\theta) = \frac{\partial\alpha}{\partial\theta}\Gamma\left(\frac{\partial V_1}{\partial x}F\right)^T + \frac{\partial V_a}{\partial\theta}\Gamma\left(\frac{\partial\alpha}{\partial x}F\right)^T \qquad (7.2.60)$$

Their role is to produce $\frac{\partial V_a}{\partial x}F\Gamma\left(\frac{\partial V_a}{\partial\theta}\right)^T$ in the aclf inequality (7.2.53). Observe that the first term in (7.2.60) incorporates $\tau_1 = \left(\frac{\partial V_1}{\partial x}F\right)^T$.

The control law $\alpha_1(x,\xi,\theta)$ in (7.2.54) is only one out of many possible control laws. Once we have shown that V_1 given by (7.2.52) is an aclf for (7.2.51), we can use, for example, the C^0 control law α_1 given by Sontag's formular (7.2.34) with $\frac{\partial V_1}{\partial(x,\xi)}g_1 = z$ and

$$\begin{aligned}\frac{\partial V_1}{\partial(x,\xi)}\tilde{f}_1(x,\xi,\theta) &= \frac{\partial V_1}{\partial(x,\xi)}\begin{bmatrix} f+F\left(\theta+\Gamma\left(\frac{\partial V_1}{\partial\theta}\right)^T\right)+g\xi \\ 0\end{bmatrix} \\ &= \frac{\partial V_1}{\partial x}(f+g\xi)+\tau_1(x,\xi,\theta)^T\left(\theta+\Gamma\left(\frac{\partial V_1}{\partial\theta}\right)^T\right)\end{aligned} \tag{7.2.61}$$

It can be shown that the following function, used as a clf in [158], is a more general aclf than (7.2.52):

$$V_1(x,\xi,\theta)=V_a(x,\theta)+\int_0^{\xi-\alpha(x,\theta)}\eta(s)ds \tag{7.2.62}$$

where η is a C^0 function such that $s\eta(s)>0$ whenever $s\neq 0$, $\eta'(0)>0$, and $\eta\notin\mathcal{L}^1((-\infty,0])\cup\mathcal{L}^1((0,+\infty])$.

The following example illustrates the use of Lemma 7.2.6.

Example 7.2.7 *Let us consider the system*

$$\begin{aligned}\dot{x}_1 &= x_2+\varphi(x_1)^T\theta \\ \dot{x}_2 &= x_3 \\ \dot{x}_3 &= u\end{aligned} \tag{7.2.63}$$

we will treat the state x_3 as an integrator added to the (x_1,x_2)-subsystem from Example 7.2.1. In the example, we have already designed an adaptive control law for the system

$$\begin{aligned}\dot{x}_1 &= x_2+\varphi(x_1)^T\theta \\ \dot{x}_2 &= x_3\end{aligned} \tag{7.2.64}$$

considering x_3 as a control input. With (7.2.18), (7.2.19), (7.2.20), *and* (7.2.22), *it can be shown that*

$$\frac{\partial V_c}{\partial x}\left[f(x)+F(x)\left(\theta+\left(\frac{\partial V_c}{\partial\theta}\right)^T\right)\right]=-c_1z_1^2-c_2z_2^2 \tag{7.2.65}$$

which means that $V_a(x_1, x_2, \theta) = V_c(x_1, x_2, te) = \frac{1}{2}(z_1^2 + z_2^2)$ *is an aclf for the system* (7.2.64) *considering* x_3 *as a control input. Therefore, Lemma 7.2.6, the function directly applicable. We define* $z = x_3 - \alpha(x, \theta)$. *By Lemma 7.2.6, the function*

$$V_1(x, \theta) = \frac{1}{2}\left(z_1^2 + z_2^2 + z_3^2\right) \tag{7.2.66}$$

is an aclf for the system (7.2.63). *With* (7.2.54) *and* (7.2.58), *we obtain*

$$\begin{aligned} \alpha_1(x, \theta) &= -z_1 - c_3 z_3 - \frac{\partial \alpha}{\partial(x_1, x_2)} \begin{bmatrix} x_2 + \varphi^T \theta \\ x_3 \end{bmatrix} + \frac{\partial \alpha}{\partial \theta} \tau_1 \\ & \quad + z_2 \varphi^T \frac{\partial \alpha}{\partial x_1} \varphi \end{aligned} \tag{7.2.67}$$

$$\tau_1(x, \theta) = \tau - \frac{\partial \alpha}{\partial x_1} \varphi z_3 \tag{7.2.68}$$

With the following adaptive control law and the parameter update law:

$$u = \alpha_1(x, \hat{\theta}) \tag{7.2.69}$$

$$\dot{\hat{\theta}} = \tau_1(x, \hat{\theta}) \tag{7.2.70}$$

it is straightforward to verify that the closed-loop adaptive system is

$$\begin{aligned} \begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \end{bmatrix} &= \begin{bmatrix} -c_1 & 1 & 0 \\ -1 & -c_2 & 1 - \frac{\partial \alpha}{\partial x_1}|\varphi|^2 \\ 0 & -1 + \frac{\partial \alpha}{\partial x_1}|\varphi|^2 & -c_3 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \\ & \quad + \begin{bmatrix} 1 \\ \frac{\partial \varphi^T}{\partial x_1}\hat{\theta} + c_1 \\ -\frac{\partial \alpha}{\partial x_1} \end{bmatrix} \varphi^T \tilde{\theta} \end{aligned} \tag{7.2.71}$$

$$\dot{\hat{\theta}} = \varphi \left[1, \quad \frac{\partial \varphi^T}{\partial x_1}\hat{\theta} + c_1, \quad -\frac{\partial \alpha}{\partial x_1} \right] \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \tag{7.2.72}$$

where z_1, z_2, z_3 *are used with* $\hat{\theta}$ *as an argument. The global stability of this system is established using the Lyapunov function* $V(x, \hat{\theta}) = V_1(x, \hat{\theta}) + \frac{1}{2}\tilde{\theta}^T \tilde{\theta}$. □

While in Lemma 7.2.6 the initial system is augmented only by an integrator, a minor modification is sufficient to obtain an analogous result for the more general system

$$\begin{aligned} \dot{x} &= f(x) + F(x)\theta + g(x)\xi \\ \dot{\xi} &= u + F_1(x,\xi)\theta \end{aligned} \tag{7.2.73}$$

Corollary 7.2.8 *The function* $V_1(x,\xi,\theta)$ *defined in* (7.2.52) *is an aclf for the system* (7.2.73) *with the control law and the tuning function given as*

$$\alpha_1(x,\xi,\theta) = \alpha_1(x,\xi,\theta) - F_1(x,\xi)\left(\theta + \Gamma\left(\frac{\partial V_c}{\partial \theta}\right)^T\right) \tag{7.2.74}$$

$$\tau_1(x,\xi,\theta) = \tau_1(x,\xi,\theta) + (\xi - \alpha)F_1(x,\xi)^T \tag{7.2.75}$$

□

A repeated application of Corollary 7.2.8 will further extend the class of nonlinear systems for this type of adaptive design. With the knowledge of V_a, τ, and α for the system (7.2.73), it is not hard to see that by applying Corollary 7.2.8 twice we can find V_2, τ_2, and α_2 for the system

$$\begin{aligned} \dot{x} &= f(x) + F(x)\theta + g(x)\xi_1 \\ \dot{\xi}_1 &= \xi_2 + F_1(x,\theta_1)\theta \\ \dot{\xi}_2 &= u + F_2(x,\xi_1,\xi_2)\theta \end{aligned} \tag{7.2.76}$$

In fact, it is clear that an n-fold application of Corollary 7.2.8 will provide us with V_n, τ_n, and α_n for the system

$$\begin{aligned} \dot{x} &= f(x) + F(x)\theta + g(x)\xi_1 \\ \dot{\xi}_1 &= \xi_2 + F_1(x,\xi_1)\theta \\ &\vdots \\ \dot{\xi}_{n-1} &= \xi_n + F_{n-1}(x,\xi_1,\ldots,\xi_{n-1})\theta \\ \dot{\xi}_n &= u + F_n(x,\xi_1,\ldots,\xi_n)\theta \end{aligned} \tag{7.2.77}$$

We will now develop a detailed design procedure for such systems.

7.3 Set-Point Regulation

With repeated use of Corollary 7.2.8, we can design an adaptive controller to globally stabilize a desired equilibrium x^e of the *parametric strict-feedback system* (6.2.1):

$$\begin{aligned}
\dot{x}_1 &= x_2 + \varphi_1(x_1)^T\theta \\
\dot{x}_2 &= x_3 + \varphi_2(x_1, x_2)^T\theta \\
&\vdots \\
\dot{x}_{n-1} &= x_n + \varphi_{n-1}(x_1, \dots, x_{n-1})^T\theta \\
\dot{x}_n &= \beta(x)u + \varphi_n(x)^T\theta
\end{aligned} \tag{7.3.1}$$

where $\theta \in \mathbb{R}^p$ is a vector of unknown constant parameters, β and

$$F = [\varphi_1, \dots, \varphi_n] \tag{7.3.2}$$

are smooth nonlinear functions taking arguments in $\mathbb{R}^n$, and $\beta(x) \neq 0$, $\forall x \in \mathbb{R}^n$.

In this section, we develop a procedure for adaptive regulation of the output $y = x_1$ to a given set-point y_s. With a constant control u^e, the first n-1 equilibrium equations of $\dot{x}^e = 0$ in (7.3.1) can be successively solved for $x_2^e, \dots, x_n^e$ as functions of x_1^e and θ:

$$\begin{aligned}
x_2^e &= -\varphi_1(x_1^e)^T\theta \\
x_3^e &= -\varphi_2(x_1^e, x_2^e)^T\theta \\
&\vdots \\
x_n^e &= -\varphi_{n-1}(x_1^e, \dots, x_{n-1}^e)^T\theta
\end{aligned} \tag{7.3.3}$$

Then the n^{th} equation $\dot{x}_n^e = 0$ yields a relationship between x_1^e, u^e, and θ. When θ is known, then $\dot{x}_n^e = 0$ can be solved for u^e needed to keep x_1^e at a desired set-point $x_1^e = y_s$. The corresponding values $x_2^e, \dots, x_n^e$ will be dictated by (7.3.3). Therefore, for each value of θ and a prescribed y_s, the equilibrium x^e and the corresponding control value u^e are uniquely defined. In the special case where $\varphi_1(0) = \cdots = \varphi_{n-1}(0) = 0$, the choice $y_s = 0$ results in the equilibrium being $x^e = 0$ for all values of θ.

Our problem now is to globally stabilize this equilibrium when θ is unknown and also to achieve set-point regulation: $x(t) \to x^e$ as $t \to \infty$.

Comparing the systems (7.3.1) and (7.2.73), we observe that if x_3 were the control variable, then Corollary 7.2.8 would provide the desired adaptive control for the subsystem made of the first two equations of (7.3.1). Therefore, we can initiate our recursive design procedure by augmenting his subsystem by the third equation, as in (7.2.76). For convenience, we will do this in a self-contained fashion, independent of Section 7.2. An additional feature of the procedure in this section in a set of error coordinates in which the stability properties of the resulting closed-loop adaptive system are clearly displayed without an explicit use of the aclf concept.

7.3.1 Design procedure

We will start by adaptively stabilizing the first equation (7.3.1) considering x_2 to be its control. At each subsequence step, we will augment the designed subsystem by one equation. At the i^{th} step, an i^{th}-order subsystem is stabilized with respect to a Lyapunov function V_i by the design of a *stabilizing function* α_i and a *tuning function* τ_i. The update law for the parameter estimate $\hat{\theta}(t)$ and the adaptive feedback control u are designed at the final step. The third step is crucial for understanding the general design procedure.

Step 1. Introducing the first two error variables

$$z_1 = x_1 - y_s \tag{7.3.4}$$

$$z_2 = x_2 - \alpha_1 \tag{7.3.5}$$

we rewrite $\dot{x}_1 = x_2 + \varphi_1(x_1)^T\theta$, the first equation of (7.3.1), as

$$\dot{z}_1 = z_2 + \alpha_1 + w_1(x_1)^T\theta \tag{7.3.6}$$

where, for uniformly with subsequence steps, we have defined the first regressor vector as

$$w_1(x_1) = \varphi_1(x_1) \tag{7.3.7}$$

Our task in this step is to stabilize (7.3.6) with respect to the Lyapunov function

$$V_1 = \frac{1}{2}z_1^2 + \frac{1}{2}\tilde{\theta}^T\Gamma^{-1}\tilde{\theta} \tag{7.3.8}$$

whose derivative along the solutions of (7.3.6) is

$$\dot{V}_1 = z_1(z_2 + \alpha_1 + w_1^T\theta) - \tilde{\theta}^T\Gamma^{-1}(\dot{\hat{\theta}} - \Gamma w_1 z_1) \tag{7.3.9}$$

We can eliminate $\tilde{\theta}$ from $\dot{V}_1$ with the update law $\dot{\hat{\theta}} = \Gamma\tau_1$, where

$$\tau_1(x_1) = w_1(x_1)z_1 \tag{7.3.10}$$

If x_2 were actual control, we would let $z_2 \equiv 0$, that is, $x_2 \equiv \alpha_1$. Then, to make $\dot{V}_1 = -c_1z_1^2$, we would choose

$$\alpha_1(x_1, \hat{\theta}) = -c_1z_1 - w_1(x_1)^T\hat{\theta} \tag{7.3.11}$$

Since x_2 is not our control, we have $z_2 \not\equiv 0$, and we do not use $\dot{\hat{\theta}} = \Gamma\tau_1$ as an update law. Instead, we retain τ_1 as our first *tuning function* and tolerate the presence of $\tilde{\theta}$ in $\dot{V}_1$:

$$\dot{V}_1 = -c_1z_1^2 + z_1z_2 + \tilde{\theta}^T(\Gamma^{-1}\dot{\hat{\theta}} - \tau_1) \tag{7.3.12}$$

The second term z_1z_2 in $\dot{V}_1$ will be cancelled at the next step. With $\alpha_1(x_1, \hat{\theta})$ as in (7.3.11), the z_1-system becomes

$$\dot{z}_1 = -c_1z_1 + z_2 + w_1(x_1)^T\tilde{\theta} \tag{7.3.13}$$

Step 2. We now consider that x_3 is the control variable in the second equation of (7.3.1). Introducing

$$z_3 = x_3 - \alpha_2 \tag{7.3.14}$$

we rewrite $\dot{x}_2 = x_3 + \varphi_2(x_1, x_2)^T\theta$ as

$$\dot{z}_2 = z_3 + \alpha_2 - \frac{\partial\alpha_1}{\partial x_1}x_2 + w_2(x_1, x_2, \hat{\theta})^T\theta - \frac{\partial\alpha_1}{\partial\hat{\theta}}\dot{\hat{\theta}} \tag{7.3.15}$$

where the second regressor vector w_2 is defined as

$$w_2(x_1, x_2, \hat{\theta}) = \varphi_2 - \frac{\partial\alpha_1}{\partial x_1}\varphi_1 \tag{7.3.16}$$

Our task in this step is to stabilize the (z_1, z_2)-system (7.3.13), (7.3.15) with respect to

$$V_2 = V_1 + \frac{1}{2}z_2^2 \tag{7.3.17}$$

whose derivative along the solutions of (7.3.13) and (7.3.15) is

$$\begin{aligned}\dot{V}_2 = \quad & -c_1 z_1^2 + z_2\left[z_1 + z_3 + \alpha_2 - \frac{\partial\alpha_1}{\partial x_1}x_2 + w_2^T\hat{\theta} - \frac{\partial\alpha_1}{\partial\hat{\theta}}\dot{\hat{\theta}}\right] \\ & +\tilde{\theta}^T\left(\tau_1 + w_2 z_2 - \Gamma^{-1}\dot{\hat{\theta}}\right)\end{aligned} \tag{7.3.18}$$

We can eliminate $\tilde{\theta}$ from $\dot{V}_2$ with the update law $\dot{\hat{\theta}} = \Gamma\tau_2$, where

$$\tau_2(x_1, x_2, \hat{\theta}) = \tau_1 + w_2 z_2 = \begin{bmatrix} w_1 & , & w_2 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \tag{7.3.19}$$

If x_3 were our actual control and, hence, $z_3 \equiv 0$, we would achieve $\dot{V}_2 = -c_1 z_1^2 - c_2 z_2^2$ by designing α_2 to make the bracketed term multiplying z_2 in (7.3.18) equal to $-c_2 z_2$, namely

$$\alpha_2(x_1, x_2, \hat{\theta}) = -z_1 - c_2 z_2 + \frac{\partial\alpha_1}{\partial x_1}x_2 - w_2^T\hat{\theta} + \frac{\partial\alpha_1}{\partial\hat{\theta}}\Gamma\tau_2 \tag{7.3.20}$$

We retain τ_2 as our second tuning function in the term $\Gamma\tau_2$ which replaces $\dot{\hat{\theta}}$ in (7.3.20). However, we do not use $\dot{\hat{\theta}} = \Gamma\tau_2$ as an update law, so that the resulting $\dot{V}_2$ is

$$\dot{V}_2 = -c_1 z_1^2 - c_2 z_2^2 + z_2 z_3 + z_2\frac{\partial\alpha_1}{\partial\hat{\theta}}(\Gamma\tau_2 - \dot{\hat{\theta}}) + \tilde{\theta}^T(\tau_2 - \Gamma^{-1}\dot{\hat{\theta}}) \tag{7.3.21}$$

The first two terms in $\dot{V}_2$ are negative definite, the third term will be cancelled at the next step, while the discrepancy between $\Gamma\tau_2$ and $\dot{\hat{\theta}}$ in the last two terms remains. By substituting (7.3.20) into (7.3.15), the (z_1, z_2)-subsystem becomes

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} -c_1 & 1 \\ -1 & -c_2 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} w_1^T \\ w_2^T \end{bmatrix} \tilde{\theta} + \begin{bmatrix} 0 \\ z_3 + \frac{\partial \alpha_1}{\alpha \hat{\theta}}(\Gamma\tau_2 - \dot{\hat{\theta}}) \end{bmatrix} \tag{7.3.22}$$

Step 3. Proceeding to the third equation in (7.3.1), we introduce

$$z_4 = x_4 - \alpha_3 \tag{7.3.23}$$

and rewrite $\dot{x}_3 = x_4 + \varphi_3(x_1, x_2, x_3)^T \theta$ as

$$\dot{z}_3 = z_4 + \alpha_3 - \frac{\partial \alpha_2}{\partial x_1} x_2 - \frac{\partial \alpha_2}{\partial x_2} x_3 + w_3(x_1, x_2, x_3, \hat{\theta})^T - \frac{\partial \alpha_2}{\partial \hat{\theta}} \dot{\hat{\theta}} \tag{7.3.24}$$

where the third regressor vector w_3 is defined as

$$w_3(x_1, x_2, x_3, \hat{\theta}) = \varphi_3 - \frac{\partial \alpha_2}{\partial x_1}\varphi_1 - \frac{\partial \alpha_2}{\partial x_1}\varphi_2 \tag{7.3.25}$$

Our task is to stabilize the (z_1, z_2, z_3)-system with respect to

$$V_3 = V_2 + \frac{1}{2} z_3^2 \tag{7.3.26}$$

whose derivative along (7.3.22) and (7.3.24) is

$$\begin{aligned} \dot{V}_3 \;\; = \;\; & -c_1 z_1^2 - c_2 z_2^2 + z_2 \frac{\partial \alpha_1}{\partial \hat{\theta}}(\Gamma\tau_2 - \dot{\hat{\theta}}) \\ & + z_3 \left[z_2 + z_4 + \alpha_3 - \frac{\partial \alpha_2}{\partial x_1} x_2 - \frac{\partial \alpha_2}{\partial x_2} x_3 + w_3^T \hat{\theta} - \frac{\partial \alpha_2}{\partial \hat{\theta}} \dot{\hat{\theta}} \right] \\ & + \tilde{\theta}^T \left(\tau_2 + w_3 z_3 - \Gamma^{-1} \dot{\hat{\theta}} \right) \end{aligned} \tag{7.3.27}$$

We can eliminate $\tilde{\theta}$ from $\dot{V}_3$ with the update law $\dot{\hat{\theta}} = \Gamma\tau_3$, where τ_3 is our tuning function

$$\tau_3(x_1, x_2, x_3, \hat{\theta}) = \tau_2 + w_3 z_3 = \begin{bmatrix} w_1, & w_2, & w_3 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \tag{7.3.28}$$

If x_3 were our actual control, we could have $z_4 \equiv 0$ and achieve $\dot{V}_3 = -c_1 z_1^2 - c_2 z_2^2 - c_3 z_3^2$ by designing α_3 to make the bracketed term multiplying z_3 equal to $-c_3 z_3$, namely

$$\alpha_3(x_1, x_2, x_3, \hat{\theta}) = -z_2 - c_3 z_3 + \frac{\partial \alpha_2}{\partial x_1} x_2 + \frac{\partial \alpha_2}{\partial x_2} x_3 - w_3^T \hat{\theta} + \frac{\partial \alpha_2}{\partial \hat{\theta}} \Gamma \tau_3 + \nu_3 \quad (7.3.29)$$

where ν_3 is a correction term yet to be chosen. Substituting (7.3.29) into (7.3.27), and noting that

$$\begin{aligned} \dot{\hat{\theta}} - \Gamma \tau_2 &= \dot{\hat{\theta}} - \Gamma \tau_3 + \Gamma \theta_3 - \Gamma \tau_2 \\ &= \dot{\hat{\theta}} - \Gamma \tau_3 + \Gamma w_3 z_3 \end{aligned} \quad (7.3.30)$$

(7.3.27) is written as

$$\begin{aligned} \dot{V}_3 &= -c_1 z_1^2 - c_2 z_2^2 + z_3 \left(\nu_3 - \frac{\partial \alpha_1}{\partial \hat{\theta}} \Gamma w_3 z_2 \right) + z_3 z_4 \\ &\quad + \left(z_2 \frac{\partial \alpha_1}{\partial \hat{\theta}} + z_3 \frac{\partial \alpha_2}{\partial \hat{\theta}} \right) (\Gamma \tau_3 - \dot{\hat{\theta}}) + \tilde{\theta}^T (\tau_3 - \Gamma^{-1} \dot{\hat{\theta}}) \end{aligned} \quad (7.3.31)$$

and the (z_1, z_2, z_3)-subsystem becomes

$$\begin{aligned} \begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \end{bmatrix} &= \begin{bmatrix} -c_1 & 1 & 0 \\ -1 & -c_2 & 1 \\ 0 & -1 & -c_3 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} + \begin{bmatrix} w_1^T \\ w_2^T \\ w_3^T \end{bmatrix} \tilde{\theta} + \begin{bmatrix} 0 \\ -\frac{\partial \alpha_1}{\partial \hat{\theta}} \Gamma w_3 z_3 \\ \nu_3 \end{bmatrix} \\ &\quad + \begin{bmatrix} 0 \\ \frac{\partial \alpha_1}{\partial \hat{\theta}} (\Gamma \tau_3 - \dot{\hat{\theta}}) \\ z_4 + \frac{\partial \alpha_2}{\partial \hat{\theta}} (\Gamma \tau_3 - \dot{\hat{\theta}}) \end{bmatrix} \end{aligned} \quad (7.3.32)$$

If x_4 were our control, we would have $z_4 = 0$, and with the update law $\dot{\hat{\theta}} = \Gamma \tau_3$, the last vector in (7.3.32) would be zero. However, the potentially destabilizing term $-\frac{\partial \alpha_1}{\partial \hat{\theta}} \Gamma w_3 z_3$ would still remain. This unmatched term must be accommodated by a choice of the correction term ν_3. From (7.3.31), the choice of ν_3 is immediate.

$$\nu_3(x_1, x_2, x_3, \hat{\theta}) = \frac{\partial \alpha_1}{\partial \hat{\theta}} \Gamma w_3 z_2 \quad (7.3.33)$$

We again postpone the decision about $\dot{\hat{\theta}}$ and do not use $\dot{\hat{\theta}} = \Gamma\tau_3$ as an update law. The resulting $\dot{V}_3$ is

$$\begin{aligned}\dot{V}_3 &= -c_1 z_1^2 - c_2 z_2^2 - c_3 z_3^3 + z_3 z_4 + \left(z_2 \frac{\partial \alpha_1}{\partial \hat{\theta}} + z_3 \frac{\partial \alpha_2}{\partial \hat{\theta}}\right)(\Gamma\tau_3 - \dot{\hat{\theta}}) \\ &\quad + \tilde{\theta}^T(\tau_3 - \Gamma^{-1}\dot{\hat{\theta}}) \end{aligned} \tag{7.3.34}$$

and the (z_1, z_2, z_3)-subsystem becomes

$$\begin{aligned}\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \end{bmatrix} &= \begin{bmatrix} -c_1 & 1 & 0 \\ -1 & -c_2 & 1 - \frac{\partial \alpha_1}{\partial \hat{\theta}}\Gamma w_3 \\ 0 & -1 + \frac{\partial \alpha_1}{\partial \hat{\theta}}\Gamma w_3 & -c_3 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} + \begin{bmatrix} w_1^T \\ w_2^T \\ w_3^T \end{bmatrix} \tilde{\theta} \\ &\quad + \begin{bmatrix} 0 \\ 0 \\ z_4 \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{\partial \alpha_1}{\partial \hat{\theta}} \\ \frac{\partial \alpha_2}{\partial \hat{\theta}} \end{bmatrix} (\Gamma\tau_3 - \dot{\hat{\theta}}) \end{aligned} \tag{7.3.35}$$

The 'system matrix' in (7.3.35) has a significant property: the skew symmetry of the nonlinear term $\frac{\partial \alpha_1}{\partial \hat{\theta}}\Gamma w_3$ achieved by the choice of ν_3 in (7.3.33). This term is analogous to the second term in (7.2.60) and the skew symmetry is crucial for stabilization.

Step i. Introducing

$$z_{i+1} = x_{i+1} - \alpha_i \tag{7.3.36}$$

we rewrite $\dot{x}_i = x_{i+1} + \varphi(x_1, \ldots, x_i)^T\theta$ as

$$\dot{z}_i = z_{i+1} + \alpha_i - \sum_{k=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_k} x_{k+1} + w_i(x_1, \ldots, x_i, \hat{\theta})^T \theta - \frac{\partial \alpha_{i-1}}{\partial \hat{\theta}} \dot{\hat{\theta}} \tag{7.3.37}$$

where the i^{th} regressor vector is defined as

$$w_i(x_1, \ldots, x_i, \hat{\theta}) = \varphi_i - \sum_{k=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_k} \varphi_k \tag{7.3.38}$$

Our objective is to stabilize the $(z_1, \ldots, z_i)$-system with respect to

$$V_i = V_{i-1} + \frac{1}{2} z_i^2 \tag{7.3.39}$$

whose derivative is

$$\begin{aligned}\dot{V}_i &= -\sum_{k=1}^{i-1} c_k z_k^2 + \left(\sum_{k=1}^{i-2} z_{k+1}\frac{\partial \alpha_k}{\partial \hat{\theta}}\right)(\Gamma\tau_{i-1} - \dot{\hat{\theta}}) \\ &\quad + z_i\left[z_{i-1} + z_{i+1} + \alpha_i - \sum_{k=1}^{i-1}\frac{\partial \alpha_{i-1}}{\partial x_k}x_{k+1} + w_i^T\hat{\theta} - \frac{\partial \alpha_{i-1}}{\partial \hat{\theta}}\dot{\hat{\theta}}\right] \\ &\quad + \tilde{\theta}^T\left(\tau_{i-1} + w_i z_i - \Gamma^{-1}\dot{\hat{\theta}}\right) \end{aligned} \tag{7.3.40}$$

We can eliminate $\tilde{\theta}$ from $\dot{V}_i$ with the update law $\dot{\hat{\theta}} = \Gamma\tau_i$, where

$$\begin{aligned}\tau_i(x_1, \ldots, x_i, \hat{\theta}) &= \tau_{i-1} + z_i w_i \\ &= \begin{bmatrix} w_1, & \ldots & w_i \end{bmatrix}\begin{bmatrix} z_1 \\ \vdots \\ z_i \end{bmatrix}\end{aligned} \tag{7.3.41}$$

Then, in the absence of z_{i+1}, we would achieve $\dot{V}_i = -\sum_{k=1}^{i} c_k z_k^2$, by designing α_i to make the bracketed term multiplying z_i equal to $-c_i z_i$, namely

$$\begin{aligned}\alpha_i(x_1, \ldots, x_i, \hat{\theta}) &= -z_{i-1} - c_i z_i + \sum_{k=1}^{i-1}\frac{\partial \alpha_{i-1}}{\partial x_k}x_{k+1} - w_i^T\hat{\theta} + \frac{\partial \alpha_{i-1}}{\partial \hat{\theta}}\Gamma\tau_i \\ &\quad + \nu_i \end{aligned} \tag{7.3.42}$$

where ν_i is a correction term yet to be chosen. Nothing that

$$\begin{aligned}\dot{\hat{\theta}} - \Gamma\tau_{i-1} &= \dot{\hat{\theta}} - \Gamma\tau_i + \Gamma\tau_i - \Gamma\tau_{i-1} \\ &= \dot{\hat{\theta}} - \Gamma\tau_i + \Gamma w_i z_i \end{aligned} \tag{7.3.43}$$

we rewrite $\dot{V}_i$ as

$$\begin{aligned}\dot{V}_i &= -\sum_{k=1}^{i-1} c_k z_k^2 + z_i\left[z_{i+1} + \nu_i - \frac{\partial \alpha_{i-1}}{\partial \hat{\theta}}(\Gamma\tau_i - \dot{\hat{\theta}})\right] \\ &\quad + \left(\sum_{k=1}^{i-2} z_{k+1}\frac{\partial \alpha_k}{\partial \hat{\theta}}\right)(\Gamma\tau_{i-1} - \dot{\hat{\theta}}) + \tilde{\theta}^T(\tau_i - \Gamma^{-1}\dot{\hat{\theta}}) \\ &= \sum_{k=1}^{i-1} c_k z_k^2 + z_i\left[z_{i+1} + \nu_i - \sum_{k=1}^{i-2} z_{k+1}\frac{\partial \alpha_k}{\partial \hat{\theta}}\Gamma w_i\right] \\ &\quad + \left(\sum_{k=1}^{i-1} z_{k+1}\frac{\partial \alpha_k}{\partial \hat{\theta}}\right)(\Gamma\tau_i - \dot{\hat{\theta}}) + \tilde{\theta}^T(\tau_i - \Gamma^{-1}\dot{\hat{\theta}}) \end{aligned} \tag{7.3.44}$$

and represent the $(z_1, \ldots, z_i)$-subsystem as

$$\begin{bmatrix} \dot{z}_1 \\ \vdots \\ \dot{z}_i \end{bmatrix} = \begin{bmatrix} -c_1 & 1 & 0 & \cdots & 0 & 0 \\ -1 & -c_2 & 1+\sigma_{23} & \cdots & \sigma_{2,i-1} & 0 \\ 0 & -1-\sigma_{23} & \ddots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 1+\sigma_{i-2,i-1} & 0 \\ 0 & -\sigma_{2,i-1} & \cdots & -1-\sigma_{i-2,i-1} & -c_{i-1} & 1 \\ 0 & 0 & \cdots & 0 & -1 & -c_i \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_i \end{bmatrix}$$

$$+ \begin{bmatrix} w_1^T \\ \vdots \\ w_i^T \end{bmatrix} \tilde{\theta} + \begin{bmatrix} 0 \\ \sigma_{2,i} z_i \\ \vdots \\ \sigma_{i-1,i} z_i \\ \nu_i \end{bmatrix} + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ z_{i+1} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{\partial \alpha_1}{\partial \hat{\theta}} \\ \vdots \\ \frac{\partial \alpha_{i-1}}{\partial \hat{\theta}} \end{bmatrix} (\Gamma \tau_i - \dot{\hat{\theta}}) \tag{7.3.45}$$

where

$$\sigma_{jk}(x, \hat{\theta}) = -\frac{\partial \alpha_{j-1}}{\alpha \hat{\theta}} \Gamma w_k \tag{7.3.46}$$

Now the correction term is chosen as

$$\nu_i(x_1, \ldots, x_i, \hat{\theta}) = \sum_{k=1}^{i-2} z_{k+1} \frac{\partial \alpha_k}{\partial \hat{\theta}} \Gamma w_i = -\sum_{k=2}^{i-1} \sigma_{k,i} z_k \tag{7.3.47}$$

Because we do not use $\dot{\hat{\theta}} = \Gamma \tau_i$ as an update law, the resulting $\dot{V}_i$ is

$$\dot{V}_i = \sum_{k=1}^{i} c_k z_k^2 + z_i z_{i+1} + \left(\sum_{k=1}^{i-1} z_{k+1} \frac{\partial \alpha_k}{\partial \hat{\theta}} \right) (\Gamma \tau_i - \dot{\hat{\theta}}) + \tilde{\theta}^T (\tau_i - \Gamma^{-1} \dot{\hat{\theta}}) \tag{7.3.48}$$

and the $(z_1, \ldots, z_i)$-subsystem becomes

$$\begin{bmatrix} \dot{z}_1 \\ \vdots \\ \dot{z}_i \end{bmatrix} = \begin{bmatrix} -c_1 & 1 & 0 & \cdots & 0 \\ -1 & -c_2 & 1+\sigma_{23} & \cdots & \sigma_{2i} \\ 0 & -1-\sigma_{23} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 1+\sigma_{i-1,i} \\ 0 & -\sigma_{2i} & \cdots & -1-\sigma_{i-1,i} & -c_i \end{bmatrix} \begin{bmatrix} z_1 \\ \vdots \\ z_i \end{bmatrix}$$
$$+ \begin{bmatrix} w_1^T \\ \vdots \\ w_i^T \end{bmatrix} \tilde{\theta} + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ z_{i+1} \end{bmatrix} + \begin{bmatrix} 0 \\ \frac{\partial \alpha_1}{\partial \hat{\theta}} \\ \vdots \\ \frac{\partial \alpha_{i-1}}{\partial \hat{\theta}} \end{bmatrix} (\Gamma \tau_i - \dot{\hat{\theta}}) \tag{7.3.49}$$

Step n. At the final step, we introduce

$$z_n = x_n - \alpha_{n-1} \tag{7.3.50}$$

and rewrite the last equation $\dot{x}_n = \beta(x)u + \varphi_n(x)^T\theta$ as

$$\dot{z}_n = \beta u + \varphi_n^T \theta - \sum_{k=1}^{n-1} \frac{\partial \alpha_{n-1}}{\partial x_k}(x_{k+1} + \varphi_k^T \theta) - \frac{\partial \alpha_{n-1}}{\partial \hat{\theta}} \dot{\hat{\theta}} \tag{7.3.51}$$

where the last regressor vector is defined as

$$w_n(x, \hat{\theta}) = \varphi_n - \sum_{k=1}^{n-1} \frac{\partial \alpha_{n-1}}{\partial x_k} \varphi_k \tag{7.3.52}$$

In this equation, the actual control input is at our disposal. We are finally in the position to design our actual update law $\dot{\hat{\theta}} = \Gamma \tau_n$ and feedback control u to stabilize the full z-system with respect to

$$\begin{aligned} V_n &= V_{n-1} + \frac{1}{2} z_n^2 \\ &= \frac{1}{2} z^T z + \frac{1}{2} \tilde{\theta}^T \Gamma^{-1} \tilde{\theta} \end{aligned} \tag{7.3.53}$$

Our goal is to make $\dot{V}_n$ nonpositive:

$$\begin{aligned}\dot{V}_n = & -\sum_{k=1}^{n-1} c_k z_k^2 + \left(\sum_{k=1}^{n-2} z_{k+1}\frac{\partial \alpha_k}{\partial \hat{\theta}}\right)(\Gamma\tau_{n-1} - \dot{\hat{\theta}}) \\ & + z_n\left[z_{n-1} + \beta u - \sum_{k=1}^{n-1}\frac{\partial \alpha_{n-1}}{\partial x_k}x_{k+1} + w_n^T\hat{\theta} - \frac{\partial \alpha_{n-1}}{\partial \hat{\theta}}\dot{\hat{\theta}}\right] \\ & + \tilde{\theta}^T\left(\tau_{n-1} + w_n z_n - \Gamma^{-1}\dot{\hat{\theta}}\right) \end{aligned} \tag{7.3.54}$$

To eliminate $\tilde{\theta}$ from $\dot{V}_n$, we choose the update law

$$\begin{aligned}\dot{\hat{\theta}} &= \Gamma\tau_n(z,\hat{\theta}) = \Gamma\tau_{n-1} + \Gamma w_n z_n \qquad (7.3.55)\\ &= \Gamma W(z,\hat{\theta})z\end{aligned}$$

where the regressor matrix W is composed of the regressor vector $w_1, \ldots, w_n$:

$$W(z,\hat{\theta}) = [w_1, \ldots, w_n] \tag{7.3.56}$$

We choose the control u to make the bracketed term multiplying z_n equal to $-c_n z_n$:

$$u = \frac{1}{\beta}\left(-z_{n-1} - c_n z_n + \sum_{k=1}^{n-1}\frac{\partial \alpha_{n-1}}{\partial x_k}x_{k+1} - w_n^T z_n + \frac{\partial \alpha_{n-1}}{\partial \hat{\theta}}\Gamma\tau_n + \nu_n\right) \tag{7.3.57}$$

where ν_n is a correction term yet to be chosen. With (7.3.57), $\dot{V}_n$ becomes

$$\dot{V}_n = -\sum_{k=1}^{n-1} c_k z_k^2 + \left(\sum_{k=1}^{n-2} z_{k+1}\frac{\partial \alpha_k}{\partial \hat{\theta}}\right)(\Gamma\tau_{n-1} - \dot{\hat{\theta}}) + z_n\nu_n \tag{7.3.58}$$

Then, noting that

$$\dot{\hat{\theta}} - \Gamma\tau_{n-1} = \Gamma\tau_n - \Gamma\tau_{n-1} = \Gamma w_n z_n \tag{7.3.59}$$

we rewrite $\dot{V}_n$ as

$$\dot{V}_n = -\sum_{k=1}^{n-1} c_k z_k^2 + z_n\left(\nu_n - \sum_{k=1}^{n-2} z_{k+1}\frac{\partial \alpha_k}{\partial \hat{\theta}}\Gamma w_n\right) \tag{7.3.60}$$

Now the correction term ν_n is chosen as

$$\nu_n(x,\hat{\theta}) = \sum_{k=1}^{n-2} z_{k+1}\frac{\partial \alpha_k}{\partial \hat{\theta}}\Gamma w_n = -\sum_{k=2}^{n-1}\gamma_{k,n}z_k \tag{7.3.61}$$

We have thus reached our goal:

$$\dot{V}_n = -\sum_{k=1}^{n} c_k z_k^2 \tag{7.3.62}$$

The overall closed-loop system is

$$\begin{aligned} \dot{z} &= A_z(z,\hat{\theta})z + W(z,\theta)^T\tilde{\theta} && (7.3.63)\\ \dot{\hat{\theta}} &= \Gamma W(z,\hat{\theta})z && (7.3.64) \end{aligned}$$

where

$$A_z(z,\hat{\theta}) = \begin{bmatrix} -c_1 & 1 & 0 & \cdots & 0 \\ -1 & -c_2 & 1+\sigma_{23} & \cdots & \sigma_{2n} \\ 0 & -1-\sigma_{23} & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 1+\sigma_{n-1,n} \\ 0 & -\sigma_{2n} & \cdots & -1-\sigma_{n-1,n} & -c_n \end{bmatrix} \tag{7.3.65}$$

The system (7.3.63) will be referred to as the *error system.* It is important to note that a major portion of the designed effort was invested into achieving

$$A_z(z,\hat{\theta}) + A_z(z,\hat{\theta})^T = -2\begin{bmatrix} c_1 & & \\ & \ddots & \\ & & c_n \end{bmatrix}, \quad \forall (z,\hat{\theta}) \in \mathbb{R}^{n+p} \tag{7.3.66}$$

which yields (7.3.62) with the simple quadratic Lyapunov function (7.3.53). We observe that, as desired, the system (7.3.63)-(7.3.64) has an equilibrium at $(z,\tilde{\theta}) = (0,0)$. The stability properties of this equilibrium will be established in Section 7.3.2.

Example 7.3.1 *In application of the tuning functions procedure, we do not need to repeat the Lyapunov argument. All we need for a specific design are the final analytical expressions provided by the procedure. Let us now illustrate this by* *designing an adaptive controller for the benchmark system from Example 7.2.7:*

$$\begin{aligned}\dot{x}_1 &= x_2 + \varphi(x_1)^T\theta \\ \dot{x}_2 &= x_3 \\ \dot{x}_3 &= u\end{aligned} \tag{7.3.67}$$

The design objective is the regulation of the output $y = x_1$ to the set-point y_s. The first three expressions provided by the procedure are the definitions (7.3.4), (7.3.5), *and* (7.3.14) *of the error variables*

$$\begin{aligned}z_1 &= x_1 - y_s \\ z_2 &= x_2 - \alpha_1(x_1, \hat{\theta}) \\ z_3 &= x_3 - \alpha_2(x_1, x_2, \hat{\theta})\end{aligned} \tag{7.3.68}$$

where α_1 and α_2 are the stabilizing functions given by (7.3.11) *and* (7.3.20):

$$\begin{aligned}\alpha_1 &= -c_1 z_1 - \varphi^T\hat{\theta} \\ \alpha_2 &= -c_2 z_2 - z_1 + \frac{\partial\alpha_1}{\partial x_1}(x_2 + \varphi^T\hat{\theta}) + \frac{\partial\alpha_1}{\partial\hat{\theta}}\tau_2\end{aligned} \tag{7.3.69}$$

The tuning functions, determined from (7.3.10), (7.3.19), *and* (7.3.28), *are*

$$\begin{aligned}\tau_1 &= z_1\varphi \\ \tau_2 &= \tau_1 - z_2\frac{\partial\alpha_1}{\partial x_1}\varphi \\ \tau_3 &= \tau_2 - z_3\frac{\partial\alpha_2}{\partial x_1}\varphi\end{aligned} \tag{7.3.70}$$

With the above expressions and the choice $\Gamma = I$, the parameter update law and the feedback control are obtained from (7.3.55) *and* (7.3.29), *respectively. They are*

$$\dot{\hat{\theta}} = \tau_3 = z_1\varphi - z_2\frac{\partial\alpha_1}{\partial x_1}\varphi - z_3\frac{\partial\alpha_2}{\partial x_1}\varphi \tag{7.3.71}$$

$$\begin{aligned}u &= -c_3 z_3 - z_2 + \frac{\partial\alpha_2}{\partial x_1}(x_2 + \varphi^T\hat{\theta}) + \frac{\partial\alpha_2}{\partial x_2}x_3 + \frac{\partial\alpha_2}{\partial\hat{\theta}}\tau_3 \\ &\quad -z_2\frac{\partial\alpha_2}{\partial\hat{\theta}}\frac{\partial\alpha_2}{\partial x_1}\varphi\end{aligned} \tag{7.3.72}$$

This completes the design of the adaptive controller for (7.3.67). *In the* $(z,\tilde{\theta})$-*coordinates the designed system is*

$$\dot{z} = \begin{bmatrix} -c_1 & 1 & 0 \\ -1 & -c_2 & 1-\frac{\partial\alpha_2}{\partial x_1}|\varphi|^2 \\ 0 & -1+\frac{\partial\alpha_2}{\partial x_1}|\varphi|^2 & -c_3 \end{bmatrix} z + \begin{bmatrix} 1 \\ -\frac{\partial\alpha_1}{\partial x_1} \\ -\frac{\partial\alpha_2}{\partial x_1} \end{bmatrix} \varphi^T\tilde{\theta} \quad (7.3.73)$$

$$\dot{\tilde{\theta}} = -\varphi\begin{bmatrix} 1, & -\frac{\partial\alpha_1}{\partial x_1}, & -\frac{\partial\alpha_2}{\partial x_1} \end{bmatrix} z \quad (7.3.74)$$

It is of interest to relate the stabilizing functions α_1 *and* α_2 *and the control law* u *to the material from Section 7.2. The stabilizing function* α_1 *has a "certainty equivalence" form. The stabilizing function* α_2 *has the term* $-\frac{\partial\alpha_1}{\partial\hat{\theta}}\tau_2$ *which accounts for parameter estimation transients, while the rest of it is in the "certainty equivalence" form. The control law* u *departs from the "certainty equivalence" form in the last two terms whose role in the same as that of* (7.2.60). *The last term in* u *is particularly important. Since* $-\frac{\partial\alpha_1}{\partial\hat{\theta}} = -\varphi^T$, *this term contributes with* $+-\frac{\partial\alpha_2}{\partial x_1}|\varphi|^2$ *in the 1system matrix' in* (7.3.73) *and achieves the skew symmetry, which is crucial for stability.* □

7.3.2 Stability and convergence

To investigate stability properties of the closed-loop adaptive system (7.3.63)-(7.3.64), we express φ_i, α_i, τ_i, and w_i in the z-coordinates. Then, by the uniform stability theorem, the global stability of the equilibrium $(z,\tilde{\theta}) = 0$ follows from the fact that the derivative $\dot{V}_n$ of V_n along the solutions of (7.3.63)-(7.3.64) is given by (7.3.62).

From LaSalle's Invariance Theorem, it further follows that the $((n+p)$-dimensional) state $(z(t),\tilde{\theta}(t))$ converges to the largest invariant set where $\dot{V}_n = 0$. This means, in particular, that $z(t) \to 0$ as $t \to 0$.

We now set out to determine M. On this invariant set, we have $z \equiv 0$ and

$\dot{z} \equiv 0$. Setting $z = 0$, $\dot{z} = 0$ in (7.3.63) we obtain $\dot{\hat{\theta}} = 0$ and

$$W(z,\hat{\theta})^T(\theta - \hat{\theta}) = 0, \quad \forall (z,\hat{\theta}) \in M \tag{7.3.75}$$

From (7.3.38) and (7.3.56), it is easily seen that

$$W(z,\hat{\theta})^T = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ -\frac{\partial \alpha_1}{\partial x_1} & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ -\frac{\partial \alpha_{n-1}}{\partial x_1} & \cdots & -\frac{\partial \alpha_{n-1}}{\partial x_{n-1}} & 1 \end{bmatrix} F(x)^T \triangleq N(z,\hat{\theta})F(x)^T \tag{7.3.76}$$

Since $N(z,\hat{\theta})$ is obviously nonsingular for all $(z,\hat{\theta}) \in M$, then (7.3.75) and (7.3.76) imply

$$F(x)^T(\theta - \hat{\theta}) = 0 \quad on\ M \tag{7.3.77}$$

Now we show that $x = x^e$ on M. Since $z_1 = x_1 - y_s$ then $x_1 = y_s = x_1^e$ on M. In view of (7.3.77), we get

$$(\theta - \hat{\theta})^T \varphi_1(x_1^e) = 0 \quad on\ M \tag{7.3.78}$$

Recall from (7.3.11) that $\alpha_1 = -c_1 z_1 - \hat{\theta}^T\varphi_1$. Therefore, on M, we have $\alpha_1 = -\hat{\theta}^T\varphi_1(x_1^e)$. Combining this with $z_2 = 0 = x_2 - \alpha_1$ and (7.3.3), we get $x_2 = x_2^e$ on M. Using (7.3.77), we obtain

$$(\theta - \hat{\theta})^T \varphi_2(x_1^e, x_2^e) = 0 \quad on\ M \tag{7.3.79}$$

Continuing in the same fashion, we prove that $x_i = x_i^e$ and $(\theta - \hat{\theta})^T\varphi_i(x_1^e, \ldots, x_i^e) = 0$ on M, $i = 1, \ldots, n$. Thus, the largest invariant set M in E is

$$\begin{aligned} M &= \left\{(z,\tilde{\theta}) \in \mathbb{R}^{n+p} \mid z = 0,\ F_e^T\tilde{\theta} = 0\right\} \\ &= \left\{(x,\hat{\theta}) \in \mathbb{R}^{n+p} \mid x = 0,\ F_e^T\hat{\theta} = F_e^T\theta\right\} \end{aligned} \tag{7.3.80}$$

where $F_e = F(x^e)$. The two equivalent expressions for M and the convergence of $(z(t), \tilde{\theta}(t))$ to M prove that $x(t) \to x^e$ as $t \to \infty$.

An important property of M is its dimension, $p-rank\{F_e\}$. When $rank\{F_e\} = p$, then $\dim M = 0$, that is, M becomes the equilibrium point $x = x^e$, $\hat{\theta} = \theta$. This means that the parameter estimates converge to their true values, so that the equilibrium $x = x^e$, $\hat{\theta} = \theta$ is globally *asymptotically* stable.

The above facts prove the following result:

Theorem 7.3.2 *The closed-loop adaptive system consisting of the plant* (7.3.1), *the controller* (7.3.57), *and the update law* (7.3.55) *has a globally stable equilibrium* $(x, \hat{\theta}) = (x^e, \theta)$. *Furthermore, its state* $(x(t), \hat{\theta}(t))$ *converges to the* $(p-rank\{F_e\})$-*dimensional equilibrium manifold* M *given by* (7.3.80), *which means, in particular, that*

$$\lim_{t\to\infty} x(t) = x^e \tag{7.3.81}$$

If $y_s = 0$ *and* $F(0) = 0$, *then* $\lim_{t\to\infty} x(t) = 0$. *The equilibrium* $x = x^e$, $\hat{\theta} = \theta$ *is globally asymptotically stable if and only if* $rank\{F_e\} = p$. □

As the dimension of M reduces, the stability properties of the adaptive system improve. The most desirable case is when M is an equilibrium point, in which case, this equilibrium is globally asymptotically stable, and the parameter estimates converge to the actual parameter values. Global asymptotic stability can be achieved with as many as $p = n$ unknown parameters. This is among the main advantages of eliminating overparametrization.

We now discuss the basic stability properties established in Theorem 7.2.12 on a simple example.

Example 7.3.3 *We consider the second order system with an unknown parameter vector* $\theta \in \mathbb{R}^p$:

$$\begin{aligned} \dot{x}_1 &= x_2 + \varphi_1(x_1)^T\theta \\ \dot{x}_2 &= u + \varphi_2(x)^T\theta \end{aligned} \tag{7.3.82}$$

The control objective is to regulate x to zero ($x_1^e = 0$). *We define the error variable*

$$\begin{aligned} z_1 &= x_1 \\ z_2 &= x_2 - \alpha_1(x_1, \hat{\theta}) \end{aligned} \tag{7.3.83}$$

The controller is designed applying (7.3.11) *and* (7.3.20) *as*

$$\begin{aligned} \alpha_1 &= -c_1 z_1 - \varphi_1(x_1)^T\hat{\theta} \\ \frac{\partial\alpha_1}{\partial x_1} &= -c_1 - \frac{\partial\varphi_1 x_1{}^T}{\partial x_1}\hat{\theta}, \qquad \frac{\partial\alpha_1}{\partial\hat{\theta}} = -\varphi_1(x_1)^T \\ u &= -z_1 - c_2 z_2 + \frac{\partial\alpha_1}{\partial x_1}x_2 - \left(\varphi_2(x)^T - \frac{\partial\alpha_1}{\partial x_1}\varphi_1(x_1)^T\right)\hat{\theta} + \frac{\partial\alpha_1}{\partial\hat{\theta}}\dot{\hat{\theta}} \end{aligned} \tag{7.3.84}$$

while the parameter update law is

$$\dot{\hat{\theta}} = \Gamma\left[\begin{array}{cc} \varphi_1, & \varphi_2 - \frac{\partial\alpha_1}{\partial x_1}\varphi_1 \end{array}\right] z \tag{7.3.85}$$

The resulting error system is

$$\dot{z} = \begin{bmatrix} -c_1 & 1 \\ -1 & -c_2 \end{bmatrix} z + \begin{bmatrix} \varphi_1^T \\ \varphi_2^T - \frac{\partial\alpha_1}{\partial x_1}\varphi_1^T \end{bmatrix}\tilde{\theta} \tag{7.3.86}$$

Now we illustrate and discuss the stability properties established by Theorem 7.3.2. From (7.3.82) we see that $x_1^e = 0$, $x_2^e = -\varphi_1(0)^T\theta$. By Theorem 7.3.2, the point

$$\begin{bmatrix} x_1 \\ x_2 \\ \hat{\theta} \end{bmatrix} = \begin{bmatrix} 0 \\ -\varphi_1(0)^T\theta \\ \theta \end{bmatrix} \tag{7.3.87}$$

is a globally stable equilibrium, and the stable of the closed-loop system converges to the equilibrium manifold

$$\begin{aligned} M \;=\; \Bigg\{ (x,\hat{\theta}) \in \mathbb{R}^{2+p} \mid & \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ -\varphi_1(0)^T\theta \end{bmatrix} \\ & , \begin{bmatrix} \varphi_1(0)^T \\ \varphi_2(0, -\varphi_1(0)^T\theta)^T \end{bmatrix}(\theta - \hat{\theta}) = 0 \Bigg\} \end{aligned} \tag{7.3.88}$$

A basic question that one would ask is: What type of figure in $\mathbb{R}^{2+p}$ is M? Our further discussion will, without loss of generality, be limited to $p \leq 2$.

In the simplest case where $\dim \theta = p = 1$, the following two possibilities exist:

- If both $\varphi_1(0) = 0$ and $\varphi_2(0,0) = 0$, then the manifold M is the subspace $x = 0 \subset \mathbb{R}^3$, that is, M is the $\hat{\theta}$-axis
- If either $\varphi_1 \neq 0$ or $\varphi_2(0, -\varphi_1(0)\theta) \neq 0$, then the manifold M is the single point $x_1 = 0$, $x_2 = -\varphi_1(0)\theta$, $\hat{\theta} = \theta$. This point is an equilibrium which is not only globally stable, but also globally *asymptotically* stable.

Next, we analyze the case $p = 2$.

- Suppose $\begin{bmatrix} \varphi_1(x_1)^T \\ \varphi_2(x_1)^T \end{bmatrix} = \begin{bmatrix} x_1^2 & e^{x_1} \\ \cos x_1 & 0 \end{bmatrix}$. Since $\begin{bmatrix} \varphi_1(0)^T \\ \varphi_2(0)^T \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ has full rank, the manifold M is the single point $x_1 = 0$, $x_2 = -\theta_2$, $\hat{\theta}_1 = \theta_1$, $\hat{\theta}_2 = \theta_2$, which is a globally *asymptotically* stable equilibrium.
- Suppose $\begin{bmatrix} \varphi_1(x_1)^T \\ \varphi_2(x_1)^T \end{bmatrix} = \begin{bmatrix} -\cos x_1 & e^{x_1} \\ \sin x_1 & 0 \end{bmatrix}$. Since $\begin{bmatrix} \varphi_1(0)^T \\ \varphi_2(0)^T \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ 0 & 0 \end{bmatrix}$, the manifold M is the linear variety $x_1 = 0$, $x_2 = \theta_1 - \theta_2$, $\hat{\theta}_2 - \hat{\theta}_1 = \theta_2 - \theta_1$. Neither of the parameter estimates in guaranteed to converge to the actual parameter value, but they are jointly converging to the line $\hat{\theta}_2 = \hat{\theta}_1 + \theta_2 - \theta_1$ in the plane $x_1 = 0$, $x_2 = \theta_1 - \theta_2$.
- Suppose $\begin{bmatrix} \varphi_1(x_1)^T \\ \varphi_2(x_1)^T \end{bmatrix} = \begin{bmatrix} x_1^2 & e^{x_1} - 1 \\ \sin x_1 & 0 \end{bmatrix}$. Since $\begin{bmatrix} \varphi_1(0)^T \\ \varphi_2(0)^T \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$, the manifold M is the plane (linear variety) $x = 0$. This is the case of the weakest convergence properties because one cannot guarantee that the parameter estimates converge to any submanifold in the plane M.

7.4 Tracking

The set-point regulation design is readily extended to the task of tracking.

The control objective is to free the output $y = x_1$ of the system (7.3.1)

$$\begin{aligned} \dot{x}_1 &= x_2 + \varphi_1(x_1)^T\theta \\ \dot{x}_2 &= x_3 + \varphi_2(x_1, x_2)^T\theta \\ &\vdots \\ \dot{x}_{n-1} &= x_n + \varphi_{n-1}(x_1, \ldots, x_{n-1})^T\theta \\ \dot{x}_n &= \beta(x)u + \varphi_n(x)^T\theta \end{aligned} \tag{7.4.1}$$

to asymptotically track the reference output $y_r(t)$ whose first n derivative are assumed to be known, bounded, and piecewise continuous.

An alternative control objective would be asymptotically track the output of a known asymptotically stable linear reference model

$$y_r = G_m(s)r(s) = \frac{k_m}{s^n + m_{n-1}z^{n-1} + \ldots + m_0} r(s) \tag{7.4.2}$$

where the denominator is Hurwitz, $k_m > 0$, and $r(t)$ is bounded and piecewise continuous. A realization which is of particular interest is

$$\begin{aligned} \dot{x}_m &= \begin{bmatrix} 0 & 1 & & 0 \\ \vdots & & \ddots & \\ 0 & 0 & & 1 \\ -m_0 & \cdots & \cdots & -m_{n-1} \end{bmatrix} x_m + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ k_m \end{bmatrix} \\ y_r &= x_{m,1} \end{aligned} \tag{7.4.3}$$

because, in this case, the derivatives of y_r are available as the states of the reference model: $y_r^{(i)} = x_{m,i+1}$, $i = 0, \ldots, n-1$.

These functions are used in the design for tracking.

$$
\begin{aligned}
z_i &= x_i - y_r^{i-1} - \alpha_{i-1} && (7.4.4)\\
\alpha_i(\bar{x}_i, \hat{\theta}, \bar{y}_r^{(i-1)}) &= -z_{i-1} - c_i z_i - w_i^T \hat{\theta} + \sum_{k=1}^{i-1}\left(\frac{\partial \alpha_{i-1}}{\partial x_k} x_{k+1} + \frac{\partial \alpha_{i-1}}{\partial y_r^{(k-1)}} y_r^{(k)}\right) \\
&\quad -k_i |w_i|^2 z_i + \frac{\partial \alpha_{i-1}}{\partial \hat{\theta}} \Gamma \tau_i + \sum_{k=2}^{i-1} \frac{\partial \alpha_{k-1}}{\partial \hat{\theta}} \Gamma w_i z_k && (7.4.5)\\
\tau_i(\bar{x}_i, \hat{\theta}, \bar{y}_r^{(i-1)}) &= \tau_{i-1} + w_i z_i && (7.4.6)\\
w_i(\bar{x}_i, \hat{\theta}, \bar{y}_r^{(i-2)}) &= \varphi_i - \sum_{k=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_k} \varphi_k && (7.4.7)
\end{aligned}
$$

for $i = 1, \ldots, n$, $\bar{x}_i = (x_1, \ldots, x_i)$ and $\bar{y}_r^{(i)} = (y_r, \dot{y}_r, \ldots, y_r^{(i)})$

Adaptive control law:

$$
u = \frac{1}{\beta(x)}[\alpha_n(x, \hat{\theta}, \bar{y}_r^{(n-1)}) + y_r^{(n)}] \tag{7.4.8}
$$

Parameter update law:

$$
\dot{\hat{\theta}} = \Gamma \tau_n(x, \hat{\theta}, \bar{y}_r^{(n-1)}) = \Gamma W z \tag{7.4.9}
$$

The closed-loop adaptive system consisting of the plant (7.4.1), the controller (7.4.8), and the update law (7.4.9) has a globally uniformly stable equilibrium at $(z, \tilde{\theta}) = 0$, and $\lim_{t\to\infty} = 0$, which means, in particular, that global asymptotic tracking is achieved:

$$
\lim_{t\to\infty}[y(t) - y_r(t)] = 0 \tag{7.4.10}
$$

Moreover, if $\lim_{t\to\infty} y_r^{(i)} = 0$, $i = 0, \ldots, n-1$, and $F(0) = -$, then $\lim_{t\to\infty} x(t) = 0$

7.5 Unknown virtual control coefficients

For the sake of clarity, the adaptive design in this chapter was presented for the class of parametric strict-feedback system. We now give an extension of the tuning function design for systems with unknown virtual control coefficient. Consider systems of the form

$$\begin{aligned}
\dot{x}_1 &= b_1 x_2 + \varphi_1(x_1)^T\theta \\
\dot{x}_2 &= b_2 x_3 + \varphi_2(x_1, x_2)^T\theta \\
&\vdots \\
\dot{x}_i &= b_i x_{i+1} + \varphi_i(x_1, \ldots, x_i)^T\theta, \qquad i = 1, \ldots, n-1 \\
&\vdots \\
\dot{x}_n &= b_n\beta(x)u + \varphi_n(x_1, \ldots, x_n)^T\theta
\end{aligned} \tag{7.5.1}$$

where, in addition to the unknown vector θ, the constant coefficients b_i are also unknown. We refer to the coefficients b_i as the 'virtual control coefficients' The occurrence of the unknown b_i-coefficients is frequent in applications ranging from electric motors and robotic manipulators to flight dynamics.

When the signs of b_i, $i = 1, \ldots, n$, are known.

We consider two special cases of (7.5.1). The extension to the general case is straightforward but tedious.

The first special case is when the only unknown virtual control coefficient is the 'high-frequency gain' b_n:

$$\begin{aligned}
\dot{x}_i &= x_{i+1} + \varphi_i(x_1, \ldots, x_i)^T\theta, \qquad i = 1, \ldots, n-1 \\
\dot{x}_n &= b_n\beta(x)u + \varphi_n(x_1, \ldots, x_n)^T\theta
\end{aligned} \tag{7.5.2}$$

For this case the modification of the tuning functions design is simple. In the design of tuning function for tracking, we only need to change the control law (7.4.8):

$$u = \frac{\hat{\varrho}}{\beta(x)}[\alpha_n(x, \hat{\theta}, \bar{y}_r^{(n-1)}) + y_r^{(n)}] \tag{7.5.3}$$

where $\hat{\varrho}$ is the estimate of $\varrho = 1/b_n$ computed as

$$\dot{\hat{\varrho}} = -\gamma sgn(b_n)\left(\alpha_n + y_r^{(n)}\right) z_n, \qquad \gamma > 0 \tag{7.5.4}$$

We are only using the knowledge of the *sign* of the unknown parameter b_n. In this simple case it is not necessary to estimate b_n itself. It can be checked that the resulting error system has the form (??) with an additional term due to $\tilde{\varrho} = \varrho - \hat{\varrho}$:

$$\dot{z} = A_z(z,\hat{\theta},t)z + W(z,\hat{\theta},t)^T\tilde{\theta} + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ -b_n\left(\alpha_n + y_r^{(n)}\right) \end{bmatrix} \tilde{\varrho} \tag{7.5.5}$$

Consider the Lyapunov function

$$V = \frac{1}{2}z^Tz + \frac{1}{2}\tilde{\theta}^T\Gamma^{-1}\tilde{\theta} + \frac{|b_n|}{2\gamma}\tilde{\varrho}^2 \tag{7.5.6}$$

Its derivative along the solutions of (7.5.4) and (7.5.5) is

$$\dot{V} \leq -c_0|z|^2 \tag{7.5.7}$$

which satisfies tracking conditions, which means, all the states are bounded and asymptotic tracking is achieved.

Now we move on to a more difficult case:

$$\begin{aligned} \dot{x}_i &= x_{i+1} + \varphi_i(x_1,\ldots,x_i)^T\theta, \qquad i = 1,\ldots,m-1,m+1,\ldots,n-1 \\ \dot{x}_m &= b_m x_{m+1} + \varphi_m(x_1,\ldots,x_m)^T\theta \\ \dot{x}_n &= \beta(x)u + \varphi_n(x_1,\ldots,x_n)^T\theta \end{aligned} \tag{7.5.8}$$

where b_m, $m < n$, is the only unknown coefficient. From step m on, the design procedure for this case differs considerably from the tracking design procedure. We now need $\hat{b}_m$ and $\hat{\varrho}$ the estimates of b_m and $\varrho = 1/b_m$. The estimate $\hat{\varrho}$ is introduced to avoid the division by $\hat{b}_m(t)$ which can occasionally take value zero. The complete design procedure is given by the following expressions (with $z_0 = 0$, $\alpha_0 = 0$, $\tau_0 = 0$):

Coordinate transformation:

$$z_i = x_i - y_r^{(i-1)} - \alpha_{i-1}, \qquad i = 1, \ldots, m \tag{7.5.9}$$

$$z_j = x_j - \hat{\varrho} y_r^{(j-1)} - \alpha_{j-1} \qquad j = m+1, \ldots, n \tag{7.5.10}$$

Regressor:

$$w_i = \varphi_i - \sum_{k=1}^{i-1} \frac{\partial \alpha_{i-1}}{\partial x_k} \varphi_k, \qquad i = 1, \ldots, n \tag{7.5.11}$$

Tuning function for $\hat{\theta}$:

$$\tau_i = \tau_{i+1} + w_i z_i, \qquad i = 1, \ldots, n \tag{7.5.12}$$

Tuning function for $\hat{b}_m$:

$$\pi_m = z_{m+1} z_m \tag{7.5.13}$$

$$\pi_j = \pi_{j-1} - \frac{\partial \alpha_m}{\partial x_m} x_{m+1} z_j, \qquad j = m+1, \ldots, n \tag{7.5.14}$$

Stabilizing functions:

$$\alpha_i(\bar{x}_i, \hat{\theta}, \bar{y}_r^{(i-1)}) = \bar{\alpha}_i, \qquad i = 1, \ldots, m-1 \tag{7.5.15}$$

$$\alpha_m(\bar{x}_m, \hat{\theta}, \bar{y}_r^{(m-1)}, \hat{\varrho}) = \hat{\varrho} \bar{\alpha}_m \tag{7.5.16}$$

$$\alpha_j(\bar{x}_j, \hat{\theta}, \bar{y}_r^{(j-1)}, \hat{b}_m, \hat{\varrho}) = \bar{\alpha}_j, \qquad j = m+1, \ldots, n \tag{7.5.17}$$

$$
\begin{aligned}
\bar{\alpha}_i &= -z_{i-1} - c_i z_i - w_i^T\hat{\theta} + \sum_{k=1}^{i-1}\left(\frac{\partial \alpha_{i-1}}{\partial x_k}x_{k+1} + \frac{\partial \alpha_{i-1}}{\partial y_r^{(k-1)}}y_r^{(k)}\right) \\
&\quad + \frac{\partial \alpha_{i-1}}{\partial \hat{\theta}}\Gamma\tau_i + \sum_{k=2}^{i-1}\frac{\partial \alpha_{k-1}}{\partial \hat{\theta}}\Gamma w_i z_k, \qquad i = 1,\ldots,m \qquad (7.5.18)
\end{aligned}
$$

$$
\begin{aligned}
\bar{\alpha}_{m+1} &= -\hat{b}_m z_m - c_{m+1}z_{m+1} - w_{m+1}^T\hat{\theta} + \sum_{k=1}^{m-1}\frac{\partial \alpha_m}{\partial x_k}x_{k+1} + \hat{b}_m\frac{\partial \alpha_m}{\partial x_m}x_{m+1} \\
&\quad + \sum_{k=1}^{m}\frac{\partial \alpha_m}{\partial y_r^{(k-1)}}y_r^{(k)} + \frac{\partial \alpha_m}{\partial \hat{\theta}}\Gamma\tau_{m+1} + \left(y_r^{(m)} + \frac{\partial \alpha_m}{\partial \hat{\varrho}}\right)\dot{\hat{\varrho}} \\
&\quad + \sum_{k=2}^{m}\frac{\partial \alpha_{k-1}}{\partial \hat{\theta}}\Gamma w_{m+1}z_k \qquad (7.5.19)
\end{aligned}
$$

$$
\begin{aligned}
\bar{\alpha}_j &= -z_{j-1} - c_j z_j - w_j^T\hat{\theta} + \sum_{k=1}^{j-1}\frac{\partial \alpha_{j-1}}{\partial x_k}x_{k+1} + \hat{b}_m\frac{\partial \alpha_{j-1}}{\partial x_m}x_{m+1} \\
&\quad + \sum_{k=1\neq m}^{j-1}\frac{\partial \alpha_{j-1}}{\partial y_r^{(k-1)}}y_r^{(k)} + \frac{\partial \alpha_{j-1}}{\partial \hat{\theta}}\Gamma\tau_j + \frac{\partial \alpha_{j-1}}{\partial \hat{b}_m}\gamma\pi_j + \left(y_r^{(j-1)} + \frac{\partial \alpha_{j-1}}{\partial \hat{\varrho}}\right)\dot{\hat{\varrho}} \\
&\quad + \sum_{k=2}^{j-1}\frac{\partial \alpha_{k-1}}{\partial \hat{b}_m}\gamma\frac{\partial \alpha_{j-1}}{\partial x_m}x_{m+1}z_k, \qquad j = m+2,\ldots,n \qquad (7.5.20)
\end{aligned}
$$

Adaptive control law:

$$u = \frac{1}{\beta(x)}[\alpha_n + \hat{\varrho}y_r^{(n)}] \qquad (7.5.21)$$

Parameter update laws:

$$
\begin{aligned}
\dot{\hat{\theta}} &= \Gamma\tau_n = \Gamma W_z && (7.5.22)\\
\dot{\hat{b}}_m &= \gamma\pi_n = \gamma\left[z_{m+1}z_m - \sum_{j=m+1}^{n}\frac{\partial \alpha_{j-1}}{\partial x_m}x_{m+1}z_j\right] && (7.5.23)\\
\dot{\hat{\varrho}} &= -\gamma sgn(b_m)(y_r^{(m)} + \bar{\alpha}_m)z_m && (7.5.24)
\end{aligned}
$$

Lengthy but straightforward calculations show that the design procedure (7.5.9) - (7.5.24) results in the closed-loop system

$$\begin{aligned}
\dot{z}_i &= -c_i z_i - \sum_{k=2}^{i-1} \sigma_{ki} z_k - z_{i-1} + \sum_{k=i+1}^{n} \sigma_{ik} z_k + w_i^T \tilde{\theta} \\
&\quad , i = 1, \ldots, m-1 && (7.5.25) \\
\dot{z}_m &= -c_m z_m - \sum_{k=2}^{m-1} \sigma_{km} z_k - z_{m-1} + \hat{b}_m z_{m+1} + \sum_{k=m+1}^{n} \sigma_{mk} z_k \\
&\quad + w_m^T \tilde{\theta} - b_m (y_r^{(m)} - \bar{\alpha}_m) \tilde{\varrho} + z_{m+1} \tilde{b}_m && (7.5.26) \\
\dot{z}_{m+1} &= -c_{m+1} z_{m+1} - \sum_{k=2}^{m} \sigma_{k,m+1} z_k - b_m z_m + z_{m+2} + \sum_{k=m+2}^{n} \sigma_{m+1,k} z_k \\
&\quad + w_{m+1}^T \tilde{\theta} - \frac{\partial \alpha_m}{\partial x_m} x_{m+1} \tilde{b}_m && (7.5.27) \\
\dot{z}_j &= -c_j z_j - \sum_{k=2}^{j-1} \sigma_{kj} z_k - z_{j-1} + z_{j+1} + \sum_{k=j+1}^{n} \sigma_{jk} z_k \\
&\quad + w_j^T \tilde{\theta} - \frac{\partial \alpha_{j-1}}{\partial x_m} x_{m+1} \tilde{b}_m, \quad i = m+2, \ldots, n && (7.5.28)
\end{aligned}$$

where σ_{ik} is defined for $k = i+1, \ldots, n$ as

$$\sigma_{ik} = \begin{cases} 0, & i = 1 \\ -\frac{\partial \alpha_{i-1}}{\partial \hat{\theta}} \Gamma w_k, & i = 2, \ldots, m+1 \\ -\frac{\partial \alpha_{i-1}}{\partial \hat{\theta}} \Gamma w_k + \frac{\partial \alpha_{i-1}}{\partial \hat{b}_m} \gamma \frac{\partial \alpha_{k-1}}{\partial x_m} x_{m+1}, & i = m+2, \ldots, n-1 \end{cases} \tag{7.5.29}$$

A Lyapunov function for this system is

$$V = \frac{1}{2} z^T z + \frac{1}{2} \tilde{\theta}^T \Gamma^{-1} \tilde{\theta} + \frac{1}{2\gamma} \tilde{b}_m^2 + \frac{|b_m|}{2\gamma} \tilde{\varrho}^2 \tag{7.5.30}$$

Its derivative along the solutions of (7.5.22)-(7.5.24) and (7.5.25)-(??),

$$\dot{V} = -\sum_{k=1}^{n} c_k z_k^2 \tag{7.5.31}$$

lead us to the same conclusion as in tracking design procedures which all the states are bounded and asymptotic tracking is achieved.

Chapter 8

Example in Aviation

8.1 Aircraft Wing Rock

Wing reck is a limit cycling oscillation in the roll angle ϕ and the roll rate $\dot{\phi}$ which can occur in high-performance aircraft with slender forebodies when flying in high angle-of-attack. Conventional methods of eliminating wing rock include a redesign of the airframe configuration and limiting of the angle-of-attack. These methods may reduce maneuverability of the aircraft. An effective method of suppressing wing rock without degrading maneuverability is using feedback control.

Several one degree-of-freedom models of wing rock have been proposed in Nguyen, Whipple, and Brandon, Hsu and Lan, and Elzebda, Nayfeh, and Mook. They are all nonlinear and contain parameters θ_i which depend on the angle-of-attack, dynamic pressure, wing reference area, wing span, roll moment of inertia, and flight velocity. We now present an adaptive controller of Monahemi, Barlow, and Krstić which allows these parameters to be unknown and eliminates te wing rock phenomenon by achieving global stabilization. The model we consider here,

$$\ddot{\phi} = \theta_1 + \theta_2\phi + \theta_3\dot{\phi} + \theta_4|\phi|\dot{\phi} + \theta_5\left|\dot{\phi}\right|\dot{\phi} \tag{8.1.1}$$

is based on wind tunnel tests at NASA Langley Research Center. In these tests, physical scaled wings were mounted on an apparatus which allows free rotation about the roll axis. These model wings are aerodynamically similar to the wings of an F-18 HARV aircraft. These are no control surface in the model (8.1.1). With ailerons modeled as first-order actuator dynamics, the state-space form of the wing-rock model is

$$\begin{aligned}\dot{\phi} &= p \\ \dot{p} &= \theta_1 + \theta_2\phi + \theta_4 |\phi| p + \theta_5 |p| p + b\delta_A \\ \tau\dot{\delta}_A &= -\delta_A + u\end{aligned} \tag{8.1.2}$$

where δ_A is the aileron deflection angle, u is the control input, τ is the aileron time constant, and b is an unknown constant parameter. Denoting $\varphi(\phi, p) = [1, \phi, p, |\phi| p, |p| p]^T$ and $\theta = [\theta_1, \theta_2, \theta_3, \theta_4, \theta_5]^T$, we rewrite (8.1.2) as

$$\begin{aligned}\dot{\phi} &= p \\ \dot{p} &= b\delta_A + \varphi(\phi, p)^T\theta \\ \dot{\delta}_A &= \frac{1}{\tau}u - \frac{1}{\tau}\delta_A\end{aligned} \tag{8.1.3}$$

This model is in the parametric strict-feedback form with unknown virtual control coefficient b, so we apply the design from the previous chapter.

Our control objective is to asymptotically track a given reference $\phi(t)$ with the roll angle ϕ. We use the error variables

$$\begin{aligned}z_1 &= \phi - \phi_r \\ z_2 &= p - \dot{\phi}_r - \alpha_1(\phi, \phi_r) \\ z_3 &= \delta_A - \hat{\varrho}\ddot{\phi}_r - \alpha_2(\phi, p, \phi_r, \dot{\phi}_r, \hat{\theta}, \hat{\varrho}),\end{aligned} \tag{8.1.4}$$

and derive the stabilizing functions

$$\begin{aligned}\alpha_1 &= -c_1 z_1 \\ \alpha_2 &= \hat{\varrho}\bar{\alpha}_2, \qquad \bar{\alpha}_2 = -z_1 - c_2 z_2 - \varphi^T\hat{\theta} - c_1(p - \dot{\phi}_r).\end{aligned} \tag{8.1.5}$$

The design procedure from Section 4.5.1 results in an adaptive control consisting of the control law

$$u = \tau[\frac{1}{\tau}\delta_A - \tilde{b}z_2 - c_3 z_3 + \frac{\partial\alpha_2}{\partial\phi}p + \frac{\partial\alpha_2}{\partial p}(\tilde{b}\delta_A + \varphi^T\hat{\theta}) \\ + \frac{\partial\alpha_2}{\partial\phi_r}\dot{\phi}_r + \frac{\partial\alpha_2}{\partial\dot{\phi}_r}\ddot{\phi}_r + \hat{\varrho}\phi_r^{(3)} + \frac{\partial\alpha_2}{\partial\hat{\theta}}\dot{\hat{\theta}} + (\ddot{\phi}_r + \bar{\alpha}_2)\dot{\hat{\varrho}}] \tag{8.1.6}$$

and the update laws

$$\dot{\hat{\theta}} = \Gamma\varphi(z_2 - \frac{\partial\alpha_2}{\partial p}z_3) \tag{8.1.7}$$

$$\dot{\hat{b}} = \gamma(z_2 - \frac{\partial\alpha_2}{\partial p}\delta_A)z_3 \tag{8.1.8}$$

$$\dot{\hat{\varrho}} = -\gamma sgn(b)(\ddot{\phi}_r + \bar{\alpha}_2)z_2. \tag{8.1.9}$$

For the resulting error system

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \\ \dot{z}_3 \end{bmatrix} = \begin{bmatrix} -c_1 & 1 & 0 \\ -1 & -c_2 & b \\ 0 & -b & -c_3 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \\ -\frac{\partial\alpha_2}{\partial p} \end{bmatrix} \varphi^T\tilde{\theta} \\ + \begin{bmatrix} 0 \\ 0 \\ z_2 - \frac{\partial\alpha_2}{\partial p}\delta_A \end{bmatrix} \tilde{b} + \begin{bmatrix} 0 \\ -b(\ddot{\phi}_r + \bar{\alpha}_2) \\ 0 \end{bmatrix} \tilde{\varrho}$$

the equilibrium $z = 0$, $\tilde{\theta} = 0$, $\tilde{b} = 0$, $\tilde{varrho} = 0$ is globally stable and, moreover, $z(t) \to 0$. This means that the wing rock phenomenon is eliminated and the tracking objective $\phi(t) - \phi_r(t) \to 0$ is achieved.

This is illustrated in Figure 8.1. Without feedback control the response to an initial condition $\phi(0) = 0.4$, $p(0) = \delta_A(0) = 0$ is the trajectory (a) which represents limit cycling oscillations typical for wing rock. They are obtained for the model (8.1.1) with wind tunnel at angle-of-attack $\alpha = 30^o$: $\theta_1 = 0$, $\theta_2 = -26.67$, $\theta_3 = 0.76485$, $\theta_4 = -2.9225$, $\theta_5 = 0$. The amplitude of the wing rock oscillations is about 35^o and the frequency is about 1Hz.

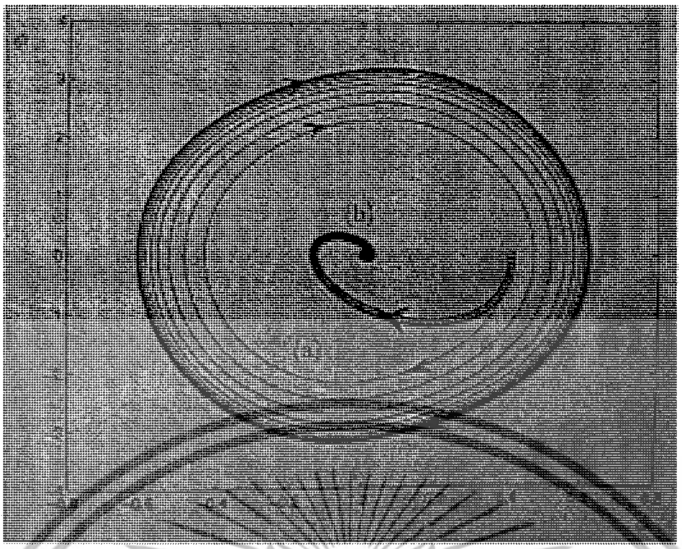

Figure 8.1: (a) Uncontrolled wing rock. (b) Suppression of wing rock by adaptive nonlinear control

The trajectory (b) in Figure 4.5 shows the wing rock suppressing effect of the adaptive controller (8.1.6)-(8.1.9) acting through the aileron with $b = 1.5$ and $\tau = 1/15$. Parameter estimates are initialized as $\hat{\theta}(0) = 1.35\theta$, $\hat{b}(0) = 1.35b$, and $\hat{\varrho}(0) = 1/\hat{b}(0)$. The controller coefficients are $c_1 = c_2 = c_3 = 5$ and the adaption gains are $\Gamma = 0.02I$, $\gamma = 0.02$. For softer regulation to the origin, we have employed an exponentially decaying reference trajectory $\phi_r(t)$ governed by the equation $(s+10)(s^2+4s+24.25)\phi_r(s) = 0$.

Chapter 9

Conclusion

9.1 Objective

Our aim is to design a nonlinear static or dynamic (adaptive) controller for systems or plants which contain nonlinearities and unknown parameters.

9.2 Problem Occurred

An interested in nonlinear control has been investigated since 1980. Twenty years later, a number of papers have been published as you have seen in the bibliography (that are only a few). Note that we have studied only a few part of nonlinear control theory which less than 10 percents of all methods. Almost textbooks contain useful theorems, lemmas, definitions, and corollaries which describe an idea of nonlinear systems. However, it is too hard to choose which we need to fulfil our project. In contrast, applications are rare and often link to another field of study, such as a mechanic of vibrations, aerodynamics, electromagnetics, etc. An aircraft wing rock is the application we chose to indicate the useful of nonlinear control system in real operations. Actually, we require a knowledge of aerodynamics to analyze overall characteristics of wing rock motion. It seems to be too difficult so we skipped those step and use the aircraft wing rock model researched by NASA as a starting point of Chapter 8.

9.3 Further Studies and Improvements

9.3.1 Nonlinear control theory

Continue from stability theories of nonlinear system, the next task should be the *describing function* method. It provides an approximation for the response of linear method. The main idea of the describing function method is it can predict a limit cycle of nonlinear systems by analyzing their frequency response. The describing function method is useful in describing a nonlinear vibration which leads to the design of a nonlinear oscillator.

9.3.2 Nonlinear controller design

Adaptive backstepping and tuning functions are Lyapunov-based designs. These recursive procedures have removed the critical relative degree restriction and provided the simplest proofs of the strongest stability properties. Along with their advantages, they have certain drawbacks. One of them is that they do not offer freedom choices of parameter update laws. For systems with many unknown parameters, a further drawback of adaptive backstepping is that the dynamic order is high (overparametrization). On the other hand, the order of the tuning functions controller is minimal, but for high-order systems, its nonlinear expression become increasingly complex. The main source of complexity is the built-in interaction between the identifier and the control law.

Both of these drawbacks are removed by *modular designs* which are more inspired by traditional *estimation-based* designs. In adaptive linear control, the estimation-based designs achieve a significant level of *modularity* of the controller-identifier pair: Any stabilizing controller can be combined with any identifier. Parameter update laws can be of either gradient or least-squares type. The controller module is capable of stabilizing the plant when all the parameters are known. This is its certainty equivalence property. The identifier module, in turn, guarantees certain boundedness properties independently of the controller module. The modularity of the estimation-based designs makes them much more flexible than the Lyapunov-based designs.

9.4 Applications

The nonlinear control theory and controller design procedures could be applied with these examples

1. Vibration Control
   - Vibration in tall building structure
   - Vibration in elevator
   - Vibration in robot arm
   - Active suspension in vehicles
   - Flexible link manipulator
2. Fuzzy control
   - Temperature and substrate feed control in biochemical processes
   - Liquid concentration control in food processes

3. Positioning

    - Positioning robot control
    - A control of levitated ball in magnetic field
    - Precise conveyor

4. Electronics

    - Decreasing of nonlinear effects in electronic components
    - Reference signal regulation and tracking
    - Nonlinear oscillator

Appendix A

Linear Algebra

We expect you to know the following theory and know how to apply it after you finish this class.

(Elementary Row Operation)

Before we define the elementary matrices, we define the elementary row operations. Let $A \in \mathbb{F}^{m \times n}$ be any matrix (not necessarily square). There are three kinds of **elementary row operations** that may be applied to the matrix A:

1. For $i = 1, 2, \ldots, m$ and $c \neq 0$, the matrix $\text{Scale}(A, i, c)$ is the matrix that results from A by multiplying the ith row by c.

2. For $i, j = 1, 2, \ldots, m$ the matrix $\text{Swap}(A, i, j)$ is the matrix results from A by exchanging the ith and jth rows.

3. For $i, j = 1, 2, \ldots, m$ with $i \neq j$ and any scalar c, the matrix $\text{Shear}(A, i, j, c)$ is the matrix that results from A by adding c times the jth row to the ith row.

Theorem A.0.1 *Each elementary row operation is reversible in the sense that it can be undone by another operation of the same kind. Specifically:*

1. *If* $B = \text{Scale}(A, i, c)$, *then* $A = \text{Scale}(B, i, c^{-1})$.

2. *If* $B = \text{Swap}(A, i, j)$, *then* $A = \text{Swap}(B, i, j)$.

3. *If* $B = \text{Shear}(A, i, j, c)$, *then* $A = \text{Shear}(B, i, j, -c)$.

(Elementary Matrices)

The elementary matrices are the simplest of all invertible matrices, except for the identity matrix. We shall see that they are the building blocks from which the invertible matrices are constructed. Here is the definition.

Definition A.0.2 *A matrix that results from the identity matrix by applying a single elementary row operation is called an* **elementary matrix**.

An elementary matrix is always a square matrix. There are three kinds.

1. **Scale**. *The matrix* $E = \text{Scale}(I, i, c)$ *is an elementary matrix for* $i = 1, 2, \ldots, m$ *and* $c \neq 0$. *It differs from the* $m \times m$ *identity matrix* $I = I_m$ *in that* $(E)_{ii} = c$ *rather than* 1.

2. **Swap**. *The matrix* $E = \text{Swap}(I, i, j)$ *is an elementary matrix for* $i, j = 1, 2, \ldots, m, i \neq j$. *It differs from the identity matrix in that*

$$(E)_{ii} = 0 \quad (E)_{ij} = 1$$
$$(E)_{ji} = 1 \quad (E)_{jj} = 0.$$

3. **Shear**. *The matrix* $E = \text{Shear}(I, i, j, c)$ *is an elementary matrix for* $i, j = 1, 2, \ldots, m, i \neq j$. *It differs from the identity matrix in that*

$$(E)_{ij} = c.$$

Theorem A.0.3 (Fundamental Theorem on Row Operations) *The matrix* *EA that results by multiplying a matrix A on the left by an elementary matrix E* *is the same as the matrix that results by applying the corresponding elementary row operation to A.*

Let the elementary matrix $E \in \mathbb{F}^{m \times m}$ result by applying some elementary row operation to the identity matrix $I = I_m \in \mathbb{F}^{m \times m}$. The Fundamental Theorem says that for any matrix $A \in \mathbb{F}^{m \times m}$, the matrix EA results by applying that same row operation to A. More precisely,

1. $E = \mathrm{Scale}(I_m, i, c) \Rightarrow EA = \mathrm{Scale}(A, i, c)$.
2. $E = \mathrm{Swap}(I_m, i, j) \Rightarrow EA = \mathrm{Swap}(A, i, j)$.
3. $E = \mathrm{Shear}(I_m, i, j, c) \Rightarrow EA = \mathrm{Shear}(A, i, j, c)$.

Theorem A.0.4 *Elementary matrices are invertible. In fact,*

1. $\mathrm{Scale}(I, i, c)^{-1} = \mathrm{Scale}(I, i, c^{-1})$.

2. $\mathrm{Swap}(I, i, j)^{-1} = \mathrm{Swap}(I, i, j)$.

3. $\mathrm{Shear}(I, i, j, c)^{-1} = \mathrm{Shear}(B, i, j, -c)$.

Corollary A.0.5 *A product of any number of elementary matrices is invertible.*

Proof

$$M = E_1 E_2 \cdots E_k \Rightarrow M^{-1} = E_k^{-1} \cdots E_2^{-1} E_1^{-1}.$$

(Reduced Row Echelon Form: RREF)

Definition A.0.6 *An $m \times n$ matrix R is in* **reduced row echelon form**, *abbreviated* **RREF**, *if and only if*

1. *all the rows that vanish identically (if any) appear below the other (nonzero) rows;*

2. *the leading entry in any row appears to the left of the leading entry of any nonzero row below;*

3. *the leading entry in any nonzero row is* 1*;*

4. *all other entries in the column of a leading entry are* 0*.*

The **leading entry** *of a row is the first nonzero entry in that row. The columns which hold the leading entries are called the* **leading columns***; the other columns are called the* **free columns***.*

Theorem A.0.7 (Gauss-Jordan Elimination) *A matrix may be transformed to a matrix in RREF by applying a sequence of elementary row operations.*

(Computing the Multiplier)

It is convenient to express the Gauss-Jordan Elimination theorem as a theorem about matrix multiplication. We can do this using the Fundamental Theorem on Row Operations.

Theorem A.0.8 (Multiplier Theorem) *For any $A \in \mathbb{F}^{m\times n}$ there is an invertible matrix M such that the matrix $R = MA$ is in RREF.*

Theorem A.0.9 *It can be proved that the reduced row echelon form R is unique: If $M_1A = R_1$ and $M_2A = R_2$ where M_1 and M_2 are invertible and R_1 and R_2 are in RREF, then $R_1 = R_2$.*

The following handy trick, called the *Multiplier Trick*, enables us to compute the multiplier matrix M as we perform elementary row operations. Form the $m \times (n+m)$ matrix $[A \;\; I_m]$. Using the block multiplication law, we have

$$M[A \;\; I_m] = [MA \;\; MI] = [MA \;\; M]$$

where $I = I_m$ is the $m \times m$ identity matrix.

Theorem A.0.10 (Multiplier Trick) *If we apply the same row operations to the matrix $[A \;\; I_m]$ that we applied to A to transform it to MA, the last m columns of the result contain the matrix M.*

Definition A.0.11 *If A is a square matrix, and if B of the same size can be found such that $AB = I$ and $BA = I$, then A is said to be* invertible *and B is called an* inverse *of A. If no such matrix B can be found, then A is said to be* singular *or* non-invertible. *The inverse of A is denoted by A^{-1}.*

Theorem A.0.12 (Laws of Inverse) *If A and B are invertible matrices, then:*

1. $(A^{-1})^{-1} = A$.

2. $(A^n)^{-1} = (A^{-1})^n$.

3. $(kA)^{-1} = \frac{1}{k}A^{-1}$, *for any nonzero scalar k.*

*4.** $(AB)^{-1} = B^{-1}A^{-1}$.

5. $(A^T)^{-1} = (A^{-1})^T$.

(How to Invert)

The Multiplier Trick provides an efficient method to

- decide if a square matrix A is invertible, and
- compute the inverse A^{-1} of an invertible matrix A.

We put A in reduced row echelon form R, simultaneously computing the invertible matrix M with $MA = R$. Of course, the matrix M is invertible. Now we apply the following.

Theorem A.0.13 (How to Invert) *Let $A, M, R \in \mathbb{F}^{n\times n}$ be square matrices with M invertible, R in RREF, and*

$$MA = R.$$

Then

- *A is invertible if and only if R is the identity matrix.*
- *If A is invertible, then $A^{-1} = M$.*

Method for Finding the Inverse

Using row operations to find A^{-1}: Construct the matrix

$$[A \mid I]$$

and apply row operations until you can the identity matrix on the left side. At that point the matrix on the right side will be the inverse of A. So the final matrix will be

$$[I \mid A^{-1}] \ .$$

Corollary A.0.14 (Factorization Theorem) *A matrix is invertible if and only if it is a product of elementary matrices.*

Corollary A.0.15 (A Criterion for Invertibility) *Suppose that A is a square matrix. Then either*

- *A is invertible, or else*
- *there is a nonzero $\vec{x}$ with $A\vec{x} = \vec{0}$.*

Proof

If R is not the identity matrix, it is easy to find a nonzero $\vec{x}$ with $R\vec{x} = \vec{0}$. It follows that $A\vec{x} = M^{-1}R\vec{x} = \vec{0}$. We will be more precise on finding such an $\vec{x}$ when we talk about the Nullspace, but for the moment consider the example

$$R = \begin{bmatrix} 1 & 0 & c_{13} \\ 0 & 1 & c_{23} \\ 0 & 0 & 0 \end{bmatrix}, \vec{x} = \begin{bmatrix} -c_{13} \\ -c_{23} \\ 1 \end{bmatrix}.$$

Then $R\vec{x} = \vec{0}$ but $\vec{x} \neq \vec{0}$ since $x_3 = 1$.

Remark A.0.16 *It is impossible that both alternative occur. If A is invertible and $A\vec{x} = \vec{0}$, then*

$$\vec{x} = A^{-1}A\vec{x} = A^{-1}\vec{0} = \vec{0}.$$

bf Systems of Equations and Invertibility

Theorem A.0.17 (Equivalent Statements) *If A is an $n \times n$ matrix, then the following statements are equivalent:*

1. *A is invertible.*
2. *$A\vec{x} = \vec{0}$ has only the trivial solution.*
3. *The reduced row-echelon form of A is I_n.*

4. $A\vec{x} = \vec{b}$ *is consistent for every column vector* $\vec{b}$.

5. $A\vec{x} = \vec{b}$ *has exactly one solution for every column vector* $\vec{b}$, *and that is* $\vec{x} = A^{-1}\vec{b}$.

Definition A.0.18 (Trace) *The* **trace** *of a square matrix* $A \in \mathbb{F}^{n\times n}$ *is the sum of its diagonal entries:*

$$\operatorname{tr}(A) = \sum_{i=1}^{n}(A)_{ii} = \sum_{i=1}^{n}\lambda_i$$

for all $\lambda_i \in \sigma(A) = \{$*all eigenvalues of* $A\}$.

Theorem A.0.19 (Trace) *Let* $A, B, C \in \mathbb{F}^{n\times n}$ *and* $a, b \in \mathbb{F}$.

1. *If* $A = B$ *then*

$$\operatorname{tr}(A) = \operatorname{tr}(B).$$

2. *The trace is linear:*

$$\operatorname{tr}(aA + bB) = a\operatorname{tr}(A) + b\operatorname{tr}(B).$$

3. *The trace of the product is invariant when the matrices in the product are commute:*

$$\begin{aligned}\operatorname{tr}(AB) &= \operatorname{tr}(BA)\\ \operatorname{tr}(ABC) &= \operatorname{tr}(CAB)\\ \operatorname{tr}(ABC) &= \operatorname{tr}(BCA)\\ \operatorname{tr}(ABA^{-1}) &= \operatorname{tr}(B).\end{aligned}$$

Definition A.0.20 (Determinant) *Let* $A \in \mathbb{F}^{n\times n}$ *be a square matrix. Then*

$$\det(A) = \sum_{j=1}^{n} a_{ij}(\mathcal{C}A)_{ij} = \sum_{j=1}^{n}(-1)^{i+j}a_{ij}D\left(A\left(i|j\right)\right), \qquad 1 \le i \le n,$$

or

$$\det(A) = \sum_{i=1}^{n} a_{ij}(\mathcal{C}A)_{ij} = \sum_{i=1}^{n}(-1)^{i+j}a_{ij}D\left(A\left(i|j\right)\right), \qquad 1 \le j \le n.$$

Theorem A.0.21 (Trace and Determinant) *The determinant of a matrix is the product of its eigenvalues. The trace of a matrix is the sum of its eigenvalues. That is*

$$\operatorname{tr}(A) = \sum_{i=1}^{n} \lambda_i,$$
$$\det(A) = \prod_{i=1}^{n} \lambda_i,$$

for all $\lambda_i \in \sigma(A)$.

Properties of the determinant function

Theorem A.0.22 *Let* A *and* B *be* $n \times n$ *matrices.*

1. *If* A *has a row or a column of zeroes, then* $\det(A) = 0$.
2. $\det(A) = \det(A^T)$.
3. $\det(AB) = \det(A)\det(B)$.
4. *A square matrix* A *is invertible if and only if* $\det(A) \neq 0$.
5. *If* A *is invertible, then* $\det(A^{-1}) = \frac{1}{\det(A)}$.
6. *If* k *is a scalar, then* $\det(kA) = k^n \det(A)$.

Determinant, invertibility and systems of linear equations

Theorem A.0.23 *If* A *is an* $n \times n$ *matrix, then the following are equivalent.*

1. *A is invertible.*
2. $\det(A) \neq 0$.
3. $A\vec{x} = \vec{b}$ *has exactly one solution for every* $n \times 1$ *matrix* $\vec{b}$.

4. *$A\vec{x} = \vec{0}$ has only the trivial solution.*

5. *The reduced row-echelon form of A is I_n.*

Definition A.0.24 (Vector Space) *A vector space V is a non-empty set of objects on which two operations are defined: addition and multiplication by scalars, and these two operations satisfy the 10 axioms. (What are those 10 axioms?)*

Definition A.0.25 (Subspace) *A subset W of a vector space V is called a subspace of V if W is itself a vector space under addition and scalar multiplication.*

Theorem A.0.26 (Subspace) *W is a subspace of V if and only if*

1. *W is a subset of V.*

2. *$\forall \vec{u}, \vec{v} \in W \quad \vec{u} + \vec{v} \in W$.*

3. *$\forall \vec{u} \in W \, \forall k \in \mathbb{F} \quad k\vec{u} \in W$.*

Definition A.0.27 (Linear Dependence:L.D./Linear Independence:L.I.) *A set of vectors $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r$ is called* linearly dependent (L.D.) *if*

$$k_1\vec{v}_1 + k_2\vec{v}_2 + \cdots + k_r\vec{v}_r = \vec{0}$$

for some scalars $k_1, k_2, \ldots, k_r$ not all zero. Otherwise $(k_1 = k_2 = \cdots = k_r = 0)$ *$\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r$ are* linearly independent (L.I.).

Definition A.0.28 (Span) $W = \operatorname{span}\{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r\}$ *if and only if* $\exists (k_1, k_2, \ldots, k_r)$ *such that*

$$\forall \vec{w} \in W, \quad \vec{w} = k_1\vec{v}_1 + k_2\vec{v}_2 + \cdots + k_r\vec{v}_r.$$

Definition A.0.29 (Basis and Dimension) *A set $\mathcal{S} = \{\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n\}$ of vectors in the vector space V is a* basis *of V if*

1. *$\mathcal{S}$ is linearly independent, and*

2. *$\mathcal{S}$ spans V.*

The number of vectors in a basis is the dimension of the vector space V. Notation: $\dim(V) = n$.

Definition A.0.30 (Nullspace or Kernel of $A : \mathrm{N}(A)$) *The **nullspace** of a matrix $A \in \mathbb{F}^{m\times n}$ is the set*

$$\mathrm{N}(A) = \left\{\vec{x} \in \mathbb{F}^{n\times 1} : A\vec{x} = \vec{0}\right\}$$

of all solutions $\vec{x}$ of the homogeneous system $A\vec{x} = \vec{0}$.

Definition A.0.31 (Rangespace or Column Space or Image of $A : \mathrm{R}(A)$) *The **range** of a matrix $A \in \mathbb{F}^{m\times n}$ is the set of all inhomogeneous terms $A\vec{y} \in \mathbb{F}^{m\times 1}$ for which the inhomogeneous system $\vec{y} = A\vec{x}$ has at least one solution $\vec{x}$. In set-theoretic notation*

$$\mathrm{R}(A) = \left\{\vec{y} = A\vec{x} : \vec{x} \in \mathbb{F}^{n\times 1}\right\}.$$

(Eigenvalues, Eigenvectors and Eigenspace)

Let $A \in \mathbb{F}^{n\times n}$ be a square matrix. If a number $\lambda \in \mathbb{F}$ and ***nonzero*** column vector $\vec{v} \in \mathbb{F}^{n\times 1}$ satisfy the equation

$$A\vec{v} = \lambda\vec{v},$$

we say that λ is an *eigenvalue* of A and that $\vec{v}$ is an *eigenvector* of A for the eigenvalue λ. (What is the geometric interpretation?) This equation can be rewritten in the form

$$(\lambda I - A)\vec{v} = \vec{0}$$

where $I = I_n$ is the $n\times n$ identity matrix. This equation is called the *eigenequation.* A square matrix is not invertible if and only if its nullspace is not the zero subspace. Hence, we may reformulate the definition as follows:

Definition A.0.32 *A number $\lambda \in \mathbb{F}$ is called an* **eigenvalue** *of the square matrix A if and only if the matrix $\lambda I - A$ is not invertible. The subspace*

$$\mathcal{E}_\lambda(A) = \mathrm{N}(\lambda I - A)$$

is called the **eigenspace** *of λ. A nonzero element of $\mathcal{E}_\lambda(A)$ is called an* **eigenvector** *of A corresponding to the eigenvalue λ.*

Theorem A.0.33 (TFSAE) *Let $A : \mathbb{F}^n \to \mathbb{F}^n$; $A : x \rightarrowtail y$ such that $y_{n\times 1} = A_{n\times n}x_{n\times 1}$. Then the following statements are equivalent.*

1. $\det A \neq 0$	*1.* $\det A = 0$
2.A is nonsingular or invertible.	*2.A is singular or non-invertible.*
3. $\exists! A^{-1}$ *such that* $AA^{-1} = A^{-1}A = I$	*3.* $\nexists A^{-1}$ *such that* $AA^{-1} = A^{-1}A = I$
4.A is injective $(1-1)$.	*4.A is not* injective (*not* $1-1$) .
5.A is surjective (*onto*).	*5.A is not* surjective (*not onto*).
6. $\ker(A) = \mathrm{N}(A) = \left\{\vec{0}\right\}$	*6.* $\ker(A) = \mathrm{N}(A) \supsetneqq \left\{\vec{0}\right\}$
7. $\mathrm{R}(A) = \mathbb{F}^n$	*7.* $\mathrm{R}(A) \subsetneqq \mathbb{F}^n$
8. $\begin{aligned} n - r = \textit{nullity of } A &= \dim(\ker(A)) \\ &= \dim(\mathrm{N}(A)) = 0 \end{aligned}$	*8.* $\begin{aligned} n - r = \textit{nullity of } A &= \dim(\ker(A)) \\ &= \dim(\mathrm{N}(A)) \gneqq 0 \end{aligned}$
9. $r = \mathrm{rank}(A) = \dim(\mathrm{R}(A)) = n$	*9.* $r = \mathrm{rank}(A) = \dim(\mathrm{R}(A)) \lneqq n$
10. $\mathrm{rref}(A) = I_n$	*10. We can have either* $\mathrm{rref}(A) = \left(\begin{array}{c\|c} I_r & C_{r\times(n-r)} \\ 0_{(n-r)\times r} & 0_{n-r} \end{array}\right)$, *where* $r < n$ *or* $\mathrm{rref}(A)$ *will have r nonzero rows (rows that have leading entries).*
11.row (column) vectors of A are L.I..	*11.row (column) vectors of A are L.D..*
12. $\forall \lambda_i \in \sigma(A), \quad \lambda_i \neq 0$	*12.* $\exists \lambda_i \in \sigma(A), \quad \lambda_i = 0$

$\square$

Bibliography

[1] Robbin, Joel W., *Matrix Algebra*, A K Peters, Ltd., MA, 1985.

[2] M. Krstic, I. Kanellakopoulos, P. Kokotovic, *Nonlinear and Adaptive Control Design*, New York : John Wiley & Sons, Inc. 1995.

[3] H. K. Khalil, *Nonlinear systems*, New Jersey : Prentice-Hall, Inc. 2002.

[4] P. Pranayanuntana, V. Riewruja, *"Nonlinear Backstepping Control Design Applied to Magnetic Ball Control"*, IEEE Tencon Proceedings, vol.III, pp. III-304 - III-307, September 2000.

[5] F. F. Kuo, *Network Analysis and Synthesis*, New York : John Wiley & Sons, Inc. 1966.

[6] C. A. Desocr, M. Vidyasagar, *Feedback Systems : Input-Output Properties*, New York : Academic Press, Inc. 1975.

[7] J. T. Wen, *"Time Domain and Frequency Domain Conditions for Strict Positive Realness"*, IEEE Trans. Automat. Contr., 33(10): pp. 988-992, October 1988.

[8] J. E. Slotine, W. Li, *Applied Nonlinear Control*, New Jersey : Prentice Hall, Inc. 1991.

[9] Z. Artstein, "Stabilization with relaxed controls," *Nonlinear Analysis*, vol. TMA-7, pp. 1163-1173, 1983.

[10] K. J. Astrom and B. Wittenmark, *Adaptive Control*, second edition, New York : Addison-Wesley, 1995.

[11] G. Bastin and G. Campion, "Indirect adaptive control of linearly parametrized nonlinear systems," *Proceeding of the 3rd IFAC Symposium on Adaptive Systems in Control, and Signal Processing*, Glasgow, UK, 1989.

[12] C. I. Byrnes and A. Isidori, "New results and examples in nonlinear feedback stabilization," *Systems in Control Letters*, vol. 12, pp. 437-442, 1989.

[13] G. Campion and G. Bastin, "Indirect adaptive state-feedback control of linearly parametrized nonlinear systems," *International Journal of Adaptive Control and Signal Processing*, vol. 35, pp. 1100-1108, 1990.

[14] J. J. Craig, *Adaptive Control of Mechanical Manipulators*, Reading, MA : Addison-Wesley, 1988.

[15] W. Dayawansa, W. M. Boothby and D. L. Elliott, "Global state and feedback equivalance of nonlinear systems," *Systems and Control Letters*, vol. 6, pp. 229-234.

[16] B. Egardt, *Stability of Adaptive Controllers*, New York : Springer-Verlag, 1979.

[17] A. Feuer, A. S. Morse, "Local stability of parameter adaptive control systems," *Proceeding of the 1978 Conference on Information Sciences and Systems*, Johns Hopkins, Baltimore, MD, pp. 107-111, 1978.

[18] G. C. Goodwin and K. S. Sin, *Adaptive Filtering Prediction and Control*, Eaglewood Cliffs, NJ : Prentice-Hall, 1984.

[19] L. R. Hunt, R. Su, and G. Meyer, "Global transformation of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 28, pp. 24-31, 1983.

[20] A. Isidori, *Nonlinear Control Systems*, Berlins : Springer-Verlag, 1989.

[21] D. H. Jacobson, *Extensions of Linear-Quadratic Control, Optimization and Matrix Theory*, New York : Academic Press, 1977.

[22] B. Jakubczyk, W. Respondek, "On linearization of systems," *Bulletin of the Polish Academy of Science, Series on Mathematical Science*, vol. 28, no.9-10, pp. 517-522, 1980.

[23] Z. P. Jiang and L. Praly, "Iterative designs of adaptive controllers for systems with nonlinear integrators," *Proceeding of the 30th IEEE Conference on Decision and Control*, Brington, UK, December 1991, pp. 2482- 2487.

[24] V. Jurdjivic and J. P. Quinn, "Controllability and Stability," *Journal of Differential Equaions*, vol. 28, pp. 381-389, 1987.

[25] I. Kanellakopoulos, *Adaptive Control of Nonlinear Systems*, Ph. D. Dissertation, University of Illinois, Urbana, 1991.

[26] I. Kanellakopoulos, P. V. Kokotovic, and R. Marino, "An extended direct scheme for robust adaptive nonlinear control," *Automatica*, vol. 27, pp. 247-255, 1991.

[27] I. Kanellakopoulos, P. V. Kokotovic, and R. H. Middleton, "Observer based adaptive control of nonlinear systems under matching conditions," *Proceeding of the 1990 American Control Conference*, San Diego, CA, pp. 549-552.

[28] I. Kanellakopoulos, P. V. Kokotovic, and R. H. Middleton, "Indirect adaptive output-feedback control of a class of nonlinear systems," *Proceeding of the 29th IEEE conference on Decision and Control*, Honolulu, HI, December 1990, pp. 2714-2719.

[29] I. Kanellakopoulos, P. V. Kokotovic, and A. S. Morse, "Systematic design of adaptive controllers for feedback linearizable systems," *IEEE Transactions on Automatic Control*, vol. 36, pp. 1241-1253, 1991.

[30] I. Kanellakopoulos, P. V. Kokotovic, and A. S. Morse, "Adaptive output feedback control of systems with output nonlinearities," pp. 495-525 , 1991.

[31] I. Kanellakopoulos, P. V. Kokotovic, and A. S. Morse, "Adaptive output feedback control of systems with output nonlinearities," *IEEE Transactions on Automatic Control*, vol. 37, pp. 1266-1282, 1992.

[32] P. V. Kokotovic and H. J. Sussmann, "A positive real condition for global stabilization of nonlinear systems," *Systems and Control Letters*, vol. 13, pp. 125-133, 1989.

[33] P. V. Kokotovic, Ed., *Foundations of Adaptive Control*, Berlin : Springer-Verlag, 1991.

[34] I. Kanellakopoulos, P. V. Kokotovic, and A. S. Morse, "Adaptive feedback linearization of nonlinear systems," pp. 311-346, 1991.

[35] M. Krstic, *Adaptive Nonlinear Control*, Ph. D. Dissertation, University of California, Santa Barbara, 1994.

[36] M. Krstic, I. Kanellakopoulos, and P. V. Kokotovic, "Adaptive nonlinear control without overparametrization," Systems and Control Letters, vol. 19, pp. 177-185, 1992.

[37] Y. D. Landau, *Adaptive Control*, New York : Marcel Dekker, 1979.

[38] R. Marino, "On the largest feedback linearizable subsystem," *Systems and Control Letters*, vol. 6, pp. 345-351, 1986.

[39] R. Marino, W. M. Boothby, and D. L. Elliot, "Geometric properties of linearizable control systems," *Mathematical Systems Theory*, vol. 18, pp. 97-123, 1985.

[40] R. Marino and P. Tomei, "Dynamic output-feedback linearization and global stabilization," *Systems and Control Letters*, vol. 17, pp. 115-121, 1991.

[41] R. Marino and P. Tomei, "Global adaptive observers for nonlinear systems via filtered transformations," *IEEE Transaction on Automatic Control*, vol. 37, pp. 1239-1245, 1992.

[42] R. Marino and P. Tomei, "Global adaptive observers and output feedback stabilization for a class of nonlinear systems," pp. 455-493, 1991.

[43] R. Marino and P. Tomei, "Global Adaptive output-feedback control of nonlinear systems, Part I : linear parametrization," *IEEE Transactions on Automatic Control*, vol. 38, pp. 17-32, 1993.

[44] R. Marino and P. Tomei, "Global adaptive output-feedback control of nonlinear systems, PartII : nonlinear parametrization," *IEEE Transactions on Automatic Control*, vol. 38, pp. 33-49, 1993.

[45] R. H. Middleton and G. C. Goodwin, "Adaptive computed torque control for rigid link manipulators," *System and Control Letters*, vol. 10, pp. 9-16, 1988.

[46] R. V. Monopoli, "Model reference adaptive control with an augmented error signal," *IEEE Transactions on Automatic Control*, vol. 19, pp. 474-484, 1974.

[47] A. S. Morse, "Global stability of parameter-adaptive control systems," *IEEE Transactions on Automatic Control*, vol. 25, pp. 433-439, 1980.

[48] K. Nam and A. Arapostathis, "A model-reference adaptive control scheme for pure-feedback nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 33, pp. 803-811, 1988.

[49] K. S. Narendra and A. M. Annaswamy, *Stable Adaptive Systems*, Engle woodcliffs, NJ : Prentice-Hall, 1989.

[50] R.Ortega and M. W. Spong, "Adaptive motion control of rigid robots : a tutorial," *Automatica*, vol. 25, pp. 877-888, 1989.

[51] J. B. Pomet and L. Praly, "Indirect adaptive nonlinear control," *Proceeding of the 27th IEEE Conference on Decision and Control*, Austin, TX, December 1988, pp. 2414-2415.

[52] J. B. Pomet and L. Praly, "Adaptive nonlinear control : an estimation based algorithm ," in *New Trends in Nonlinear Control Theory*, J. Descusse, M. Fliess, A. Isidori, and D. Leborgne, Eds., Spinger-Verlag, New York, 1989.

[53] J. B. Pomet and L. Praly, "Adaptive nonlinear regulation : estimation from the Lyapunov equation," *IEEE Transactions on Automatic Control*, vol. 37, pp. 729-740, 1992.

[54] A. Saberi, P. V. Kokotovic, and H. J. Sussmann, "Global stabilization of partially linear composite systems," *SIAM J. Control Opt.*, 1990, vol. 28, pp. 1491-1503.

[55] S. S. Sastry and M. Bodson, *Adaptive Control : Stability, Convergence and Robustness*, Englewood Cliffs, NJ: Prentice-Hall, 1989.

[56] S. S. Sastry and A. Isidori, "Adaptive control of linearizable systems," *IEEE Transactions on Automatic Control*, vol. 34, pp. 1123-1131, 1989.

[57] J. J. E. Slotine and W. Li, "On the adaptive control of robot manipulators," *International Journal of Robotics Research*, vol. 6, pp. 49-59, 1987.

[58] J. J. E. Slotine and W. Li, "Adaptive manipulator control : a case study," *IEEE Transactions on Automatic Control*, vol. 33, pp. 995-1003, 1988.

[59] E. D. Sontag, "A Lyapunov-like characterization of asymptotic controllability," *SIAM Journal of Control and Optimization*, vol. 21, pp. 462-471, 1983.

[60] E. D. Sontag and H. J. Sussmann, "Further comments on the stability of the angular velocity of a rigid body," *System and Control Letters*, vol. 12, pp. 437-442, 1988.

[61] D. Taylor, P. V. Kokotovic, R. Marino and I. Kanellakopoulos, "Adaptive regulation of nonlinear systems with unmodeled dynamics," *IEEE Transactions on Automatic Control*, vol. 34, pp. 405-412, 1989.

[62] A. R. Teel, R. R. Kadiyala, P. V. Kokotovic and S. S. Sastry, "Indirect techniques for adaptive input-output linearization of nonlinear systems," *International Journal of Control*, vol. 53, pp. 193-222, 1991.

[63] J. Tsinias, "Sufficient Lyapunov-like conditions for Stabilization," *Mathematics of Control, Signal and Systems*, vol. 2, pp. 343-357, 1989.