Glenn-Ole Kaasa

Nonlinear output-feedback control applied to electro-pneumatic clutch actuation in heavy-duty trucks

Doctoral thesis
for the degree of philosophiae doctor

Trondheim, March 2006

Norwegian University of Science and Technology
Faculty of Engineering Science and Technology
Department of Energy and Process Engineering
Telemark University College
Faculty og Technology

NTNU
Norwegian University of Science and Technology

Doctoral thesis
for the degree of philosophiae doctor

Faculty of Engineering Science and Technology
Department of Energy and Process Engineering
Telemark University College
Faculty og Technology

ISBN 82-471-7857-5 (printed version)
ISBN 82-471-7856-7 (electronic version)
ISSN 1503-8181

Doctoral theses at NTNU, 2006:55

Printed by NTNU-trykk

Nonlinear output-feedback control applied to electro-pneumatic clutch actuation in heavy-duty trucks

Glenn-Ole Kaasa

A doctoral thesis submitted to
The Norwegian University of Science and Technology (NTNU)
for the degree of
philosophiae doctor (Ph.D.)

To
Linn, and Ole

Abstract

The thesis deals with the problem of output-feedback position tracking control of an electro-pneumatic actuator. The considered application is a single-acting pneumatic cylinder operated by a three-way proportional valve, which is used for clutch actuation in automated manual transmissions (AMT), and clutch-by-wire (CBW) solutions on heavy-duty trucks.

The high compressibility of air in combination with nonlinear flow and friction characteristics, complicate the control design. Moreover, the pneumatic actuator operates against a highly nonlinear clutch compression spring which constitutes the main load of the actuator. These strong nonlinearities in the system motivates the use of nonlinear control techniques which are capable of explicitly handling nonlinearities. An additional requirement, which further complicate the design, is that only the position is measured and available for feedback control, *i.e.*, an output-feedback control problem must be solved.

A literature study on the modeling and control of electro-pneumatic actuators, reveals that particular properties of existing models exclude the application of existing output-feedback solutions available in the nonlinear control literature. This work provides a unified treatment of the modeling of electro-pneumatic actuators in the context of nonlinear and adaptive control, and introduces some modifications which makes the resulting design model applicable for solutions available for output-feedback control of nonlinear systems. In particular, improved models of the flow rate of flow control valves are proposed.

The modeling work is summarized in a 6th-order dynamic model of the electro-pneumatic clutch actuator, consisting of the actuator position, velocity, friction (seal) deflection, pressures of both chambers, and the valve spool position, as dynamic states. The resulting model is fully feedback linearizable with relative degree four, and can thus be expressed in input-output form where constructive procedures for (adaptive) output-feedback control utilizing high-gain observers are available. Furthermore, the model is in pure-feedback form, which makes it applicable for a nonlinear state-feedback control design by a backstepping approach, and for output-feedback control by an observer-based backstepping approach, provided that an asymptotic observer is available. In this thesis, the latter approach is pursued in the design of an output-feedback tracking controller for the electro-pneumatic system.

It is shown that the model of the unmeasured states can be used as an open-loop nonlinear observer for the electro-pneumatic actuator. Analyzing the stability properties of this open-loop observer, which does not include the unstable integrator of the position y, it is shown to be asymptotically stable, which again establishes

the detectability property of the system. This shows that pneumatic actuators have some inherent stability properties which enable the design of simple nonlinear observers.

Based on this results, two nonlinear observers are proposed, where both are compatible with output-feedback control by an observer-based backstepping approach: a full-order observer and a simpler reduced-order observer. The observers combine closed-loop estimation with linear output-injection terms for the estimation of the main states and open-loop estimation of the remaining states. The observers are asymptotically stable as long as the estimated states remain within the region of validity of the model.

As a robust re-design of the observers, smooth saturation of the state estimates is introduced to constrain the observer dynamics to the feasible region of the state-space, by which ensures global uniform stability properties even if the unsaturated observer states enter the non-feasible region in state-space during initial transients. The differentiability of the introduced smooth saturation ensures compatibility with observer backstepping. The initial transients of the observer are further improved by projecting its non-saturated observer states by a discontinuous projection to a small boundary layer around the region of normal operation. Since the discontinuous projection is only active for estimates which are fully saturated, the smoothness of the saturated estimates is preserved. Hence, a control law can be designed by a backstepping approach using the observer with saturated estimates, and then implemented using the observer with combined saturation and projection, without introducing discontinuities in the control input, thus, preserving the stability properties of the closed-loop system. The performance of the observers are validated by simulations, and experimentally on the test rig.

Based on the reduced-order observer, a robust output-feedback tracking controller is designed by a recursive observer-based backstepping procedure in four steps. An alternative approximate backstepping design is also presented, where the backstepping procedure is simplified for the last two steps using high-gain observers to estimate, rather than calculate analytically, the derivative of the stabilizing function designed at the previous step. Theoretically, the (approximate) backstepping controller achieves exponential (practical) tracking according to a prescribed tracking precision, which can be made arbitrary accurate by sufficiently high feedback gain in the observer and controller.

The output-feedback backstepping controller can be tuned using four main design parameters: The observer gains are set according to the design bandwidth λ_o, and the feedback gains and scaling of control law according to the two parameters c_c and ν_c. The parameters of the reference model are determined according to a design bandwidth λ_r, which can be viewed as the design bandwidth of the closed-loop tracking controller since it determines the time-constant τ_r of the tracking of the reference input r.

A main strength of the output-feedback backstepping controller, is the high tracking performance achieved experimentally. The experimental results shows that the maximum achievable bandwidth of the controller is limited by unmodeled valve dynamics. The controller achieves accurate tracking of the filtered reference tra-

jectory y_r for arbitrary reference inputs r, provided that the bandwidth is chosen according to $\lambda_r < 50$. This corresponds to a time-constant of $\tau_r = 0.100\,\mathrm{s}$ for this 5th-order reference model. For comparison, what is achievable using a properly tuned PID controller while requiring the same accuracy, is tracking according to a time-constant of $\tau_r = 1.500\,\mathrm{s}$.

Preface

This doctoral thesis is based primarily on research conducted in the period December 1999 through March 2003, when I held a research scholarship as a PhD candidate, partially funded as part of a research project through the VARP program by the Norwegian Research Council and by Kongsberg Automotive ASA (KA), in cooperation with the Norwegian University of Science and Technology (NTNU), and Telemark University College (HiT).

My research has included work on different types of prototype pneumatic clutch actuators and valve configurations, and experimental testing has been conducted both in the laboratory and with in-vehicle implementations. For inclusion in this thesis, I have focused on an actuator configuration using a proportional valve, and all experimental results and simulations presented in this thesis are related to a particular application; a pull-type prototype actuator mounted on a test rig clutch at the laboratory at Kongsberg Automotive.

I have chosen to focus on a configuration using a proportional valve, and not multiple pairs of on-off valves, which is the most cost-effective solution for series production. This choice has been made of two main reasons: The first is that the use of a proportional valve (or servo valve) is the most common configuration for electro-pneumatic servo systems in general. Hence, by considering a proportional valve configuration, it makes the results on the modeling and control of the clutch actuation system presented in this thesis, more generalized, as they readily carry over to a wide range of applications of electro-pneumatic actuators using proportional or servo valves. The second reason, which perhaps is more important, is that KA's industrialized solution using on-off valves has not yet been set into series production. Hence, in order to let KA be the first to bear the fruits from my efforts in this area, I have decided to postpone the publication of the results with on-off valves, which in many ways are extensions of the results obtained using a proportional valve. Of this reason, I have also decided not to include in the thesis, the results from adaptive designs, where the parameters of the clutch load characteristic are estimated on-line by the output-feedback controller.

In the writing of this thesis, I have adopted an informal and personal style from the US that I have come to like, but which is somewhat uncommon for Ph.D. theses in Europe. This means that I often write in plural form using the personal pronoun 'we' when describing work and results which are entirely my own.

Several people have contributed to this research in different ways. First of all, I would like to thank my advisors Professor Peter J. Chapple at Department of Energy and Process Engineering, NTNU, and Professor Bernt Lie at Department of

Electrical Engineering, Information Technology, and Cybernetics, HiT. I am grateful for their dedicated interest in my research. I would like to thank Bernt, in particular, for thorough and valuable feedback on my work, and for help with arranging my research stay in the US.

During my doctoral work, I have been affiliated with the Cybernetics Research Group (CyneRG) at HiT, led by Bernt Lie, together with Bernt's other Ph.D. students, Tor Anders (with whom I shared office), Marta, and Beathe. Thank you all for a rewarding time, both professionally and personally.

I am thankful to Professor Brad Paden at University of California, Santa Barbara (UCSB) for inviting me for a research stay the first half of 2002 — a visit which has greatly influenced my work. Particularly, am I grateful to Professor Petar V. Kokotović at UCSB, who through his encouraging and enthusiastic lecturing in his adaptive control courses, opened a door into the exciting world of nonlinear and adaptive control. This has had a strong influence on my work, and has definitively brought my understanding of control theory to a higher level. I would also like to thank my fellow students from the laboratory at Engineering II at UCSB and Professor Kokotović's courses: Lasse, Ove, Niklas, Makan, Mihailo, Yonggang, Dragan, Emre, Aruna, and all of you who contributed to make my stay an unforgettable memory.

During the spring 2002, I was also fortunate to have the possibility to work with Professor Masanori Takahashi at Ariake National College of Technology in Japan, applying his results in adaptive control to the pneumatic clutch system, which resulted in a joint publication for the AdCONIP'02 conference in Japan, 2002. Also, Professor Per-Åge Krogstad at Department of Applied Mechanics, Thermodynamics and Fluid Dynamics, NTNU, is thanked for taking the time to review some of my work on flow modeling. Furthermore, I owe a thank to the staff at the library at HiT, particularly Patricia Floor, for excellent help with obtaining various literature for my research, and Per Bjørnaas at NTNU, for support whenever my laptop seemed to fail me.

From April 2003 I was employed by Kongsberg Automotive, where I have concluded the writing of the thesis. I am indebted to my colleagues at KA for providing a stimulating work environment, and for letting me draw from their well of knowledge. Particularly, I would like to thank Inge André Haraldstad Johansen, previous fellow student at NTNU, now colleague, and Morten Gunnerud, also colleague at the R&D group at KA. Thanks should also go to Olav Volldal, CEO at KA, and Bjørn Iversen, R&D manager at the time, for making the Ph.D. project possible, and to Lucien Lenerand, Manager Concept Development, for seeing the benefit of an extension of the Ph.D. project. Finally, acknowledgements should go to the U.S.- Norway Fulbright Foundation, who through a research scholarship made my rewarding research stay at UCSB possible, together with funding from Kongsberg Automotive ASA and the Norwegian Research Council.

Not least, I would like to thank my wife Linn for the understanding and the sacrifices she has made during periods when there has been much work and little time left for her. She, and our son Ole, have provided a new dimension to life, for which I am utmost grateful.

Notation

For time-varying variables, like inputs, outputs, and states, the time-argument t is generally dropped, except when time-dependence should be emphasized, or it is unclear from the context. Function arguments are often dropped whenever no confusion might occur, and sometimes they are replaced by a dot, like *e.g.* $f(\cdot)$ instead of $f(x_1, x_2, x_3)$, to indicate dependence on arguments.

Bold font types are used for vectors and matrices to distinguish from scalars. For vectors, the inequality operators ($>$, $\geq$, $<$, $\leq$), means an elementwise application, *i.e.*, $[a, b, c]^T < [2, 6, 11]^T \iff a < 2 \wedge b < 6 \wedge c < 11$. Similarly, the inverse operator (x^{-1}) refers to application to each element when applied to a vector, *i.e.*, $\mathbf{x}^{-1} = [x_1^{-1}, x_2^{-1}, \cdots, x_n^{-1}]^T$. Furthermore, we let the operator $\times$ denote elementwise multiplication, for example for the vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, we have

$$\mathbf{q} = \mathbf{x} \times \mathbf{y} = [x_1 y_1, x_2 y_2, \cdots, x_n y_n]^T, \tag{1}$$

where $\mathbf{q} \in \mathbb{R}^n$.

Signal measures

For a scalar $x \in \mathbb{R}$, the operator $|\cdot|$ denotes the absolute value. For a vector $\mathbf{x} \in \mathbb{R}^n$, we let $|\mathbf{x}|$ denote the ℓ_2 (Euclidian) vector norm, defined below. The ℓ_1 norm is defined as

$$|\mathbf{x}|_1 \triangleq |x_1| + \cdots + |x_n|, \tag{2}$$

the ℓ_2 (Euclidean) norm as

$$|\mathbf{x}|_2 \triangleq \sqrt{x_1^2 + \cdots + x_n^2}, \tag{3}$$

and the ℓ_∞ norm as

$$|\mathbf{x}|_\infty \triangleq \max_{1 \leq i \leq n} |x_i|. \tag{4}$$

The weighted Euclidian (Frobenius) norm is defined as

$$|\mathbf{x}|_{\mathbf{P}} \triangleq \sqrt{\mathbf{x}^T \mathbf{P} \mathbf{x}} \tag{5}$$

where $\mathbf{P}$ is a weighting matrix.

We let $\|\mathbf{x}(t)\|_1$, $\|\mathbf{x}(t)\|_2$, and $\|\mathbf{x}(t)\|_\infty$ denote, respectively, the $\mathcal{L}_1$, $\mathcal{L}_2$, and $\mathcal{L}_\infty$ norms which are used to characterize upper bounds on a time-varying vector signal $\mathbf{x}(t) \in \mathbb{R}^n$. The $\mathcal{L}_1$ norm is defined as

$$\|\mathbf{x}(t)\|_1 \triangleq \int_0^\infty |\mathbf{x}(t)|_1 \, dt = \int_0^\infty (|x_1(t)| + \cdots + |x_n(t)|) \, dt, \tag{6}$$

the $\mathcal{L}_2$ norm as

$$\|\mathbf{x}(t)\|_2 \triangleq \int_0^\infty |\mathbf{x}(t)|_2 \, dt = \int_0^\infty \sqrt{x_1(t)^2 + \cdots + x_n(t)^2} dt, \tag{7}$$

and the $\mathcal{L}_\infty$ norm as

$$\|\mathbf{x}(t)\|_\infty \triangleq \sup_{t \geq 0} |\mathbf{x}(t)|_\infty = \sup_{t \geq 0} \max_{1 \leq i \leq n} |x_i(t)| . \tag{8}$$

If $\|\mathbf{x}(t)\|_1$ exists, *i.e.*, it is finite, we say that $\mathbf{x}(t)$ is *integrable*, if $\|\mathbf{x}(t)\|_2$ exists, $\mathbf{x}(t)$ is *square integrable*, and if an upper bound $\|\mathbf{x}(t)\|_\infty$ exists, $\mathbf{x}(t)$ is *bounded*.

Terminology

Some commonly used expressions from the modeling and control terminology (and the meaning in which the author make use of them), are summarized below:

affine: An expression $f(x, y)$ is said to be *affine* in its variable x, if it is a linear function of x, *i.e.*, $f(x, y) = g(y) \cdot x$. Furthermore, we say that $f(x, y)$ is *piecewise affine* in x, if it is a piecewise linear function of x. The notation is extended to the multivariable case. We say that an expression $f(\mathbf{x}, \boldsymbol{\theta})$ is *parameter-affine* if it is linear in its parameter vector $\boldsymbol{\theta} \in \mathbf{R}^p$, *i.e.*, $f(\mathbf{x}, \boldsymbol{\theta}) = g_1(\mathbf{x}) \theta_1 + \cdots + g_p(\mathbf{x}) \theta_p = \mathbf{g}(\mathbf{x})^T \boldsymbol{\theta}$.

Hurwitz matrix: A square real matrix is Hurwitz if all its eigenvalues have negative real parts, *i.e.*, a linear system $\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}$, where $\mathbf{A}$ is Hurwitz, is exponentially stable.

matching condition: A term (function, uncertainty, *etc.*) is said to satisfy the *matching condition* if it is in the range of the control input, that is, it enters in the state equation at the same point as the control input, making it possible to directly cancel it by the control.

mechanistic: We use the word *mechanistic* in the meaning "based on the common laws of nature", *i.e.*, we refer to a model as mechanistic, when it is based on physical laws, with parameters of physical meaning. In this sense, mechanistic is the antonym to *empirical*, since an empirical model refers to a model based merely on observations rather than theory (or physical laws), where the parameters in general have no physical interpretation.

output-feedback: By *output-feedback* control we refer to the case when only the output is available for feedback to the controller, *i.e.*, only the output is measured. By *state-feedback* control (or partial state-feedback control), we refer to the case when all the states of the dynamic system (or some of them) are measured, thus, available for feedback to the controller.

parametrization: A *parametrization* of a function refers to a mathematical expression of the function in terms of variables and parameters (*i.e.*, not a look-up table). When a function is expressed in a parameter-affine form, it is referred to as *linear parametrization*. For example, a 2nd-order polynomial $f(x) = a_0 + a_1 x + a_2 x^2$, is a linear parametrization of the function $f(x)$ in the parameters a_0, a_1, and a_2.

smooth: We use the term *smooth* to indicate that a function is differentiable, *i.e.*, a function $f(x)$ is (sufficiently) smooth in x, if it is (sufficiently many times) differentiable with respect to x. For example, the continuous function $f(x) = |x|$ has a non-smooth breakpoint for $x = 0$, thus, is not smooth, and its derivative is undefined for $x = 0$.

tracking: By *tracking* we refer to the case when the objective of the controller is to make the output y of the system track a time-varying reference trajectory, *e.g.* given by $\mathbf{z}_r \triangleq [y_r, \dot{y}_r, \cdots, y_r^{(n-1)}]^T$. Tracking is thus a more general (and difficult) control task than *regulation* (or *stabilization)*, which refers to the special case where the reference is a fixed set-point, *i.e.*, simply $\mathbf{z}_r \triangleq [y_r, 0, \cdots, 0]^T$.

Mathematical symbols

For a compact notation, various mathematical symbols are used, summarized in the table below.

Symbol	Description, meaning
$\times$	elementwise multiplication operator
$\equiv$	identically equal
$\approx$	approximately equal
$\triangleq$	defined equal
$\forall$	for all
$\rightarrow$	tends to
$\in$	belongs to
$\exists$	there exists
$\subset$	subset of
$\wedge$	and
$\vee$	or
$\mathbf{0}_{m\times n}$	matrix of zeros of dimension $m \times n$
$\lambda(\mathbf{A})$	eigenvalue(s) of the matrix $\mathbf{A}$
$\lambda_{\max}(\mathbf{A})$	maximum eigenvalue of the matrix $\mathbf{A}$
$\lambda_{\min}(\mathbf{A})$	minimum eigenvalue of the matrix $\mathbf{A}$
$\mathbf{e}_j$	jth unit basis vector (*e.g.*, $\mathbf{e}_1 = [1, 0, \cdots]^T$)
$\hat{x}$	estimate of x
$\dot{x}$	$\triangleq \frac{d}{dt}(x)$, first-order time-derivative of x
$\ddot{x}$	$\triangleq \frac{d^2}{dt^2}(x)$, second-order time-derivative of x
$x^{(n)}$	$\triangleq \frac{d^n}{dt^n}(x)$, nth-order time-derivative of x
$\operatorname{sgn}(x)$	$\triangleq \begin{cases} 1, & x > 0 \\ 0 & x = 0 \\ -1 & x < 0 \end{cases}$, signum function
$\operatorname{sat}(x)$	$\triangleq \begin{cases} 1, & x > 1 \\ x & x \in [-1, 1] \\ -1 & x < -1 \end{cases}$, saturation function
$\operatorname{dist}(\mathbf{x}, \Omega)$	$\triangleq \inf_{\mathbf{z}\in\Omega} \lvert\mathbf{x} - \mathbf{z}\rvert$, shortest distance from a point $\mathbf{x}$ to a set Ω
$\max(\cdot)$	maximum
$\min(\cdot)$	minimum
$\sup(\cdot)$	supremum, the least upper bound
$\inf(\cdot)$	infimum, the greatest lower bound

Acronyms and abbreviations

A summary of some commonly used acronyms and abbreviations are given in the table below.

Abbreviation	**Meaning**
AMT	Automated Manual Transmission
CBW	Clutch-by-Wire
SISO	Single-Input Single-Output
NN	Neural Network
RBF	Radial Basis Function
AS	Asymptotically Stable
ES	Exponentially Stable
ISS	Input-to-State Stable
GUB	Globally Uniformly Bounded
GUS	Globally Uniformly Stable
GUAS	Globally Uniformly Asymptotically Stable
GUES	Globally Uniformly Expnentially Stable
CLF	Control Lyapunov Function

Contents

Chapter 1

Introduction

The topic of this doctoral thesis is the design of a tracking control system for electro-pneumatic clutch actuation in heavy-duty trucks. This introductory chapter provides the background and an overview of the problem, a review of existing work in the literature, and outlines how the problem has been solved in this thesis.

First, Section 1.1 describes the background for this project, and outlines its main challenges. The next two sections provide a review of existing published work; Section 1.2 on mathematical modeling of electro-pneumatic actuators, and Section 1.3 on control of electro-pneumatic actuators, also describing some applicable results from nonlinear control theory not yet employed for control of pneumatic actuators. An outline of the organization of the remaining parts of the thesis is given in Section 1.4.

1.1 Background

1.1.1 Electro-pneumatic clutch actuation

An increasing demand for improved driving comfort has resulted in an increased effort from vehicle manufacturers to develop cost-effective automated solutions. In heavy-duty trucks, electro-pneumatic actuators are used to automate the clutch and gear shift operation of *manual transmissions* with friction disc clutches. Such systems are usually referred to as *automated manual transmissions* (AMT). In heavy-duty trucks, AMT systems are the preferred choice over *automatic transmissions* with hydraulic clutches, mainly because automatic transmissions designed for high torque transfer, are expensive, and have a considerable power loss compared to manual transmissions. Furthermore, the possibility to electronically control the clutch actuation — inherent in AMT systems — offers a higher degree of flexibility with respect to engine and transmission control which can be used to reduce engine emissions, fuel consumption, and minimize wear of clutch and transmission.

Because pressurized air is available on the truck, pneumatic actuators are the preferable choice over hydraulic actuators (used in AMT systems on personal cars), which would require an additional hydraulic power unit in order to provide the necessary hydraulic pressure supply.

Compared to hydraulic actuators, pneumatic actuators are inherently difficult to control, mainly due to the high compressibility and nonlinear flow characteristics of air. The automation of the shift and select actuation of the gear box requires only an on-off function, and is performed by simple open-loop control of the pneumatic actuators, with the positioning of the actuator piston performed by mechanical constraints. The automation of the clutch actuation, on the other hand, requires both smooth and precise engagement and disengagement of the clutch in order to assure smooth speed control, and low wear of the clutch and transmission. This requires the control system to be able to perform high-performance tracking control of the pneumatic actuator piston position. This is a particularly difficult control task which is not properly accomplished in existing pneumatic clutch actuation systems, which exhibit a relatively poor tracking performance primarily due to an inability of the control system to properly compensate for the strong nonlinearities in the system. Consequently, there lies a significant potential in utilizing advances in nonlinear control theory in order to improve the tracking performance of pneumatic clutch actuation systems.

1.1.2 Mathematical modeling

The design of a nonlinear control system for improved and robust tracking performance of electro-pneumatic clutch actuation systems, requires a mathematical model of the system which is suited for control design. The development of a model for control design is always a compromise between accuracy and simplicity, *i.e.*, in order to minimize the complexity of the resulting control design the model should not be more detailed than required by the specific control task. However, the model must be accurate enough to make possible an observer design for reconstruction of the unmeasured states, and to provide precise feedforward compensation of nonlinearities. In addition, the model should have certain mathematical properties which accommodate a nonlinear control design. The pneumatic system has a physical structure which lets it be expressed in the so-called *pure-feedback form* in state-space, which makes it applicable for a recursive *integrator backstepping* design. In order to make the model suited for exact backstepping, the nonlinear functions — or *nonlinearities* — in the model must be sufficiently *differentiable* (or *smooth*). Furthermore, it is advantageous that uncertain nonlinearities in the model are *linearly parametrizable* in order to facilitate parameter estimation, and make possible existing constructive adaptive designs where the uncertain nonlinearities, *i.e.* the parameters of the nonlinear functions, are estimated on-line. In particular, this applies to the modeling of the clutch compression spring — referred to as the *clutch load characteristic* — which is a strongly nonlinear function of the actuator position, representing the most significant nonlinearity in the system.

Existing models of specific system components and nonlinearities in electro-pneumatic actuators are not well suited for the constructive methods currently available for nonlinear and adaptive model-based control; that is, recursive designs by a backstepping approach, or designs based on a transformation of the model to an input-output form. In most applications where electro-pneumatic actuators are

used, the model of the *flow control valve* constitutes the most important nonlinearity in system. Existing models of flow control valves are either highly accurate models developed for simulation which are too complex for control design; or simplified explicitly invertible models, which for many control valves provide less accuracy than desired, and that are not differentiable. Furthermore, the *friction*—both in the actuator and in the external load—constitutes another important nonlinearity in the system which must be accurately modeled in order to achieve high-performance tracking. The modeling of friction in pneumatic actuators has until recently been limited to static models, which are inadequate for high-precision model-based control, because important dynamic properties of friction in pneumatic actuators are then disregarded. Summarizing, in order to utilize recent advances in nonlinear and adaptive control theory for improved control of electro-pneumatic actuators, a required first step is the development of a model of the electro-pneumatic system which is suited for nonlinear and adaptive control.

1.1.3 Control design

The primary objective of the control system is to achieve robust high-precision tracking control of the electro-pneumatic clutch actuator. The high compressibility of air in combination with the strong inherent nonlinearities of electro-pneumatic actuators — like the air flow characteristics and dynamic friction — makes the control design a particularly difficult task. Additionally, in the clutch actuation application, the electro-pneumatic actuator operates against a strongly nonlinear clutch compression spring, which is a static nonlinear function of the clutch position referred to as the *clutch load characteristic*. This nonlinear function constitutes an additional, and most significant, nonlinearity in the system, making high-performance tracking control impossible without nonlinear compensation of some sort. This motivates the application of *nonlinear control techniques* which are capable of explicitly handling nonlinearities in the system.

An important and severe requirement for the control design, is that only the position is measured and available for feedback control, as economic considerations preclude the use of additional sensors to measure all, or some, of the remaining system states. Consequently, in lack of full-state measurements, we pursue an *output-feedback* design by following a state-variable approach which requires the synthesis of an *observer* to reconstruct the unmeasured states for use by the control law.

An obvious and ultimate objective of the control system, is that it must be *robust* in the sense that it effectively deals with *uncertainties* in the design model, and attenuates possible *disturbances*, such as changes in the friction and clutch load characteristic due to temperature changes, ageing, and wear of the clutch.

An additional desired objective of the control system, is that the controller should be *universal*, in the sense that it should be able to perform high-performance tracking control of the electro-pneumatic actuator with any type of clutch, without the need of a manual change, or tuning, of its parameters. In essence, this most likely requires the design of a self-tuning, or an adaptive controller, which is able to identify the strongly nonlinear clutch load characteristic, either by an initialization routine after

assembling, or by on-line adaptation.

1.2 Literature review — mathematical modeling

Only a few published papers have been found that address the modeling and control of electro-pneumatic actuators applied to clutch actuation, and none of these gives a complete and uniform treatment of the modeling of the system. Research that should be mentioned is work by Tanaka *et al.*, who consider torque control by clutch actuation in automated manual transmissions by use of an electro-pneumatic pressure proportional valve[1], see e.g. [93] and the references therein. In the papers [40] and [42], Kaasa *et al.* employ slightly different models of the electro-pneumatic clutch actuator for experimental implementation of an Extended Kalman filter for state estimation, and for the simulation and analysis of an adaptive tracking controller, respectively. Xiang and Wikander are other researchers who have considered pneumatic clutch actuation in particular. See the technical report [103] cited in Xiang's thesis [102] on the control of pneumatic actuators.

The modeling of pneumatic actuators in general, on the other hand, has received a great amount of attention during the last decades, and a vast number of papers on the subject have been published. The basic theory of the modeling and control of pneumatic actuators can be found in the fluid power text books by Blackburn *et al.* [9], or Anderson [4]. The early work by Shearer [86] has been frequently referenced in research papers. Reethof and Shearer's work on the modeling of pneumatic actuators, summarized in [9], was further extended by Jebar in his thorough thesis [37], providing an analysis of the dynamics of pneumatic cylinder actuators which is substantiated by extensive experimental validation. An assortment of the most interesting literature is reviewed below, grouped into the modeling of the dynamics of the air states in the pneumatic chambers — referred to as the *air dynamics*, the *flow rate* of control valves and restrictions, the *friction* in the pneumatic actuator and load, and briefly the modeling of *static nonlinearities* in general.

1.2.1 Dynamics of the pneumatic chambers

The physical mechanisms of the thermodynamic properties *e.g.* of air in the pneumatic cylinder chambers, are well understood. A vast number of published papers address the modeling of pneumatic systems, which also includes the modeling of the dynamics of the pneumatic chambers, referred to as the *air dynamics*. By applying some reasonable assumptions, an accurate full-state dynamic model of the air states (*e.g.*, pressure and temperature) in the pneumatic chamber can be derived based on simple thermodynamics using an empirical heat transfer model. This full-order model of the air dynamics — with pressure and temperature as state variables — is derived by Jebar in his comprehensive thesis on the design of pneumatic actuator systems [37]. The model is consequently referred to as the *Jebar model*. In [20],

[1] The valve has a pressure sensor with an internal pressure control loop which regulates the pressure according to the pressure set-point given as the valve's input signal.

Det *et al.* proposed a modification of the Jebar model by introducing a pressure and temperature dependent convective heat coefficient in the empirical heat transfer model, motivated by research on combustion engine modeling. In [13] and [42], the full-order Jebar model, applied to the modeling of pneumatic cylinder actuators, is presented in state-space form with pressure and temperature as state variables. A derivation of the full-order Jebar model — in a form with the chamber energy and mass as state variables — is given by Maré *et al.* in [64].

For analysis and control design, a simplified reduced-order model of the air dynamics is preferred. The classical reduced-order model of the air dynamics is obtained by neglecting heat transfer, and assuming that the temperature is constant and equal to the temperature of the inlet flow. The air dynamics is then given as a 1st-order nonlinear equation for the pressure dynamics. This model was first presented by Shearer in [86, Part I]. Variations of Shearer's model is obtained by viewing the ratio of specific heats ($\kappa \triangleq c_p/c_v$) as a polytropic exponent n, and letting the chamber temperature be given as a function of the chamber pressure according to a polytropic process. In this case, the choice $n = 1$ is referred to as the *isothermal* reduced-order model, while with the choice $n = 1.4$, the model is referred to as the *adiabatic* reduced-order model. In the thesis work [97], Virvalo presents an extensive experimental validation of Shearer's reduced-order model applied to heavy pneumatics, reporting highly accurate results with commercially available pneumatic cylinder actuators. Shearer's reduced-order model is the model most researchers have used to model the air dynamics of pneumatic actuators, and practically all published work dealing with nonlinear model-based control of pneumatic actuators employ this model, see e.g. [13], [56], [69], [79], [80], [96], [98], [104].

1.2.2 Flow rate modeling

The conventional approach to the modeling of the flow rate characteristic of a pneumatic component, is by use of the theoretically derived equation for isentropic compressible flow through a simple orifice, referred to as the *isentropic orifice flow equation*. A rigorous treatment of the fundamental compressible flow theory can be found in student texts on fluid mechanics, *e.g.* the textbook by White [101]. A fluid power approach to the theory is given by Blackburn *et al.* in [9]. In the ISO standard [35], a simple, but rather accurate approximation of the isentropic orifice flow equation has been standardized for the determination of the flow rate characteristic of pneumatic components. This simplified model, referred to as the *standardized* orifice *flow equation*, consists of an elliptic approximation to the theoretically derived *pressure ratio function* which describes the flow rate's dependence on the pressure ratio over the restriction. For most pneumatic components, this elliptic approximation provides a closer fit to measurements than the theoretically derived pressure ratio function because it allows a tuning of the effective *critical pressure ratio* B, which for a given restriction geometry, uniquely determines the pressure dependence of the modeled flow rate.

The flow equation is used to construct a model of the flow rate characteristics of flow control valves, thus, it should provide an accurate description of the flow

rate over the full range of valve openings. The elliptic pressure ratio function of the standardized orifice flow equation provides an accurate description for simple restrictions where frictional effects are small, however, for flow (or leakage) through smaller clearances, the effect of friction becomes increasingly important, and the accuracy of the equation reduces. Furthermore, in the formulation of the elliptic pressure ratio function, the parameter B appears highly nonlinearly, which makes parameter estimation difficult and not suited for adaptive control. Thus, aspects of the flow modeling which still remains to be properly solved, is the development of a parametrization of the pressure ratio function which allows modeling of the flow through small clearances with improved accuracy, in a form suited for subsequent application of existing tools for nonlinear and adaptive control.

Flow control valves

Flow control valves used for actuator control are either three-way valves connected to one cylinder chamber, supply, and exhaust reservoirs, or five-way valves connected to both cylinder chambers, supply, and exhaust. With a few exceptions, most flow control valves are classified as sliding type valves, e.g. spool, sliding plate, or rotary-plug valves, referring to the principle for operation of the valve. In the following, we discuss the modeling of a spool type valve, which applies to the modeling of most types of sliding valves. Typical flow control valves used with pneumatic actuators are proportional valves and servo valves, where the main difference lies in the type of actuation device. Pneumatic *proportional valves* are actuated by a an electro-magnetic *force motor*, which is either direct acting on the spool of the flow stage, or operates a pilot spool on two-stage valves. The principal characteristic of proportional valves is that the spool position is proportional to the valve input in steady-state, where the positioning is performed either by balancing the solenoid force against the spring force, or by a control loop with feedback from a sensor measuring the spool position. The force motor is usually a proportional solenoid which provides a force proportional to the coil current in steady-state. For valves with spool feedback the force motor can alternatively be an on-off solenoid. Pneumatic *servo valves* are usually two-stage valves where an electro-magnetic *torque motor* actuates the pilot stage of the valve which pneumatically amplifies the positioning of the spool in the main flow stage.

An important characteristic of flow control valves is their steady-state response from valve input to spool position, referred to as the *input–spool characteristic*. For proportional valves, the input–spool characteristic is ideally linear, but without spool feedback it will always be encumbered with hysteresis resulting from friction, and the disturbing effect of flow forces which depend on the flow rate. For pneumatic servo valves, the input–spool characteristic is usually considerably nonlinear.

The flow rate as a function the position of the spool is referred to as the *spool–flow characteristic* of the valve. The spool–flow characteristic is characterized by its spool and port design, which determines the geometry of the flow path. The spool lap is the main factor which determines the flow properties in the *null region* of the valve. The null region may be defined as the region of spool positions in the neighborhood the center position where there is leakage flow through both the

supply and exhaust ports. Outside the null region the flow is primarily orifice flow either through the supply port, or the exhaust port, which is proportional to the position of the spool. The flow properties in the null region represents a significant nonlinearity in the valve. Ideally, the spool of proportional valves is designed with zero lap in order to provide a linear proportional spool–flow characteristic. However, manufacturing tolerances introduce an inevitable leakage flow in the null region. To compensate for manufacturing tolerances and reduce the leakage flow in the null region, proportional valves are often designed with a slight overlap, resulting in a dead zone nonlinearity in the spool–flow characteristic. Most servo valves (and some proportional valves), on the other hand, are designed with underlap, resulting in a considerable leakage flow in the null region.

A good description of the modeling of the steady-state flow rate characteristic of a three-way pneumatic proportional valve is given in the first of the referenced papers by Shearer [86], and in [9]. The description applies for the modeling of the spool–flow characteristic of sliding type valves in general. The flow rate is modeled according to the isentropic orifice equation, assuming that the effective restriction area has a linear dependence on the spool for each valve orifice, and includes the modeling of leakage flow in the null region by introducing an underlap in the model. In the above mentioned paper, Shearer demonstrates a close fit to measurements of a zero-lapped sliding plate valve. Though providing an accurate description of zero- or under-lapped proportional valves, the model is not well suited for control because the introduced underlap causes flow through both flow paths in the null region, which complicates the computation of the inverse. Another example of an accurate mechanistic model of the spool–flow characteristic of a five-way spool valve including leakage in the null region, is derived by Mo in the paper [66]. The model is more general than Shearers model and is validated experimentally to be highly accurate for the modeling of a spool valve. Like Shearer's model, however, it is well suited for simulation, but too complex for control.

Simplifying Shearer's model by assuming a zero underlap, results in a model which is *piecewise input-affine*, *i.e.*, the input u appears affinely in the form

$$w = g\left(p, \operatorname{sgn} u\right) \cdot u, \tag{1.1}$$

where p is the chamber pressure, and where u is the spool position which is regarded as the input. Because of the affine form, the model is explicitly invertible, making it particularly suited for control design. However, because the model disregards leakage flow in the null region, it is primarily suited for the modeling of high-performance zero-lapped valves with precise manufacturing tolerances such that the null region is negligibly small. This model has been used by several researchers for model-based nonlinear control employing full state-feedback. Common for these applications, is their use of robust control techniques which suppress model uncertainties, like the disregarded leakage in the null region. Pandian *et al.* employ the model in successful implementations of sliding mode control applied to the control of both vane-type rotary actuators and cylinder actuators in [73], [74], [75]. Wang *et al.* report good experimental results in a robust nonlinear design applied to the control of a cylinder actuator in [99], [100]. In [44], Keller & Isermann employ a model in this piecewise

input-affine form using the elliptic approximation of the ISO standard for the modeling of the pressure ratio function, demonstrating good experimental results with a model-based nonlinear adaptive scheme applied to a pneumatic cylinder actuator. In addition, several researchers employ this simplified piecewise input-affine model in simulation studies of nonlinear control of pneumatic actuators, see *e.g.* [12], [21], or the more recent papers [2] and [1] by Acarman *et al.*

In the paper [10], Bobrow and McDonell point out the discrepancy of modeling a jet-pipe type Moog servo valve using the input-affine model of a proportional valve. This is primarily due to a considerable leakage in the null region of the valve combined with a nonlinear input–spool characteristic, inherent in most pneumatic servo valves. In [79], Richard proposes a model structure which consists of an orifice flow term modeled by the orifice flow equation with a fictitious flow area as a function of the input, and an additional leakage term representing a fixed clearance in the valve. The model is in the form

$$w = g_o\left(p, \operatorname{sgn}\phi\left(u\right)\right) \cdot \phi\left(u\right) + g_l\left(p\right), \tag{1.2}$$

where the fictitious area function $\phi\left(u\right)$ — referred to as the *input nonlinearity* of the valve — appears in a piecewise affine form, thus, this model structure is referred to as being in an *input nonlinearity-affine form*. The input nonlinearity $\phi\left(u\right)$ can alternatively be modeled as a flow conductance function, and is in essence a lumped function approximating the total steady-state input nonlinearity of the valve, *i.e.*, representing both the nonlinear leakage flow in the null region and the nonlinear input–spool characteristic. This form is convenient for control because the nonlinear input function $\phi\left(u\right)$ is straightforward to parametrize in a form which is explicitly invertible, thus, resulting in a flow model which is explicitly invertible. A main drawback, however, is that a model in the form (1.2) is not capable of providing an accurate description of the leakage flow in the null region, since this requires the modeling of the flow through each of the flow paths separately due to the nonlinear flow characteristics of compressed air. Hence, a model in the form (1.2) is not suited for accurate modeling of zero- or underlap valves where the leakage flow in the null region is significant.

The form (1.2) is well suited for the modeling of overlap valves where the leakage flow in the null region is small. For example in the paper [80], Richard & Scavarda implements a feedback linearization controller applied to the tracking control of a pneumatic actuator, using a model in the form (1.2) to model an overlap servo valve. Further validation of a model in this form for the modeling of overlapped servo valves has been demonstrated in the recent papers by Lee *et al.* [56] and by Maré *et al.* in [64]. Maré *et al.* employ an experimentally obtained nonlinear function of the effective conductance instead of the effective area, and employ the ISO standardized elliptic approximation instead of the isentropic pressure ratio function. In the models presented in the above mentioned papers [79], [80], [56], [64], the input nonlinearity is implemented as a look-up table which is obtained experimentally for the given application, and a parametrization of the nonlinear function is not provided.

In [85], a detailed experimental study of a three-way Servotronic proportional valve was conducted, and a highly accurate flow rate characteristic was obtained.

Purely empirical parametrizations of this characteristic, in the form of multivariable polynomial approximations, were proposed in [8]. In this paper, Belgharbi *et al.* present a highly accurate model for simulation, and a simplified model for control which is in the above mentioned input nonlinearity-affine form (1.2). In this simplified empirical model, the input nonlinearity becomes a piecewise linear function with a small discontinuity for the spool in its center position due to a slight overlap in the valve. The model is employed for tracking control of pneumatic actuators in work by Brun *et al*, for example in the paper [13]. A main drawback with the simplified parametrization in the form (1.2) presented in [8], is its discontinuity in the spool position at the origin, and its pure empirical nature which requires extensive measurements in order to identify all the parameters of the model.

The currently most general parametrization of a model in the form (1.2) is proposed by Xiang and Wikander, who use the model in approximate feedback linearization approaches for the control of pneumatic actuators, see for example the paper [102, Paper A]. Their model is partially mechanistic in the sense that the pressure dependence is modeled by the (mechanistic) elliptic pressure ratio function, and the lumped input nonlinearity is parametrized by a dead zone function multiplied by an empirical polynomial approximation.

1.2.3 Friction modeling

Friction in mechanical servo systems is a complex phenomenon, which traditionally is considered hard to model accurately. The first real survey on the modeling of friction for control purposes, was conducted by Armstrong-Hélouvry *et al.* in 1994 [6]. In many ways, this work initiated an extensive amount of research on the subject, which has accelerated the development of new friction models suitable for model-based friction compensation.

Several works on the modeling of friction in pneumatic actuators has been published in the literature, where most are concerned with static friction models only. Schroeder and Singh analyze the most interesting of these static friction models in [84], where seven models are compared and validated from experiments with a pneumatic cylinder actuator. In most pneumatic actuators dynamic frictional effects are important, and a dynamic friction model is required in order to achieve an accurate description of the friction in pneumatic actuators. In particular, the dynamic pre-sliding deflection of the seal between the piston and cylinder wall, plays an important role for accurate friction modeling due to its high elasticity.

In the paper [15], Canudas de Wit *et al.* present a dynamic friction model — referred to as the *LuGre model* (Lund-Grenoble), which captures most of the qualitative properties of friction reported in the literature. The model is simple, and well suited for model-based control. A drawback with the model is that it results in a non-physical drift in position in the pre-sliding phase, *i.e.*, it does not render true stiction. An improved model, compared to the LuGre model, was published by Dupont *et al.* in [22]. This model is referred to as the *Elasto-plastic model*, and employs a switching function in order to render plastic deformation (which results in drift) only for load forces above a certain limit. That is, the model renders pure

elastic deformation, thus, true stiction, below this limit. In [91], Swevers *et al.* introduce a more elaborate model — referred to as the *Leuven model*, which demonstrates improved pre-sliding properties compared to the LuGre model (presumably also compared to the Elasto-plastic model). By introducing a hysteresis function with nonlocal memory, the Leuven model renders true stiction without the use of a switching function. Some modifications of the Leuven model was later proposed by Lampaert *et al*, in [54]. In the paper [70], Nouri *et al.* employ the Leuven dynamic friction model in an experimental study on the modeling of a pneumatic actuator.

In [31], Hsieh & Pan provide the most complete and accurate description to date of the pre-sliding friction properties (in the author's view), and present a highly accurate model of the pre-sliding friction properties of mechanical systems. The model has many similarities to the above mentioned Leuven model, but is more complex.

1.2.4 Static nonlinearities

From a parameter estimation point of view, it is beneficial that nonlinear functions in the system model are parametrized in forms where the parameters appear in an affine fashion. With respect to off-line parameter fitting from measurements, parameter-affine models make possible parameter estimation by *convex optimization* (for example, linear least-squares parameter fit to measurements). Furthermore, existing constructive tools for nonlinear adaptive control, comprise only systems with uncertain nonlinearities which can be parametrized in a parameter-affine form.

A smooth nonlinear function can in general be modeled, or approximated, by a weighted sum of simple basis functions, where increased complexity of the nonlinearity, simply requires a larger number of basis functions in order to meet a required accuracy. This is for example exploited in (one-layer) neural network models, which are composed of a large number of simple basis functions. Parametrizations using a weighted sum of basis functions results in models which are parameter-affine, making possible parameter estimation by a convex optimization approach. A preferable choice of basis functions, are the bell-shaped smooth *normalized Gaussian*, or the *B-spline* basis functions. For a reference to the theory of neural network models, see *e.g.* the textbook [67] by Nelles, and for a reference on B-splines, see the textbook [19, Chapter 7] by Cheney. Another application of this type of smooth parameter-affine models is proposed by Johansen (see *e.g.* [38]), where normalized Gaussian basis functions are utilized for the construction of smooth Lyapunov functions for performance and stability analysis of nonlinear systems, in general.

1.3 Literature review — output-feedback tracking control

1.3.1 Introduction

For linear systems, tools for analysis and control are well developed, and the problem of output-feedback tracking of single-input single-output (SISO) systems may

be solved by various techniques; either a *state-variable approach* using *static (memoryless) state-feedback* combined with an *observer* to reconstruct the unmeasured states, or a direct *polynomial approach* where a *dynamic output-feedback* control law is designed based on an input-output representation of the system. A detailed treatment of linear system theory can be found in graduate level texts, like *e.g.* [43], [81] or [17]. In the sections below, we briefly outline the basics of these two approaches for output-feedback tracking control of SISO linear systems, mainly in continuous-time.

For linear systems, the so-called *separation principle* holds, which means that the output-feedback control problem can be separated into the design of a state-feedback controller and observer, independently. For linear time-invariant systems, the stability of the resulting closed-loop system is given by the union of the state-feedback and observer poles. This indirect approach is referred to as an *observer-based output-feedback* design, and is based on a model of the system in state-space form. First, a linear state-feedback controller may be designed by various techniques, such as pole placement control or linear quadratic (LQ) optimal control, assuming that all states are measured. Next, provided that the system is observable, the design of an observer is solved by using a copy of the system and adding linear output correction terms, providing exponentially converging estimates of the unmeasured states (with arbitrary fast rate of convergence). This type of observer for a linear deterministic[2] system was introduced by Luenberger [60], and is therefore often referred to as a Luenberger observer. The output-feedback problem is then solved by replacing the unmeasured states of the state-feedback controller with the estimates provided by the observer. A description of full-order and reduced-order observer design for linear systems is given *e.g.* in [43, Ch. 4], [81, Ch. 15], [17, Ch. 8] or [57, Ch. 37].

The alternative direct approach of designing a *dynamic output-feedback* controller for a linear system is solved by first expressing the system in an appropriate input-output form, *e.g.*, as a transfer function in polynomial form. A dynamic output-feedback controller may then be designed by *e.g. pole placement control* where a dynamic feedback is designed such that the resulting characteristic equation of the closed-loop system coincides with a desired characteristic polynomial. Alternatively, a *model reference control* approach may be pursued, where the dynamic feedback is designed so that the closed-loop system matches a desired reference model. A plain summary of the design and analysis of pole placement and model reference control can be found in [33, Ch. 7.3 and Ch. 6.3], and in more detail in [17, Ch. 9].

For nonlinear systems, tools for control and analysis are still at an evolutionary stage, therefore, nonlinear control has not yet become a common engineering tool. For this reason, and because tools for control and analysis of linear systems are constructive and powerful, the most common approach is to apply linear control techniques based on a locally linearized model of the nonlinear system. This approach is usually best suited for systems which can be characterized as "weakly" nonlinear. When nonlinearities in the system become significant, a common ap-

[2] In a stochastic setting with normal noise distributions, the well known *Kalman-Bucy filter* provides the optimal solution to the state estimation problem.

proach in order to extend the applicability and improve the performance of linear control, is by applying so-called *gain scheduling* techniques. Gain scheduling refers to the concept of interpolating (or scheduling the parameters of) a family of linear controllers over a set of operating points in the state-space. In this way, a time-varying linear controller may be designed using the wide range of design tools that are available for linear systems. For a review of gain scheduling techniques, see *e.g.* the survey paper by Rugh & Shamma (2000) [82].

If the nonlinearities in the system are essential, a control design based on locally linearized models may lead to poor performance, or even instability of the closed-loop controlled nonlinear system. In these cases, a pure nonlinear control design is likely to provide superior performance and robustness compared to linear methods (even when using gain scheduling). In nonlinear control theory, the problem of output-feedback tracking is considered one of the most challenging. First of all, the *separation principle* holds only for a limited class of nonlinear systems[3]. That is, for nonlinear systems in general, the design of a state-feedback controller where the states are replaced with asymptotically convergent estimates obtained from an observer, does not imply stability of the combined output-feedback solution. In order to guarantee stability of an observer-based output-feedback design, the state-feedback controller must be robust (*i.e.*, input-to-state stable) with respect to state estimation errors. Another challenge with output-feedback control of nonlinear systems, is the design of an asymptotically convergent observer, which is solved only for restricted classes of nonlinear systems. Therefore, constructive output-feedback designs for nonlinear systems are available only for restricted classes of systems. A detailed account of the most significant results within the field of output-feedback control of nonlinear systems can be found in the thesis by Maggiore (2000) [61]. Furthermore, an overview of constructive nonlinear control, including results in output-feedback control, is reviewed in the historical survey by Kokotović & Arcak (2001) [51].

In the remaining of this literature review, we first attempt to give an overview of some of the main approaches applied to control of electro-pneumatic actuators. Next, we review the most interesting constructive techniques that are currently available for nonlinear control design, restricting ourselves to results which are applicable to observer-based output-feedback control of electro-pneumatic actuators.

1.3.2 Control of electro-pneumatic actuators

A large amount of research has been published on the problem of position tracking control of electro-pneumatic actuators. In order to avoid an exhaustive presentation, we attempt to outline only the main approaches which have been considered in the literature, implying that some research will not be mentioned in this review. Too complete the overview on existing work, see also the literature reviews provided in the thesis works [97] (1995) by Virvalo, and [102] (2001) by Xiang, on the control of electro-pneumatic actuators.

[3] Assuming that the *separation principle* holds is often a practical solution to many design problems even when it can not be rigorously established.

Due to the inherent low stiffness and low damping of pneumatic actuators, the conventional proportional plus differential (PD) output-feedback control achieves unsatisfactorily poor (read: slow) tracking performance for most applications, thus, it not discussed here. Other research which (perhaps unfairly) are omitted in this review, are applications of model predictive control (MPC), and neural network (NN), and Fuzzy control. Main reasons for this are that MPC is computationally highly demanding since it requires the solution of a nonlinear optimization problem at each sample, and that the main advantage of NN and Fuzzy control, which is to adapt empirically to uncertain systems, are not needed since pneumatic systems are well understood, and accurate mechanistic models exist.

Hence, with a reservation to unfair omission of research, the main strategies applied to tracking control of electro-pneumatic actuators may roughly be grouped into the three categories:

Linear control Designs based on linearized models of the pneumatic actuator, which allow the application of linear control techniques.

Feedback linearizing control A nonlinear control design performed in two steps: First, a change of coordinates is found such that all nonlinearities satisfy the *matching condition* (that is, they appear in the same equation as the control input). Second, a control is designed which cancels all nonlinearities and makes the system linear for a redefined control input such that linear control techniques can be applied to the feedback linearized system.

Sliding mode (or variable structure) control A robust nonlinear design methodology where the nonlinear control problem is partitioned into two design phases: the design of a *sliding manifold* which defines some ideal motion of the system, and the design of a control which forces convergence to this manifold (referred to as *sliding mode*), where switching terms in the controller compensates for model uncertainties (or imprecisions) satisfying the matching condition.

Linear control

A common approach is to describe the dynamics of the actuator by a simplified 4th-order nonlinear model with the actuator position, the actuator velocity, and the pressures of the two chambers as dynamic states, *i.e.*, $\mathbf{x} = [y, v, p_A, p_B]^T$. The valve dynamics is then assumed to be negligibly fast, and the friction is assumed to be a static function of the velocity. Based on this nonlinear dynamics, a locally linearized model is obtained by linearizing about the equilibrium point $\mathbf{x}^* = [y^*, 0, p_A^*, p_B^*]^T$, with y^* usually taken as the mid-stroke position of the actuator, and where p_A^* and p_B^* are equilibrium pressures corresponding to the chosen operating point y^*, *i.e.*, $p_A^* = p_A^*(y^*)$, $p_B^* = p_B^*(y^*)$. The resulting 4th-order linear model is strictly minimum-phase with relative degree three, *i.e.*, it has one strictly stable zero. A detailed description of this model in state-space form can be found *e.g.* in the papers [83], [58], [20], [11]. For control design, the model is usually further simplified by averaging the time constant of the two pressure states such that the system can be

expressed in the 3rd-order normal form

$$\frac{d}{dt}\begin{bmatrix} y \\ \dot{y} \\ \ddot{y} \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -\omega_n^2 & -2\zeta_n\omega_n \end{bmatrix}\begin{bmatrix} y \\ \dot{y} \\ \ddot{y} \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ K\omega_n^2 \end{bmatrix} u, \tag{1.3}$$

where y is the position as a deviation variable from the equilibrium position y^*, the variable $\dot{y}$ is the velocity, and the two pressures are replaced by the acceleration variable $\ddot{y}$, that is, the new state vector is $\mathbf{x} \triangleq [y, \dot{y}, \ddot{y}]^T$. The parameter K is the steady-state gain, ω_n is the resonance frequency (also called the natural frequency), and ζ_n is the damping coefficient. The dynamics (1.3) can alternatively be expressed as a transfer function from input u to the position y, *i.e.*, in the input-output form

$$G(s) \triangleq \frac{y(s)}{u(s)} = \frac{K\omega_n^2}{s(s^2 + 2\zeta_n\omega_n s + \omega_n^2)}. \tag{1.4}$$

Remark 1 *In the reduced-order model (1.3), the state which was rendered unobservable by the transformation to the normal form, was simply cancelled due to averaging the two pressure states. This zero-pole cancellation can be allowed because the 4th-order system is strictly minimum-phase, thus, the cancelled unobservable state is strictly stable.*

The 3rd-order linear model, given by (1.3) or (1.4), has been used by several researchers as basis for linear control designs. For a detailed description of the model, including expressions for K, ζ_n, and ω_n, see *e.g.* the references [20], [96], [97], [80], [13]. Characteristic properties of pneumatic actuators which makes control difficult (compared to hydraulic actuators), are that the damping coefficient ζ_n typically is very low due to low viscous friction, and that the resonance frequency ω_n is low due to the high compressibility of air, and of course that both ζ_n and ω_n varies significantly as functions of the position y.

Based on the above 3rd-order linear model with states $\mathbf{x} \triangleq [y, \dot{y}, \ddot{y}]^T$, the conventional linear state-feedback tracking controller is given by

$$u = -K_p(y - y_r) - K_v\dot{y} - K_a\ddot{y}, \tag{1.5}$$

where y_r is the desired reference position, and K_p, K_v, K_a, are the controller feedback gains for the position, the velocity, and the acceleration, respectively. When the parameters K, ζ_n and ω_n of the model is known, the feedback gains can be determined by various linear control techniques, such as pole placement, steady-state linear quadratic (LQ) optimal control, or common performance criterion like the ITAE (integral of the time-weighted absolute error) criterion, *etc.* In most cases, additional fine-tuning of feedback gains may be required in order to achieve satisfactory performance for a given application.

The application of linear control techniques for the control of heavy pneumatics using commercial components has been experimentally studied by Virvalo. In his thesis [97] (1995), Virvalo implements the linear control law (1.5) for tracking control, and investigates a wide range of applicable linear techniques for the determination of the feedback gains K_p, K_v and K_a utilizing the 3rd-order linear model

(1.3). Virvalo also considers gain scheduling of the feedback gains K_p, K_v, and K_a as functions of the position y in order to improve performance over the entire operating range of the system. The application of the control law (1.5) with gain scheduled feedback gains, has also been investigated experimentally by several other researchers, see *e.g.* the work by Thomasset *et al.* [96], Richard & Scavarda [80], and Brun *et al.* [13].

In order to obtain an output-feedback solution, the velocity $\dot{y}$ and the acceleration $\ddot{y}$ must be estimated for use by the control law (1.5). Estimates of $\dot{y}$ and $\ddot{y}$, which are the 1st and 2nd-order derivative of the measured position y, may be obtained by (filtered) numerical differentiation of y, or by the use of an observer. Numerical differentiation is the most common approach for the estimation of $\dot{y}$ and $\ddot{y}$. A major drawback with this approach is that it is vulnerable to measurement noise, particularly in the 2nd-order derivative. Furthermore, by introducing filtration of the numerical derivatives, one introduces a time-delay in the estimates which may significantly degrade the performance of the closed-loop system. The alternative to numerical differentiation is to design an observer. The full-order Luenberger observer for the system (1.3) is given as

$$\frac{d}{dt}\begin{bmatrix}\hat{y}\\ \widehat{\dot{y}}\\ \widehat{\ddot{y}}\end{bmatrix} = \begin{bmatrix}k_1\\ k_2\\ k_3\end{bmatrix}(y-\hat{y}) + \begin{bmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 0 & -\omega_n^2 & -2\zeta_n\omega_n\end{bmatrix}\begin{bmatrix}\hat{y}\\ \widehat{\dot{y}}\\ \widehat{\ddot{y}}\end{bmatrix} + \begin{bmatrix}0\\ 0\\ K\omega_n^2\end{bmatrix}u, \tag{1.6}$$

where $[\hat{y}, \widehat{\dot{y}}, \widehat{\ddot{y}}]^T$ are the estimates of $[y, \dot{y}, \ddot{y}]^T$, and $[k_1, k_2, k_3]^T$ are the observer gains of the output error injection term $(y - \hat{y})$ which may be determined by pole placement of the resulting closed-loop observer error dynamics. Alternatively, a reduced-order observer may be designed, which estimates only the unmeasured states $\dot{y}$ and $\ddot{y}$.

For estimation of $\dot{y}$ and $\ddot{y}$, Virvalo [97] investigates experimentally both the use of numerical differentiation, the reduced-order and the full-order Luenberger observer (1.6). The application of the reduced-order Luenberger observer for the system (1.3) has also been investigated by Hong & Yongxian *e.g.* in [30].

The application of dynamic output-feedback control based on input-output models in polynomial form (which does not require an observer to recover the unmeasured states), has been investigated by some researchers. The results found in the literature mainly consider discrete-time designs, which are convenient for digital implementation. The continuous-time transfer function (1.4) is then discretized, and expressed in the polynomial form

$$H(z) = \frac{B(z)}{A(z)} \triangleq \frac{y(z)}{u(z)} = \frac{b_2z^2 + b_1z + b_0}{(z-1)(z^2 + d_1z + d_0)} = \frac{b_2z^2 + b_1z + b_0}{z^3 + a_2z^2 + a_1z + a_0}. \tag{1.7}$$

A dynamic output-feedback controller for the discrete system (1.7) is given in polynomial form as

$$R(z)\,u = -S(z)\,y + T(z)\,r. \tag{1.8}$$

where R, S, and T are polynomials of z. Designing a minimum-degree pole placement controller with no zero cancellation for (1.7), the polynomials must be $R(z) =$

$z^2 + r_1 z + r_0$, $S(z) = s_2 z^2 + s_1 z + s_0$ and $T(z) = t_2 z^2 + t_1 z + t_0$. The polynomial $T(z)$ is a design factor. A good choice is to take $T(z) = t_2 z^2 = t_2 A_0(z)$, where t_2 is chosen to provide a unity steady-state gain, while $A_0(z) = z^2$ is pole dynamics which is cancelled by the controller in the resulting closed-loop system. By defining a desired closed-loop pole polynomial $A_d(z)$, the coefficients of $R(z)$ and $S(z)$ are found by solving the so-called *Diophantine equation*[4]

$$A(z) R(z) + B(z) S(z) = A_0(z) A_d(z). \tag{1.9}$$

The discrete-time implementation of the dynamic controller (1.8) is thus given as

$$u_k = -r_0 \cdot u_{k-2} - r_1 \cdot u_{k-1} - s_0 \cdot y_{k-2} - s_1 \cdot y_{k-1} - s_2 \cdot y_k + t_2 \cdot r_k \tag{1.10}$$

where subscript k denotes the sample at time $t_k \triangleq k \cdot \Delta T$ (where ΔT is the sample time). Experimental implementation of (1.8) is among others investigated by Virvalo, see [97]. In [87], Shih & Huang applies the dynamic output-feedback pole placement controller in combination with on-line estimation of the model parameters (*i.e.*, the coefficients of $A(z)$ and $B(z)$), using a recursive least squares identifier.

Feedback linearizating control

The first application of a pure nonlinear control scheme for position tracking control of electro-pneumatic actuators, was the use of *input-output feedback linearization* by Richard & Scavarda (1989) [79]. This work has been further refined and validated by several members of Scavarda's research group, see *e.g.* the papers [96], [80], [13]. As their basis for design, the dynamics of the actuator was described by the simplified 4th-order model (or variants with minor differences)

$$\begin{aligned} \dot{y} &= v \\ \dot{v} &= \frac{A_A}{M} p_A - \frac{A_B}{M} p_B - \frac{1}{M} f_f(v) \\ \dot{p}_A &= -A_A \frac{1}{V_A(y)} v p_A + R T_0 \frac{1}{V_A(y)} w_v(p_A, u) \\ \dot{p}_B &= A_B \frac{1}{V_B(y)} v p_B + R T_0 \frac{1}{V_B(y)} w_v(p_B, -u). \end{aligned} \tag{1.11}$$

where the valve dynamics was assumed to be negligibly fast. In order to make possible input-output linearization, the friction $f_f(v)$ was required to be a differentiable function of the velocity, *e.g.* $f_f(v) = Dv$. Furthermore, to facilitate the computation of the inverse of the valve flow $w_v(p, u)$ with respect to u, the model of the valve flow was confined to the (piecewise input nonlinearity-affine) form

$$w_v(p, u) = g_o(p, \operatorname{sgn} \phi(u)) \cdot \phi(u) + g_l(p), \tag{1.12}$$

[4]The Diophantine equation may be formulated as a linear equation in the form $\mathbf{A}\mathbf{x} = \mathbf{b}$, where $\mathbf{x}$ contains the unknown coefficients of $R(z)$ and $S(z)$ which is straightforwardly computed by inverting the *Sylvester matrix* $\mathbf{A}$.

where $g_l(p)$ is a leakage term, and $\phi(u)$ is an invertible input nonlinearity. With the redefined control input $\bar{u} \triangleq \phi(u)$, the system (1.11) is in the piecewise input-affine form

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}) + \mathbf{g}(\mathbf{x}, \operatorname{sgn}\bar{u}) \cdot \bar{u} \\ y &= h(\mathbf{x}), \end{aligned} \tag{1.13}$$

with the state vector $\mathbf{x} \triangleq [y, v, p_A, p_B]^T$, and measured output y. Provided that the friction term $f_f(v)$ is one time differentiable, the system (1.13) is input-output linearizable, with a well defined relative degree $\rho = 3$, for all physically realizable pressures (*i.e.*, $p_A, p_B > 0$). Hence, the model is in a form which is applicable for a straightforward use of input-output feedback linearization, following the approach described *e.g.* in the textbooks by Slotine & Li [89] or Isidori [34].

The system is thus transformable by a change of coordinates $\mathbf{z} = \boldsymbol{\phi}(\mathbf{x})$ and $\zeta = \mu(\mathbf{x})$ to the *normal form*

$$\begin{aligned} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= z_3 \\ \dot{z}_3 &= a(\mathbf{x}) + b(\mathbf{x}, \operatorname{sgn}\bar{u}) \cdot \bar{u} \\ \dot{\zeta} &= q(\mathbf{z}, \zeta) \\ y &= z_1 \end{aligned} \tag{1.14}$$

where the states $\mathbf{z} = [z_1, z_2, z_3]^T \triangleq [y, \dot{y}, \ddot{y}]^T$ are the measured output and its derivatives, and ζ the state of the unobservable internal dynamics. A linearizing state feedback is then

$$\bar{u} = \frac{1}{b(\mathbf{x}, \operatorname{sgn}\bar{u})} \left[-a(\mathbf{x}) + \alpha \right], \tag{1.15}$$

which makes the $\mathbf{z}$-subsystem a triple integrator for the new control term α. Given the tracking reference trajectory $\mathbf{z}_r = [y_r, \dot{y}_r, \ddot{y}_r]^T$, the poles of the dynamics of the closed-loop tracking error $\tilde{\mathbf{z}} \triangleq \mathbf{z} - \mathbf{z}_r = [\widetilde{y}, \dot{\widetilde{y}}, \ddot{\widetilde{y}}]^T$ can be arbitrary placed with the feedback

$$\alpha(\tilde{\mathbf{z}}) = -K_y \widetilde{y} - K_v \dot{\widetilde{y}} - K_a \ddot{\widetilde{y}}. \tag{1.16}$$

Furthermore, it can be established that the *zero dynamics* $\dot{\zeta} = q(\mathbf{z}_r, \zeta)$ is stable for all bounded reference trajectories $\mathbf{z}_r$, which is required for the feedback linearizing control law (1.15)–(1.16) to be feasible.

Remark 2 *In the work [79] [96], [80], and [13], the state of the zero dynamics is taken to be $\zeta = p_B$. This means that the control $\bar{u}$ appears as a variable in the internal dynamics,* i.e., *$\dot{\zeta} = q(\mathbf{z}, \zeta, \bar{u})$. Consequently, the system in new coordinates $\mathbf{z}$ and ζ, is not in the proper normal form (1.14) where ζ is carefully chosen so that its governing dynamics becomes independent of the input $\bar{u}$. Hence, in order to guarantee stability, one must prove that the dynamics governing $\zeta = p_B$, is stable for all possible inputs $\bar{u}$ and reference trajectories $\mathbf{z} = \mathbf{z}_r$. The authors proves this only for the special case when the reference is a set-point,* i.e., *$\mathbf{z}_r = [y_r, 0, 0]^T$.*

The feedback linearizing control (1.15)–(1.16) yields superior tracking performance compared to linear control approaches, mainly because the nonlinearities in the system are explicitly compensated for. However, like model-based linear control approaches, careful modeling and accurate tuning of model parameters is required. A possible drawback with the feedback linearizing control law, is that it relies on cancellation of the nonlinearities $a(\mathbf{z}, \zeta)$ and $b(\mathbf{z}, \zeta)$. For nonlinear systems in general, such a cancelling control law is not robust. Additionally, a cancelling control law is not optimal in the sense that it wastes control effort by failing to recognize stabilizing nonlinearities, *i.e.*, nonlinearities which contribute to push the states towards the reference trajectory. Another possible drawback with the input-output feedback linearization approach, is the necessity to transform the system to the normal form, which imposes constraints on the model used for design; to allow a continuos state transformation to the normal form, the model nonlinearities must be differentiable. An alternative feedback linearizing approach for control of electro-pneumatic actuators is proposed by Xiang in his thesis work (2001) [102], which is referred to as block-oriented approximate feedback linearization. Here, Xiang considers approximate cancellation, allowing the use of a design model with nonsmooth nonlinearities, like *e.g.* discontinuous friction. Other work which should be mentioned, is an input-state feedback linearization approach considered by Kimura *et al* in [49].

No results have been found in the literature which consider output-feedback control using a feedback linearizing control law in combination with an observer for the estimation of the unmeasured states. In most published papers which address control by a feedback linearization approach, partial state-feedback is pursued, where both pressures and the position are measured, while a velocity estimate (and possibly an acceleration estimate) is obtained by filtered numerical differentiation. Hence, output-feedback control by pursuing a feedback linearization approach is still an open problem, consisting of two parts: First, the design a nonlinear observer which provides asymptotically convergent estimates of all the unmeasured states, and second, to establish under which conditions the feedback linearizing controller with observer is stable with respect to initial state estimation errors. The design of a nonlinear observer for electro-pneumatic actuators is discussed in the last section on sliding mode control below.

Sliding mode control

Sliding mode control is perhaps the most popular and successful nonlinear control technique which has been applied to tracking control of electro-pneumatic actuators. The design of a sliding mode controller is a general methodology which is performed in two steps: First, a *sliding manifold* is constructed for a subsystem not containing the actual control input. The design is simply a reduced-order control problem for the considered subsystem, where one of the states is assumed to be the control input. Next, the actual control law is designed so that the system states are forced to converge to the sliding manifold in finite time and remain there for all future time. This is a scalar control problem which is achieved by introducing switching terms

in the controller, where convergence can be made robust to bounded disturbances satisfying the matching condition. The dynamics of a sliding mode controlled system is usually characterized by two phases; the convergence to the sliding manifold which is referred to as the *reaching phase*, and the motion on the manifold once reached, which is referred to as the *sliding phase*. In the sliding phase — when the system is maintained on the sliding manifold — the system is said to be in *sliding mode*, characterized by infinitely fast switching of the control input.

Successful applications of sliding mode control for tracking control of electro-pneumatic actuators have been proposed by several researchers. A constructive approach to the design of a sliding mode controller for systems transformable to the normal form, is described in the textbook by Slotine & Li [89]. This approach is described below to illustrate sliding mode tracking control of electro-pneumatic actuators, which is the approach also applied by Bouri *et al.* [12] using the 4th-order nonlinear model (1.11) transformed to the normal form (1.14). Sliding mode designs based directly on 4th-order model (1.14) is, among others, proposed by Acarman & Hatipoğlu [1] and Drakunov *et al.* [21]. Pandian *et al.* [73], [74], [75], propose a similar approach based on a simplified 3rd-order design model where the two pressure states are replaced with a single differential pressure state.

In this section, we illustrate the sliding mode control approach, based on the 4th-order model (1.11) transformed to the normal form (1.14), as described in [12]. Recalling that the tracking error is defined as $\tilde{\mathbf{z}} \triangleq \mathbf{z} - \mathbf{z}_r = [\widetilde{y}, \dot{\widetilde{y}}, \ddot{\widetilde{y}}]^T$, the tracking error dynamics of the open-loop system (1.14) is given as

$$\begin{aligned} \dot{\tilde{z}}_1 &= \tilde{z}_2 \\ \dot{\tilde{z}}_2 &= \tilde{z}_3 \\ \dot{\tilde{z}}_3 &= a(\mathbf{x}) + b(\mathbf{x}, \operatorname{sgn}\bar{u})\,\bar{u} - \dddot{y}_r + \delta(t) \\ \dot{\zeta} &= q(\mathbf{z}, \zeta), \end{aligned} \tag{1.17}$$

where the additional term $\delta(t)$ may represent any bounded uncertainty such as modeling imprecisions or bounded disturbances, and the zero dynamics $\dot{\zeta} = q(\mathbf{z}_r, \zeta)$ is stable for all possible $\mathbf{z}_r$. The design of an appropriate sliding manifold is straightforward considering the subsystem

$$\begin{aligned} \dot{\tilde{z}}_1 &= \tilde{z}_2 \\ \dot{\tilde{z}}_2 &= \tilde{z}_3. \end{aligned}$$

Assuming $\tilde{z}_3$ to be the control input, a pole placement controller with critically damped poles placed at $-\lambda$ is given by

$$\tilde{z}_3 = -2\lambda\tilde{z}_2 - \lambda^2\tilde{z}_1.$$

The sliding manifold is thus defined as

$$s(\tilde{\mathbf{z}}) = \tilde{z}_3 + 2\lambda\tilde{z}_2 + \lambda^2\tilde{z}_1 = 0 \tag{1.18}$$

such that on the sliding manifold, *i.e.*, $s(\tilde{\mathbf{z}}) = 0$, the tracking error dynamics is described by $s(\tilde{\mathbf{z}}) = \ddot{\widetilde{y}} + 2\lambda\dot{\widetilde{y}} + \lambda^2\widetilde{y} = 0$.

The design of a control which forces convergence to the sliding manifold $s(\tilde{\mathbf{z}}) = 0$ is a scalar control problem, where the control is taken in the form

$$\bar{u} = \bar{u}_{eq} - \bar{U}\,\mathrm{sgn}(s)\,. \tag{1.19}$$

Here, $\bar{u}_{eq}$ is referred to as the nominal *equivalent control* which is designed to cancel all known terms such that in the absence of uncertainty, taking $\bar{u} = \bar{u}_{eq}$ would give $\dot{s} = 0$ and sliding mode ($s \equiv 0$) would be maintained once it was reached. The switching term $-\bar{U}\,\mathrm{sgn}(s)$ is designed to ensure robust convergence to sliding mode, even in the presence of a bounded uncertainty δ satisfying $|\delta| \leq D$. Hence, taking the equivalent control as

$$\bar{u}_{eq} = \frac{1}{b}\left(-a + \ddot{y}_r - 2\lambda\tilde{z}_3 - \lambda^2\tilde{z}_2\right), \tag{1.20}$$

and the magnitude of the switching term as

$$\bar{U} = \frac{D + c}{b},$$

the time-derivative of s is governed by

$$\begin{aligned}
\dot{s} &= a + b\bar{u} - \ddot{y}_r + \delta + 2\lambda\tilde{z}_3 + \lambda^2\tilde{z}_2 \\
&= \delta - b\bar{U}\,\mathrm{sgn}(s) \\
&= \delta - (D + c)\,\mathrm{sgn}(s) \\
&\Downarrow \\
\dot{s} &= \begin{cases} \leq -c & , \quad s > 0 \\ \geq c & , \quad s < 0 \end{cases}
\end{aligned}$$

which proves convergence to sliding mode ($s = 0$) in finite time. Since the switching term can be designed to guarantee convergence to the manifold for any bounded uncertainty δ with sufficiently high $\bar{U}$, the main feature of the continuous component u_{eq} is that it cancels all known terms in the dynamics of s, in order to reduce the magnitude $\bar{U}$ of the switching term required to ensure convergence to sliding mode, $s = 0$.

Notice the similarity between the equivalent control law (1.20) and the feedback linearizing control law (1.15). In order to reduce the magnitude $\bar{U}$ compared to the case with $u_{eq} = 0$, the continuous control part u_{eq} can be taken as any controller which approximately stabilizes $s = 0$. In the work [90] by Surgenor & Vaughan, the linear tracking controller (1.5) is adopted as the equivalent control, *i.e.*,

$$u_{eq} = -K_p\tilde{z}_1 - K_v\tilde{z}_2 - K_a\tilde{z}_3. \tag{1.21}$$

An example of a sliding mode design based on the linearized model (1.3) in the normal form is illustrated in the paper [96] by Thomasset *et al.*

A possible drawback with the sliding mode controller (1.19) is that it results in a discontinuous control law due to the switching term $\bar{U}\,\mathrm{sgn}(s)$. In applications with high friction, where the controller is implemented with sufficiently high sampling

rate, this may be advantageous as high-frequency switching in the control introduces a *dither* which eliminates or reduces stiction. In most applications, however, a discontinuous control law is undesirable as it introduces chattering in the control input which may wear out the valve. In these cases, the discontinuous sliding mode controller (1.19) is replaced with the continuous approximation

$$\bar{u} = \bar{u}_{eq} - \bar{U} \,\mathrm{sat}\left(\frac{s}{\Phi}\right), \tag{1.22}$$

where Φ is referred to as the boundary layer thickness. The modified continuous control (1.22) now only guarantees convergence to the boundary layer $|s| \leq \Phi$ in finite time. Inside the boundary layer $|s| \leq \Phi$, the controller is simply a local high-gain controller which will maintain tracking within a small neighborhood of sliding mode, $s = 0$.

The inherent strong robustness properties of sliding mode control, makes it particularly suited for observer-based control. Unlike the case of exact feedback linearization, establishing stability of the resulting output-feedback solution is in most cases constructive because the last design step is scalar. However, the extension of full state-feedback sliding mode control to the case of using output-feedback only, requires the design of an asymptotically convergent observer for the unmeasured states, which in general is not a trivial task.

For the full 4th-order system (1.11), the unmeasured states are $\dot{y}$, p_A, and p_B. With an asymptotically convergent observer available for these states, the problem of output-feedback tracking would be solved using either of the mentioned (partial) state-feedback sliding mode designs presented in [12], [1], or [21]. However, there are very few published papers that address the problem of nonlinear state estimation using the 4th-order system (1.11). One of the few constructive solutions to observer design for nonlinear systems is the so-called Extended Kalman filter (EKF), which is a straightforward extension of the renowned Kalman filter for stochastic systems[5]. In [18], Chen *et al.* attempt to design an asymptotic nonlinear observer for all the states of the nonlinear system (1.11) by application of the Extended Kalman filter, in discrete-time. The paper indicates that good experimental results are obtained. However, no evaluation of the observability properties of the system is provided, hence, the design is based on the indirect assumption that the Jacobian — the linearization along the estimated state trajectory — of (1.11) is observable from the output y under all conditions. Furthermore, even if the observability assumption is satisfied, it is not guaranteed that the Extended Kalman filter is indeed asymptotically stable under all conditions for the nonlinear system (1.11). This is one of the main drawbacks of the EKF, which, in general, only guarantees asymptotic stability locally, and no quantitative region of attraction is provided by design. Hence, asymptotic stability in a given region must be established separately. Another drawback

[5]The Kalman filter provides the optimal solution — the minimum of the mean square of the estimation error — to the state estimation problem for linear time-varying systems subjected to uncorrelated stochastic disturbances and measurement noise. For the fundamental theory on stochastic systems and filtering, see *e.g.* [36].

with the EKF, is that the implementation of the EKF is computationally demanding, mainly because the Ricatti equation is solved on-line, which also requires the continuous computation of the Jacobian.

An important aspect with respect to observer design for electro-pneumatic actuators, is that the observability properties of (1.11) with y being measured, has not been rigorously established. A common approach to determine the observability properties of the system has been to linearize (1.11) about some steady-state operating point in state space, which may lead to the conclusion that only one of the pressures, p_A and p_B, are observable simultaneously from y. This apparent lack of full observability of (1.11), which is false in the main operating range of the actuator, has possibly been a main reason for the slow progress in research on output-feedback control and nonlinear observer design for electro-pneumatic actuators using the 4th-order model (1.11). On the other hand, if an observer is designed on an assumption of full observability (like in the case of simply applying the Extended Kalman filter and not checking the observability conditions), one runs the risk that the output-feedback correction term in the observer may cause unobservable states to diverge and make the output-feedback solution unstable.

Remark 3 *That the nonlinear system (1.11) has at least three observable states, is proven in a nonlinear sense by the existence of a state transformation to the normal form (1.14) which is invertible. Furthermore, since the unobservable internal dynamics of (1.14) is asymptotically stable, the system (1.11) is at least* detectable, i.e.*, the possibly unobservable state is asymptotically stable. However, this insight on observability and detectability of the nonlinear system (1.14) which is provided by the transformation to the normal form, has not been commented in any paper found in the literature.*

One way to avoid the problem of observability, is by approximating the system by a 3rd-order model, which makes the system clearly observable from y (see Remark 3), which again greatly simplifies the observer design. Following this approach, with a 3rd-order model in the states $[y, \dot{y}, \Delta p]^T$ as in [73], [74], [75], Takemura *et al.* [92] propose a reduced-order Luenberger-type observer for estimation of the velocity $\dot{y}$ and the differential pressure Δp. The observer is combined with the sliding mode design proposed in [73] and [75] to obtain one of the few nonlinear output-feedback solutions proposed for electro-pneumatic actuators. The output-feedback solution is validated experimentally with reasonably good results.

1.3.3 Output-feedback control theory

In this section, the most interesting (in the author's view) constructive procedures which are currently available for nonlinear control are reviewed, constrained to results applicable for output-feedback control of electro-pneumatic actuators. A comprehensive review of existing constructive nonlinear control theory is provided by Kokotović and Arcak in their historical survey paper [51] (2001). Their review describes a close to exponential growth of results within the field of nonlinear state-feedback control during the two last decades, while the progress in output-feedback

control has been relatively slow. In the output-feedback case, the main challenges are the development of an asymptotically convergent *observer* for the unmeasured states, and the lack of a general *separation principle* for nonlinear systems. As a result, constructive procedures for output-feedback control are developed only for particular classes of nonlinear systems. A detailed review of the output-feedback control literature can be found in the thesis by Maggiore (2000) [61]. An output-feedback solution relies on the design of an observer to recover the unmeasured states, and various approaches to nonlinear observer design are reviewed *e.g.* in the theses by Johansson (2001) [39] and Rajamani (1995) [78].

Provided that some smoothness requirements are satisfied (basically on the friction and valve flow models), the full nonlinear model of the electro-pneumatic actuator is fully *feedback linearizable*. An important property of electro-pneumatic actuators, is that their full nonlinear model is in *pure-feedback form* in its natural coordinates. In the first section below, we first describe some results for *feedback linearizable* nonlinear systems, the most general class in the nonlinear control literature for which constructive procedures for nonlinear output-feedback tracking control are available. Next, we discuss nonlinear systems in *pure-feedback form*, which are applicable for nonlinear state-feedback control design by a *backstepping* approach, and for output-feedback control by an *observer-based backstepping* approach, provided that an asymptotically convergent observer is available. In the last section, we provide a review of a few selected approaches for nonlinear observer design, applicable for electro-pneumatic actuators.

In the review below, we limit our discussion to the single-input single-output (SISO) case and approaches relevant for electro-pneumatic actuators, and we roughly sketch the main ideas and refer explicitly to only a few selected references. For a more complete overview over existing approaches, see the literature reviews in [61], [39], and the survey [51].

Feedback linearizable systems

Most output-feedback solutions considered in the literature are based on high-gain observers to estimate all, or some of, the unmeasured states of the system. A constructive design of high-gain observers which robustly estimates the derivatives of the output, are available for a general class of nonlinear systems which are fully input-output linearizable. The main feature of the observer is that it uses high gain in its output-injection term to achieve an arbitrary large region of attraction with respect to vanishing pertubations, in addition to arbitrary fast rate of convergence, and arbitrary attenuation of the effect of bounded disturbances. See *e.g.* the tutorial paper by Khalil [48], and the references therein. A drawback with this high-gain observer, is that it estimates the states of the model transformed to the input-output form, *i.e.*, where the states are the output and its derivatives. In order to obtain estimates of the original states, an invertible observability mapping between the output-derivatives and the original states is required to be explicitly known. In his thesis [61], Maggiore removes this inconvenient and unpractical restriction by proposing a high-gain observer for feedback linearizable systems which operates in the original states, hence, which do not require a preceding transformation of the

system to an input-output form with the output and its derivatives as states.

One of the first general classes of nonlinear systems which was considered in the output-feedback control literature, is *feedback linearizable* systems of *minimum-phase*, which can be expressed in the *normal form*

$$\begin{aligned} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= z_3 \\ &\cdots \\ \dot{z}_{r-1} &= z_r \\ \dot{z}_r &= a(\mathbf{z}, \boldsymbol{\zeta}) + b(\mathbf{z}, \boldsymbol{\zeta}) \cdot u \\ \dot{\boldsymbol{\zeta}} &= q(\mathbf{z}, \boldsymbol{\zeta}), \end{aligned} \tag{1.23}$$

where $\mathbf{z} = [z_1, z_2, \cdots, z_r]^T \triangleq [y, \dot{y}, \cdots, y^{(r-1)}]^T \in \mathbb{R}^r$, with y being the output, u the input, and $\boldsymbol{\zeta} \in \mathbb{R}^{n-r}$ the states of the internal dynamics, while n is the order, and r the relative degree of the system. Since the state $\boldsymbol{\zeta}$ is carefully devised so that the control input u does not appear in the internal dynamics $\dot{\boldsymbol{\zeta}} = q(\mathbf{z}, \boldsymbol{\zeta})$, it can not be controlled, thus, the system (1.23) is required to be minimum-phase, or more precisely, the *zero dynamics* $\dot{\boldsymbol{\zeta}} = q(\mathbf{z}_r, \boldsymbol{\zeta})$ is required to be stable (or bounded) for all possible reference trajectories $\mathbf{z} = \mathbf{z}_r$. In the partial-state feedback case where $\boldsymbol{\zeta}$ is measured (in addition to y), a high-gain observer can be designed to estimate the $(r-1)$ derivatives of the output y, with regional (and even semi-global) asymptotic convergence. A partial state-feedback design then follows a separation approach: First, a full state-feedback controller is designed to be robust to errors in $\mathbf{z}$. Secondly, as $\mathbf{z}$ is not measured, the designed full-state controller is implemented using the estimates $\hat{\mathbf{z}}$ of $\mathbf{z}$ obtained by a high-gain observer. Since $\hat{\mathbf{z}}$ converges asymptotically to $\mathbf{z}$, the tracking performance of the full state-feedback controller is asymptotically recovered by partial state-feedback of $\boldsymbol{\zeta}$ and y only. A key element of the design, is to make the full state-feedback control globally bounded in $\hat{\mathbf{z}}$, in order to prevent the *peaking* (initial transients with large magnitudes) of the high-gain observer from destabilizing the resulting output-feedback solution.

Remark 4 *Provided the model of the electro-pneumatic actuator is sufficiently smooth, the normal form (1.23) (possibly non-affine in the input u), is obtained by differentiating the output y until the input u appears in the equation. Typically, a model of the electro-pneumatic actuator consists of the position y, the velocity v, and the two pressures p_A and p_B, and then, the dynamic order is $n = 4$, and the relative degree $r = 3$. Increasing the dynamic accuracy of the model, it could be further extended to include states of the valve dynamics, friction dynamics, and possibly also temperature dynamics. Then, the relative degree r would be increased by the order of the valve dynamics, while the other dynamics would enter in the internal dynamics,* i.e., *only increasing the order n, but not the relative degree r.*

The fact that $\boldsymbol{\zeta}$ is unobservable from y, represents a major obstacle for an output-feedback solution when $\boldsymbol{\zeta}$ is not measured. This means that $\boldsymbol{\zeta}$ is unavailable for

control, such that the controller design is limited to a partial state-feedback using $\mathbf{z}$ only. A high-gain observer is then used to robustly estimate the output derivatives of y. Since $\boldsymbol{\zeta}$ is not measured, it represents an unknown bounded disturbance in the observer such that only approximate asymptotic convergence is achieved, hence, the performance achieved with the underlying state-feedback controller is only approximately recovered in the output-feedback case. An example of output-feedback stabilization[6] by the above outlined approach, is proposed by Mahmoud and Khalil in the paper [62], which extends the results of [45] by Khalil, to also include systems with zero dynamics.

Remark 5 *In the above described approach, the state $\boldsymbol{\zeta}$ of the internal dynamics represents an unknown bounded disturbance, since it is not estimated. Hence, only approximate tracking is possible to achieve using a continuous high-gain observer and continuous control. It should be noted, however, that in the case when the internal dynamics is asymptotically stable, a copy of the system would provide asymptotically convergent estimates of $\boldsymbol{\zeta}$. Thus, an asymptotically convergent observer could be straightforwardly designed, despite of $\boldsymbol{\zeta}$ not being observable from y.*

The largest class of nonlinear systems, for which constructive procedures for nonlinear output-feedback tracking control are available, is fully feedback linearizable systems which can be expressed in the *perturbed input-output form*

$$y^{(n)} = \bar{a}\left(\mathbf{z}, \mathbf{u}, w\left(t\right)\right) + \bar{b}\left(\mathbf{z}, \mathbf{u}, w\left(t\right)\right) u^{(n-r)}, \tag{1.24}$$

where n is the order and r is the relative degree of the system. Here, the nonlinear functions $\bar{a}\left(\cdot\right)$ and $\bar{b}\left(\cdot\right)$ may depend on the output and its derivatives $\mathbf{z} \triangleq [y, \dot{y}, \cdots, y^{(n-1)}]^T \in \mathbb{R}^r$, and also on the input and its derivatives $\mathbf{u} \triangleq [u, \dot{u}, \cdots, u^{(n-r-1)}] \in \mathbb{R}^{n-r}$ when the relative degree is lower than the order of the system ($r < n$), and w is a bounded uncertainty which can represent modeling errors and time-varying disturbances. In the form (1.24), the state vector is the derivatives of the output y, hence, provided that the input u and its derivatives are known, a high-gain observer can be designed to provide asymptotically converging estimates of all the states for the nominal (unperturbed) system without uncertainties, *i.e.*, $w\left(t\right) = 0$. The high-gain observer can be designed with a region of attraction which can be made arbitrary large, by sufficiently large gain. Furthermore, in the presence of the bounded uncertainty $w\left(t\right)$, the high-gain observer recovers the output-derivatives with arbitrary accuracy, by sufficiently large gain. Hence, with an asymptotically stable observer available to robustly estimate all of the system states, the main obstacle to an output-feedback solution is removed. A new difficulty with this approach is that the full state-feedback controller must be designed to provide, not only the input u, but also its derivatives $\dot{u}, \cdots, u^{(n-r-1)}$ in order to be able to estimate all the states of (1.24) using a high-gain observer. To solve this problem, a series of integrators are augmented on the input side of the system (1.24), whose states

[6] The term *tracking* refers to the most general case where the reference is a time-varying trajectory, while the term *regulation*, or *stabilization*, refers to the special case when the reference is a fixed set-point.

are denoted $\mathbf{u} \triangleq [u_1, u_2, \cdots, u_{n-r}]^T = [u, \dot{u}, \cdots, u^{(n-r-1)}]$, such that the augmented system is given by the state-space model

$$\begin{aligned}
\dot{z}_1 &= z_2 \\
\dot{z}_2 &= z_3 \\
&\vdots \\
\dot{z}_n &= \bar{a}(\mathbf{z}, \mathbf{u}, w(t)) + \bar{b}(\mathbf{z}, \mathbf{u}, w(t))\,\nu \\
\dot{u}_1 &= u_2 \\
\dot{u}_2 &= u_3 \\
&\vdots \\
\dot{u}_{n-r} &= \nu.
\end{aligned} \tag{1.25}$$

Remark 6 *Provided that the electro-pneumatic actuator is modeled sufficiently smooth, the model can be expressed in the input-output form (1.25) (however, not necessarily affine in u^{n-r}) by successive differentiation of the output.*

Viewing $\nu \triangleq u^{(n-r)}$ as the control input of (1.25), this corresponds to using a dynamic controller of order $(n-r)$ to control $\mathbf{z}$. As opposed to the design based on the normal form (1.23), we no longer need to impose a partial state-feedback restriction on the controller design of ν: The state $\mathbf{z}$ contains the derivatives of the output y which can be robustly estimated using a high-gain observer, while $\mathbf{u}$ is the state of the dynamic compensator which is readily available for feedback upon integration of ν. Following this approach (in a reversed manner), Oh & Khalil [71] design a high-gain observer for the tracking error ($\tilde{\mathbf{z}} \triangleq \mathbf{z} - \mathbf{z}_r$), and then a globally bounded sliding mode controller in the observer coordinates which forces the estimated observer tracking error to zero. Since the convergence of the actual tracking error follows by the convergence of the high-gain observer, semi-global tracking is achieved which is robust to the bounded, time-varying uncertainty $w(t)$, *i.e.*, modeling errors and time-varying disturbances satisfying the matching condition. The input ν is discontinuous. However, there is no chattering in the actual control u since it is obtained by $(n-r)$ times integration of ν. For the augmented system (1.25) without the uncertainty w, Aloliwi & Khalil [3] propose (as an extension of Khalil's paper [46]) an adaptive output-feedback tracking controller, with unknown parameters $\boldsymbol{\theta}$ appearing in (1.24) with $\bar{a}(\cdot) = f_0(\cdot) + \mathbf{f}(\cdot)^T \boldsymbol{\theta}$ and $\bar{g} = g_0(\cdot) + \mathbf{g}_1(\cdot)^T \boldsymbol{\theta}$. Based on the augmented system (1.25), Mahmoud & Khalil proposed in [63] the full generalization of [62] to the case of robust output-feedback tracking control utilizing high-gain observers.

Systems in feedback form — integrator backstepping

For systems in *pure-feedback form* with only feedback paths in the dynamics, like

$$\begin{aligned}
\dot{q}_1 &= f_1(q_1, q_2) \\
\dot{q}_2 &= f_2(q_1, q_2, q_3) \\
\dot{q}_3 &= f_3(q_1, q_2, q_3, u),
\end{aligned} \tag{1.26}$$

a nonlinear controller may be designed recursively by an *integrator backstepping* approach. Backstepping is a flexible design procedure, where the nonlinear control design becomes constructive by breaking down the control problem into scalar problems at each step. A particular strength of backstepping, is that it can be used to effectively deal with *structured uncertainties* which do not satisfy a *matching condition* with the control input u (which is a common restriction in robust control designs). With *adaptive backstepping* we may design asymptotic tracking controllers in the case where the uncertainty is unknown constant parameters which appear in an affine form, *i.e.*, pure-feedback systems where the nonlinear functions can be factored as $f_i(\cdot) = f_{0i}(\cdot) + \boldsymbol{\phi}_i(\cdot)^T \boldsymbol{\theta}$, where $f_{0i}(\cdot)$ and $\boldsymbol{\phi}_i(\cdot) \in \mathbb{R}^p$ are known, and $\boldsymbol{\theta} \in \mathbb{R}^p$ is a vector of p uncertain parameters. Adaptive backstepping, and backstepping in general, is treated in the textbook by Krstić *et al.* [52]. With *robust backstepping* we may design robust tracking controllers in the presence of time-varying disturbances, bounded by known functions, *i.e.*, with nonlinearities in the form $f_i(\cdot) = f_{0i}(\cdot) + \mathbf{q}_i(\cdot)^T \mathbf{w}(t)$, where $\mathbf{w}(t)$ is a bounded uncertainty, while $f_{0i}(\cdot)$ and $\mathbf{q}_i(\cdot)$ are known functions satisfying the pure-feedback structure. Robust backstepping is also briefly treated in [52], and in the textbook by Freeman & Kokotović [26]. Robust and adaptive backstepping techniques can effectively be combined to deal with systems with a combination of uncertain parameters and time-varying disturbances, *i.e.*, with nonlinearities in the form $f_i(\cdot) = f_{0i}(\cdot) + \boldsymbol{\phi}_i(\cdot)^T \boldsymbol{\theta} + \mathbf{q}_i(\cdot)^T \mathbf{w}(t)$, as proposed by Freeman *et al.* [27].

One of the main purposes of backstepping is the construction of a Control Lyapunov Function (CLF) which is used for control design. Briefly stated, a CLF is a Lyapunov function candidate where a stabilizing control law exists which would render the CLF a Lyapunov function for the closed-loop system with a negative definite time-derivative. Some extensions of the backstepping procedure is based on modifications of the recursive construction of a CLF. By introducing a flattened CLF, Freeman & Kokotović [25] and [26, Section 5.3.2] removed an undesirable high-growth property of robust terms in the control law, resulting from a backstepping design using a quadratic CLF. Furthermore, Freeman & Praly [28] extended the backstepping procedure to systems with actuator magnitude and rate constraints. Another important extension of the backstepping procedure, with considerable practical importance because it guarantees stability margins, is the possibility to render the design *inverse optimal*. Inverse optimal and locally optimal backstepping design procedures are proposed by Ezal in his thesis [24]. See also [26, Section 5.3.2], Krstić & Li [53], or Pan *et al.* [72] and the references therein.

Remark 7 *The physical structure of electro-pneumatic actuators confines its model to the pure-feedback form in its physical coordinates. The backstepping techniques found in the control literature usually consider systems in the* strict-feedback form*, that is, a lower-triangular form where the nonlinearities are restricted to the structure* $f_i(q_1, \cdots, q_{i+1}) = a_i(q_1, \cdots, q_i) + b_i q_{i+1}$*,* i.e.*, affine in* q_{i+1}*. Provided that an implicit function restriction is imposed on the dependence of* $f_i(\cdot)$ *on* q_{i+1}*, most results for systems in the strict-feedback form carries over to systems in the more general pure-feedback form, however, the results are then no longer global, see [52, Section 4.5.3].*

The most general class of feedback systems where backstepping applies, comprising most models of electro-pneumatic actuators, is the pure-feedback form with internal dynamics and uncertainties bounded by known functions. This form can be expressed as

$$\begin{aligned}
\dot{q}_1 &= f_1(q_1, q_2, \boldsymbol{\zeta}) + \mathbf{g}_1(q_1, q_2, \boldsymbol{\zeta})^T \mathbf{w}(t) \\
\dot{q}_2 &= f_2(q_1, q_2, q_3, \boldsymbol{\zeta}) + \mathbf{g}_2(q_1, q_2, q_3, \boldsymbol{\zeta})^T \mathbf{w}(t) \\
&\vdots \\
\dot{q}_r &= f_r(q_1, \cdots, q_r, u, \boldsymbol{\zeta}) + \mathbf{g}_r(q_1, \cdots, q_r, u, \boldsymbol{\zeta})^T \mathbf{w}(t) \\
\dot{\boldsymbol{\zeta}} &= \mathbf{F}(\mathbf{q}, \boldsymbol{\zeta}, u) + \mathbf{G}(\mathbf{q}, \boldsymbol{\zeta}, u) \mathbf{w}(t)
\end{aligned} \tag{1.27}$$

where n is the order of the system, r is the relative degree, with the r first states contained in the vector $\mathbf{q} = [q_1, \cdots, q_r]^T$, the remaining $(n-r)$ states of the internal dynamics contained in $\boldsymbol{\zeta}$, and $\mathbf{w}(t) \in \mathbb{R}^p$ is a vector of p time-varying bounded disturbances. The nominal dynamics $f_i(\cdot)$ and the disturbance gains $\mathbf{g}_i(\cdot)$, $i = 1, \cdots, r$, the nominal dynamics $\mathbf{F}(\cdot) \in \mathbb{R}^{n-r}$, and the disturbance gain matrix $\mathbf{G}(\cdot) \in \mathbb{R}^{(n-r)\times p}$ of the internal dynamics, are known functions. The internal $\boldsymbol{\zeta}$-dynamics is written in a compact vectorized form for simplicity, however, it is required to be in a form which does not violate the pure-feedback property of the $\mathbf{q}$-dynamics.

An additional strength of the backstepping procedure is that, provided an asymptotic observer is available, a robust output-feedback design is constructive by an observer-based backstepping approach. Denoting the estimated states $\hat{\mathbf{q}} = [\hat{q}_1, \hat{q}_2, \cdots, \hat{q}_r]$ and $\hat{\boldsymbol{\zeta}}$, a backstepping design is performed on the dynamics of the $(q_1, \hat{q}_2, \cdots \hat{q}_r, \hat{\boldsymbol{\zeta}})$-system. The observer error $q_2 - \hat{q}_2$ then appears in the design as a time-varying disturbance, which is systematically counteracted by robust terms in the backstepping design — referred to as *nonlinear damping*. By observer backstepping, an output-feedback solution is derived, which is robust to observer errors, and where the performance of an equivalent full state-feedback backstepping design is asymptotically recovered as the observer error $q_2 - \hat{q}_2$ converges to zero. Hence, for systems in the pure-feedback form, the only obstacle to an output-feedback solution by a backstepping approach, is the design of an asymptotically convergent observer for the system, which in general is not a trivial task.

Nonlinear observer design

The problem of reconstructing states which are not measured directly, is referred to as an observer or state estimation problem in the control literature. A general theory on nonlinear observer design does not exist, and constructive designs are available only for particular classes of nonlinear systems. An overview of the nonlinear observer literature is provided in the theses by Maggiore (2000) [61] and Johanson (2001) [39]. Several approaches to nonlinear observer design are collected in the textbook (1999) [68], edited by Nijmeijer & Fossen. In this section, we review a few nonlinear observer designs which are interesting with respect to electro-pneumatic actuators.

Consider the nonlinear system

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{f}(\mathbf{x}, u) \\ y &= \mathbf{c}^T \mathbf{x}, \end{aligned} \tag{1.28}$$

where $\mathbf{x} \in \mathbb{R}^n$ is the state vector, and $u, y \in \mathbb{R}$ are the input and the measured output, respectively.

If the system dynamics and the input u is known, an *open-loop observer* — also referred to as a *ballistic observer* — for the states $\mathbf{x}$, is given by

$$\dot{\hat{\mathbf{x}}} = \mathbf{f}(\hat{\mathbf{x}}, u), \tag{1.29}$$

where $\hat{\mathbf{x}}$ is the estimated state. When the input u is known, and the system dynamics, given by $\mathbf{f}(\mathbf{x}, u)$, is asymptotically stable, the estimate $\hat{\mathbf{x}}$ of the open-loop observer (1.29) converges asymptotically towards the actual state $\mathbf{x}$. In other words, (1.29) is an asymptotic observer for (1.28) where the estimation error $\tilde{\mathbf{x}} \triangleq \mathbf{x} - \hat{\mathbf{x}}$ converges asymptotically to zero. Due to its simplicity, an open-loop observer could be a good choice for nonlinear systems, whose dynamics can be established to be asymptotically stable[7].

For certain classes of nonlinear systems which are uniformly observable and not required to be asymptotically stable, a Luenberger-type observer with linear output-injection may provide asymptotic estimates of the state $\mathbf{x}$ of (1.28). This Luenberger-type observer is given by

$$\begin{aligned} \dot{\hat{\mathbf{x}}} &= \mathbf{f}(\hat{\mathbf{x}}, u) + \mathbf{k} \cdot (y - \hat{y}) \\ \hat{y} &= \mathbf{c}^T \hat{\mathbf{x}}, \end{aligned} \tag{1.30}$$

where $\mathbf{k} \in \mathbb{R}^n$ is a vector of output-injection feedback gains which one attempt to choose to make the resulting observer error asymptotically stable. Rajamani [77] (1993) develops necessary and sufficient conditions under which the Luenberger-type observer (1.30) is asymptotically stable for nonlinear systems whose dynamics can be written in the form

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \boldsymbol{\phi}(\mathbf{x}, u) \\ y &= \mathbf{c}^T \mathbf{x}, \end{aligned} \tag{1.31}$$

where $\left(\mathbf{A}, \mathbf{c}^T\right)$ is an observable pair, and where $\boldsymbol{\phi}$ is a Lipschitz nonlinearity in $\mathbf{x}$, satisfying

$$|\boldsymbol{\phi}(\mathbf{x}, u) - \boldsymbol{\phi}(\hat{\mathbf{x}}, u)| \leq L |\mathbf{x} - \hat{\mathbf{x}}| \tag{1.32}$$

for some positive constant L. Rajamani shows that the stability of the observer error dynamics depends both on the eigenvalues and eigenvectors of $\left(\mathbf{A} - \mathbf{k}\mathbf{c}^T\right)$. Briefly stated, the observer gain vector $\mathbf{k}$ must be chosen such that the asymptotic stability of the linear part of the observer error dynamics, given by $\left(\mathbf{A} - \mathbf{k}\mathbf{c}^T\right)$, dominates the destabilizing effect of $\boldsymbol{\phi}(\mathbf{x}, u) - \boldsymbol{\phi}(\hat{\mathbf{x}}, u)$, which appears as a vanishing perturbation in the observer error dynamics.

[7] Recently, a general approach to the construction of a smooth Lyapunov function for nonlinear systems, was proposed by Johansen, see *e.g.* [38]. This approach can be used to establish an estimate of the region of attraction, in which exponential stability can be guaranteed.

Utilizing the sliding mode observer

$$\begin{aligned} \dot{\hat{\mathbf{x}}} &= \mathbf{f}\left(\hat{\mathbf{x}}, u\right) + \mathbf{k} \cdot \left(y - \hat{y}\right) + \mathbf{l} \cdot \operatorname{sgn}\left(y - \hat{y}\right) \\ \hat{y} &= \mathbf{c}^T \hat{\mathbf{x}}, \end{aligned} \tag{1.33}$$

proposed by Slotine *et al.* [88] (1987), the class of systems considered by [77] is extended to include systems with nonlinearities satisfying

$$\left|\boldsymbol{\phi}\left(\mathbf{x}, u\right) - \boldsymbol{\phi}\left(\hat{\mathbf{x}}, u\right)\right| \leq L_0 + L_1 \left|\mathbf{x} - \hat{\mathbf{x}}\right|, \tag{1.34}$$

for some positive constants, L_0 and L_1. The sliding mode observer [88] has strengthened convergence and disturbance attenuation properties compared to (1.30), and achieves asymptotic convergence even in the presence of bounded non-vanishing disturbances.

1.4 Thesis outline

The main contributions, and organization of the thesis, are outlined below.

1.4.1 Contributions

The main contributions from this work are within the fields:

- Modeling for nonlinear and adaptive control
- Nonlinear observer design
- Nonlinear output-feedback tracking control
- Experimental implementation and validation.

The first part of the thesis deals with the mathematical modeling of the electro-pneumatic clutch actuation system for nonlinear and adaptive output-feedback control. Main objectives have been the development of improved models of particular model parts, *i.e.*, friction and flow rate models, with respect to accuracy, and properties which are advantageous for control, *i.e.*, parameter-affinity and differentiability. Modeling of nonlinearities in parameter-affine form allows parameter identification to be solved as a convex optimization problem, and is required for Lyapunov-based adaptive control design. Differentiability of system nonlinearities enables exact backstepping, and is also required for the system to be fully feedback linearizable such that it can be expressed in the input-output form to which existing solutions for output-feedback control using high-gain observers apply.

This work provides a unified treatment of the modeling of electro-pneumatic actuators in general, and introduces some modifications which make the resulting design model applicable for solutions available for nonlinear and adaptive output-feedback control. In particular, a smooth dynamic friction model, a generalized flow rate equation, and valve flow models with improved accuracy, are introduced.

The modeling is summarized in a 6th-order dynamic model of the electro-pneumatic clutch actuator. This model is fully feedback linearizable with relative degree four, and in pure-feedback form, which makes it applicable for backstepping. The most important nonlinearities of the model, *i.e.*, the valve flow and clutch load model, are linearly parametrizable, hence, suited for online identification by an adaptive control design in the case of full state-feedback.

Utilizing the smooth design model derived in the first part of the thesis, the second part addresses the design of an output-feedback tracking control system for the electro-pneumatic clutch actuator.

The first part of this output-feedback problem, is that of designing an observer for the unmeasured states of the system. We first propose to use the design model as an open-loop nonlinear observer for the unmeasured states. Analyzing the stability properties of this observer, we demonstrate that the pneumatic actuator has some inherent stability properties which enables the design of simple nonlinear observers. Based on this result, we propose two simple nonlinear observers for the electro-pneumatic actuators — a full-order observer and a reduced-order observer — both compatible with output-feedback control by an observer-based backstepping approach. The performance of the observers are validated by simulations and experimentally on a test rig.

The second part of the output-feedback problem, is that of designing an observer-based tracking controller. Based on the proposed reduced-order observer, a robust output-feedback tracking controller is designed by an observer backstepping approach. The controller achieves asymptotic practical tracking in the presence of bounded disturbances (*e.g.* modeling errors), which means tracking within a prescribed precision which can be made arbitrary small by sufficiently high gain in the controller. The performance of the output-feedback tracking controller is validated by simulations and experimentally on a laboratory test rig.

1.4.2 Organization of the thesis

The remaining chapters are divided into four parts: Part I, *Mathematical Modeling*, which addresses the modeling of the electro-pneumatic clutch actuation system, Part II, *Control Design*, which addresses the design of an output-feedback tracking controller for the actuator, Part III, *Thesis Conclusions*, and Part IV *Appendices*.

Part I on mathematical modeling includes Chapters 2–6:

Chapter 2: Provides a brief description of the clutch actuation system and the laboratory test rig.

Chapter 3: Addresses the modeling of the motion dynamics, including subsections on the modeling of the nonlinear load characteristic, the dynamic friction, and hardstop forces.

Chapter 4: Reviews the full and reduced-order equations describing the air dynamics of the pneumatic chambers.

Chapter 5: Addresses the modeling of the static flow rate characteristics of pneumatic restrictions in a nonlinear and adaptive control setting. The chapter includes novel parametrizations of flow rate equation for fixed restrictions, and the flow rate of flow control valves, and a simplified model of the valve dynamics.

Chapter 6: Summarizes, in state-space form, the full dynamic model of the electro-pneumatic clutch actuator, and outlines its characteristic properties which are important with respect to observer and control design.

Part II includes Chapters 7–9 on control design:

Chapter 7: Recapitulates some technical tools and terminology, used in the subsequent chapters on observer and control design.

Chapter 8: Addresses the problem of nonlinear observer design for electro-pneumatic actuators. Three simple observers are proposed, where all are compatible with output-feedback control by an observer-based backstepping approach, are proposed. Their stability and convergence properties are analyzed, and validated by simulations and experimentally on the test rig.

Chapter 9: Addresses nonlinear output-feedback control of electro-pneumatic actuators. Utilizing the proposed reduced-order observer, a particular design based on observer backstepping, is proposed. The performance of the controller is validated by simulations and experimentally on the test rig.

Part III summarizes the thesis conclusions in Chapter 10:

Chapter 10: Concludes the results of the thesis.

Part IV includes the Appendices A–C:

Appendix A: Reviews and discusses a general approach for the modeling of smooth static nonlinearities in parameter-affine form using the bell-shaped Gaussian or B-Spline basis functions.

Appendix B: Briefly reviews the linear least squares approach for parameter estimation of parameter-affine static models, and the nonlinear least squares optimization approach for fitting of the parameters of non-affine models.

Appendix C: Provides a detailed derivation of the full-order dynamics of the pneumatic cylinder chambers.

Most of the results included in this thesis have not yet been published. A primary reason for this has been customer projects of strategic importance for KA, on the development of electro-pneumatic actuators for AMT systems. In these projects, the results from the Ph.D. work have been key competitive factors, that have motivated

to delay the publication of results. The application of the results from the Ph.D. in these projects, has resulted in a patent application on a control strategy for pneumatic clutch actuation, submitted early October 2005.

A paper describing the design model of Chapter 6, was presented at the *3rd FPNI-PhD Symposium 2004* [41]. Related work based on Chapter 8 on nonlinear observer design for pneumatic clutch actuators, is the M.Sc. thesis by Knutsen [50]. The work on nonlinear observer design is currently being refined into a journal paper, planned submitted by January 2006. Another journal paper which compares the cancelling backstepping design presented in Chapter 9, with a non-cancelling design which avoids cancellation of the inherent stabilizing nonlinearities of the pneumatic clutch actuator, is planned submitted by June 2006.

Part I

Mathematical Modeling

Chapter 2

System Description

A brief description of the physical parts of the clutch actuation system, is given below. A drawing of a prototype clutch actuator from Kongsberg Automotive, is given in Figure 2.1.

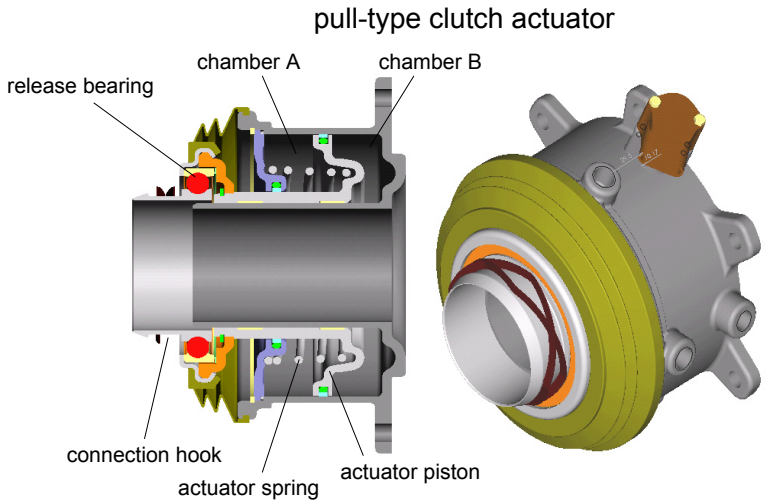

Figure 2.1: Drawing of a pull-type clutch actuator.

Clutch: The functional objective of the friction clutch is to provide a disengagement of the engine from the vehicle (*e.g.,* during transmission shifting), and to control the torque transmitted from the engine to the vehicle during clutch engagement. Torque is transmitted by the friction plates which are compressed by a clutch compression spring.

Actuator: A pneumatic cylinder actuator is used to disengage the clutch. The actuator illustrated in Figure 2.1, is a direct acting pull-type actuator, where the piston is connected to the clutch compression spring by a hook which pulls at the release bearing.

Valve: Electro-pneumatic flow control valve(s) are used to control the air flow to and from chamber A of the actuator. In the prototype system considered in this work, a closed center three-way *proportional valve* is used. An alternative valve configuration is by use of multiple pairs of *on-off valves*, where each pair consists of a *fill valve* controlling the flow from the supply into chamber A, and a *vent valve* to escape air to exhaust.

Sensor: A sensor is required in order to provide measurements of the position of the actuator for feedback control.

Electronic Control Unit (ECU): The ECU is a digital computer system on the vehicle which is used to implement the control system.

A simplified schematic diagram of a pneumatic clutch actuator is given in Figure 2.2. The system consists of a direct acting pull-type prototype actuator, controlled by a closed center three-way *proportional valve*, with an *Electronic Control Unit* (ECU) which implements the control system, and a *position sensor* measuring the actuator position. The load and friction forces, denoted f_l and f_f respectively, are represented by arrows acting on the connection hook of the actuator piston.

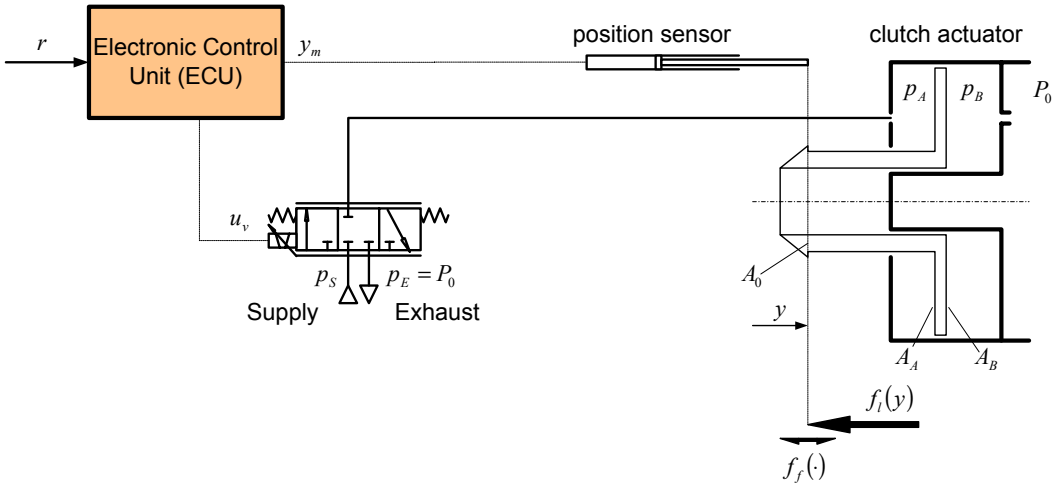

Figure 2.2: A schematic diagram of the pneumatic clutch actuator PCA.

The position y of the clutch actuator is measured by a position sensor, and the measured position y_m is fed back to the *Electronic Control Unit* (ECU) which computes the valve control input u_v according to a given reference signal r. The valve distributes the air flow in and out of chamber A by opening for flow from the *supply reservoir* (with pressure p_S), or escaping air to the *exhaust reservoir* (usually at atmospheric pressure $p_E = P_0$). By controlling the air flows, the valve controls the pressure p_A of chamber A. The back-chamber B is connected to atmosphere (with pressure P_0) through a duct, referred to as the *outlet restriction* of chamber B. For certain actuators, the flow resistance in the outlet restriction is so high that

actuator movements are causing a significant dynamic back-pressure p_B in chamber B. The resulting force provided by the pneumatic actuator is given as

$$f_a = A_0 P_0 + A_A p_A - A_B p_B.$$

The position y of the clutch actuator is consequently controlled by manipulating the valve input u_v to generate the necessary pressure p_A in chamber A to balance the resultant *load force* f_l, and compensate for the *friction force* f_f, and the back-pressure p_B. The load and friction forces acting on the pneumatic clutch actuator are briefly described below.

The position dependent load force f_l, referred to as the *clutch load characteristic*, is a lumped spring force composed of the clutch compression spring which compresses the clutch friction discs, and a counteracting (much weaker) actuator spring. The clutch compression spring is a stiff and highly nonlinear diaphragm spring which constitutes the main part of the load characteristic, while the actuator spring is a linear coil spring.

The resulting load characteristic may differ significantly from truck to truck, depending on various factors: The clutch compression spring is designed according to the type and required torque capacity of the clutch, which obviously depends on the type of truck. In general, there is a significant difference between the load characteristic of a new and a worn clutch, because the point of attack of the clutch compression spring changes with wear of the clutch. The clutch load characteristic also depends on whether the clutch actuator is of push or pull type, or whether the actuator operates the clutch directly or indirectly by a lever. Typically, push type actuators have a stronger nonlinear characteristic and a higher force level than pull type actuators. The exchange ratio of the lever is usually about 2, thus, the lever operated actuators have about twice the operating range, and consequently, half the force level compared to the directly acting actuators.

The actuator seals and the mechanical contact between moving parts in the actuator and clutch, constitute a considerable friction force in the system. In Figure 2.2, these friction forces are lumped together in the single resultant *friction force* f_f acting on the connection hook of the actuator.

2.1 Laboratory test rig setup

All the experimental results presented in this thesis are obtained with an early prototype pneumatic clutch actuator mounted on a Scania clutch at the laboratory at Kongsberg Automotive's research department at Kongsberg, in Norway. The actuator on the test rig is a *concentric* actuator of *pull type*, which means that the actuator has a concentric placement inside the clutch housing, and that the actuator is pulling directly on the clutch release bearing.

The actuator is equipped with a Servotronic proportional valve from Joucomatic, and sensors providing measurements of the actuator's position y, acceleration $\ddot{y}$,

cylinder chamber pressures, p_A and p_B, and supply pressure p_S. A dSPACE real-time computer system together with Matlab/Simulink is used for data acquisition and control.

Figure 2.3 shows a photo of the prototype test rig and data acquisition system in the R&D Laboratory of Kongsberg Automotive's R&D department at Kongsberg.

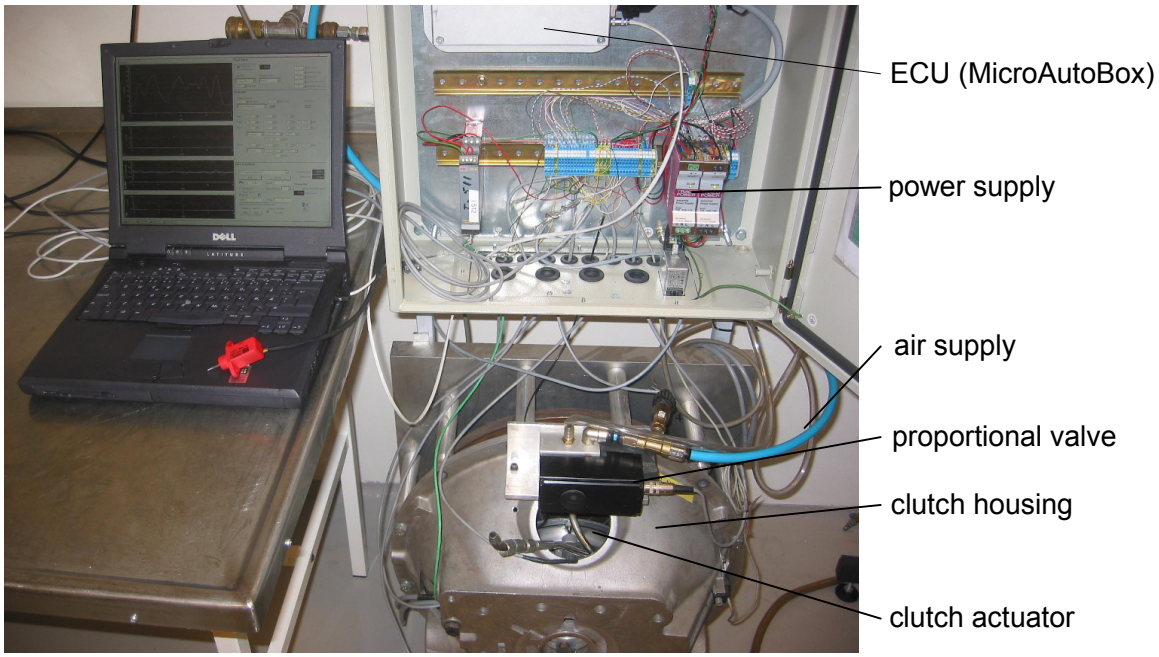

Figure 2.3: Photo of the test rig used for experimental testing, consisting of clutch, clutch actuator, valve, sensors, ECU, and laptop used for data acquisition and control of experiments.

2.1.1 Observer and controller implementation

The simulation models of the electro-pneumatic actuator, the observers, and the output-feedback controller, designed in Chapters 8 and 9, were implemented in Simulink/Matlab using continuous-time integrators, and simulated using a variable step solver.

For real-time implementation of the observer and controller, a real-time control system from dSPACE was used, where the observer and controller were implemented in discrete-time using a 3rd-order explicit Runge-Kutta fixed step solver with sample time $\Delta T = 2.0\,\text{ms}$. The required C-code was generated by automatic code generation in Matlab, and downloaded to a real-time operating system running on a MicroAutoBox from dSPACE. Virtual instrument control panels were developed in a dSPACE Control Desk program, which were used to control the experiment via a laptop connected to the MicroAutoBox. The instrument panels developed for the observer and controller are depicted in Figure 2.4 and 2.5, respectively. The main functions of the instrument panels were to perform data acquisition, adjust parame-

ters, and to plot the estimated and measured variables on-line. They also provided a possibility to adjust selected observer and controller parameters and initial values, and choose control mode (open-loop or closed loop), and type of reference or control input: square wave, sine wave, clutch sequence or manual input by means of a hand-held position sensor.

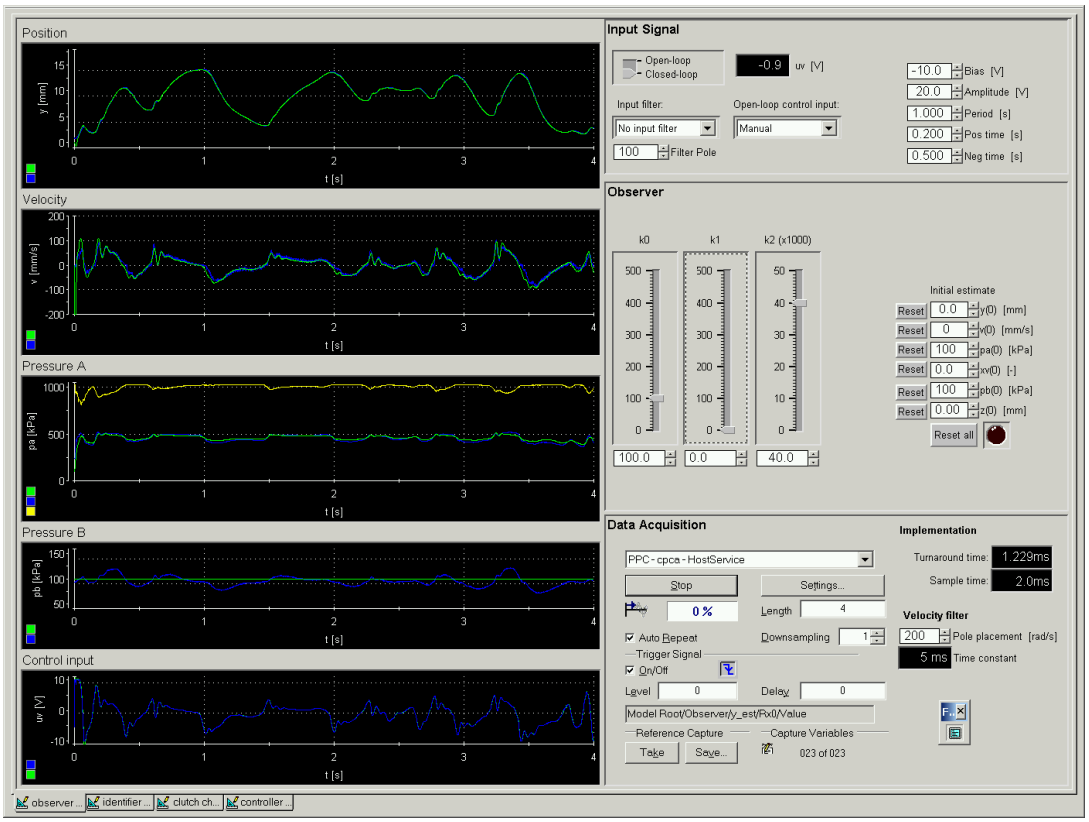

Figure 2.4: A screen dump of the observer instrument panel.

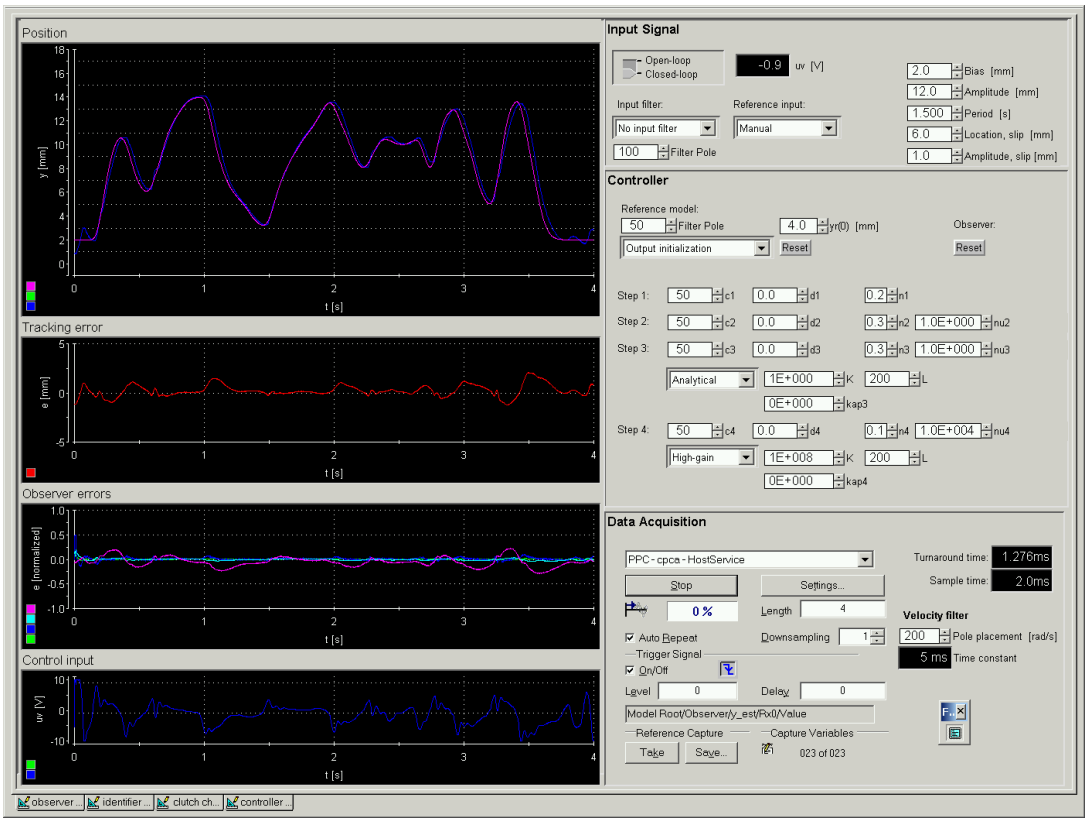

Figure 2.5: A screen dump of the controller instrument panel.

Chapter 3

Motion Dynamics

The *motion dynamics* of the actuator piston is expressed by the equation of motion, known as *Newton's 2nd law*. The resulting force acting on the actuator piston is composed of the clutch load force f_l, the friction force f_f, and the resultant actuator pressure force $(A_0P_0 + A_Ap_A - A_Bp_B)$. In addition, the stroke of the actuator piston is limited by the physical length of the cylinder, where the physical constraints give rise to the hardstop forces f_h. Furthermore, the mechanical coupling between the actuator and clutch compression spring is assumed to be stiff, and the inertia of moving parts is lumped to the mass of the piston. Thus, the motion dynamics is described by

$$M\frac{d^2y}{dt^2} = A_0P_0 + A_Ap_A - A_Bp_B - f_l(y) - f_f(\cdot) - f_h(\cdot), \tag{3.1}$$

where y, p_A, p_B are the piston position and pressures, and indices A and B are referring to chambers A and B, respectively; M is the effective mass of moving parts; A_A and A_B are the piston areas. The area $A_0 = A_B - A_A$ is an area of the piston on the chamber A side of the actuator which is subjected to atmospheric pressure P_0. The mathematical modeling of the static load force f_l, the dynamic friction force f_f, and the hardstop force f_h, are addressed in the following sections.

3.1 Clutch load characteristic

The clutch load characteristic is a nonlinear function of the clutch position, *i.e.*, $f_l = f_l(y)$ (also referred to as a static nonlinearity). For any type of clutches and actuator configurations, the clutch load characteristic $f_l(y)$ can be modeled in the parameter affine form

$$f_l(y) = \boldsymbol{\theta}_l^T \cdot \boldsymbol{\phi}_l(y), \tag{3.2}$$

where $y \in \mathcal{Y} \triangleq [y_{\text{lb}}, y_{\text{ub}}]$ is the clutch actuator position, and $f_l \in \mathcal{F}_l \subset \mathbb{R}$ is the modeled load force. The regressor $\boldsymbol{\phi}_l(y) = [\phi_{l\,1}(y), \phi_{l\,2}(y), \cdots, \phi_{l\,p}(y)]^T \in \mathbb{R}^p$ is a vector of basis functions, which is weighted by the parameter vector $\boldsymbol{\theta}_l = [\theta_{l\,1}, \theta_{l\,2}, \cdots, \theta_{l\,p}]^T \in \Theta_l \subset \mathbb{R}^p$. We use *normalized Gaussian basis functions*, defined

according to

$$\phi_{l,i}(y) = \frac{\mu_i(y)}{\sum_{j=1}^{p} \mu_j(y)} \tag{3.3}$$

$$\mu_i(x) = \mathrm{e}^{-\frac{1}{2}w_i^2(x-c_i)^2}, \tag{3.4}$$

where $\mathbf{w} = [w_1, w_2, \cdots w_p]^T$ is a vector of scaling parameters, and $\mathbf{c} = [c_1, c_2, \cdots, c_p]^T$ a vector of offset parameters. Equation (3.3) provides a normalization of the standard Gaussian function given by (3.4), and the parameters w_i determine the widths of the basis functions $\phi_{l,i}(y)$, thus, the degree of smoothness of the modeled output $f_l(y)$, and the parameters c_i determine the location (or centers) of the basis functions $\phi_{l,i}(y)$. See Appendix A for further discussion of the choice of the parameters w_i and c_i.

With properly defined scaling vector $\mathbf{w}$ and center vector $\mathbf{c}$, the load model (3.2) using normalized Gaussian basis functions has the following desirable properties:

- Provided that $\mathbf{c}$ and $\mathbf{w}$ are viewed as fixed, non-tunable parameter vectors, the model is affine in its tunable parameter vector $\boldsymbol{\theta}_l$, which is advantageous for parameter estimation, and necessary to make possible on-line identification by existing constructive Lyapunov-based adaptive controller designs.

- The interpretation of each parameter is good, in the sense that there is a close relation between the weighting θ_i of the ith basis function $\phi_{l,i}(y)$ and the modeled output $f_l(y)$ at the center $y = c_i$. This is mainly due to the *unity property*

$$\sum_{i=1}^{p} \phi_{l,i}(y) \equiv 1, \tag{3.5}$$

  which is a result of the normalization.

- Each basis function in the interior of $\mathcal{Y}$ ($\phi_{l,i}(y)$, $i = 2, \cdots, p-1$) has *exponential local support* (defined in Appendix A), which means that each basis function is practically zero outside a finite subset of $\mathcal{Y}$ around its center. Such basis functions have superior numerical properties with respect to to parameter identification, hence, are desirable for adaptive controller designs.

- The modeled load force is infinitely smooth, *i.e.*, it is infinitely differentiable.

The measured and modeled clutch load characteristic of the test rig clutch actuator, is plotted in Figure 3.1. The full operating range of a new clutch of this actuator configuration is between 0 and 15 mm. As the clutch wears, the clutch load characteristic will typically move to the left and increase in magnitude. The operating range of a worn clutch of the considered configuration is typically -10 to 5 mm at the end of its lifetime. Consequently, for the model to be valid for the entire lifetime of the clutch, the input range of the modeled load characteristic should be $\mathcal{Y} = [-10, 15]$ mm. The parameter vector $\boldsymbol{\theta}_l$ can be fitted to the

mean of a quasi-static load characteristic obtained from measurements on the test rig. This characteristic may be obtained by one slow disengagement and subsequent engagement of the clutch while measuring the actuator pressures. From the measured actuator pressures, it is possible to compute the corresponding actuator load force. This measured load force results in a characteristic with hysteresis, where it is possible to extract the mean load force. The hysteresis can be attributed to friction which is modeled separately (see section below). In Figure 3.1, the extracted mean of the measured load characteristic, denoted $f_{l,m}(y)$, is plotted together with the modeled clutch load characteristic $f_l(y)$, where the parameters $\boldsymbol{\theta}_l$ are fitted to the measurements by a least squares method (see Appendix B). The corresponding set of normalized Gaussian basis functions $\boldsymbol{\phi}_l(y)$, is plotted in the lower pane in the figure. The figure also shows a typical example of the load characteristic of a worn clutch, $f_{l,w}(y)$, and some conservative estimates of upper and lower bounds, discussed in the subsection below.

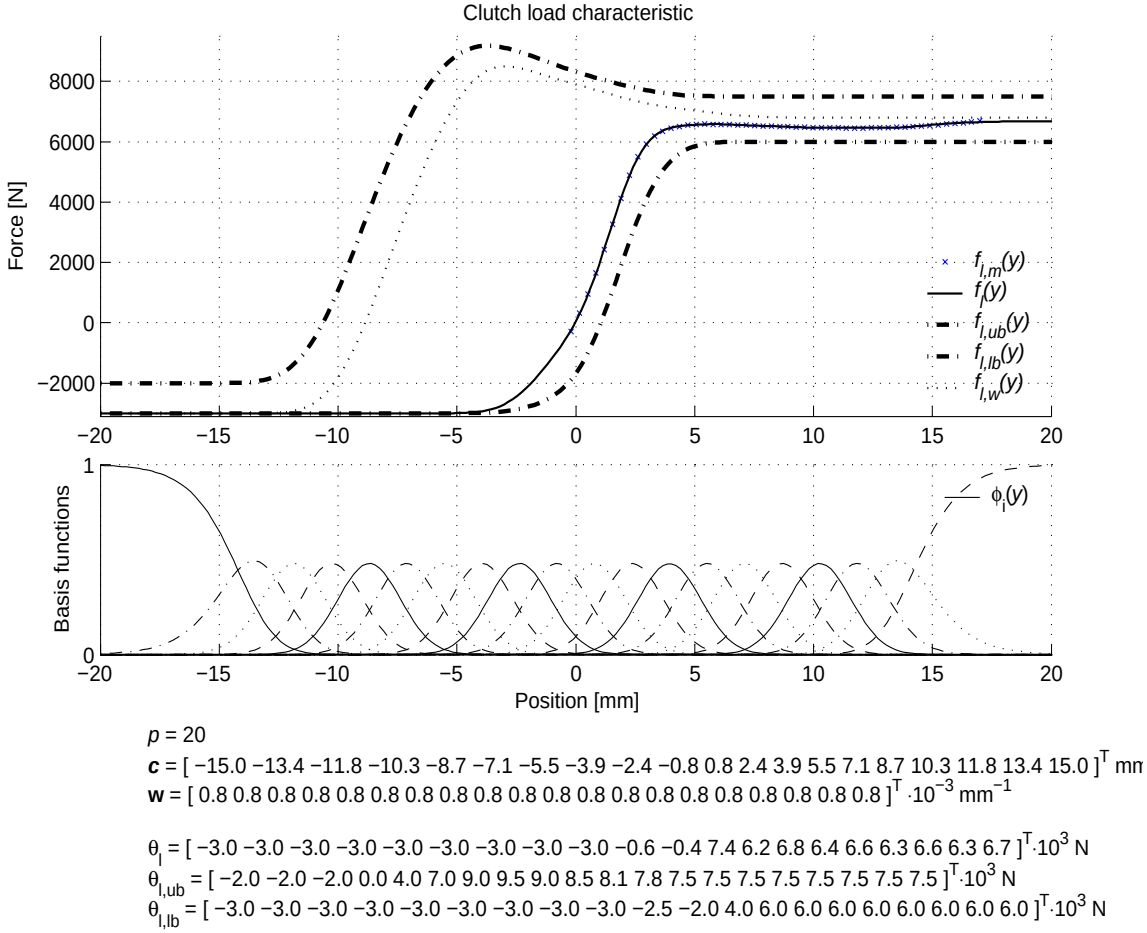

Figure 3.1: The modeled clutch load characteristic of the test rig (push-type actuator).

Remark 8 *The clutch load characteristic can alternatively be modeled using B-spline basis functions, as illustrated in Appendix A, Figures A.7–A.8 with cubic*

(3rd-order) B-spline basis functions. The modeling capabilities are very similar to the normalized Gaussian basis functions. A difference is that the B-spline functions have a finite degree of differentiability with degree that increases with the order of the function. An advantage over normalized Gaussian functions, is that each B-spline function has a precise region of local support (instead of exponentially local support). This means that the B-spline functions can be used to define a set of basis functions with a precise region of support. This makes the B-spline functions better suited to define customized parametrizations, where the smoothness of the modeled output may vary over the region of support, which again can be used to minimize the number of necessary basis functions for a given accuracy.

3.1.1 Uncertainty modeling

The load characteristic will change considerably during the life-time of the truck due to wear of the clutch. When the modelled load characteristic is used for model-based compensation, it therefore needs to be updated as the clutch wears. This can be done continuously using an adaptive control scheme, or by offline parameter estimation at scheduled intervals. In both cases, to obtain a robust solution, it will be necessary to compute bounds on $f_l(y)$ which are valid through the entire lifetime of the clutch. The given parametrization of the clutch load characteristic provides a simple way to implement such upper and lower bounds. That is, the upper and lower bounds of the load characteristic can be implemented as

$$\begin{aligned} f_{l,\mathrm{ub}}(y) &\triangleq \boldsymbol{\theta}_{l,\mathrm{ub}}^T \boldsymbol{\phi}_l(y), \\ f_{l,\mathrm{lb}}(y) &\triangleq \boldsymbol{\theta}_{l,\mathrm{lb}}^T \boldsymbol{\phi}_l(y), \end{aligned} \tag{3.6}$$

where $\boldsymbol{\theta}_{l,\mathrm{ub}}$ and $\boldsymbol{\theta}_{l,\mathrm{lb}}$ are the parameter vectors for the upper and lower bound, respectively. Hence, a parametric uncertainty model of the load characteristic is given by (3.2) and the parameter set

$$\Theta_l \triangleq \{\forall \boldsymbol{\theta}_l : \boldsymbol{\theta}_{l,\mathrm{lb}} \leq \boldsymbol{\theta}_l \leq \boldsymbol{\theta}_{l,\mathrm{ub}}\}, \tag{3.7}$$

i.e., we have that $\boldsymbol{\theta}_l \in \Theta_l, y \in \mathcal{Y} \Longrightarrow f_{l,\mathrm{lb}}(y) \leq f_l(y) \leq f_{l,\mathrm{ub}}(y)$.

Remark 9 *The clutch load characteristic of the test rig application is a new clutch of push-type, which has a rather mild nonlinear characteristic. The clutch load force of push-type clutches, however, exhibits a stronger nonlinear characteristic. An example of the modeled clutch load characteristic of a push-type clutch, is plotted for both a new and a worn clutch in Figure 3.2. Notice that for both push-type and pull-type clutches, the nonlinear form of the clutch characteristic becomes more pronounced as the clutch wears. Notice also that the push-type clutch characteristic is less steep than the pull-type characteristic, thus, can be modeled accurately with fewer basis functions. The "measured" clutch load characteristics $f_{l,m}$ and $f_{l,w}$ in the figure, which are used for fitting of parameters, are provided by the clutch manufacturer as look-up tables. These characteristics are theoretic curves computed from the geometry of the clutch compression spring and its configuration,* i.e., *they are not actually measured.*

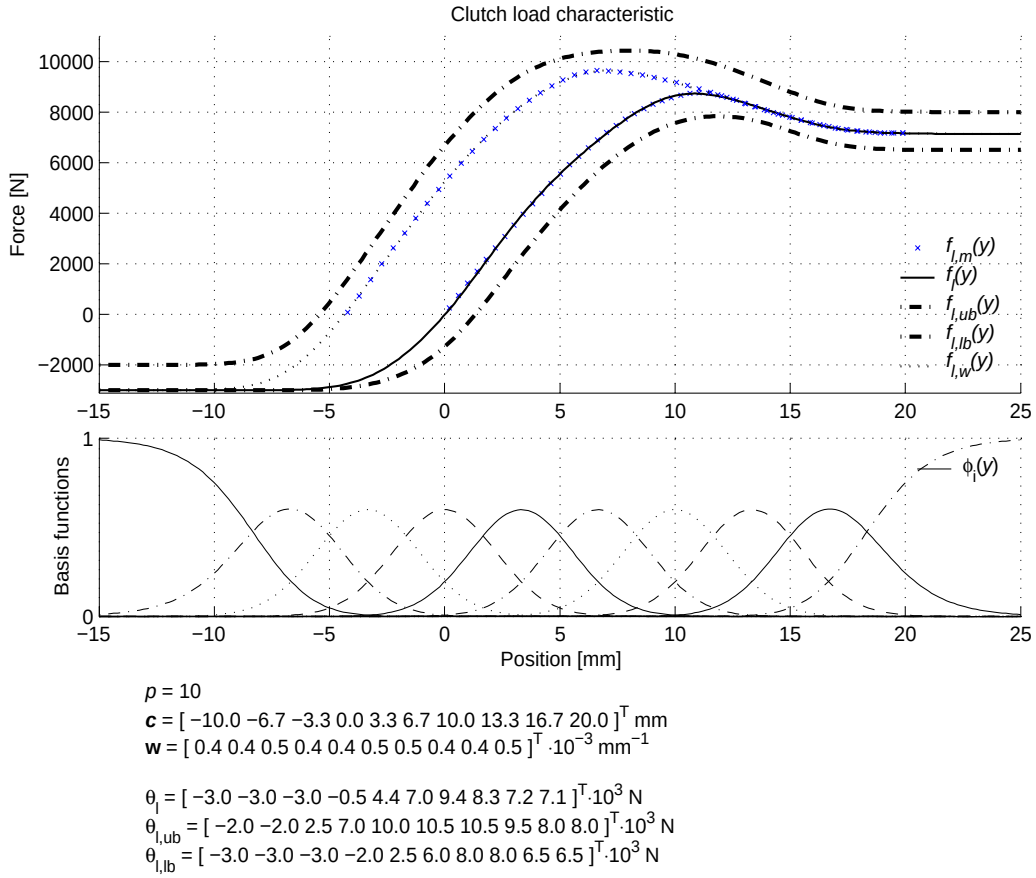

Figure 3.2: Example of the clutch load characteristic of a push-type clutch.

3.2 Friction

In the clutch actuator application the resultant friction force is considerable. Hence, friction modeling and friction compensation plays an important role with respect to high-performance position control of the actuator.

The properties of friction are distinguished by two friction regimes, referred to as the *pre-sliding* and *sliding* regime. *Pre-sliding friction* refers to the friction force during the small displacement (or micro-slip) occurring in the contact material prior to sliding, *i.e.*, when there is still zero relative velocity between the piston seal and the cylinder wall. *Sliding friction* refers to the case when the contact surfaces are moving relatively to each other, *i.e.*, when the piston seal slides on the cylinder wall. Pre-sliding and sliding friction are briefly described below:

Pre-sliding friction: When subjected to a load force which is below the friction break-away force (also referred to as the stiction force), there will be a pre-sliding displacement of the asperities in contact, in particular, an elasto-plastic deformation of the asperities. This phenomenon is intuitive since all materials exhibit strain when subjected to stress. Hence, pre-sliding displacement occurs for all kinds of materials in contact. However, the softer the material

is, the more pronounced will the pre-sliding displacement be. The pre-sliding displacement is composed of an elastic and a plastic deformation. The elastic deformation will behave like a nonlinear spring-damper system, while the plastic deformation is characterized by a creep motion and work hardening of the asperities in contact. By creep motion we mean a continuous deformation of the material, which due to work hardening will cause a decreasing deformation rate. For a complete and thorough description of the pre-sliding friction phenomenon, see [31].

Remark 10 *Due to the elasticity in the seal between piston and cylinder, the pre-sliding friction properties of pneumatic actuators plays an important role for accurate friction modeling*[1].

Static sliding friction: The friction at steady-state sliding, that is, at constant velocities, is referred to as *static sliding friction*. The static sliding friction may be described in terms of *viscous* and *dry friction*, where viscous friction is due to the shear stress in the fluid separating two moving surfaces, while dry friction is due to the abrasive contact between surface asperities in contact. The static sliding friction is static in the sense of being a pure function of the velocity, and possibly chamber pressures, due to the use of lip seals, or actuator position, due to changing misalignment of the piston axis along the actuator stroke. Hence, a static friction model is a function of the actuator states, and does not contain any internal dynamic states.

Dynamic sliding friction: In addition to pure static properties, sliding friction exhibits certain dynamic phenomena. These are *frictional lag* and *varying break-away force*. For a detailed description, see e.g. [5]. *Frictional lag*, or frictional memory, is simply a time-delay in the corresponding change in friction after a change in the velocity. The effect of frictional lag is probably most significant in applications with a highly viscous fluid, or lubricant. *Varying break-away force*, or rising static friction, refers to the phenomenon that the friction force level for which sliding occurs — the break-away force, or stiction force — increases with the dwell time, *i.e.*, the time spent at rest (in stiction). In the case of dry surfaces, one explanation of this phenomenon is that cold welding occurs between the contact asperities when the relative motion between the surfaces reaches zero, thus increasing the friction. Another explanation in the case of lubricated surfaces is that lubrication needs time to flow away, thus, contributes to keep the surfaces separated a short time after reaching zero relative velocity.

[1] For example for the high-precision clutch servo system, a significant part of the operaton is within the pre-sliding regime, thus, pre-sliding friction constitutes an important part of the friction model.

3.2.1 Static friction models

Pure *viscous friction* is usually modeled as proportional to velocity, sometimes including a viscous drag component, such as *e.g.*,

$$f_v(v) = D_v v + D_d |v| v, \tag{3.8}$$

where v is the velocity, D_v is the viscous damping coefficient, and D_d is the drag damping coefficient.

The *dry friction* is better known by the common *Coulomb* friction model, which states that the friction is proportional to the normal force of contact, *i.e.*,

$$F_C = \mu N, \tag{3.9}$$

where μ is a friction coefficient and N is the normal force compressing the surfaces. A common phenomenon of dry friction, not captured by the Coulomb friction model, is the *stiction effect*. The stiction effect refers to the phenomenon that the dry friction force is higher at rest than when sliding. For unlubricated surfaces, the stiction effect appears as a discontinuous drop in the dry friction force once sliding occurs, however, physically, the transition from stiction to sliding must be continuous. A simplified, but appealing explanation of this phenomenon is to view the fluid as a lubricant, where the relative movement of the two surfaces will cause the available fluid to build a lubrication film (on a microscopic level) which will try to separate the two surfaces. This separation depends on the relative velocity, and varies from zero separation at low or zero velocities, to full separation for velocities which are sufficiently high to build a full fluid film between the surfaces. Consequently, the friction will be high for zero velocity, and decrease as a fluid film builds between the surfaces, reducing the normal force on the asperities in contact. This velocity dependence gives rise to a decreasing force characteristic at low velocities, known as the Striebeck effect. Assuming negligible position and pressure dependence, the absolute value of the dry friction force — including the Striebeck effect — is often modeled in the form

$$f_d(v) = F_C + (F_s - F_C)\, \mathrm{e}^{-|v/v_S|^\sigma}, \tag{3.10}$$

where F_C is the Coulomb friction level, F_s is the stiction force level (also referred to as the break-away force), and v_S is referred to as the Striebeck velocity. The parameter σ determines the characteristic of the Striebeck curve, and is usually taken in the range $\sigma \in [1, 2]$.

In many applications of pneumatic actuators, the viscous and dry friction forces exhibit a significant dependence on the pressures in the cylinder chambers mainly due to the use of lip seals. This is usually captured by modeling the viscous friction coefficient D_v, and the dry friction coefficients F_C and F_s as functions of the pressures in the chambers.

The viscous friction force $f_v(\cdot)$, omitting the viscous drag term, can thus be modeled according to

$$\begin{aligned} f_v(v, p_A, p_B) &= D_v(p_A, p_B) \cdot v & (3.11)\\ D_v(p_A, p_B) &= D_{v0} + \beta_A \cdot (p_A - p_0)^{\rho_A} + \beta_B \cdot (p_B - p_0)^{\rho_B}, & (3.12) \end{aligned}$$

where the pressure dependence of the viscous friction function $D_v(p_A, p_B)$ is determined by the scaling factors $\beta_A, \beta_B > 0$ and the exponents $\rho_A, \rho_B > 0$, while D_{v0} is a nominal viscous friction coefficient defined for a nominal actuator pressure p_0.

In a similar manner, the dry friction force $f_d(\cdot)$, including the Striebeck effect, can be modeled according to

$$f_d(v, p_A, p_B) = F_C(p_A, p_B) + R_s \cdot F_C(p_A, p_B)\, \mathrm{e}^{-|v/v_S|^\sigma} \tag{3.13}$$

$$F_C(p_A, p_B) = F_{C0} + \alpha_A \cdot (p_A - p_0)^{\gamma_A} + \alpha_B \cdot (p_B - p_0)^{\gamma_B}, \tag{3.14}$$

where $R_s \triangleq (F_{s0} - F_{C0})/F_{C0}$ is the fraction of friction increase due to stiction, the pressure dependence of the Coulomb friction function $F_C(p_A, p_B)$ is determined by the empirical parameters $\alpha_A, \alpha_B, \gamma_A, \gamma_B > 0$, and F_{C0} and F_{s0} are nominal Coulomb and stiction force, respectively, defined for actuator pressures at p_0.

3.2.2 The LuGre dynamic friction model

A dynamic friction model which captures most of the qualitative properties of friction, while being well suited for model-based control, is the *LuGre* (Lund-Grenoble) dynamic friction model proposed by Canudas de Wit *et al.* in the paper [15]. The model is recapitulated below.

The dynamic pre-sliding deflection state of the *LuGre* dynamic friction model can be given as

$$\dot{z} = v - \frac{K_z}{f_d} |v|\, z, \tag{3.15}$$

where the resulting dynamic friction force is given by

$$f_f(v, z) = D_v v + K_z z + D_{\dot{z}} \dot{z}(v, z). \tag{3.16}$$

Here, z is the friction state which for pneumatic actuators may be interpreted as a pre-sliding seal deflection, where the parameter K_z is the deflection stiffness, and $D_{\dot{z}}$ is the deflection damping coefficient. Furthermore, f_d is the dry friction force, and D_v is the viscous damping coefficient, which in general are functions of actuator velocity and chamber pressures as in Equations (3.13)–(3.14), and (3.11), respectively.

Properties of the LuGre friction model

From (3.15), with $\dot{z} = 0$, the steady-state deflection z^* becomes

$$z^* = \frac{f_d}{K_z} \frac{v}{|v|} = \frac{f_d}{K_z} \operatorname{sgn}(v). \tag{3.17}$$

Hence, from (3.16), the resulting steady-state friction force is given as

$$f_f^*(v, f_d) = f_d \operatorname{sgn}(v) + D_v v. \tag{3.18}$$

The friction characteristic f_f^* given by (3.18), is referred to as the *static sliding friction characteristic* of the dynamic friction model (3.15–3.16).

To gain more insight into the dynamic friction model, it is instructive to show that the pre-sliding deflection must be finite. The proof below is recapitulated from [15].

Proposition 11 (Finite pre-sliding deflection, [15]) *The pre-sliding deflection $z(t)$ is finite. That is, the solution $z(t)$ of the pre-sliding dynamics (3.15) is globally uniformly bounded (GUB) for all bounded initial values $z(0)$.*

Proof. Consider the function

$$V = \frac{1}{2}z^2,$$

which time-derivative along the trajectories of $z(t)$ becomes

$$\begin{aligned}
\dot{V} &= z \cdot \dot{z} \\
&= z \cdot \left(v - \frac{K_z}{f_d} |v| z \right) = vz - \frac{K_z}{f_d} |v| z^2 \\
&= -|v||z| \left(-\operatorname{sgn}(v) \operatorname{sgn}(z) + \frac{K_z}{f_d} |z| \right) \\
&\leq -|v||z| \left(-1 + \frac{K_z}{f_d} |z| \right).
\end{aligned}$$

Hence,

$$|z(t)| > \frac{\|f_d(t)\|_\infty}{K_z} \Longrightarrow \dot{V} < 0,$$

which by LaSalle-Yoshizawa's Theorem [52, Theorem A.8] proves that all solutions $z(t)$ of (3.15) converges to the invariant set

$$\Omega_z \triangleq \{\forall z : \ |z(t)| \leq \|f_d(t)\|_\infty / K_z\}.$$

■

Remark 12 *Note that the upper bound $\|f_d(t)\|_\infty$ on the dry friction force is simply the level of the stiction force. For example, with the dry friction characteristic given by (3.10), the upper bound is given by the constant stiction force F_s,* i.e., *$\|f_d(t)\|_\infty = F_s$.*

We define the parameter

$$Z_{\max} \triangleq \|f_d(t)\|_\infty / K_z, \tag{3.19}$$

which we denote the *max*imum *pre-sliding deflection*. From the proof of Proposition 11, we see that by choosing the initial deflection $z(0)$ less than the maximum pre-sliding deflection $Z_{\max}$, it will remain so ever after, *i.e.*,

$$\begin{aligned}
|z(0)| &\leq Z_{\max} \\
&\Downarrow \\
z(t) &\leq Z_{\max} \quad \forall t \geq 0.
\end{aligned}$$

Remark 13 *Since pre-sliding displacement larger than $Z_{\max}$ is not physically justified, the initial deflection $z(0)$ should always be chosen less than $Z_{\max}$.*

The following proposition states the passivity properties of the LuGre friction model with the velocity v as input and the friction force f_f as output. The proof of the proposition follows an approach using Lyapunov arguments, rather than the standard integral definition of passivity as in [7, 23].

Proposition 14 (Passivity) *Consider the LuGre friction model, given by (3.15)–(3.16), with the velocity v as input and the friction force f_f as output. The model has the following passivity properties:*

i) *$f_d \in [F_C, F_s]$: With the dry friction modeled to include the Striebeck effect, the dynamic friction model is passive from v to f_f if*

$$D_v \geq D_{\dot{z}} \frac{F_s - F_C}{F_C},$$

for $\forall K_z, D_{\dot{z}} \geq 0$, and strictly passive (with excess of passivity) if the above inequality is strict.

ii) *$f_d = F_C$: With the dry friction modeled as a constant Coulomb friction force, the dynamic friction model is passive from v to f_f for*

$$D_v \geq 0.$$

for $\forall K_z, D_{\dot{z}} \geq 0$, strictly passive if $D_v > 0$.

Proof. Consider the scalar function

$$V = \frac{K_z}{2} z^2,$$

which time-derivative becomes

$$\begin{aligned} \dot{V} &= K_z z \dot{z} \\ &= K_z z \left(v - \frac{K_z}{f_d} |v| z \right). \end{aligned}$$

Now add and substract $v f_f = D_v v^2 + K_z v z + D_{\dot{z}} v \dot{z}$, and rewrite

$$\begin{aligned} \dot{V} &= K_z z \left(v - \frac{K_z}{f_d} |v| z \right) + v f_f - D_v v^2 - K_z v z - D_{\dot{z}} v \left(v - \frac{K_z}{f_d} |v| z \right) \\ &= v f_f - D_v v^2 - \frac{K_z^2}{f_d} |v| z^2 - D_{\dot{z}} v \left(v - \frac{K_z}{f_d} |v| z \right). \end{aligned}$$

Using that

$$|z| \leq \frac{\|f_d\|_\infty}{K_z} = \frac{F_s}{K_z} = \frac{F_C}{K_z} + \frac{F_s - F_C}{K_z},$$

we obtain

$$\begin{aligned}\dot{V} &\leq vf_f - D_v v^2 - \frac{K_z^2}{f_d}|v|z^2 - D_{\dot{z}}\left(1 - \frac{K_z z}{f_d}\frac{|v|}{v}\right)v^2 \\ &\leq vf_f - D_v v^2 - \frac{K_z^2}{f_d}|v|z^2 - D_{\dot{z}}\left(1 - \frac{F_C}{f_d}\operatorname{sgn} v\right)v^2 + D_{\dot{z}}\frac{F_s - F_C}{f_d}\operatorname{sgn} v v^2 \\ &\leq vf_f - D_v v^2 - \frac{K_z^2}{f_d}|v|z^2 + D_{\dot{z}}\frac{F_s - F_C}{f_d}\operatorname{sgn} v v^2.\end{aligned}$$

Hence, with

$$D_v \geq D_{\dot{z}}\frac{F_s - F_C}{F_C} \geq D_{\dot{z}}\frac{F_s - F_C}{f_d},$$

we get

$$\dot{V} \leq vf_f - \frac{K_z^2}{f_d}|v|z^2,$$

which proves passivity from v to f_f. Moreover, with

$$D_v \geq \varepsilon + D_{\dot{z}}\frac{F_s - F_C}{F_C},$$

where $\varepsilon > 0$ is an arbitrary small constant. Then,

$$\dot{V} \leq vf_f - \frac{K_z^2}{f_d}|v|z^2 - \varepsilon v^2,$$

which proves strict passivity from v to f_f in the case when the inequality is strict: $D_v > D_{\dot{z}}(F_s - F_C)/F_C$. ■

3.2.3 A simplified smooth dynamic friction model

For our control design, we employ a modified version of the LuGre dynamic friction model, assuming simple static friction characteristics, and by introducing a smooth approximation of the pre-sliding deflection dynamics. The pre-sliding deflection dynamics (3.15) is modified according to

$$\dot{z} = v - \frac{K_z}{F_C}|v|_s z, \tag{3.20}$$

where $\varepsilon_0 > 0$ is an arbitrary small design parameter, and the dry friction characteristic is taken as a constant Coulomb friction, $f_d = F_C$. In order to make the model applicable for subsequent application of backstepping techniques, the non-smooth absolute value term $|v|$ has been replaced with the square root term

$$|v|_s \triangleq \sqrt{v^2 + \varepsilon_0^2}, \tag{3.21}$$

which is a smooth approximation to the absolute value term $|v|$. The resulting smooth friction force has the same form

$$f_f(v, z) = D_v v + K_z z + D_{\dot{z}}\dot{z}(v, z), \tag{3.22}$$

where the viscous friction coefficient D_v is taken as constant. This means that we have neglected the dependence on actuator position, chamber pressures, and also the Striebeck effect in our dynamic friction model, in order to obtain a simple model for control design.

Properties of the smoothed LuGre friction model

The introduced smooth approximation $|v|_s$ represents a smooth upper bound on the absolute operator $|v|$, which can be made arbitrary accurate by a reduction of the design constant ε_0, *i.e.*,

$$\lim_{\varepsilon_0 \to 0} |v|_s = \lim_{\varepsilon_0 \to 0} \sqrt{v^2 + \varepsilon_0^2} = |v|. \tag{3.23}$$

From (3.20), the steady-state deflection z^* becomes

$$z^* = \frac{F_C}{K_z} \frac{v}{\sqrt{v^2 + \varepsilon_0^2}} \tag{3.24}$$

$$= \frac{F_C}{K_z} \operatorname{sgn}_s(v), \tag{3.25}$$

where $\operatorname{sgn}_s(v)$ is a smooth signum function, defined as

$$\operatorname{sgn}_s(v) \triangleq \frac{v}{\sqrt{v^2 + \varepsilon_0^2}}. \tag{3.26}$$

The modified steady-state characteristic (3.24) is a smooth approximation to the steady state characteristic (3.17) of the original LuGre friction model, which is discontinuous at $v = 0$. From (3.24), the *static sliding friction characteristic* of the modified dynamic friction model (3.20)–(3.22), becomes

$$f_f^*(v) = F_C \frac{v}{\sqrt{v^2 + \varepsilon_0^2}} + D_v v \tag{3.27}$$

$$= F_C \operatorname{sgn}_s(v) + D_v v. \tag{3.28}$$

The modified dynamic friction model, given by (3.20)–(3.22), results in a smooth static sliding friction characteristic, but which does not give stiction for zero velocity, *i.e.*, $f_f^*(0) = 0$.

It is straightforward to show that the pre-sliding deflection for the modified dynamics is still finite, and Proposition 11 still holds. However, the maximum finite deflection is slightly reduced, as shown by the following proof:

Proof. The time-derivative of $V = \frac{1}{2} z^2$ along the trajectories of $z(t)$ becomes

$$\dot{V} = z \cdot \left(v - \frac{K_z}{F_C} |v|_s z \right)$$

$$= -|v|_s |z| \left(\frac{v}{|v|_s} \operatorname{sgn} z - \frac{K_z}{F_C} z \operatorname{sgn} z \right),$$

which gives,

$$|z(t)| > \frac{F_C}{K_z}\left|\frac{v}{|v|_s}\right| \Longrightarrow \dot{V} < 0.$$

Noting that

$$\frac{v}{|v|_s} = \frac{v}{\sqrt{v^2+\varepsilon_0^2}} \in \langle -1, 1\rangle,$$

by LaSalle-Yoshizawa's Theorem [52, Theorem A.8], all solutions $z(t)$ of (3.20) converge to the invariant set

$$\Omega_z \triangleq \{\forall z : \ |z(t)| < F_C/K_z\}.$$

■

The maximum pre-sliding deflection (3.19), becomes

$$Z_{\max} = F_C/K_z, \tag{3.29}$$

since the maximum dry friction force is the Coulomb friction, $\|f_d\|_\infty = F_C$. From the above proof, it follows that by choosing the initial deflection to satisfy $z(0) < Z_{\max}$, the maximum pre-sliding deflection $Z_{\max}$ represents an upper bound on the pre-sliding deflection state, $z(t) < Z_{\max}, \ \forall t \geq 0$.

The passivity properties stated by Proposition 14 still holds for the smoothed LuGre model, which is straightforward to assess.

3.3 Mechanical constraints

The mechanical constraints of the actuator are modeled simply as nonlinear spring–damper forces, which becomes active when the piston hits its end-stroke. The hard-stop force representing the upper and lower mechanical constraints, is thus modeled according to

$$f_h(y, v) = K_h \cdot \mu_h(y) + D_h v \cdot \rho_h(y), \tag{3.30}$$

where y and v are the cylinder position and velocity, respectively, K_h is the spring stiffness reflecting the elasticity of the parts in contact (thus, the value of K_h is typically very high), and D_h is a damping coefficient which reflects the plastic deformation that occurs locally in the materials during a hardstop. The functions $\mu_h(y)$ and $\rho_h(y)$, respectively, are a *smooth dead zone function*, and a *smooth indicator function*, which become "active" when the piston y runs into its end-stroke. Both functions $\mu_h(y)$ and $\rho_h(y)$, are plotted in Figure 3.3, and their construction is addressed in the subsection below.

3.3.1 Smooth dead zone function

A smooth dead zone function with unity slope $\mu_k(\cdot)$ and smoothing interval $[-\varepsilon_h, \varepsilon_h]$, may be devised according to

$$\mu_h(y) = \varepsilon_h \cdot g\left(\frac{y - y_{\text{ub}}}{\varepsilon_h}\right) - \varepsilon_h \cdot g\left(\frac{y_{\text{lb}} - y}{\varepsilon_h}\right) \tag{3.31}$$

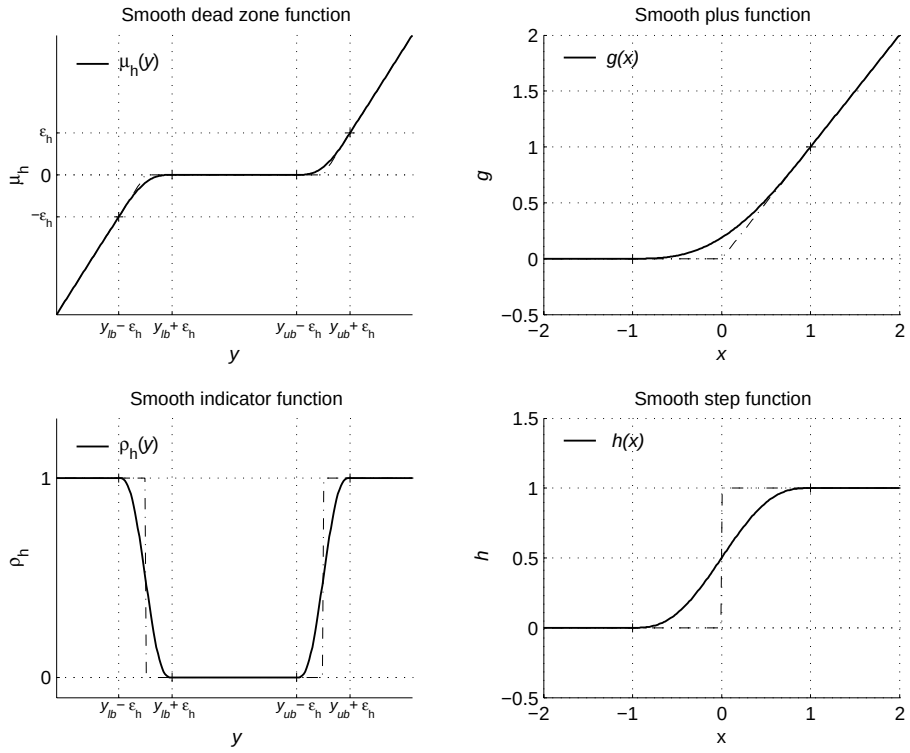

Figure 3.3: The smooth dead zone and indicator functions used for the construction of the hardstop force $f_h(y, v)$.

where y_{ub} and y_{lb} are the upper and lower bound on the cylinder position y, and $g(\cdot)$ is a *smooth plus function* with smoothing interval $[-1, 1]$, like *e.g.* (3.32) below. In (3.31), $\varepsilon_h > 0$ is referred to as the smoothing width at the break points of the dead zone $\mu_h(y)$, and is typically chosen to be small compared to the operating range of y.

The smooth plus function $g(\cdot)$ may be constructed as a piecewise defined polynomial function — referred to as a spline function, which can be made arbitrary smooth by using a polynomial of sufficiently high order. An example of a smooth plus function $g(x)$, which is two times differentiable, is given by

$$g(x) = \begin{cases} x, & x > 1 \\ \frac{3}{16} + \frac{1}{2}x + \frac{3}{8}x^2 - \frac{1}{16}x^4, & |x| \leq 1 \\ 0, & x < -1 \end{cases}, \tag{3.32}$$

where the polynomial coefficients are computed to satisfy smoothness at the connecting points $x = \{-1, 1\}$.

3.3.2 Smooth indicator function

In a similar manner, a smooth indicator function $\rho_h(\cdot)$ with smoothing interval $[-\varepsilon_h, \varepsilon_h]$, may be devised according to

$$\rho_h(z) = \left[h\left(\frac{y - y_{\text{ub}}}{\varepsilon_h}\right) - h\left(\frac{y_{\text{lb}} - y}{\varepsilon_h}\right) \right], \tag{3.33}$$

where $h(\cdot)$ is a *smooth step function* with smoothing interval $[-1, 1]$, which can be constructed as a spline function. A step function $h(x)$ which is two times differentiable, is given by

$$h(x) = \begin{cases} 1, & x > 1 \\ \frac{1}{2} + \frac{15}{16}x - \frac{5}{8}x^3 + \frac{3}{16}x^5, & |x| \leq 1 \\ 0, & x < -1 \end{cases}, \tag{3.34}$$

where the polynomial coefficients are computed to satisfy smoothness at the connecting points $x = \{-1, 1\}$.

Summarizing: the above smooth functions $\mu_h(y)$ and $\rho_h(y)$, are two times differentiable ($\mu_h, \rho_h \in C^2$), and have precisely defined smoothing regions, $[y_{\text{lb}} - \varepsilon_h, y_{\text{lb}} + \varepsilon_h]$ and $[y_{\text{ub}} - \varepsilon_h, y_{\text{ub}} + \varepsilon_h]$, where ε_h is referred to as the smoothing width of the break points.

3.4 Summary

In this chapter we have addressed the modeling of the motion dynamics of the electro-pneumatic clutch actuator. The forces acting on the actuator are composed of the pressures in the two chambers, the clutch compression spring, friction forces in the cylinder and clutch, and mechanical constraints. The clutch compression spring — which is a diaphragm spring with a highly nonlinear position–load characteristic which constitutes the main load of the actuator — is modeled using a parameter-affine parametrization utilizing normalized Gaussian basis functions. A smooth modification of a simple 1st-order dynamic friction model — known as the *LuGre* model — is proposed for the modeling of the resulting friction force in the cylinder actuator and clutch, including the friction force arising from seal deflections. The mechanical constraints of the actuator are modeled simply as nonlinear spring–damper forces, which become active when the piston reach its end-stroke. These spring–damper forces are referred to as hardstop forces, and are constructed using a smooth plus function and a smooth step function, created as customized spline functions.

Chapter 4

Air Dynamics

A detailed derivation — based on simple thermodynamics — of the full-order model of the dynamics of air in a pneumatic cylinder actuator with pressure and temperature as state variables, is given in Appendix C. In Section 4.1, we apply this full-order model for the modeling of the air dynamics of the two actuator chambers. In Section 4.2, we describe the reduced-order isothermal model of the pressure dynamics which we use for our controller-observer design.

4.1 Pressure & temperature dynamics

Figure 4.1[1] shows a schematic diagram of the flow control valve and pneumatic actuator, and indicates the control volume used for the derivation of the air dynamics. We review the assumptions applied to the derivation of the full-state air dynamics, below:

A4.1) At the attainable pressures, air behaves like an *ideal gas* obeying the ideal gas *equation of state* (see Appendix C, Equation (C.1)) with negligible error.

A4.2) The specific heats c_p and c_v of air are assumed to be constant, *i.e.*, not functions of temperature (or pressure[2]). For the attainable temperature range for this application, the deviations are insignificant, see *e.g.* [16, Section 3.7, pp. 182-134].

A4.3) The energy change in the fluid due to elevation is negligible.

A4.4) The thermodynamic properties are uniformly distributed (homogenous) within the control volume, *i.e.*, "perfectly mixed". This is reasonable due to the small dimensions of the system, and lets us avoid a complex distributed problem formulation.

[1] Duplicate of Figure C.1 in Appendix C

[2] For an ideal gas, the internal energy u and enthalpy h vary only with temperature. Furthermore, the specific heats c_v and c_p will in general vary with temperature: $c_v = c_v(T)$ and $c_p = c_p(T)$. However, this temperature dependence is insignificant for this application.

A4.5) The flow through pipes, valves and restrictions in the system is assumed to be *isentropic*[3]. That is, *frictionless* flow is assumed, and the effect of heat transfer on the flow (*adiabatic* flow) is disregarded.

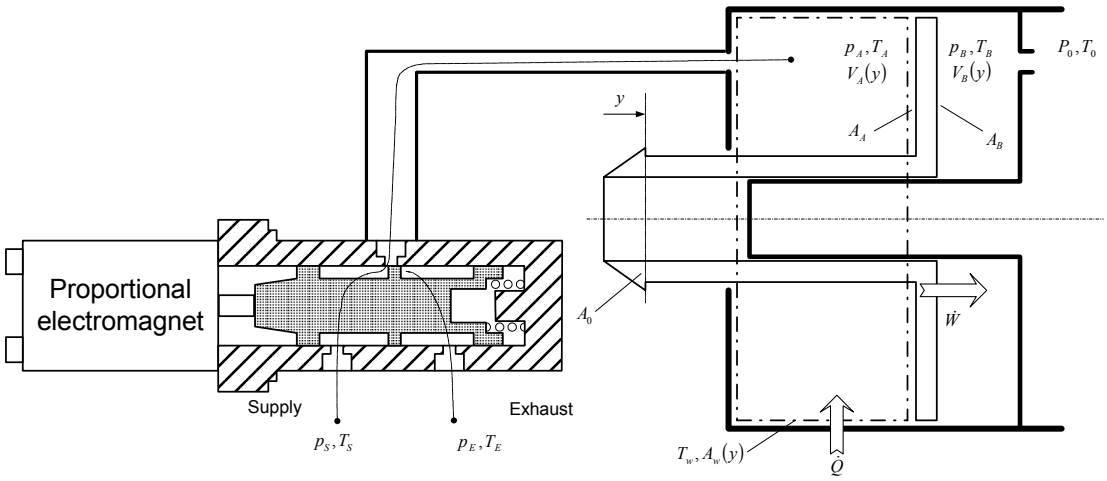

Figure 4.1: Schematic diagram illustrating the pneumatic subsystem and control valve.

4.1.1 Chamber *A*

The thermodynamic state of the air in the pneumatic chamber is given by two state variables, where it is natural to choose the two measurable variables pressure and temperature. The full-order air dynamics of chamber A is according to (C.22) and (C.28), expressed as

$$\frac{dp_A}{dt} = -\frac{\kappa A_A v}{V_A(y)} p_A + \frac{\kappa R T_{in,A}}{V_A(y)} w_{in,A} - \frac{\kappa R T_A}{V_A(y)} w_{out,A} + \frac{(\kappa-1)\cdot H_w A_{w,A}(y)}{V_A(y)} (T_w - T_A), \tag{4.1}$$

$$\begin{aligned} \frac{dT_A}{dt} &= -\frac{(\kappa-1)\cdot A_A v}{V_A(y)} T_A + \frac{(\kappa T_{in,A} - T_A)\cdot R T_A}{p_A V_A(y)} w_{in,A} - \frac{(\kappa-1)\cdot R T_A^2}{p_A V_A(y)} w_{out,A} \\ &\quad + \frac{(\kappa-1)\cdot T_A H_w A_{w,A}(y)}{p_A V_A(y)} (T_w - T_A). \end{aligned} \tag{4.2}$$

An illustration of the control volume used for the derivation is given in Figure 4.1. In the above equations, p_A and T_A are the dynamic pressure and temperature states of the chamber; the parameters κ and R are the ratio of specific heats and gas constant of air, respectively; H_w is the empirical convective heat coefficient, and T_w is the actuator cylinder wall temperature. The chamber volume $V_A(y)$, and the effective wall area of heat transfer $A_{w,A}(y)$, are functions of the actuator position y, given by

[3]This is a common approximation in compressible fluid analysis when the system dimensions are small, [101].

(4.3) and (4.4) below. The variables $w_{in,A}$ and $w_{out,A}$, are the air flows in and out of chamber A, where $T_{in,A}$ is the temperature associated with the inlet flow. For chamber A, the inlet flow $w_{in,A}$ is flow from the supply reservoir, *i.e.*, $T_{in,A} = T_S$.

The chamber volume is given by

$$V_A(y) = V_{A0} + A_A y, \tag{4.3}$$

where V_{A0} is the chamber volume when $y = 0$. The effective wall area of heat transfer is given by

$$A_{w,A}(y) = A_{w,A0} + L_w y, \tag{4.4}$$

where $A_{w,A0}$ is effective area of heat transfer for $y = 0$, and L_w is the inner perimeter of the cylinder wall.

4.1.2 Chamber *B*

The dynamic equations governing chamber B are similar. From (C.22) and (C.28), the full-order air dynamics of chamber B can be expressed as

$$\frac{dp_B}{dt} = \frac{\kappa A_B v}{V_B(y)} p_B + \frac{\kappa R T_{in,B}}{V_B(y)} w_{in,B} - \frac{\kappa R T_B}{V_B(y)} w_{out,B} + \frac{(\kappa - 1) \cdot H_w A_{w,B}(y)}{V_B(y)} (T_w - T_B), \tag{4.5}$$

$$\begin{aligned} \frac{dT_B}{dt} &= \frac{(\kappa - 1) \cdot A_B v}{V_B(y)} T_B + \frac{(\kappa T_{in,B} - T_B) \cdot R T_B}{p_B V_B(y)} w_{in,B} - \frac{(\kappa - 1) \cdot R T_B^2}{p_B V_B(y)} w_{out,B} \\ &\quad + \frac{(\kappa - 1) \cdot T_B H_w A_{w,B}(y)}{p_B V_B(y)} (T_w - T_B), \end{aligned} \tag{4.6}$$

where p_B and T_B are the pressure and temperature states of chamber B, and the chamber volume $V_B(y)$ and the effective wall area of heat transfer $A_{w,B}(y)$, are given by (4.7) and (4.8) below. The variables $w_{in,B}$ and $w_{out,B}$, are the air flows in and out of chamber B, where $T_{in,B}$ is the temperature associated with the inlet flow. For chamber B, the inlet flow $w_{in,B}$ is flow from the exhaust reservoir, *i.e.*, $T_{in,B} = T_E$.

The volume of chamber B is given by

$$V_B(y) = V_{B0} - A_B y, \tag{4.7}$$

where V_{B0} is the chamber volume when $y = 0$. The effective wall area of heat transfer is given as

$$A_{w,B}(y) = A_{w,B0} - L_w y, \tag{4.8}$$

where $A_{w,B0}$ is effective area of heat transfer for $y = 0$.

4.2 Reduced-order isothermal model

For control design purposes, it is desirable to simplify the model so that it is not more detailed than what is required by the particular control task. For the air dynamics of

the pneumatic actuator, the pressures may be considered to be the main variables, since they enter as inputs in the motion dynamics. The temperatures, on the other hand, are usually considered to be less important since both the pressure dynamics and the air flow rate characteristic has a relatively low sensitivity to temperature variations.

There is a strong static coupling between the pressure and temperature of a pneumatic chamber, and it is reasonable to approximate the temperature as a static function of the pressure, with the pressure as the only dynamic state in the model. A common assumption for combustion engine modeling, is to assume that the thermodynamics is governed by a closed thermodynamic process — a polytropic process, where the static relation between the pressure and temperature is given by

$$T = T_0 \left(\frac{p_0}{p} \right)^{\frac{1-n}{n}}, \tag{4.9}$$

which is characterized by the polytropic coefficient n. For example, for an *adiabatic process*, the polytropic exponent is taken as the ratio of specific heats, *i.e.*, $n = \kappa$ ($\kappa = 1.4$ for air). In the simplest case, we may assume an *isothermal process* by taking $n = 1$.

4.2.1 Pressure dynamics of the cylinder chambers

We apply the following assumptions to the full-order air dynamics in order to arrive at the reduced-order model of the pressure dynamics which we use for our control design:

A4.6) We assume isothermal conditions ($n = 1$), which means that the chamber temperature is constant. Furthermore, we assume that all temperatures equals the standardized atmospheric reference condition (T_0) given by the ISO standard [35], *i.e.*, $T_S = T_E = T_A = T_B = T_0$.

A4.7) We assume a constant supply pressure p_S, and assume that the exhaust pressure p_E equals a constant atmospheric pressure. That is, we neglect transients in the supply pressure p_S, and changes in the ambient pressure P_0.

Assumption A4.6 is equivalent to assuming infinite heat transfer with the ambient and reservoir temperatures equal to T_0. The main justification of A4.6 is that the pressures dynamic's sensitivity to temperature changes, is small. Assumption A4.7 is often more questionable with respect to the supply pressure. For example, when the supply reservoir is kept at an approximately constant pressure using an accumulator (buffer tank) with a relay controlled compressor, the supply pressure will exhibit slower fluctuations due to long-lasting air consumption and subsequent refilling by the compressor. Furthermore, it is likely that the supply pressure will exhibit transient variations when subjected to sudden and large changes in the air consumption (w_v).

Applying the above assumptions, the equations describing the pressure dynamics of the two chambers reduce to

$$\dot{p}_A = -\frac{A_A v}{V_A(y)} p_A + \frac{RT_0}{V_A(y)} w_v, \tag{4.10}$$

$$\dot{p}_B = \frac{A_B v}{V_B(y)} p_B + \frac{RT_0}{V_B(y)} w_r, \tag{4.11}$$

where $w_v \triangleq w_{in,A} - w_{out,A}$ is the resultant valve flow rate into chamber A, while $w_r \triangleq w_{in,B} - w_{out,B}$ is the resultant flow into chamber B from atmosphere through the outlet restriction. The modeling of the flow rate characteristics for w_v and w_r, is addressed in Chapter 5 below.

4.3 Summary

In this chapter we have addressed the modeling of the air dynamics of the pneumatic cylinder chambers. The full-order air dynamics with both the pressure and temperature as dynamics states, is reviewed, and common assumptions are applied in order to arrive at the reduced-order isothermal model with only the pressure as a dynamics state.

Chapter 5

Flow Control Valve

In this chapter, we address the modeling of the flow rate characteristic of pneumatic restrictions, and the modeling of the flow control valve. First, we provide a brief characterization of the types of flow, and review the equations which conventionally have been used to describe the flow rate of pneumatic restrictions. A generalized affine parametrization of the flow rate characteristic of pneumatic restrictions is developed, based on a novel parametrization of the pressure ratio function. We utilize this generalized flow rate equation to construct a compact bidirectional[1] model of the flow rate of the orifice restriction of chamber B. The model is general in the sense that it can be applied to accurate modeling of a fixed pneumatic restriction of any type, and the model is piecewise linearly parametrizable in both its tunable parameters as a result of using the proposed flow rate equation.

Next, the generalized flow rate equation is utilized for the parametrization of the static spool–flow characteristic of flow control valves. First, we develop a simple piecewise input-affine model, very similar to the model proposed by Richard [79]. Next, we develop two novel parametrizations of the spool–flow characteristic of flow control valves of sliding type, based on an individual description of the flow through each flow path of the valve, thus, providing improved accuracy compared to existing models for valves with significant leakage flow in the null region, which is the case of most servo valves. The first parametrization is an input-invertible model developed for nonlinear control by a *feedback linearization* approach, and the other is a smooth and piecewise linearly parametrizable model suited for nonlinear and adaptive control by a *backstepping* approach. Finally, we briefly describe the modeling of the dynamics of a proportional valve.

This section is organized as follows: A brief review of flow rate modeling of pneumatic restrictions is given in Section 5.1. The generalized flow rate equation is presented in Section 5.2, and in Section 5.3, we utilize this generalized flow rate equation to construct a model of the bidirectional flow rate of a fixed restriction. Finally, in Section 5.4, we address the modeling of the flow rate characteristic of flow control valves, and the modeling of the dynamics of a proportional valve.

[1] By *bidirectional*, we mean the the model describes flow in both directions, *i.e.*, either positive or negative flow depending on the direction of the pressure drop over the restriction.

5.1 Flow rate modeling–a brief review

5.1.1 Flow characterization

The characterization of the type of flow in a pneumatic component is important with respect to flow rate modeling because it determines the pressure dependence of the flow rate. The type of flow depends upon many factors, but the primary parameter is the dimensionless *Reynolds number*, Re. The Reynolds number can be given as

$$\mathrm{Re} = \frac{\rho v D}{\mu}, \tag{5.1}$$

where the variables ρ, v, and μ, are the density, the average fluid velocity, and the coefficient of viscosity of the fluid, respectively, while $D \triangleq 4A/P$ is the *hydraulic diameter*, where A and P are the area and perimeter of the cross-section, respectively [101]. The flow can be characterized by three flow regimes which are roughly indicated by the value of the Reynolds number: For low Re the flow is smooth and steady (*laminar*), for high Re the flow is fluctuating and agitated (*turbulent*), while for intermediate Re there is a change-over from laminar to turbulent flow which is denoted *transition* flow [101, Chapter 6]. Furthermore, when a fluid flows at speeds in the range of its local speed of sound[2] the effect of density changes becomes significant and the flow is termed *compressible flow*. For gases in general, the speed of sound is low, hence, the effect of compressibility is important in most cases of flow rate modeling of pneumatic components. This means that the type of flow in pneumatic components may range from *compressible turbulent flow* for high Re to *incompressible laminar flow* for low Re.

With respect to flow rate modeling, it is convenient to categorize the restriction in a pneumatic component as either an *orifice restriction*, or a *clearance restriction*, depending on the amount of flow resistance in the flow path. Most pneumatic components such as orifices and short ducts, fall into the category of an *orifice restriction*, which is characterized by relatively *low* flow resistance, where frictional effects are small. Due to the low speed of sound and low viscosity of air, the type of flow through an orifice flow restriction is for most normal operating conditions characterized as *compressible* and *turbulent*. For flow through simple orifices, the effect of friction is usually negligible, hence, when heat transfer is negligible, the flow rate can be accurately described by the theoretically derived equation for isentropic[3] flow of a compressible fluid through a converging nozzle [101, ch. 9]. In these cases, we refer to the flow as *isentropic orifice flow*.

The effect of friction on the flow of a compressible fluid is theoretically difficult to analyze, and deriving an explicit expression for the mass flow rate is only possible for special cases. For high-speed flow in short ducts, it is reasonable to assume adiabatic flow, which results in equations which require numerical iteration for the calculation of the mass flow. For long ducts, however, we may assume isothermal flow which lets us derive an explicit expression for the mass flow, which we refer

[2] The speed of sound of a fluid is given by $a = \sqrt{\kappa RT}$.

[3] When there is no heat transfer (*adiabatic* conditions), *frictionless flow* implies *isentropic* flow.

to as *isothermal compressible flow*. See [101, Chapter 9] for an exact description. In [9], an equation for the flow of a compressible fluid through a capillary passage, referred to as *capillary compressible flow*, is derived for adiabatic flow by assuming that momentum effects in the flow is negligible. This capillary flow equation is for example used in [86], [9], and [14], to model compressible flow through capillaries connected to auxiliary tanks, introduced to improve the performance of pneumatic servo actuators.

Very small clearances, of the type encountered between the sleeve and the spool in a spool valve, or between the piston and the wall of a cylinder actuator, may be referred to as a *clearance restriction*. Since the flow resistance of a clearance restriction is typically very *high*, the flow is for most normal conditions relatively slow and viscous, such that it can be characterized as *incompressible* and *laminar*. In these cases, the flow rate can be described by the common incompressible laminar flow equation, and we refer to the flow as *incompressible laminar flow*.

5.1.2 The isentropic orifice flow equation

The theoretical equation for steady-state air flow through an orifice is based on the assumption of isentropic flow of a compressible fluid through a nozzle. In the conventional equation used to describe the flow rate of pneumatic components [9], a discharge coefficient is introduced to account for flow contractions. The resulting equation is referred to as the *isentropic orifice flow equation*, which can be expressed as

$$w = C_d A_r \sqrt{\frac{\kappa}{R} \cdot \left(\frac{2}{\kappa+1}\right)^{\frac{\kappa+1}{\kappa-1}}} \cdot \omega_0(p_l/p_h) \cdot \frac{p_h}{\sqrt{T_h}}. \tag{5.2}$$

Here, w is the mass flow through the restriction, T_h and p_h are the upstream temperature and pressure, respectively, and p_l is the downstream pressure (the subscripts "h" and "l" refer to the *high* and the *low* pressure reservoirs, respectively). The function $\omega_0(\cdot)$ is a normalized function of the pressure ratio p_l/p_h over the restriction — referred to as the *pressure ratio function* — given by (5.3) below. The parameter C_d is the discharge coefficient, which is a lumped parameter that accounts for flow contractions and possibly minor frictional effects, A_r is the smallest cross-sectional area of the restriction, κ is the ratio of specific heats, and R the gas constant of the fluid ($\kappa = 1.4$ and $R = 287\,\mathrm{J/(kg\,K)}$ for air). See *e.g.* [4] for a thorough discussion of the discharge coefficient of pneumatic components.

The theoretically derived *isentropic pressure ratio function* $\omega_0(\cdot)$ for flow through an isentropic restriction is

$$\omega_0(r) \triangleq \frac{w}{w^*} = \begin{cases} 1 & , r \in [0, B_0] \\ \sqrt{\dfrac{r^{\frac{2}{\kappa}} - r^{\frac{\kappa+1}{\kappa}}}{\dfrac{\kappa-1}{2} \cdot \left(\dfrac{2}{\kappa+1}\right)^{\frac{\kappa+1}{\kappa-1}}}} & , r \in \langle B_0, 1] \end{cases}, \tag{5.3}$$

where

$$B_0 \triangleq \left(\frac{p_l}{p_h}\right)^* = \left(\frac{2}{\kappa+1}\right)^{\frac{\kappa}{\kappa-1}} \tag{5.4}$$

denotes the *critical pressure ratio* of the isentropic restriction, and $B_0 = 0.528$ for air[4]. For pressure ratios below critical, *i.e.* $p_l/p_h < B_0$, a further lowering of the downstream pressure p_l does not result in an increase in the mass flow rate w, thus the flow is said to be *choked*. Under choked conditions, the velocity in the smallest restriction area (A_r) equals the speed of sound. See *e.g.* Blackburn *et al.* [9, Section 3.3], or White [101, Chapter 9], for the derivation of the isentropic orifice flow equation given by (5.2)–(5.4).

5.1.3 ISO standardized orifice flow equation

In the ISO standard [35], an approximation to the above isentropic orifice flow equation has been standardized for the determination of the flow rate characteristics of pneumatic components. The *standardized orifice flow equation* is given as

$$w = \rho_0\sqrt{T_0}C \cdot \omega_e(p_l/p_h)\,\frac{p_h}{\sqrt{T_h}}, \tag{5.5}$$

where the capacity of the restriction is characterized by the sonic conductance C, defined at a common *reference condition*, given here by the density ρ_0 and the temperature T_0. The isentropic pressure ratio function $\omega_0(r)$ (5.3) is approximated by the simpler elliptic function

$$\omega_e(r) \triangleq \begin{cases} 1 & ,r \in [0,B] \\ \sqrt{1-\left(\frac{r-B}{1-B}\right)^2} & ,r \in \langle B,1] \end{cases}, \tag{5.6}$$

where B is the effective critical pressure ratio of the restriction[5]. The sonic conductance C, and the critical pressure ratio B, are viewed as mechanistic parameters in the sense that they have clear physical meanings, and their values can be derived from physical laws in the ideal case of isentropic flow. The sonic conductance corresponding to the conventional orifice flow equation (5.2) is given as

$$C = \frac{C_d A_r}{\rho_0\sqrt{T_0}}\sqrt{\frac{\kappa}{R}\left(\frac{2}{\kappa+1}\right)^{\frac{\kappa+1}{\kappa-1}}}, \tag{5.7}$$

while the critical pressure ratio for isentropic flow is given by (5.4).

The definition of the sonic conductance originates from the way the flow rate of pneumatic components is traditionally measured, and is in essence a measure of the flow capacity of the pneumatic component. The flow rate of pneumatic components is usually measured in terms of the volumetric flow rate q [m^3/s] ($w = \rho q$

[4]The superscript asterisk ($*$) denotes *sonic*, or *critical*, flow conditions.

[5]For air with $B = B_0$, the deviation between $\omega_e(r)$ and $\omega_0(r)$ is less than 0.2%, thus, hardly significant.

[kg/s]), with the outlet and the upstream temperature at atmospheric conditions[6]. The sonic conductance C [$\mathrm{m}^3/(\mathrm{s\,Pa})$] is defined as the proportionality constant of the volumetric flow rate for choked flow, measured at a defined common *reference condition* for air, which by the ISO standard is $\rho_0 \triangleq 1.185\,\mathrm{kg/m}^3$, $T_0 \triangleq 293\,\mathrm{K}$ and $P_0 \triangleq 1.00 \times 10^5\,\mathrm{Pa}$. That is, the volumetric flow rate can be expressed as $q = C \cdot p_h$ when the flow is choked, and $T_h = T_0$, $p_l = P_0$, $T_l = T_0$ and $\rho_l = \rho_0$.

The value of the critical pressure ratio B indicates the pressure ratio for which choked flow occurs. Due to frictional effects, the value of B for most pneumatic components is lower than the isentropic value given by (5.4). This means that when B is allowed to take values below its isentropic value, the validity of the *standardized orifice flow equation* (5.5) is broader than just describing *isentropic orifice flow*, in the sense that it is able to describe (non-isentropic) frictional effects to a certain extent. For example, when modeling orifice flow according to (5.5), increased frictional effects due to *e.g.* complex geometry of the flow path, are captured as a reduction in the effective critical pressure ratio parameter B.

5.1.4 Incompressible leakage flow equation

The transition from laminar to turbulent flow is indicated by the critical Reynolds number, which is approximately $\mathrm{Re}_{crit} \approx 2300$, but may vary significantly with the geometry of the restriction [101, ch. 6]. For low Reynolds numbers, $\mathrm{Re} \ll \mathrm{Re}_{crit}$, the flow is characterized as incompressible and laminar. Here, we refer to this type of flow as *incompressible laminar flow*, which can be modeled simply as

$$w_c = \rho_0 C_c \cdot (p_h - p_l)\,, \tag{5.8}$$

where w_c is the clearance flow, and C_c is a lumped *clearance flow constant*, which is a characteristic of the geometry of the clearance, the viscosity of the fluid, and the wall roughness.

It is interesting to note that we can rewrite (5.8) as

$$w_c = \rho_0 C_c \cdot \omega_c(p_l/p_h)\, p_h, \tag{5.9}$$

with the pressure ratio function defined as

$$\omega_c\,(r) \triangleq 1 - r, \quad r \in [0, 1]\,, \tag{5.10}$$

which is in a form similar to the orifice flow equations (5.2) and (5.5).

5.1.5 Compressible leakage flow equations

For high-speed flow through small clearances, the effect of compressibility may be significant so that the assumption of incompressible flow is invalid. The analysis of compressible flow with friction is complicated, and a simple explicit equation for high-speed flow of a compressible viscous fluid through small clearances does not

[6] Due to the high compressibility of gases, the density of the gas and consequently the measured volumetric flow rate, varies significantly with the pressure at which it is measured.

exist. For example by assuming adiabatic conditions, we have to iteratively solve for the flow rate. However, by assuming isothermal conditions, we are able to derive an explicit equation for the flow [101]. This expression can be written as

$$w_c = A_r \sqrt{\frac{1-(p_l/p_h)^2}{fL/D + 2\ln p_h/p_l}} \frac{p_h}{\sqrt{T_h}}, \tag{5.11}$$

where A_r is the cross-sectional restriction area, L is the length of the clearance, $D \triangleq 4A_r/P$ is the hydraulic diameter, and f is the viscous friction coefficient. An interesting point is that when the length is long compared to the hydraulic diameter D, this equation has a form very similar to the elliptic pressure ratio function with $B = 0$. That is,

$$fL \gg D \Longrightarrow fL/D + 2\ln p_h/p_l \approx fL/D, \tag{5.12}$$

which is reasonably accurate for pressure ratios p_l/p_h above a certain limit, for example, $p_l/p_h > 0.1$. In this case, by defining $C_c \triangleq A_r\sqrt{D/(fL)}/\rho_0$ as a lumped conductance parameter, we have

$$w_c \approx \rho_0 C_c \sqrt{1-(p_l/p_h)^2} \frac{p_h}{\sqrt{T_h}}. \tag{5.13}$$

Hence, the equation for isothermal frictional compressible flow becomes identical to the standardized orifice flow equation with $B = 0$.

An alternative analytic solution of frictional compressible flow can be derived by assuming that the momentum effects of the fluid flow are negligible. This equation is used to describe compressible flow through capillary passages, which is referred to as *compressible capillary flow*. See *e.g.* Blackburn *et al.* [9]. This equation can be expressed in the form

$$w_c = \rho_0 C_c \cdot \left(p_h^2 - p_l^2\right) \frac{1}{T_h}, \tag{5.14}$$

where C_c is a lumped *capillary flow constant* which depends on the geometry of the restriction.

5.2 Generalized flow equation

In accordance with the description in the previous section[7], the flow through pneumatic restrictions in general, can be described by an equation in the form

$$w = \rho_0 \sqrt{T_0} C \cdot \omega\left(\frac{p_l}{p_h}\right) \frac{p_h}{\sqrt{T_h}}, \quad p_h \geq p_l, \tag{5.15}$$

where w is the mass flow rate through the restriction, T_h and p_h are the upstream temperature and pressure, respectively, and p_l is the downstream pressure. Subscripts "h" and "l" refers to the *high* and the *low* pressure reservoirs, respectively.

[7]For the reader's convenience, this section includes some duplication of the equations from Section 5.1.

The capacity of the restriction is characterized by its conductance C, defined at a common *reference condition* for air, given by the density ρ_0 and the temperature T_0. The *pressure ratio function* $\omega(r) \in [0,1]$, is a normalized function of the pressure ratio $r = p_l/p_h$ over the restriction, which determines the pressure dependence of the flow rate. Thus, the parametrization of the pressure ratio function $\omega(r)$ in the above flow rate equation (5.15), is the main factor which determines the accuracy of the flow rate description of a given restriction.

5.2.1 Pressure ratio function

The pressure characteristic of a restriction is strongly linked to the type of restriction, more precisely, the degree of flow resistance in the flow path. We may roughly group pneumatic restrictions into two categories (see Section 5.1 above):

Orifice restrictions Restrictions which are characterized by relatively low flow resistance where frictional effects are negligible small, like short pipes or ducts.

Clearance restrictions Restrictions characterized by high flow resistance where frictional effects are important, *e.g.* small clearances of the type encountered between the sleeve and the spool in valves, or between the piston and the wall in a cylinder actuator.

The flow rate characteristic of orifice restrictions is in most cases accurately described by the ISO standardized orifice flow equation (5.5), which is an approximation of the theoretically derived equation for isentropic compressible flow (5.2). This equation is in the form (5.15), with theoretically derived isentropic pressure ratio function $\omega(r)$ approximated as a quarter of an ellipse according to

$$\omega_e(r) \triangleq \begin{cases} 1 & , r \in [0, B] \\ \sqrt{1 - \left(\frac{r-B}{1-B}\right)^2} & , r \in \langle B, 1] \end{cases}, \tag{5.16}$$

where B is the effective critical pressure ratio of the restriction. This equation is valid when the flow can be characterized as approximately isentropic, *i.e.*, when frictional effects are small.

For clearance restrictions, the effect of friction is important, and the accuracy of the elliptic pressure ratio function is reduced. When the flow resistance is *high*, the flow is for normal conditions relatively slow and viscous, such that the flow rate can be described by the common incompressible laminar flow equation (5.8). This equation can be rewritten in the form (5.15) with the pressure ratio function defined as

$$\omega_c(r) \triangleq 1 - r, \quad r \in [0,1], \tag{5.17}$$

where the subscript "c" refers to "clearance" flow.

Here, we propose a piecewise parameter-affine parametrization of the pressure ratio function $\omega(r)$, which has a validity that encompasses the full range of possible restriction types, from simple orifice restrictions to small clearance restrictions. This

parametrization is based on the elliptic pressure ratio function (5.16), and the linear pressure ratio function (5.17) for incompressible laminar flow. This generalized piecewise parameter-affine pressure ratio function is given as

$$\omega_a(r) = \Omega_0(r) + b \cdot \Omega_1(r, \operatorname{sgn} b), \quad b \in [-1, 1], \tag{5.18}$$

where $r \in [0, \infty\rangle$ is the pressure ratio, and b is a critical pressure ratio–like parameter which uniquely determines the characteristic of $\omega_a(r)$. The basis functions $\Omega_0(r)$ and $\Omega_1(r, \operatorname{sgn} b)$, plotted in Figure 5.1, are constructed from the incompressible flow pressure ratio function (5.17), and the upper and lower bounds of the elliptic pressure ratio function (5.16), according to

$$\Omega_0 \triangleq \begin{cases} \sqrt{1-r^2} & , r \in [0,1] \\ 0 & , r > 1 \end{cases} \tag{5.19}$$

$$\begin{aligned} \Omega_1(r,+1) &\triangleq -\Omega_0(r) + \begin{cases} 1 & , r \in [0, B_0] \\ \sqrt{1-\left(\frac{r-B_0}{1-B_0}\right)^2} & , r \in \langle B_0, 1] \\ 0 & , r > 1 \end{cases} \\ \Omega_1(r,-1) &\triangleq \Omega_0(r) - \begin{cases} 1-r & , r \in [0,1] \\ 0 & , r > 1 \end{cases} \end{aligned} \tag{5.20}$$

where $B_0 \triangleq (p_l/p_h)^* = 0.528$ is the isentropic critical pressure ratio for air. Note that $\Omega_0(r)$ and $\Omega_1(r, \operatorname{sgn} b)$ are defined also for pressure ratios $r > 1$, which means that $\omega_a(r)$ is defined for $\forall r \in [0, \infty\rangle$. In Figure 5.1, the basis functions $\Omega_0(r)$ and $\Omega_1(r, \operatorname{sgn} b)$, and the affine function $\omega_a(r)$, are plotted for different values of the critical pressure ratio–like parameter b. The upper darkest shaded area in the figure represents the range of the elliptic function $\omega_e(r)$ for $B \in [0, B_0]$, which is the same as the range of the parameter-affine function $\omega_a(r)$ for $b \in [0, 1]$, where $\Omega_1 = \Omega_1(r, +1)$. The lower shaded area represents the range of ω_a for $b \in [-1, 0]$, where $\Omega_1 = \Omega_1(r, -1)$.

The main achievement of the introduced parametrization of the pressure ratio function is that we obtain a generalized flow equation which can be used to accurately model pneumatic restrictions ranging from simple orifices to small clearances. Another important property of this parametrization is the piecewise affinity of its only tunable parameter b. This is advantageous in the case when no *a priori* knowledge about b exists, and we may need to design an adaptive controller which identifies the pressure characteristic of the restriction on-line.

Remark 15 *In this work, we have put effort into deriving a parameter-affine parametrization of the pressure ratio function with as few parameters as possible. An alternative and simple approach, would be by use of more general basis functions, like* e.g. *normalized Gaussian functions, which can be made arbitrary accurate for any pressure ratio characteristic simply by increasing the number of basis functions, thus also the number of parameters.*

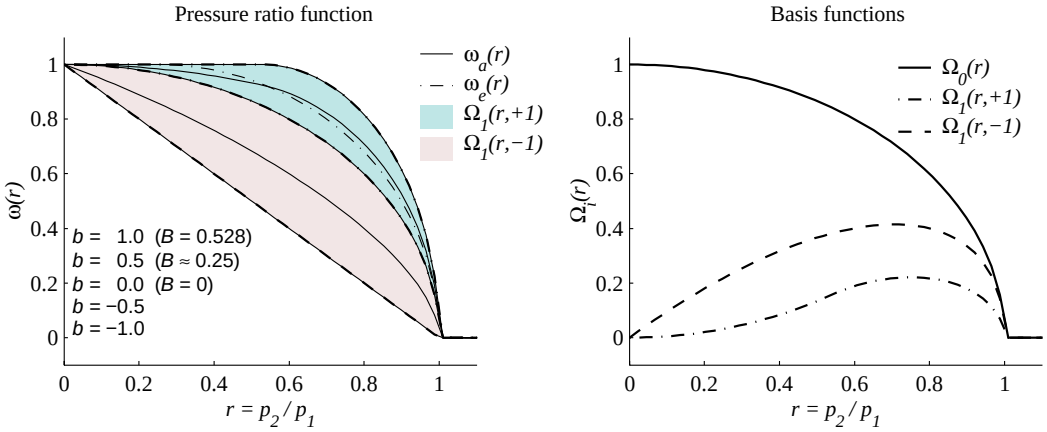

Figure 5.1: Left: The properties of the piecewise parameter-affine pressure ratio function $\omega_a(r)$ are compared with the elliptic function $\omega_e(r)$. The upper darkest shaded area represents the output range of $\omega_e(r)$, while the entire shaded area represents the range of $\omega_a(r)$. Right: The corresponding basis functions.

5.3 Outlet restriction

The flow path from chamber B to exhaust E (atmosphere) of the pneumatic clutch actuator, can be modeled as a fixed orifice restriction, like it is illustrated in Figure 5.2. We define the flow rate as positive ($w_r > 0$) for filling of chamber B, *i.e.*, flow from exhaust E to chamber B for $p_E > p_B$, and negative ($w_r < 0$) for venting to atmosphere when $p_E < p_B$.

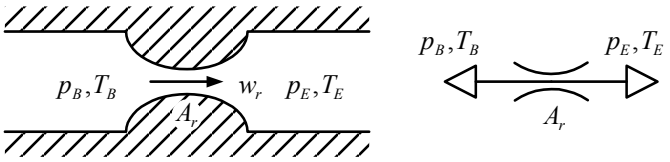

Figure 5.2: Schematic drawing of the outlet restriction of chamber B.

Utilizing the generalized flow equation (5.15), the outlet restriction of chamber B can be modeled according to $w_r = w_{in} - w_{out}$ as

$$w_r = \rho_0 \sqrt{T_0} C_r \cdot \omega_r\left(p_B/p_E\right) \frac{p_E}{\sqrt{T_E}} - \rho_0 \sqrt{T_0} C_r \cdot \omega_r\left(p_E/p_B\right) \frac{p_B}{\sqrt{T_B}}. \tag{5.21}$$

where C_r is the conductance of the restriction, referred to the standardized *reference condition* of air, given by ρ_0 and T_0, and the physical variables p_E, T_E, p_B, T_B are the pressures and temperatures of the exhaust E and chamber B, as indicated in Figure 5.2. The obtain a model which is piecewise linearly parametrizable, the pressure ratio function $\omega_r(r)$ is modeled according to (5.18)–(5.20) with the critical pressure ratio–like parameter b_r.

5.3.1 Linear parametrization

In preparation for off-line parameter estimation, or on-line adaptive control where both C_r and b_r are allowed to be unknown, we note that by substituting with the expression for $\omega_r(\cdot)$ given by (5.18), the flow rate model can be expressed in the parameter-affine vector form

$$w_r = \boldsymbol{\theta}_r^T \cdot \boldsymbol{\phi}_r(p_E, T_E, p_B, T_B), \tag{5.22}$$

by defining the parameter and regressor vector as

$$\boldsymbol{\theta}_r \triangleq \rho_0\sqrt{T_0}\begin{bmatrix} C_r \\ C_r b_r \end{bmatrix}, \quad \boldsymbol{\phi}_r \triangleq \begin{bmatrix} \Omega_0(p_B/p_E)\frac{p_E}{\sqrt{T_E}} - \Omega_0(p_E/p_B)\frac{p_B}{\sqrt{T_B}} \\ \Omega_1(p_B/p_E, \operatorname{sgn} b_r)\frac{p_E}{\sqrt{T_E}} - \Omega_1(p_E/p_B, \operatorname{sgn} b_r)\frac{p_B}{\sqrt{T_B}} \end{bmatrix}. \tag{5.23}$$

Viewing the temperatures T_E, T_B, and pressure p_E as physical variables (that are not tunable), the above model is said to be piecewise linearly parametrized in its tunable parameters (C_r and b_r).

Remark 16 *Using the generalized pressure ratio function (5.18)–(5.20), the leakage between the two chambers can in most cases be accurately modeled as a fixed restriction with the same model as the outlet restriction of chamber* B, i.e., *according to (5.21).*

5.3.2 Simplified partially linear parametrization

With the objective of simplifying for control, we apply assumptions A4.6–7 from the derivation of the pressure dynamics in Section 4.2 also for the flow rate model. That is, we take all temperatures and the exhaust pressure to be equal to the reference condition, *i.e.*, we take $T_B = T_E = T_0$, and $p_E = P_0$. With application of assumptions A4.6–7, the resulting flow rate model of a pneumatic restriction can be expressed in the partially parameter-affine form

$$w_r = \rho_0 C_r \cdot \psi_r(p_B), \tag{5.24}$$

by defining the *restriction flow function*

$$\psi_r \triangleq \omega_r\left(\frac{p_B}{P_0}\right)P_0 - \omega_r\left(\frac{P_0}{p_B}\right)p_B. \tag{5.25}$$

This is a convenient formulation in situations where reasonable accurate estimates of b_r (or alternatively B_r) exists, while the capacity C_r is uncertain. The nonlinear flow function $\psi_r(p_B)$ is then known, and the parameter C_r appears in an affine form, which makes the model particularly suited for parameter estimation or adaptive control designs where C_r is estimated on-line.

For the clutch actuator, we know that the outlet restriction of chamber B can be characterized as an orifice restriction since it is a short duct with relatively low flow resistance, but in many cases, we do not know *a priori* the exact dimensions and geometry of the restriction. Then we know that the flow rate with reasonable accuracy can be modeled with $b_v = 1$, but the flow capacity C_r of the restriction is uncertain.

Remark 17 *For the modeling of the outlet restriction in the form (5.24), we do not utilize the parameter-affinity of the proposed parametrization of the pressure ratio function given by (5.18)–(5.20). In this case, the simpler standardized elliptic function given by (5.16) could be used, which for most orifice restrictions achieves approximately the same degree of accuracy. For clearance restrictions, however, the elliptic function (5.16) is less accurate than the affine parametrization (5.18)–(5.20).*

5.4 Proportional valve

A three-way valve is used to control the flow to the pneumatic chamber A. Figure 5.3 shows a schematic cross-sectional drawing of a proportional spool valve, which is connected to chamber A, the supply reservoir S, and the exhaust E. The flow through the valve consists of the flow from the supply to chamber A, and the flow from chamber A to exhaust. Additionally, there will be internal leakage from the supply port through the clearances in the valve to the exhaust port.

Our objective is to obtain a model of the flow rate characteristic of the valve — referred to as the *spool–flow characteristic* — which is suitable for control. The model should also be able include any pipe flow resistance in the description, when the valve is connected to the different chambers through short pipes or ducts.

In this section, the generalized flow rate equation is utilized for the parametrization of the static spool–flow characteristic of flow control valves. First, we describe the presumptions we have applied to the flow rate modeling. Then, we present a simple piecewise input-affine flow model which is based on the assumption of an ideally proportional spool–flow characteristic, similar to the model most researchers use for control design because it is easily inverted. Then, we propose two novel models for control with improved accuracy. The first of these is developed for subsequent application of nonlinear control by a feedback linearization approach, which is input-invertible. The other is developed for nonlinear and adaptive control by a backstepping approach, and is smooth and piecewise linearly parametrizable in all its tunable parameters. Finally, we address the modeling of the dynamics of the spool position of proportional valves.

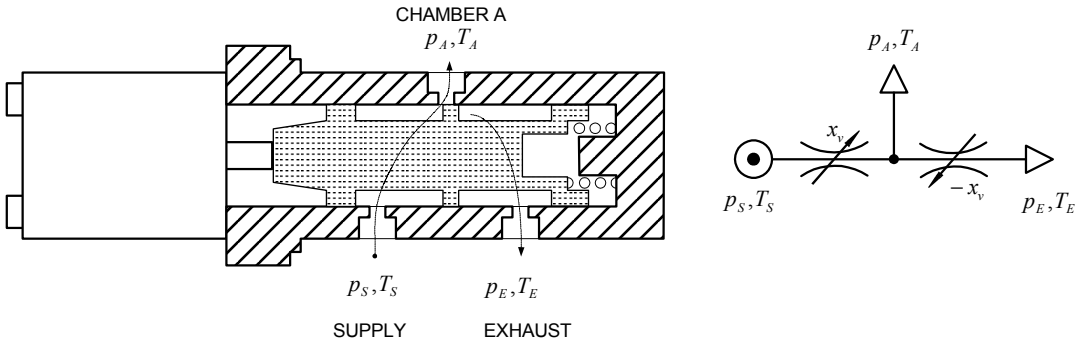

Figure 5.3: Schematic drawing of cross-section and flow diagram of the spool valve.

The position of the spool determines the cross-sectional area of each valve orifice. In the following, we refer to the spool position x_v as the input to the flow rate model, and we will refer to a normalized spool position $x_v \in [-1, 1]$, such that $x_v = 1$ refers to the valve fully open to supply, and $x_v = -1$ fully open to exhaust. We denote the flow from the supply reservoir to chamber A, as supply flow w_S, and the flow from chamber A to exhaust, as exhaust flow w_E. Furthermore, we apply the following (obvious) assumptions for the valve models:

A5.1) The concerned pressures are always greater than absolute vacuum:

$$p_E, p_A, p_S > 0.$$

A5.2) The chamber pressure is always within the pressure range of the supply and the exhaust pressure, *i.e.*, $p_S \geq p_A \geq p_E$.

A5.3) The geometry of the supply and the exhaust orifices are symmetrical about the center position of the spool ($x_v = 0$).

Assumption A5.1 is trivial, and is physically always satisfied. Assumption A5.2 is for most pneumatic applications practically always satisfied, and its implication is that it simplifies the formulation of a valve model because the supply and the exhaust flows can then always be regarded as positive flows, *i.e.*, $w_S, w_E \geq 0$. Assumption A5.3 is a reasonable assumption for a three-way spool valve, and is convenient since it means that the parameters of each valve orifice can then be regarded as identical.

We have pursued to develop models which are (partially) *mechanistic*, in the sense that they are based on physical laws and states, with a set of parameters that are physically meaningful, and which are functions of all the physical variables. To achieve this, and be able to develop an accurate description of the leakage in the null region of the valve, we model the flow paths individually, and then combine the modeled supply flow w_S and the exhaust flow w_E to obtain a resulting valve flow model according to $w_v = w_S - w_E$. Following this approach, it is straightforward to extend the models to describe the flow rate of, for example, of a four-way, or a five-way valve. For most control designs, the physical variables p_S, T_S, p_E, T_E and T_A are assumed to be constant parameters. In accordance with this, and for notational simplicity, we will treat only p_A and x_v as function arguments in the following discussion, so that the valve model can be written in the form $w_v = g(p_A, x_v)$.

5.4.1 A piecewise input-affine flow model

Assuming that the air flow through the valve is composed of a *leakage* term w_l which is independent of x_v, and an orifice flow term w_o which is piecewise proportional to x_v, the valve flow model can be expressed in the *piecewise input-affine* form

$$w_v = w_l + w_o = g_l(p_A) + g_o(p_A, \operatorname{sgn} x_v) \cdot x_v. \tag{5.26}$$

This form is justified for the ideal case when the valve port orifices have *zero overlap*, and where the leakage is due to a *fixed clearance* in the valve.

The leakage flow w_l is assumed to be independent of x_v, and is therefore characterized by a constant *leakage conductance* (or capacity) C_l, and is composed of the leakage into chamber A from the supply port, and the leakage out of chamber A through the exhaust port. Using the *generalized flow equation* (5.15), the resulting *leakage flow* can be modeled according to $w_l = w_{l,S} - w_{l,E}$ as

$$w_l = \rho_0 \sqrt{T_0} C_l \cdot \omega_l \left(p_A/p_S\right) \frac{p_S}{\sqrt{T_S}} - \rho_0 \sqrt{T_0} C_l \cdot \omega_l \left(p_E/p_A\right) \frac{p_A}{\sqrt{T_A}}, \tag{5.27}$$

where the pressure ratio function $\omega_l(r)$ is given by (5.18)-(5.20) with the parameter b_l.

For the modeling of the spool dependent orifice flow term $w_o = g_o\left(p_A, \operatorname{sgn} x_v\right) \cdot x_v$, we express the variable conductance c_o of each valve orifice as piecewise proportional to x_v according to

$$\begin{aligned} c_o &= \begin{cases} C_o x_v, & x_v \geq 0, \\ 0, & x_v < 0 \end{cases} \\ &= C_o x_v \cdot \chi\left[x_v \geq 0\right], \end{aligned} \tag{5.28}$$

where C_o is the *orifice conductance* for the valve orifices fully open ($x_v = \{-1, 1\}$). For a compact notation, we have defined the *indicator function* $\chi[X]$ of the event X (as in [94]):

$$\chi[X] \triangleq \begin{cases} 1, & X \text{ is true}, \\ 0, & \text{else}. \end{cases} . \tag{5.29}$$

Like the leakage flow w_l, the *orifice flow* w_o is modeled according to $w_o = w_{o,S} - w_{o,E}$ as

$$w_o = \rho_0 \sqrt{T_0} C_o \cdot \omega_o \left(p_A/p_S\right) \frac{p_S}{\sqrt{T_S}} \chi\left[x_v \geq 0\right] \cdot x_v + \rho_0 \sqrt{T_0} C_o \cdot \omega_o \left(p_E/p_A\right) \frac{p_A}{\sqrt{T_A}} \chi\left[x_v < 0\right] \cdot x_v, \tag{5.30}$$

where the pressure ratio function $\omega_o(r)$ is given by (5.18)-(5.20) with the parameter b_o.

Remark 18 *The pressure ratio functions $\omega_l(r)$, and $\omega_o(r)$, can alternatively be modeled by the simpler elliptic function (5.16), see Remark 17.*

The resulting flow rate model is given as $w_v = w_l + w_o$, and can be expressed in the partially parameter-affine form

$$w_v = \rho_0 \sqrt{T_0} C_l \cdot \psi_l\left(p_A\right) + \rho_0 \sqrt{T_0} C_o \cdot \psi_o\left(p_A, \operatorname{sgn} x_v\right) \cdot x_v, \tag{5.31}$$

by defining the *leakage flow function* ψ_l and the *orifice flow function* ψ_o as

$$\psi_l \triangleq \omega_l \left(p_A/p_S\right) \frac{p_S}{\sqrt{T_S}} - \omega_l \left(p_E/p_A\right) \frac{p_A}{\sqrt{T_A}}, \tag{5.32}$$

$$\psi_o \triangleq \omega_o \left(p_A/p_S\right) \frac{p_S}{\sqrt{T_S}} \chi\left[x_v \geq 0\right] + \omega_o \left(p_E/p_A\right) \frac{p_A}{\sqrt{T_A}} \chi\left[x_v < 0\right]. \tag{5.33}$$

The above piecewise input-affine model (5.31) is in a form suitable for a Lyapunov-based adaptive control design, where the two characteristic parameters C_l and C_o,

that appear linearly, can be identified on-line by the adaptive controller if b_l and b_o are known. The parameters ρ_0, and T_0 are known parameters, defined by the ISO standard.

Remark 19 *Since the parameters b_l and b_o of the pressure ratio functions $\omega_l(r)$ and $\omega_o(r)$ appear in an affine form, the flow model (5.31) is linearly parametrizable with respect to all its tunable parameters (C_l, b_l, C_o, b_o). However, in most cases, the accuracy of the input-affine model is crude, and the effect of tuning b_l and b_o is more or less negligible on the overall accuracy of the model. Hence, it makes little sense to adapt these parameters on-line. Typical choices which usually provides sufficient accuracy, are $b_l = 0$ and $b_o = 1$, which are equivalent to the elliptic pressure ratio function $\omega_e(r)$ with critical pressure ratios $B_l = 0$ and $B_o = B_0 = 0.528$, respectively.*

5.4.2 An input-invertible valve flow model

A valve flow model in the input-affine form (5.26) which is used by many researchers, is convenient for control because the model is explicitly input-invertible, *i.e.*, we are able to solve for the input as an explicit function of the output, $x_v = g^{-1}(w_v)$. A model which is input-invertible, is not required for a control design, but for certain nonlinear control techniques, such as feedback linearization, it facilitates the design and implementation of the controller.

For most flow control valves, the constraints that are imposed by limiting the model to the piecewise input-affine form (5.26) significantly limits the accuracy of the model, particularly in the null region of the valve where the flow rate exhibits a significant nonlinear dependence on x_v. Improved accuracy can be obtained by modeling the variable orifice conductance of the individual supply and exhaust port as a nonlinear function of the spool position x_v, *i.e.*, $c_{o,S}(x_v)$ and $c_{o,E}(x_v)$, respectively. Due to symmetry of the valve, the conductance function can be given by a single function according to $c_{o,S} = c_o(x_v)$, and $c_{oE} = c_o(-x_v)$. In this section, we introduce a parametrization of the variable conductance which incorporates an overlap X_δ (or overlap for negative values, $X_\delta < 0$), and a smoothed overlap between the supply and the exhaust flows in the null region, defined by $[-X_k, X_k]$. The model is in the form

$$w_v = w_l + w_o = g_l(p_A) + g_o(p_A, x_v), \tag{5.34}$$

where the orifice flow function $g_o(p_A, x_v)$ is explicitly input-invertible, *i.e.*, we can solve for $x_v = g_o^{-1}(p_A, w_o)$. The leakage flow term $w_l = g_l(p_A)$ is identical to the leakage term of the input-affine model, given by (5.27). Utilizing the generalized flow equation (5.15) in combination with a model of the variable orifice conductance, we are able to accurately describe the orifice flow term w_o over the full range of spool positions. The resulting orifice flow is modeled according to $w_o = w_{o,S} - w_{o,E}$, and is then given as

$$w_o = \rho_0\sqrt{T_0}\cdot c_o(x_v)\,\omega_o\left(\frac{p_A}{p_S}\right)\frac{p_S}{\sqrt{T_S}} - \rho_0\sqrt{T_0}\cdot c_o(-x_v)\,\omega_o\left(\frac{p_E}{p_A}\right)\frac{p_A}{\sqrt{T_A}}, \tag{5.35}$$

where $\omega_o(\cdot)$ is the pressure ratio function of the orifice flow w_o, and $c_o(\cdot)$ is the orifice conductance function of each flow path as a function of the position of the spool. The pressure ratio function $\omega_o(\cdot)$ is parametrized according to (5.18)–(5.20), which is uniquely characterized by the parameter b_o. The parametrization of the variable orifice flow conductance $c_o(\cdot)$ is discussed in the subsection below.

Orifice flow conductance function

The orifice flow conductance function $c_o = c_o(x_v)$ is conveniently modeled in the form

$$c_o = C_o \cdot \mu(x_v), \tag{5.36}$$

where $\mu(x_v) \in [0,1]$ so that C_o is the *orifice flow conductance* for the orifice port fully open, *i.e.*, $x_v = \{-1, 1\}$. The normalized conductance function $\mu(x_v)$ is modeled as a customized spline function, composed of a linear and a quadratic polynomial term according to

$$\mu = \begin{cases} n_0 + n_1 x_v, & x_v \in \langle X_k, 1] \\ m_0 + m_1 x_v + m_2 x_v^2, & x_v \in [-X_k, X_k] \\ 0, & x_v \in [-1, -X_k\rangle \end{cases}. \tag{5.37}$$

The conductance function $c_o(x_v)$, and its normalized function $\mu(x_v)$, are plotted in Figure 5.4. The function $\mu(x_v)$ is uniquely characterized by the overlap parameter X_δ,and the junction knot X_k, together with the requirement that $\mu(x_v)$ be both continuous and smooth at $-X_k$ and continuous at X_k. The parameter X_δ represents the actual *overlap* of the valve, while the parameter X_k determines the effective *null region* $[-X_k, X_k]$, where leakage flow through both the supply and exhaust ports are possible. The spline polynomial coefficients are expressed in terms of X_δ and X_k as

$$\begin{array}{lll} n_0 = -\frac{X_\delta}{1-X_\delta}, & n_1 = -\frac{1}{1-X_\delta}, & \\ m_0 = \frac{Y_k}{4}, & m_1 = \frac{Y_k}{2X_k}, & m_2 = \frac{Y_k}{4X_k^2}, \end{array} \tag{5.38}$$

where the linear and the quadratic polynomials are jointed together at the knot (X_k, Y_k), with Y_k given by

$$Y_k \triangleq \mu(X_k) = \frac{X_k - X_\delta}{1 - X_\delta}. \tag{5.39}$$

Remark 20 *By using a quadratic polynomial in the null region, the modeled conductance function $c_o(x_v)$ will contain a non-smooth breakpoint at X_k. This non-smooth breakpoint "softens" for decreasing values of X_δ, and vanishes completely for $X_\delta = 0$. Hence, the non-smoothness is negligible when X_δ is small compared to the full stroke of x_v. A function $c_o(x_v)$ that is smooth for all X_δ is easily obtained,* e.g. *by using a cubic polynomial in the null region, however, this results in a rather messy expression for the inverse $x_v = g^{-1}(p_A, w_o)$, which is impractical for implementation in a nonlinear controller. Hence, we have sacrificed smoothness at X_k for a simpler expression of the inverse.*

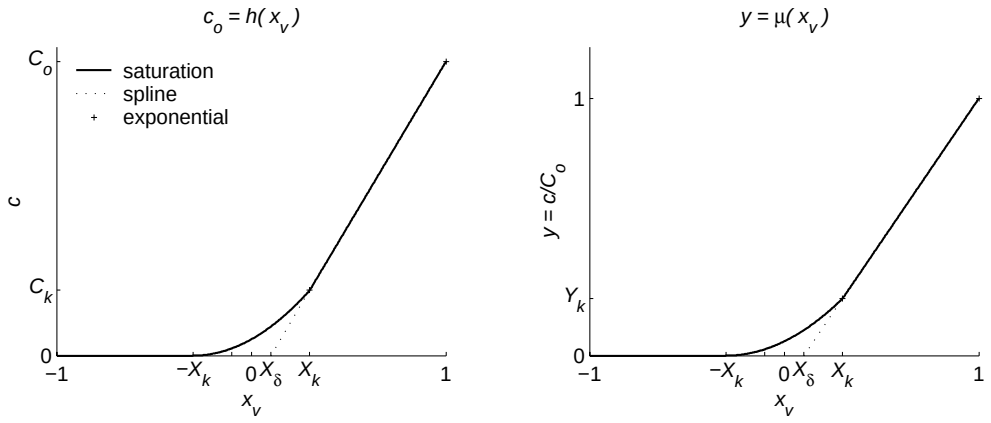

Figure 5.4: The modeled orifice conductance function $c_o = h(x_v)$, and corresponding basis function $\mu(x_v)$.

The resulting input-invertible valve flow model

Summarizing, the developed input-invertible valve model is given by

$$\begin{aligned} w_v &= w_l + w_o \\ &= \rho_0\sqrt{T_0}C_l \cdot \omega_l(p_A/p_S)\frac{p_S}{\sqrt{T_S}} - \rho_0\sqrt{T_0}C_l \cdot \omega_l(p_E/p_A)\frac{p_A}{\sqrt{T_A}} \\ &\quad +\rho_0\sqrt{T_0}\cdot c_o(x_v)\,\omega_o\left(\frac{p_A}{p_S}\right)\frac{p_S}{\sqrt{T_S}} - \rho_0\sqrt{T_0}\cdot c_o(-x_v)\,\omega_o\left(\frac{p_E}{p_A}\right)\frac{p_A}{\sqrt{T_A}}, \end{aligned} \tag{5.40}$$

where the pressure ratio functions $\omega_l(\cdot)$ and $\omega_o(\cdot)$ are parametrized according to (5.18)–(5.20), uniquely characterized by the parameters b_l and b_o, respectively. The variable flow conductance function $c_o(\cdot)$ is given by (5.36)–(5.37) — as discussed in the previous subsection — which is characterized by the orifice flow conductance C_o, the overlap X_δ, and the parameter X_k which determines the null region $[-X_k, X_k]$ of the valve.

The resulting input-invertible flow model (5.40), is uniquely characterized by only six parameters, where a set of physically meaningful parameters are C_l, b_l, C_o, b_o, X_δ, and X_k. Furthermore, the model consists of the physical variables p_S, T_S, p_A, T_A, p_E, T_E, and the known parameters ρ_0 and T_0 of the defined ISO standard *reference condition* of air. It should be noted, that by neglecting the leakage term $w_l = g_l(p_A)$, assuming a negligible null region by setting $X_k = 0$, and using the isentropic pressure ratio function (5.3) for $\omega_o(r)$, the model becomes identical to the model proposed by Shearer [86] (see the literature review in Section 1.2 of the introduction.

Simplified partially affine parametrization

Simplifying for control, we take $T_A = T_S = T_E = T_0$, and $p_E = P_0$ according to assumptions A4.6–7 from the derivation of the reduced-order isothermal air dynamics

in Section 4.2. The flow rate model can then be compactly expressed in the partially parameter-affine form

$$w_v = \rho_0 C_l \cdot \psi_l(p_A) + \rho_0 C_o \cdot \psi_o(p_A, x_v), \tag{5.41}$$

with the *leakage flow function* $\psi_l(p_A)$ defined according to (5.32), and with the *orifice flow function* defined as

$$\psi_o \triangleq \mu(x_v) \cdot \omega_o\left(\frac{p_A}{p_S}\right) p_S - \mu(-x_v) \cdot \omega_o\left(\frac{P_0}{p_A}\right) p_A. \tag{5.42}$$

With the parameters b_l, b_o, X_k, X_δ known, the nonlinear flow functions $\psi_l(p_A)$ and $\psi_o(p_A, x_v)$ are known, and the leakage and orifice flow capacities C_l and C_o appear in an affine form, which makes the formulation (5.41) particularly suited for parameter estimation of C_l and C_o, on-line by an adaptive controller design, or off-line from measurements. The input-invertible flow model (5.41), with its parameters b_l, b_o, C_l, C_o X_k and X_δ fitted to measurements of the flow rate characteristic of the considered Servotronic proportional valve, is plotted in Figure 5.5.

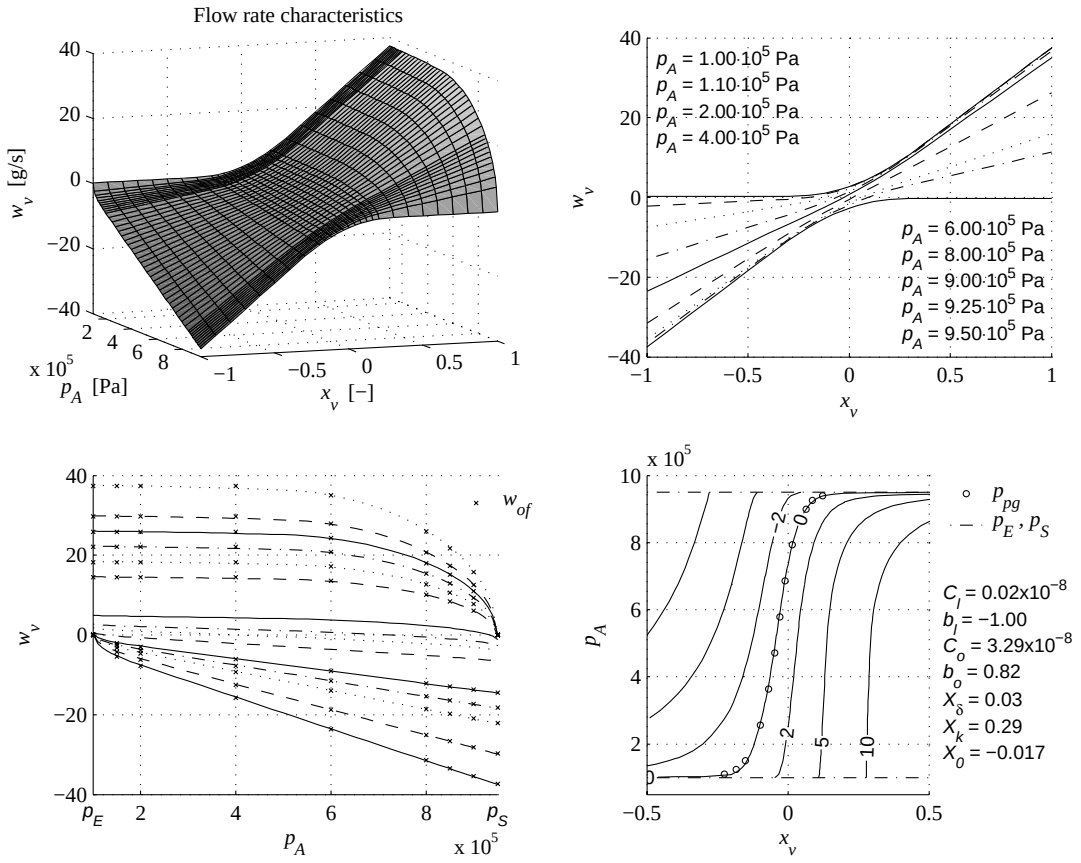

Figure 5.5: The input-invertible valve flow model (5.41) with its parameters fitted to measurements of the flow rate characteristic of the considered Servotronic proportional valve.

Remark 21 *The pressure ratio function $\omega_o(r)$ of the orifice flow term, can alternatively be modeled by the simpler elliptic function (5.16). So can the pressure ratio function $\omega_l(r)$ of the leakage term, however, with reduced accuracy for most types of flow control valves. See Remark 17.*

Tabulated characteristics of the measured sonic conductance C and the critical pressure ratio B as a function of valve inputs, were provided by the valve manufacturer. The measurements are given for a set of valve inputs in the orifice flow region of each single orifice, *i.e.*, for $|x_v| \geq X_k$. From these tabulated measurements of C and B, we create a set of data points of the resulting flow rate characteristic in the orifice flow regions of the spool ($|x_v| \geq X_k$) by use of the ISO standard orifice equation (5.5). We refer to this set of data points as the *orifice flow measurements*, denoted with the subscript "of" in Figure 5.5. The valve manufacturer also provides a plot of the *pressure gain* curve for zero flow[8] of the valve for $p_S = 7 \cdot 10^5\,\mathrm{Pa}$. However, since the pressure p_A is measured in our test rig, we use the experimentally obtained curve which is obtained on the test rig by measuring the equilibrium pressure p_A^* for each corresponding spool position x_v^*. We refer to this set of data points as the (zero-flow) *pressure gain measurements*, indicated by the subscript "pg" in the figure.

The orifice flow measurements and the pressure gain measurements were used to fit the parameters of the model, marked in the figure in the two lower panes with '$\times$' and 'o', respectively. The parameters of the model were fitted to the measurements by use of the function `lsqnonlin` in Matlab, which employs a nonlinear search algorithm in order to find a set of parameters that minimizes some scalar function of the modeling errors. See Appendix B for a brief description of parameter estimation from measurements.

The normalized steady-state spool position of the valve is given as

$$x_v^* = K_v u_v + X_0, \tag{5.43}$$

where $u_v \in [-10, 10]\,\mathrm{V}$ is the actual valve control input, $K_v = 1/10\,\mathrm{V}^{-1}$ is the steady-state gain, and $X_0 = K_v U_{v0}$ is a small spool offset in the valve.

In Figure 5.5, the flow rate characteristic is plotted for the full operating range of the spool, $x_v \in [-1, 1]$, and the full range of chamber pressures $p_A \in [P_0, P_S]$, with supply pressure $P_S = 9.5 \cdot 10^5\,\mathrm{Pa}$, and exhaust pressure $P_0 = 1.0 \cdot 10^5\,\mathrm{Pa}$. The temperatures are assumed to be equal a constant room temperature of $25\,^\circ C$, *i.e.*, $T_0 = T_S = T_A = T_E = 297\,\mathrm{K}$. The fitted parameters are printed the right of the plotted pressure gain characteristic.

Computation of the inverse $x_v = g^{-1}(p_A, w_o)$

When using the flow rate model for nonlinear control, we may need to compute x_v from a given w_o in the control law, and preferably we want to do this by having an

[8] In general, a *pressure gain* curve is a contour line of the flow rate characteristics in the $x_v - p_A$ plane, *i.e.*, a curve representing a constant flow rate. Thus, the *pressure gain* curve for zero flow refers to the contour line for $w_v = 0$, *i.e.*, the set $\{(x_v, p_A) : w_v = 0\}$.

explicit expression for the inverse according to

$$x_v = g_o^{-1}(p_A, w_o). \tag{5.44}$$

This is straightforwardly achieved for the proposed valve model. First we note that we can rewrite the orifice flow part ($w_o = g_o(p_A, x_v)$) of (5.41) as

$$w_o = C_o\psi_S(p_A)\,\mu(x_v) - C_o\psi_E(p_A)\,\mu(-x_v), \tag{5.45}$$

by defining the pressure functions for the supply and the exhaust flow according to

$$\psi_S \triangleq \rho_0\sqrt{T_0}\omega_o(p_A/p_S), \tag{5.46}$$
$$\psi_E \triangleq \rho_0\sqrt{T_0}\omega_o(p_E/p_A), \tag{5.47}$$

respectively. In order to calculate the inverse $g_o^{-1}(p_A, w_o)$ we divide the domain of $g_o(x_v)$ into the *supply region* $[X_k, 1]$ where only ψ_S is nonzero, the *null region* $[-X_k, X_k]$ where both ψ_S and ψ_E are nonzero, and the *exhaust region* $[-1, -X_k]$ where only ψ_E is nonzero. Furthermore, it is convenient to define the corresponding indicator functions

$$\begin{aligned}\chi_S &\triangleq \chi[X_k \le x_v \le 1],\\ \chi_c &\triangleq \chi[-X_k \le x_v \le X_k],\\ \chi_E &\triangleq \chi[-1 \le x_v \le -X_k],\end{aligned} \tag{5.48}$$

respectively, where the general indicator function $\chi[\cdot]$ is defined by (5.29) in the previous section. In this way, we can rewrite (5.45) as

$$\begin{aligned}w_o &= (C_o\psi_S n_0 + C_o\psi_S n_1 x_v)\cdot\chi_S(x_v)\\ &\quad + \left(C_o m_0\cdot(\psi_S - \psi_E) + C_o m_1\cdot(\psi_S + \psi_E)\,x_v + C_o m_0\cdot(\psi_S - \psi_E)\,x_v^2\right)\cdot\chi_c(x_v)\\ &\quad + (C_o\psi_E n_0 - C_o\psi_E n_1 x_v)\cdot\chi_E(x_v),\end{aligned} \tag{5.49}$$

where the argument is dropped in $\psi_S(p_A)$ and $\psi_E(p_A)$ for simplicity of notation. In Figure 5.6, the orifice flow function $w_o = g_o(p_A, x_v)$ and its corresponding inverse $x_v = g_o^{-1}(p_A, w_o)$ is illustrated for some fixed pressures p_A. The figure is plotted with exaggerated values of X_δ in order to better illustrate the resulting non-smooth knots at $-X_k$ and X_k.

The inverse is computed by considering each region separately. In the supply and exhaust regions, the orifice flow function is linear with respect to x_v, and the inverse is given by

$$x_v = X_\delta + \frac{1 - X_\delta}{C_o}\psi_S^{-1} w_o \quad \wedge \quad x_v \in [X_k, 1] \tag{5.50}$$

$$x_v = -X_\delta + \frac{1 - X_\delta}{C_o}\psi_E^{-1} w_o \quad \wedge \quad x_v \in [-1, -X_k]. \tag{5.51}$$

In the null region, the orifice flow function is quadratic with respect to x_v, hence, the inverse becomes the square root expression

$$x_v = \frac{X_k}{\psi_S - \psi_E}\left(2\sqrt{\psi_S\psi_E + C_k^{-1}(\psi_S - \psi_E)\,w_o} - \psi_S - \psi_E\right) \quad \wedge \quad x_v \in [-X_k, X_k], \tag{5.52}$$

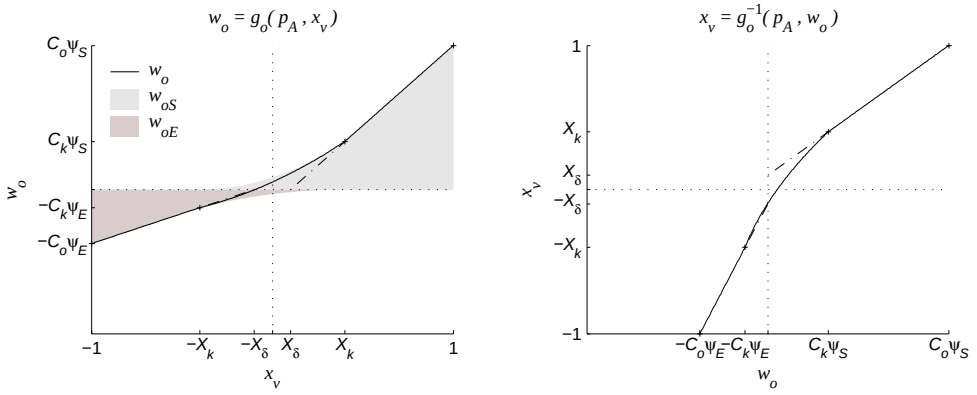

Figure 5.6: An illustration of the orifice flow model $w_o = g_o(p_A, x_v)$ and its inverse $x_v = g_o^{-1}(p_A, w_o)$.

where $C_k \triangleq C_o Y_k$. Due to the term $(\psi_S - \psi_E)^{-1}$, Equation (5.52) appear to be not well defined for $\psi_S - \psi_E = 0$, since it is indeterminate (0/0). However, its limit exists and is a straight line through the origin in which can be found by l'Hôpital's rule to be

$$x_v = C_k^{-1}\psi_S^{-1} w_o. \tag{5.53}$$

5.4.3 A smooth parameter-affine valve flow model

The valve flow model need not be explicitly input-invertible to be applicable for a nonlinear control design; it is sufficient that the model is *one-to-one*. Utilizing an integrator backstepping approach for the design of a controller, we may circumvent the need to compute the inverse of the flow rate. In this case, it is advantageous to develop a model which is smooth so that its derivative becomes continuous, which is required in order to design a smooth control law by integrator backstepping.

We develop a model in a similar manner as for the input-invertible model, by utilizing the generalized flow equation (5.15) in combination with a model of the variable conductance. Modeling the flow paths through the supply and exhaust ports individually, as general nonlinear functions of the spool x_v, *i.e.*, $c_S(x_v)$ and $c_E(x_v)$, we are able to accurately describe the flow rate over the full range of spool positions. Utilizing the symmetry of the valve, we define a single nonlinear function which applies for both the supply and the exhaust flows according to $c_S = c_v(x_v)$, and $c_E = c_v(-x_v)$. An accurate description of the spool–flow characteristic of flow control valves can thus be obtained by modeling the resulting flow rate according to $w_v = w_S - w_E$, in the form

$$w_v = \rho_0\sqrt{T_0}\cdot c_v(x_v)\,\omega_v\left(\frac{p_A}{p_S}\right)\frac{p_S}{\sqrt{T_S}} - \rho_0\sqrt{T_0}\cdot c_v(-x_v)\,\omega_v\left(\frac{p_E}{p_A}\right)\frac{p_A}{\sqrt{T_A}}, \tag{5.54}$$

where $\omega_v(\cdot)$ is the pressure ratio function, and $c_v(\cdot)$ is the variable conductance of each flow path as a function of the position of the spool. The pressure ratio function $\omega_v(\cdot)$ is parametrized according to (5.18)–(5.20), which is uniquely characterized by

the parameter b_v. The parametrization of the variable flow conductance function $c_v(\cdot)$ is discussed in the subsection below.

Flow conductance function

The variable conductance of each flow path as a function of the position of the spool is conveniently modeled in the form

$$c_v = C_v \cdot \mu(x_v), \tag{5.55}$$

where $\mu(x_v) \in [0,1]$ so that C_v represents the *valve flow conductance* for the valve port fully open, *i.e.*, $x_v = \{-1, 1\}$. The normalized conductance function $\mu(x_v)$ is parametrized in the parameter-affine form

$$\mu(x_v) = \boldsymbol{\theta}_\mu^T \cdot \boldsymbol{\phi}_\mu(x_v), \tag{5.56}$$

where the regressor $\boldsymbol{\phi}_\mu(x_v) = [\phi_{\mu 1}(x_v), \phi_{\mu 2}(x_v), \cdots, \phi_{\mu p}(x_v)]^T \in \mathbb{R}^p$ is a vector of p basis functions which is weighted by the parameter vector $\boldsymbol{\theta}_\mu = [\theta_{\mu 1}, \theta_{\mu 2}, \cdots, \theta_{\mu p}]^T \in \mathbb{R}^p$. The basis functions $\phi_{\mu,i}(x_v)$ are modeled using *normalized Gaussian basis functions*, defined according to

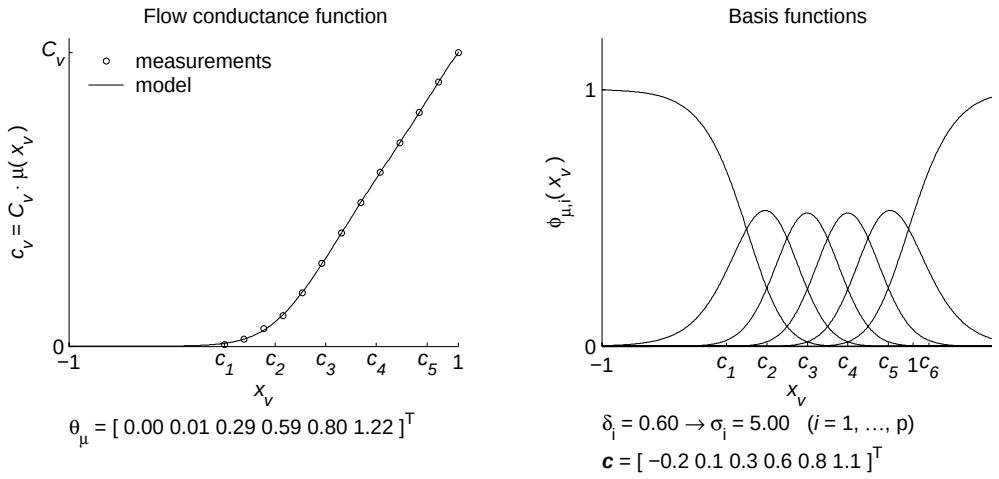

Figure 5.7: The modeled conductance function $c_v = C_v \cdot \mu(x_v) = C_v \boldsymbol{\theta}_\mu^T \cdot \boldsymbol{\phi}_\mu(x_v)$ and corresponding basis fuctions $\boldsymbol{\phi}_\mu(x_v)$ fitted to some tabulated data of the valve.

$$\phi_{\mu,i}(x_v) = \frac{\psi_i(x_v)}{\sum_{j=1}^{p} \psi_j(x_v)} \tag{5.57}$$

$$\psi_i(y) = \mathrm{e}^{-\frac{1}{2}\sigma_i^2 (y - c_i)^2}, \tag{5.58}$$

where $\boldsymbol{\sigma} = [\sigma_1, \sigma_2, \cdots \sigma_p]^T$ is a vector of scaling parameters, and $\mathbf{c} = [c_1, c_2, \cdots, c_p]^T$ a vector of offset parameters. Equation (5.57) provides a normalization of the standard Gaussian functions given by (5.58), and the parameters σ_i determines the

widths, and the parameters c_i determines the location (or centers), of the corresponding basis function ψ_i. In general, the accuracy of the modeled nonlinearity is improved by increasing the number of basis functions, p. The modeled conductance function $\mu(x_v) = \boldsymbol{\theta}_\mu^T \cdot \boldsymbol{\phi}_\mu(x_v)$ and its corresponding basis functions are illustrated in Figure 5.7. Notice in the figure, that we have not placed basis functions over the entire input domain $x_v \in [-1, 1]$ of $\mu(x_v)$, but with centers ranging from $c_1 = -0.2$ to $c_6 = 1.1$. With this choice we found that $p = 6$ functions was sufficient to provide a highly accurate fit to the tabulated conductance characteristic (marked with circles '$\circ$' in the figure).

Piecewise linear parametrization

With the introduced affine parametrization of the nonlinear conductance function, the resulting flow rate model is piecewise linearly parametrizable, *i.e.*, the model can be expressed in a form where all its characteristic parameters appear piecewise linearly. That is, the model can be expressed in the form

$$w_v = \boldsymbol{\theta}_v^T \cdot \boldsymbol{\phi}_v \left(p_S, T_S, p_A, T_A, p_E, T_E, x_v, \operatorname{sgn} b_v\right), \tag{5.59}$$

with the parameter and regressor vectors defined as

$$\begin{aligned} \boldsymbol{\theta}_v &\triangleq \rho_0 \sqrt{T_0} \begin{bmatrix} C_v \boldsymbol{\theta}_\mu \\ C_v b_v \boldsymbol{\theta}_\mu \end{bmatrix} \qquad (5.60) \\ \boldsymbol{\phi}_v &\triangleq \begin{bmatrix} \boldsymbol{\phi}_\mu(x_v) \cdot \Omega_0\left(\frac{p_A}{p_S}\right) \frac{p_S}{\sqrt{T_S}} - \boldsymbol{\phi}_\mu(-x_v) \cdot \Omega_0\left(\frac{p_E}{p_A}\right) \frac{p_A}{\sqrt{T_A}} \\ \boldsymbol{\phi}_\mu(x_v) \cdot \Omega_1\left(\frac{p_A}{p_S}, \operatorname{sgn} b_v\right) \frac{p_S}{\sqrt{T_S}} - \boldsymbol{\phi}_\mu(-x_v) \cdot \Omega_1\left(\frac{p_E}{p_A}, \operatorname{sgn} b_v\right) \frac{p_A}{\sqrt{T_A}} \end{bmatrix}, \end{aligned}$$

where we have substituted with the right-hand sides of (5.18) for the pressure ratio function $\omega_v(\cdot)$, and (5.7) for the flow conductance function $\mu(x_v)$ in the flow rate model (5.54).

This form is particularly suited for parameter estimation of all the model parameters, either on-line by an adaptive controller design, or off-line from measurements, since $C_v\boldsymbol{\theta}_\mu \in \mathbb{R}^p$ and $C_v b_v \boldsymbol{\theta}_\mu \in \mathbb{R}^p$ appear linearly. Note that the dimension of the resulting parameter vector $\boldsymbol{\theta}_v \in \mathbb{R}^{2p}$, where p is the number of parameters in the conductance function (5.56). Hence, with $p = 6$ (which is sufficient to provide a highly accurate fit to the Servotronic flow rate characteristic), the total number of parameters is 12.

Simplified partially affine parametrization

Simplifying for control, we take $T_A = T_S = T_E = T_0$, and $p_E = P_0$ according to assumptions A4.6–7 from the derivation of the pressure dynamics in Section 4.2. The flow rate model can then be expressed in the compact, partially parameter-affine form

$$w_v = \rho_0 C_v \cdot \psi_v\left(p_A, x_v\right), \tag{5.61}$$

by defining the *valve flow function*

$$\psi_v \triangleq \mu(x_v) \cdot \omega_v\left(\frac{p_A}{p_S}\right) p_S - \mu(-x_v) \cdot \omega_v\left(\frac{P_0}{p_A}\right) p_A. \tag{5.62}$$

When the parameters b_v and $\boldsymbol{\theta}_\mu$ are known, the nonlinear flow function $\psi_v(p_A, x_v)$ is known, and the valve flow capacity C_v appears in an affine form which makes the formulation (5.61) particularly suited for parameter estimation of C_v.

Remark 22 *Like for the flow rate model of the outlet restriction, the pressure ratio function $\omega_v(\cdot)$ can alternatively be modeled using the simpler elliptic parametrization (5.16), however, with reduced accuracy for the flow rate in the null region of the valve.*

Since the model can be formulated in the piecewise parameter-affine form (5.59), we use the function `lsqlin` in Matlab for parameter fitting, which is a convex optimization routine that finds the set of parameters providing the least squares fit to the given measurements. See Appendix B for a brief description of its use for parameter estimation. In Figure 5.8, the flow rate characteristic is plotted for the full operating range of the spool, $x_v \in [-1, 1]$, and the full range of chamber pressures $p_A \in [p_E, p_S]$. With a constant supply pressure $p_S = P_S = 9.5 \cdot 10^5$ Pa, and a constant exhaust pressure $p_E = P_0 = 1.0 \cdot 10^5$ Pa. The temperatures are assumed to be equal a constant room temperature of 25 °C, *i.e.*, $T_0 = T_S = T_A = T_E = 297$ K. The fitted parameters are printed on the right of the plotted pressure gain characteristic.

5.4.4 Valve dynamics

In this section we briefly review the modeling of the *input–spool dynamics* of a *proportional valve* actuated by an electro-magnetic force motor with a proportional input–force characteristic (usually referred to as a *proportional solenoid*). We then present a simplified reduced-order model of the dynamics which we use for control design.

Motion dynamics of the spool

The dynamics of the spool is governed by the equation of motion, and can thus be expressed as

$$M_v \ddot{x}_v = -K_s x_v + K_m i - f_f(\cdot) - f_h(x_v, \dot{x}_v), \tag{5.63}$$

where i is the current of the solenoid coil, M_v is the mass of the spool, K_s the resulting spring stiffness of two centering springs, and K_m is the proportionality constant of the solenoid coil. The friction forces are represented by $f_f(\cdot)$, and the mechanical constraints of the spool by the hardstop force $f_h(\cdot)$. See Chapter 3 for the modeling of friction and hardstop forces.

In (5.63), we have assumed that the spool is centered by two linear coil springs, and that the electro-magnetic proportional solenoid has an ideally linear characteristic such that the electro-magnetic force becomes proportional to the coil current, so that it can be expressed as $f_m = K_m i$.

Remark 23 *The 2nd-order dynamics (5.63) is expected to be significantly underdamped due to the low viscous friction of air.*

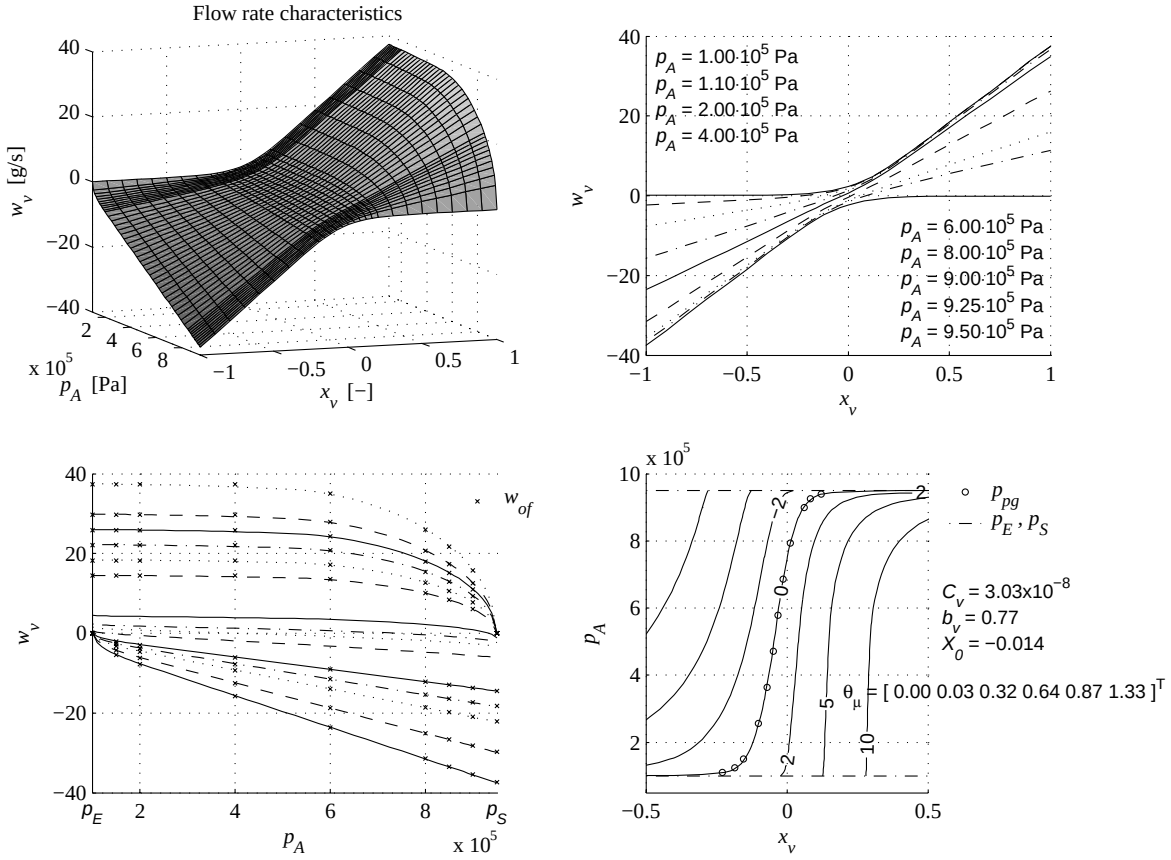

Figure 5.8: The smooth piecewise parameter-affine valve flow model (5.59) fitted to measurements.

In order to obtain a simplified, reduced-order model for control, we assume that the mass of the spool is negligible small, *i.e.*,

$$M_v \approx 0,$$

and assume that the friction force f_f can be given as

$$f_f = D_{\dot{x}_v} \dot{x}_v, \tag{5.64}$$

where $D_{\dot{x}_v}$ is a viscous friction coefficient. Furthermore, we neglect the hardstop force $f_h(\cdot)$. The dynamics of the spool (5.63) then reduces to the first-order dynamics

$$D_{\dot{x}_v} \dot{x}_v = -K_s x_v + K_m i. \tag{5.65}$$

Remark 24 *The reduced-order dynamics (5.65) is expected to represent a reasonable approximation to the full-order dynamics (5.63) for low frequency inputs,* i.e., *for frequencies well below the resonance frequency of the 2nd-order dynamics (5.63). However, for inputs with high-frequency components (such as* e.g. *a step input), the 1st-order dynamics (5.65) will exhibit a damped response, while the actual 2nd-order dynamics (5.63) will exhibit an oscillatory (underdamped) response.*

Dynamics of the electrical coil circuit

The dynamics of the electrical circuit is approximated by

$$L\frac{di}{dt} = -Ri + u_v, \tag{5.66}$$

where u_v is the applied voltage, L is the conductance, and R the electrical resistance of the solenoid coil circuit. The coil dynamics is linear, with the time constant $\tau_i = L/R$. Most proportional solenoids, however, are controlled by an internal feedback from the coil current in order to improve of the transient performance. This feedback is usually taken as

$$u_v = K_i \cdot (i_d - i), \tag{5.67}$$

where K_i is the proportional feedback gain, and i_d is the current demand signal which usually is given by a corresponding voltage signal u as

$$i_d = K_u u, \tag{5.68}$$

with an arbitrary scaling constant K_u. The coil dynamics is thus given as

$$L\frac{di}{dt} = -(R + K_i) \cdot i + K_i K_u u, \tag{5.69}$$

with the improved time constant $\overline{\tau}_i = L/(R + K_i)$.

In most cases with current feedback, the feedback gain K_i is typically very high, giving a coil dynamics which is negligible fast compared to the dynamics of the other (pneumatic and mechanical) states of the pneumatic actuator. To clearly see this, we can rewrite (5.69) as

$$\frac{L}{K_i}\frac{di}{dt} = -\frac{R + K_i}{K_i} i + K_u u. \tag{5.70}$$

With $K_i \gg R$, we can approximate

$$\frac{L}{K_i}\frac{di}{dt} \approx -i + K_u u, \tag{5.71}$$

where the time constant $\tau_i = L/K_i \ll 1$. Hence, in order to obtain a simplified reduced-order model for control design, it is usually reasonable to assume

$$i = K_u u. \tag{5.72}$$

Simple reduced-order valve dynamics

Combining the static approximation (5.72) of the coil dynamics, and the reduced-order dynamics (5.65) of the motion dynamics of the spool, the valve dynamics reduces to

$$\frac{D_{\dot{x}_v}}{K_s}\frac{dx_v}{dt} = -x_v + \frac{K_m K_u}{K_s} u, \tag{5.73}$$

which is linear with the time-constant $\tau_v = D_{\dot{x}_v}/K_s$.

Equation (5.73) serves as a justification to describe the dynamics of a pneumatic proportional valve by the first-order linear dynamic model

$$\frac{dx_v}{dt} = -\frac{1}{\tau_v}x_v + \frac{1}{\tau_v}\,\mathrm{sat}\left(K_v u_v + K_v U_{v0}\right), \tag{5.74}$$

where τ_v is the time-constant of the dynamics, $u_v \in [-U_v, U_v]$ is the control input, and K_v is the steady-state gain which is chosen such that the input $K_v u_v$ is normalized, *i.e.*, for $\forall u_v \in [U_v, U_v] \Longrightarrow K_v u_v \in [-1, 1]$. The parameter U_{v0} is introduced to represent the *zero-point drift* in the valve — a small offset between the valve input u_v and the corresponding steady-state spool position x_v — which is due to temperature variations[9]. A saturation of the summed input $K_v \cdot (u_v + U_{v0})$ is introduced to guarantee that $x_v \in [-1, 1]$ when the initial state is chosen to satisfy $x_v(0) \in [-1, 1]$.

Most high-performance valves have spool feedback which greatly improves the response time and steady-state accuracy of the valve. Depending on the applied control technique, the feedback control law of proportional valves with spool feedback, more or less, provides a closed-loop response which is damped and approximately linear, which means that the dynamics can be well approximated by the simple model (5.74). Hence, the dynamics (5.74) can also be used to model proportional valves with spool feedback.

In order to obtain a smooth model, we may modify (5.74) according to

$$\frac{dx_v}{dt} = -\frac{1}{\tau_v}x_v + \frac{1}{\tau_v}\cdot\pi_u\left(K_v u_v + K_v U_{v0}\right), \tag{5.75}$$

where $\pi_u(\cdot)$ is a smooth saturation function with saturation limits $[x_{\mathrm{lb}}, x_{\mathrm{ub}}] = [-1, 1]$ and smoothing width ε_π, defined as follows.

A smooth saturation function with general saturation limits may be constructed according to

$$\pi\left(x, x_{\mathrm{lb}}, x_{\mathrm{ub}}, \varepsilon_\pi\right) \triangleq x + \varepsilon_\pi \cdot g\left(\frac{-x + x_{\mathrm{lb}}}{\varepsilon_\pi}\right) - \varepsilon_\pi \cdot g\left(\frac{x - x_{\mathrm{ub}}}{\varepsilon_\pi}\right), \tag{5.76}$$

where $\varepsilon_\pi > 0$ is an arbitrary small design constant, referred to as the smoothing width of the breakpoints of $\pi(\cdot)$, and where $g(\cdot)$ is smooth plus function with unity slope and smoothing interval $[-1, 1]$. An example of $g(\cdot)$ constructed as a spline function, is given by (3.32) in Section 3.3, which is reviewed below for the reader's convenience

$$g(x) = \begin{cases} x, & x > 1 \\ \frac{3}{16} + \frac{1}{2}x + \frac{3}{8}x^2 - \frac{1}{16}x^4, & |x| \le 1 \\ 0, & x < -1 \end{cases}.$$

The smooth saturation function (5.76) and the above smooth plus function, is plotted in Figure 5.9.

[9] In Figures 5.5 and 5.8, this zero-point offset is represented as $X_0 \triangleq K_v U_{v0}$.

For brevity of notation, we may omit the parameters x_{lb}, x_{ub} and ε_π as function arguments, and let $\pi_u(x) \triangleq \pi(x, -1, 1, \varepsilon_\pi)$, *i.e.*, a smooth saturation function with unity saturation limits $[x_{\mathrm{lb}}, x_{\mathrm{ub}}] = [-1, 1]$. The smooth saturation function $\pi_u(x)$, is thus a smooth approximation of the standard saturation function $\mathrm{sat}(x)$, which can be made arbitrary accurate by reducing the smoothing width ε_π, *i.e.*,

$$\lim_{\varepsilon_\pi \to 0} \pi_u(x) = \mathrm{sat}(x).$$

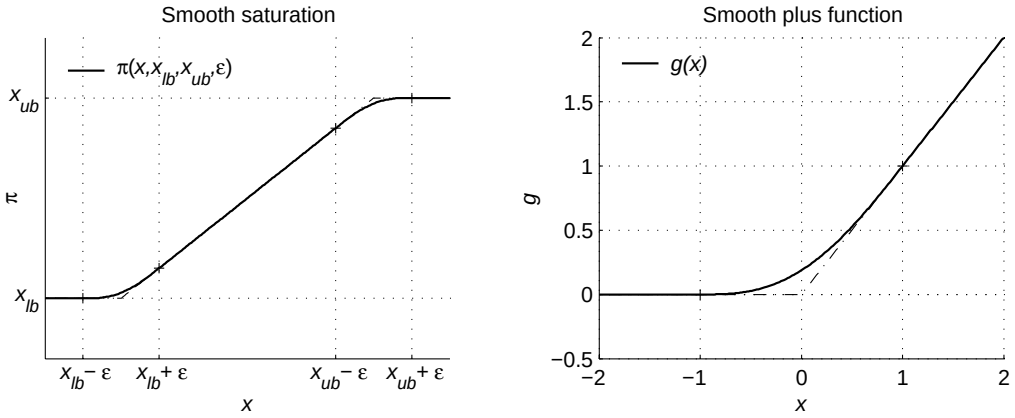

Figure 5.9: Left: The smooth saturation function $\pi(\cdot)$ with upper and lower bounds x_{ub} and x_{lb}, and smoothing width ε. Right: The smooth plus function.

Unmodelled valve dynamics

When applying the reduced-order dynamics (5.74) for the modeling of the spool dynamics of proportional valves, it is important to be aware that this simplified model is not accurate for inputs with high-frequency content. That is, for fast changes in the input u_v, the model results in unmodeled dynamics. That is, the reduced-order model (5.74) is not able to reproduce the oscillating transient behavior of 2nd-order underdamped systems, see Remark 24.

This discrepancy of the model is most pronounced for valves without spool feedback, however, it also applies to certain valves with simple spool feedback, *i.e.*, using standard PID feedback control. For example, with proportional feedback $u = K_p \cdot (x_d - x_v)$ applied to (5.73), the model structure remains the same, only with a smaller time constant τ_v. Another example of unmodeled dynamics arises when using the reduced-order model (5.74) to model proportional valves with spool feedback with integral action. This is the case with the considered Servotronic proportional valve, which exhibits a slowly converging overshoot for fast changes in the control signal, apparently due to slow integral action in the spool positioning loop.

From a control point of view, this unmodeled dynamics is likely to put a significant limit on the achievable bandwidth of a tracking controller for the electropneumatic actuator. The obvious alternative would be to model the valve dynamics

using a more elaborate dynamic model of higher order, however, this increases the complexity of the controller design for the electro-pneumatic actuator.

5.5 Summary

The modeling of the air flow rate in fixed restrictions and flow control valves in the context of nonlinear and adaptive control, is considered in this chapter. A detailed summary of the work is provided below.

A generalized, piecewise parameter-affine parametrization of the *flow rate characteristic* of pneumatic restrictions, is developed. This generalized flow rate equation is constructed from basis functions using the standardized orifice flow equation, and the equation for incompressible laminar fluid flow. In addition to its physical pressure and temperature variables, the equation is uniquely characterized by a conductance parameter C, and a critical pressure ratio–like parameter b. The novelty of this parametrization, is that it is piecewise affine in b, unlike the ISO standardized equation where the critical pressure ratio B appears nonlinearly, and that its validity ranges from isentropic nozzle flow to incompressible laminar flow, *i.e.*, it is generalized in the sense that it encompasses most pneumatic restrictions.

The generalized flow rate equation is utilized to construct a model of the bidirectional flow rate through a fixed pneumatic restriction, which has a validity range that encompasses the full range of possible restriction types, from simple orifice restrictions to small clearance restrictions. The resulting model is in piecewise parameter-affine form, uniquely characterized by two parameters C_r and b_r, where the parameter C_r describes the capacity of the restriction, while b_r determines the pressure dependence of the flow rate.

The generalized flow rate equation is utilized for the parametrization of the static spool–flow characteristic of flow control valves. First, we develop a simple piecewise input-affine model, very similar to the most commonly used model in the literature. Next, we develop two novel parametrizations of the spool–flow characteristic of flow control valves of sliding type, where the first is explicitly input-invertible, and the other is (piecewise) fully linearly parametrizable and differentiable. Both models are based on an individual description of the flow through each flow path of the valve, thus, providing improved accuracy compared to existing models, particularly for valves with significant leakage flow in the null region. The two models are described below:

An input-invertible valve flow model: An accurate parametrization of the static spool–flow characteristic of a flow control valve is developed for subsequent application of nonlinear control by a *feedback linearization* approach. The model is *input-invertible* in the sense that the spool position (input) can be expressed explicitly as a function of the flow rate, and it is *mechanistic* in the sense that all its parameters have a physical meaning. The model is completely characterized by only six parameters, composed of a fixed restriction leakage term, characterized by C_l and b_l, an orifice flow term, characterized by C_o, b_o,

and incorporates an overlap X_δ (or underlap for $X_\delta < 0$), and a smoothed characteristic in the null region, characterized by its width X_k.

A smooth parameter-affine valve flow model: An accurate piecewise parameter-affine and smooth model of the static spool–flow characteristic is developed for subsequent application of nonlinear and adaptive control by a *backstepping* approach. The model is based on a parameter-affine parametrization of the spool–conductance characteristic of each valve port, utilizing normalized Gaussian basis functions. The resulting model is differentiable, hence, suited for an exact backstepping design where the valve dynamics is included in the design. Furthermore, the model can be expressed in a piecewise parameter-affine form which makes it particularly suited for an adaptive design where all its parameters may be *a priori* unknown, and estimated on-line. The model is characterized by a conductance parameter C_v representing the flow capacity of fully open valve ports, a pressure ratio-like parameter b_v, and a vector $\boldsymbol{\theta}_\mu \in \mathbb{R}^p$ of the parameter-affine spool–conductance nonlinearity.

The modeling of the motion dynamics of the valve spool and the dynamics of the coil current of a proportional valve, are briefly reviewed, and serves as a justification as well as a clarification of the underlying presuppositions for approximating the valve dynamics by a simple linear model.

Chapter 6

Model for Control Design

In this chapter, we combine the work on mathematical modeling from Chapters 3–5, in a smooth design model for the electro-pneumatic clutch actuation system. We describe the model in state-space form, discuss its region of validity, and outline some important properties with respect to nonlinear and adaptive control.

6.1 Design model in state-space form

6.1.1 Motion dynamics

The position and velocity states y and v are governed by the equation of motion (3.1), as discussed in Chapter 3. The static clutch load characteristic $f_l(y)$ is parametrized according to (3.2) using normalized Gaussian basis functions, as it is illustrated in Figure 3.1. The friction forces in the cylinder actuator and clutch is modeled using the modified smooth version of the 1st-order dynamic LuGre model, where the pre-sliding deflection state z and the friction force $f_f(v, z)$, are governed by (3.20)–(3.22), respectively. Furthermore, the hardstop force $f_h(y, v)$, representing the mechanical constraints of the actuator, is modeled according to (3.30), utilizing smooth plus and step functions, which may be constructed as spline functions with required smoothness.

6.1.2 Air dynamics

The air dynamics of the two pneumatic chambers is modeled assuming isothermal conditions so that the temperatures can be taken as constant and the pressure becomes the only dynamic state. The pressure states p_A and p_B of the two actuator chambers are thus governed by the common reduced-order isothermal pressure dynamics, given by (4.10) and (4.11), respectively.

6.1.3 Flow rate characteristics and valve dynamics

The flow rate of the fixed outlet restriction of the back-chamber (chamber B) of the actuator, is modeled according to (5.24)–(5.25) utilizing the affine parametrization

of the pressure ratio function, given by (5.18)–(5.20). The flow rate characteristic of the flow control valve is modeled using the developed smooth valve flow model, given by (5.61)–(5.62. In this smooth and piecewise linearly parametrizable flow model, the spool–conductance characteristic is parametrized according to (5.56) using normalized Gaussian basis functions, as illustrated in Figure 5.7, and the pressure ratio function is parametrized according to (5.18)–(5.20). Finally, the valve dynamics is modeled using the simple smooth 1st-order model, given by (5.75) which utilizes the smooth saturation function (5.76).

6.1.4 State-space model

The complete 6th-order state-space model is given as

$$\begin{aligned}
\dot{y} &= v \\
\dot{v} &= \frac{A_0 P_0}{M} + \frac{A_A}{M} p_A - \frac{A_B}{M} p_B - \frac{1}{M} f_f(v,z) - \frac{1}{M} f_l(y) - \frac{1}{M} f_h(y,v) \\
\dot{p}_A &= -A_A \frac{1}{V_A(y)} v p_A + \rho_0 T_0 R C_v \frac{1}{V_A(y)} \psi_v(p_A, x_v) \\
\dot{x}_v &= -\frac{1}{\tau_v} x_v + \frac{1}{\tau_v} \pi_u (K_v u_v + K_v U_{v0}) \\
\dot{p}_B &= A_B \frac{1}{V_B(y)} v p_B + \rho_0 T_0 R C_r \frac{1}{V_B(y)} \psi_r(p_B) \\
\dot{z} &= v - \frac{K_z}{F_C} |v|_s z,
\end{aligned} \tag{6.1}$$

where the dynamic states are y, v, p_A, x_v, p_B and z; the manipulated input is u_v, and the measured output is the position y; the function $f_f(v,z)$ is the friction force, $f_l(y)$ is the nonlinear clutch load force, and the $f_h(y,v)$ is the hardstop force; the functions $V_A(y)$ and $V_B(y)$ are the chamber volumes which are positive linear functions of y given by (4.3) and (4.7); $\psi_v(p_A, x_v)$ and $\psi_r(p_B)$ are the valve and restriction flow functions, modeled according to (5.62) and (5.25), respectively; and $\pi(\cdot)$ is the smooth saturation function given by (5.76). The valve flow function $\psi_v(p_A, x_v)$ and the restriction flow function $\psi_r(p_B)$, are plotted in Figures 6.1 and 6.2, respectively.

Remark 25 *The pressure ratio functions $\omega_v(r)$ and $\omega_r(r)$ used to construct $\psi_v(p_A, x_v)$ and $\psi_r(p_B)$, respectively, could alternatively be modeled using the ISO standardized elliptic pressure ratio function (5.16), which is simpler. In this case, for the construction of $\psi_v(\cdot)$ and $\psi_r(\cdot)$, the elliptic function (5.16) must be redefined so that it is valid also for $r \geq 1$, like*

$$\omega(r) \triangleq \begin{cases} 1 & , r \leq B \\ \sqrt{1 - \left(\frac{r-B}{1-B}\right)^2} & , r > B \\ 0 & , r \geq 1 \end{cases} \quad \wedge \quad r \in [0, \infty\rangle\,, \tag{6.2}$$

where B is the critical pressure ratio. An approximate relation between the critical pressure ratio B, and the critical pressure ratio-like parameter b, is approximately $B \approx 0.528 \cdot b$.

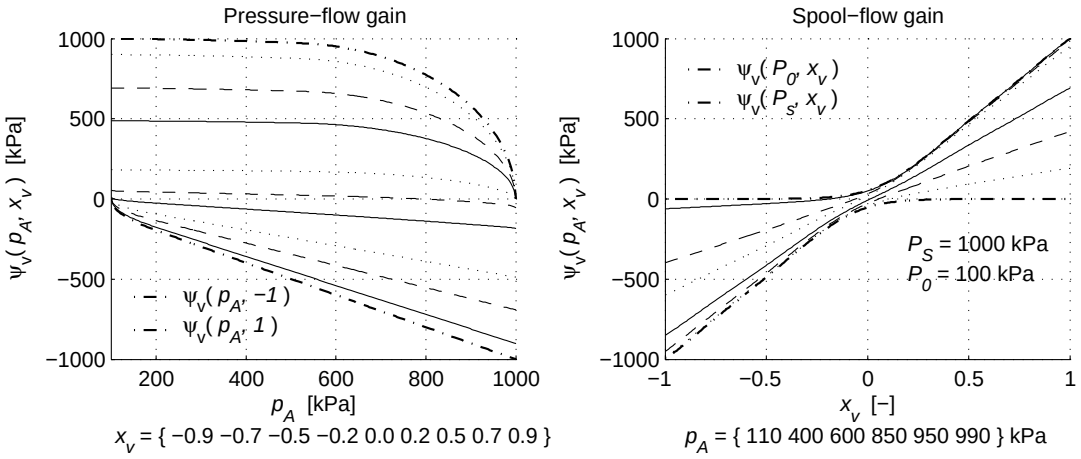

Figure 6.1: The valve flow nonlinearity $\psi_v\left(p_A, x_v\right)$ of the proportional valve.

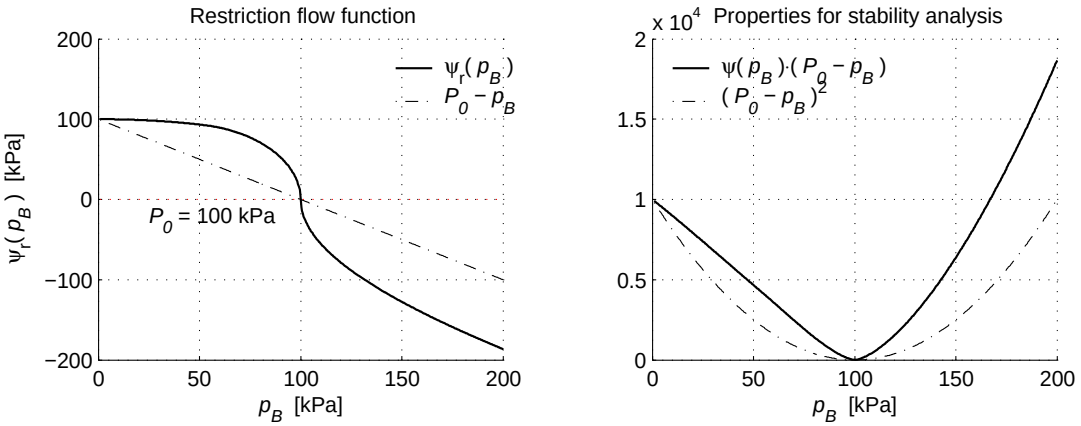

Figure 6.2: The compressible flow nonlinearity $\psi_r\left(p_B\right)$ of the outlet restriction.

Region of validity

Due to the assumptions applied to the modeling, and because some states are only physically feasible (or meaningful) within certain ranges, the model is valid only in a subset of the full state-space. More precisely, the *region of validity* of the model is the set $\mathcal{X}_0 \subset \mathbb{R}^6$ defined by

$$\mathcal{X}_0 \triangleq \left\{\forall \mathbf{x} : \mathbf{x}_{\min} \leq \mathbf{x} \leq \mathbf{x}_{\max}\right\}, \tag{6.3}$$

where $\mathbf{x} \triangleq [y, v, p_A, x_v, p_B, z]^T$ is the full state vector, and $\mathbf{x}_{\min} \triangleq [y_{\min}, v_{\min}, p_{A\,\min}, x_{v\,\min}, p_{B\,\min}, z_{\min}]^T$ and $\mathbf{x}_{\max} \triangleq [y_{\max}, v_{\max}, p_{A\,\max}, x_{v\,\max}, p_{B\,\max}, z_{\max}]^T$, are the minimum and maximum feasible values of the states in $\mathbf{x}$. The set $\mathcal{X}_0$ is also referred to as the *feasible region* of the model.

The physically feasible ranges of each of the state variables are identified in the following:

Actuator postion, y: The feasible range of the actuator position is physically limited by the mechanical constraints of the clutch and actuator. The mechanical constraints are given by the lower and upper bounds y_{lb} and y_{ub}, and is modeled by the hardstop force $f_h(y, v)$ which also emulates the flexibility in the mechanical constraints. Letting δ_h denote an upper bound on the mechanical flexibility of these constraints, the feasible range of actuator positions are

$$\begin{aligned} y_{\max} &= y_{\text{ub}} + \delta_h \\ y_{\min} &= y_{\text{lb}} - \delta_h. \end{aligned}$$

Note that the basis functions of the modeled clutch load characteristic (3.2) must be chosen to cover the full feasible range of actuator positions (as illustrated in Figure 3.1).

Actuator velocity, v: The model is valid for all real velocities, thus, the feasible velocity range is given by

$$\begin{aligned} v_{\max} &= \infty \\ v_{\min} &= -\infty. \end{aligned}$$

However, for a given actuator configuration, an upper bound on the physically attainable velocity during normal operation, is easily assessed.

Pre-sliding deflection, z: The dynamics of the pre-sliding deflection state is defined for all real values. However, since the pre-sliding deflection is explained as a seal flexibility property, it is only physically justified for finite deflections with an upper bound $Z_{\max}$, defined by (3.19), which is given by the level of the stiction force and the stiffness of the seals:

$$\begin{aligned} z_{\max} &= Z_{\max} \triangleq F_C/K_z \\ z_{\min} &= -Z_{\max} \triangleq -F_C/K_z. \end{aligned}$$

Pressure, p_A: As stated by Assumption A5.2, the valve flow function $\psi_v(p_A, x_v)$ relies on the assumption that the controlled chamber pressure p_A is constrained from below by the exhaust (atmospheric) pressure P_0, and from above by the supply pressure P_S:

$$\begin{aligned} p_{A\max} &= P_S \\ p_{A\min} &= P_0. \end{aligned}$$

Pressure, p_B: The model is only physically meaningful for a positive pressure p_B, as stated by Assumption A5.1 in Chapter 5. Hence the feasible pressure range of the back-chamber is

$$\begin{aligned} p_{B\max} &= \infty \\ p_{B\min} &= 0. \end{aligned}$$

Valve position, x_v: In the valve flow function $\psi_v(p_A, x_v)$, we have assumed a normalized spool position x_v, where the basis functions of the flow conductance function $c_v(x_v)$ are chosen to cover the full range $x_v \in [-1, 1]$. Hence, the range of feasible valve spool positions are given by

$$\begin{aligned} x_{v\max} &= 1 \\ x_{v\min} &= -1. \end{aligned}$$

6.2 Model properties

6.2.1 Pure-feedback form

The system (6.1) is structurally in the so-called *pure-feedback form* in state-space, for which a constructive nonlinear design by an *integrator backstepping* approach applies. Denoting the states and control input

$$\mathbf{q} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} \triangleq \begin{bmatrix} y \\ v \\ p_A \\ x_v \end{bmatrix}, \quad \boldsymbol{\zeta} = \begin{bmatrix} \zeta_1 \\ \zeta_2 \end{bmatrix} \triangleq \begin{bmatrix} p_B \\ z \end{bmatrix}, \quad u \triangleq K_v u_v,$$

the dynamics (6.1) can be expressed in the form

$$\begin{aligned} \dot{q}_1 &= f_1(q_2) \\ \dot{q}_2 &= f_2(q_1, q_2, q_3, \zeta_1, \zeta_2) \\ \dot{q}_3 &= f_3(q_1, q_2, q_3, q_4) \\ \dot{q}_4 &= f_4(q_4, u) \\ \dot{\zeta}_1 &= g_1(q_1, q_2, \zeta_1) \\ \dot{\zeta}_2 &= g_2(q_2, \zeta_2). \end{aligned} \tag{6.4}$$

Neglecting the internal dynamics, *i.e.*, the $\boldsymbol{\zeta}$-subsystem, the system is clearly in the pure-feedback form:

$$\begin{aligned} \dot{q}_1 &= f_1(q_2) \\ \dot{q}_2 &= f_2(q_1, q_2, q_3) \\ \dot{q}_3 &= f_3(q_1, q_2, q_3, q_4) \\ \dot{q}_4 &= f_4(q_4, u). \end{aligned} \tag{6.5}$$

Since $\boldsymbol{\zeta}$ appears exclusively in the dynamics of the q_2-subsystem, and since the dynamics of the $\boldsymbol{\zeta}$-subsystem depends on the first two states $q_1$and q_2 only, the complete $(\mathbf{q}, \boldsymbol{\zeta})$-system is in pure-feedback form.

6.2.2 Feedback linearizability

Relative degree

All nonlinearities are sufficiently smooth, so that the model (6.1) is sufficiently differentiable with respect to time. It can be shown that in the set

$$\mathcal{X}_u \triangleq \{\forall \mathbf{x} \in \mathcal{X}_0 : P_0 + \varepsilon \leq p_A \leq P_S - \varepsilon\}, \tag{6.6}$$

for some (arbitrary small) constant $\varepsilon > 0$, the system (6.1) has a well-defined *relative degree* equal to 4. This means that the output y is separated from the input u_v by four integrators, *i.e.*, we need to differentiate the output y four times for the control u_v to appear in the dynamics. For $p_A = \{P_0, P_S\}$, the relative degree is undefined for some spool positions because the control vanish, *i.e.*, the system looses controllability in one direction. More precisely, a loss of controllability occurs in the cases

$$\begin{aligned} \textbf{i)} \quad & p_A = P_0 \quad \wedge \quad x_v \leq -X_k \\ \textbf{ii)} \quad & p_A = P_S \quad \wedge \quad x_v \geq X_k, \end{aligned}$$

where $[-X_k, X_k]$ is the null region of the valve (defined in Chapter 5). In either of these cases, the valve flow and the spool–flow gain becomes zero, *i.e.*,

$$\psi_v(p_A, x_v) = 0 \quad \wedge \quad \frac{\partial \psi_v(p_A, x_v)}{\partial x_v} = 0,$$

which means that changes in the spool position x_v (indirectly the control u) do not cause a flow, thus, the pressure p_A cannot be controlled by the input u. The spool–flow gain of the valve flow function $\psi_v(p_A, x_v)$ is plotted in Figure 6.1 (page 97).

This loss of controllability for $p_A = \{P_0, P_S\}$ has an obvious physical cause: **(i)** if the chamber pressure p_A equals atmospheric pressure P_0, flow from the chamber to atmosphere is no longer possible (because flow is only possible to a lower pressure), hence, the system is uncontrollable for spool positions in the exhaust region of the valve. Likewise, **(ii)** if the chamber pressure equals supply pressure, $p_A = P_S$, flow from the supply into the chamber is no longer possible, hence, the system is uncontrollable for spool positions in the supply region.

Normal and input-output forms

In the set $\mathcal{X}_u$ defined above, which covers the full feasible region of the state-space except $p_A = \{P_0, P_S\}$, the system (6.1) has a well defined relative degree $r = 4$, which implies that it is *input-output feedback linearizable* in $\mathcal{X}_u$. Hence, there exists a well defined change of coordinates $\mathbf{\Phi}(\mathbf{x}) = [\mathbf{z}, \boldsymbol{\zeta}]^T$ for $\forall \mathbf{x} \in \mathcal{X}_u$, which transforms the system to the *normal form*

$$\begin{aligned} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= z_3 \\ \dot{z}_3 &= z_4 \\ \dot{z}_4 &= a(\mathbf{z}, \boldsymbol{\zeta}) + b(\mathbf{z}, \boldsymbol{\zeta}) \cdot u \\ \dot{\boldsymbol{\zeta}} &= g(\mathbf{z}, \boldsymbol{\zeta}) \end{aligned}$$

with the main state vector $\mathbf{z} = [z_1, z_2, z_3, z_4]^T \in \Omega_z \subset \mathbb{R}^4$, and the states of the *internal dynamics* $\boldsymbol{\zeta} = [\zeta_1, \zeta_2]^T \in \Omega_\zeta \subset \mathbb{R}^2$, where $\Omega_z \times \Omega_\zeta = \boldsymbol{\Phi}(\mathcal{X}_u)$. It can further be shown that the internal dynamics is input-to-state stable with respect to $\mathbf{z} = \mathbf{0}$, thus, the system is *minimum phase*.

Since both $a(\cdot)$ and $b(\cdot)$ are sufficiently differentiable, the system is *fully feedback linearizable* for $\forall \mathbf{x} \in \mathcal{X}_u$, and the system can be transformed to the *input-output form*

$$y^{(n)} = \bar{a}(\mathbf{q}, \mathbf{u}) + \bar{b}(\mathbf{z}, \mathbf{u}) \cdot u^{(2)},$$

where the nonlinear functions $\bar{a}(\cdot)$ and $\bar{b}(\cdot)$ may depend on the output and its derivatives $\mathbf{q} \triangleq [y, \dot{y}, \cdots, y^{(5)}]^T \in \mathbb{R}^6$, and the input and its first-order derivative $\mathbf{u} \triangleq [u, \dot{u}] \in \mathbb{R}^2$.

6.2.3 Linear parametrization

A particular effort has been put into modeling nonlinearities in parameter-affine forms, with a minimum number of parameters. As a result, the model is particularly suited for parameter estimation and adaptive control: The nonlinear clutch load characteristic $f_l(y)$ given by (3.2) is affine in its tunable parameter vector $\boldsymbol{\theta}_l$, where the scaling and center vectors $\mathbf{w}$ and $\mathbf{c}$ of $\boldsymbol{\phi}_l(y)$ are viewed as fixed parameters, *i.e.*, not tunable. As discussed in Section 5.3 and Section 5.4, the nonlinear flow functions $\psi_r(p_B)$ and $\psi_v(p_A, x_v)$ of the restriction and valve, can be expressed in the piecewise parameter-affine forms (5.22)–(5.23), and (5.59)–(5.60), with their affine tunable parameters being $\boldsymbol{\theta}_r = [\rho_0 C_r, \rho_0 C_r b_r]^T$ and $\boldsymbol{\theta}_v = [\rho_0 C_v \boldsymbol{\theta}_\mu, \rho_0 C_v b_v \boldsymbol{\theta}_\mu]^T$, respectively. Furthermore, the saturation function $\pi(\cdot)$ satisfies

$$\pi(x) = x, \quad \forall |x| \leq 1 - \varepsilon_\pi,$$

where the smoothing width is negligible small, *i.e.*, $\varepsilon_\pi \ll 1$. Since the control input term is physically limited to $u \triangleq K_v u_v \in [-1, 1]$, and because $K_v U_{v0} \ll 1$, both the offset parameter U_{v0} and the control input u_v appear ian affine form in practically the full operating range of the valve, *i.e.*, for $|K_v U_{v0} + K_v u_v| \leq 1 - \varepsilon_\pi$. Consequently, the valve dynamics is *linearly parametrizable* in its tunable parameters, and *input-affine* in the practical operating range of the valve.

The inverse of the chamber volumes,

$$\begin{aligned} V_A(y)^{-1} &= \frac{1}{V_{A0} + A_A y} \\ V_B(y)^{-1} &= \frac{1}{V_{B0} - A_B y}, \end{aligned}$$

are obviously not linearly parametrizable. However, the inverse of the volumes V_A and V_B may be accurately approximated by a 2nd-order polynomial, *i.e.*,

$$V(y)^{-1} = a_0 + a_1 y + a_2 y^2,$$

which is obvious from the plots of $V_A(y)^{-1}$ and $V_B(y)^{-1}$ in Figure 6.3. Hence, using this approximation, a complete linear parametrization of all the tunable parameters in the design model (6.1), may be obtained. Table 6.1, summarizes typical values of the parameters of the model.

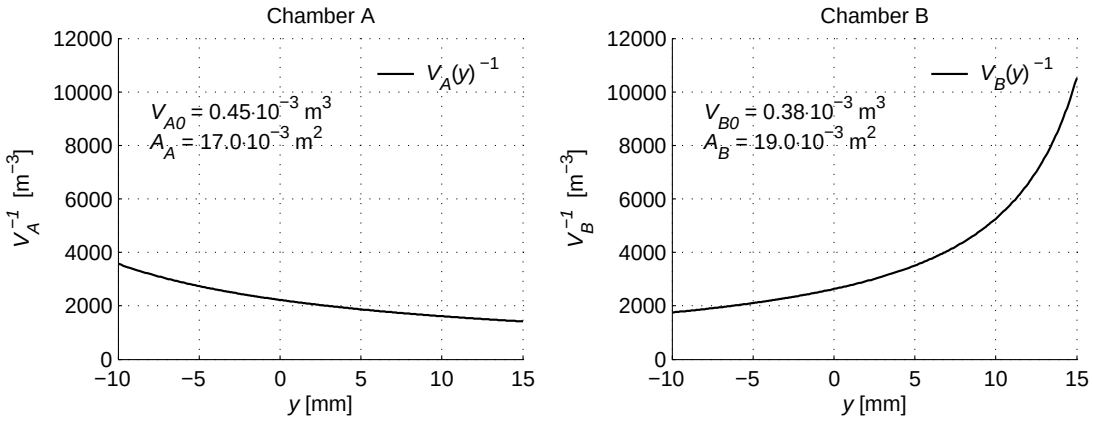

Figure 6.3: The inverse chamber volumes, $V_A^{-1}(y)$ and $V_B^{-1}(y)$.

6.3 Summary

The work on modeling of the electro-pneumatic clutch actuator in Chapters 3–5, is in this chapter recapitulated in a 6th-order design model. The model consists of the actuator position, velocity, friction (seal) deflection, pressures of both chambers, and the valve spool position, as dynamic states. In the region of normal operation, the resulting model is smooth, linearly parametrizable, and fully feedback linearizable with relative degree four, which means that it can be transformed to a (parameter-affine) input-output form where constructive procedures for (adaptive) output-feedback control utilizing high-gain observers are available. Furthermore, the model is in pure-feedback form, which makes it applicable for a nonlinear state-feedback design by a backstepping approach, and for output-feedback by an observer-based backstepping approach, provided that an asymptotic observer is available.

Table 6.1: Model parameters used in the full 6th-order model of the pneumatic clutch actuator.

Parameter	Unit	Value	Parameter	Unit	Value
K_z	$\mathrm{N/m}$	$400 \cdot 10^3$	T_0	K	293
$D_{\dot{z}}$	$\mathrm{N\,s/m}$	$5 \cdot 10^3$	P_0	Pa	$1 \cdot 10^5$
ε_0	$\mathrm{m/s}$	$0.01 \cdot 10^{-3}$	P_S	Pa	$10 \cdot 10^5$
M	kg	20	R	$\mathrm{J/(kg\,K)}$	288
A_A	$\mathrm{m^2}$	$1.70 \cdot 10^{-2}$	ρ_0	$\mathrm{kg/m^3}$	1.185
A_B	$\mathrm{m^2}$	$1.86 \cdot 10^{-2}$	τ_v	s	$15 \cdot 10^{-3}$
A_0	$\mathrm{m^2}$	$0.16 \cdot 10^{-2}$	K_v	$1/\mathrm{V}$	0.1
V_{A0}	$\mathrm{m^3}$	$0.45 \cdot 10^{-3}$	U_{v0}	V	0.01
V_{B0}	$\mathrm{m^3}$	$0.38 \cdot 10^{-3}$	b_v	—	0.8
D_v	$\mathrm{N\,s/m}$	$5 \cdot 10^3$	b_r	—	0.6
F_C	N	200	C_v	$\mathrm{m^3/(Pa \cdot s)}$	$3.0 \cdot 10^{-8}$
K_h	$\mathrm{N/m}$	$10 \cdot 10^6$	C_r	$\mathrm{m^3/(Pa \cdot s)}$	$1.0 \cdot 10^{-8}$
D_h	$\mathrm{N\,s/m}$	$5 \cdot 10^3$	ε_h	m	$0.05 \cdot 10^{-3}$

Part II

Control Design

Chapter 7

Technical Preliminaries

This chapter recapitulates some technical preliminaries, tools, and terminology, used in the subsequent chapters on observer and controller design.

7.1 Technical lemmas

We frequently use *completion of squares* to obtain upper bounds on sign-indefinite terms.

Lemma 26 (Completion of squares) *For any real-valued scalar variables x, y, and $\varepsilon > 0$, we have*

$$xy \leq \frac{\varepsilon}{2} x^2 + \frac{1}{2\varepsilon} y^2. \tag{7.1}$$

For vectors $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$, note that $\mathbf{x}^T\mathbf{y} \leq \left|\mathbf{x}^T\mathbf{y}\right| \leq |\mathbf{x}|\,|\mathbf{y}|$.

We use the *Mean Value Theorem* (see *e.g.* [47]) to rewrite differences of nonlinear functions in time-varying affine forms.

Lemma 27 (Mean Value Theorem) *Let the scalar function $f(\mathbf{x})$ with inputs $\mathbf{x} \in \mathbb{R}^n$ be continuously differentiable on the open set $\mathcal{S} \in \mathbb{R}^n$. For any two points $\mathbf{x}_1$ and $\mathbf{x}_2$ in $\mathcal{S}$, where the line segment joining $\mathbf{x}_1$ and $\mathbf{x}_2$ are also in $\mathcal{S}$, there exists a point $\mathbf{x}^* \in \mathcal{S}$ such that*

$$f(\mathbf{x}_2) - f(\mathbf{x}_1) = \left(\frac{\partial f}{\partial \mathbf{x}}\right)^T \Bigg|_{\mathbf{x}^* \in \mathcal{S}} (\mathbf{x}_2 - \mathbf{x}_1). \tag{7.2}$$

The following convergence lemma is useful to establish ISS bounds [52, Lemma C.5].

Lemma 28 (ISS bound) *Suppose the variable $v(t) \in \mathbb{R}^+$ satisfies the differential inequality*

$$\dot{v} \leq -cv + dw(t)^2, \tag{7.3}$$

for some constants $c > 0$ and $d > 0$, and $w(t) \in \mathbb{R}^+$ is a bounded input ($w \in \mathcal{L}_\infty$). Then

$$v(t) \leq v(0)\, e^{-c\cdot t} + \frac{d}{c} \|w(t)\|_\infty^2. \tag{7.4}$$

7.2 Smooth saturation

Motivated by the use of smooth saturation of parameter estimates in nonlinear adaptive control (proposed by Teel in [95], see also *e.g.* [55, 105, 107]), we will use smooth saturation to obtain global robustness of the observers and output-feedback control system presented in Chapters 8 and 9. Because the saturation is smooth, it preserves differentiability, and with that, compatibility with backstepping.

We use a smooth saturation function defined as

$$\pi\left(x, x_{\mathrm{lb}}, x_{\mathrm{ub}}, \varepsilon\right) \triangleq x + \varepsilon \cdot g\left(\frac{-x + x_{\mathrm{lb}}}{\varepsilon}\right) - \varepsilon \cdot g\left(\frac{x - x_{\mathrm{ub}}}{\varepsilon}\right), \tag{7.5}$$

where $\varepsilon > 0$ is an arbitrary small design constant referred to as the smoothing width of $\pi\left(\cdot\right)$, and $g\left(\cdot\right)$ is a *smooth plus function* with unity slope and smoothing interval $[-1, 1]$. For brevity of notation, we may omit the constant parameters as arguments and write $\pi\left(x\right) = \pi\left(x, x_{\mathrm{lb}}, x_{\mathrm{ub}}, \varepsilon\right)$. The smooth plus function $g\left(\cdot\right)$ is constructed as a spline function according to

$$g\left(x\right) = \begin{cases} x, & x > 1 \\ \frac{5}{32} + \frac{1}{2}x + \frac{15}{32}x^2 - \frac{5}{32}x^4 + \frac{1}{32}x^6, & |x| \leq 1 \\ 0, & x < -1 \end{cases}, \tag{7.6}$$

which is three times differentiable ($g\left(x\right) \in C^3$). The smooth saturation function (7.5) and the above smooth plus function (7.6), is plotted in Figure 5.9.

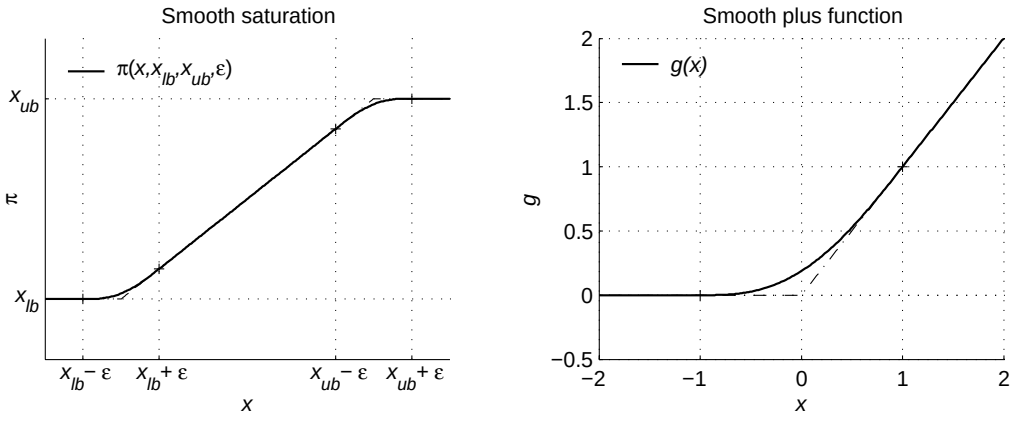

Figure 7.1: Left: The smooth saturation function $\pi\left(\cdot\right)$ with upper and lower bounds x_{ub} and x_{lb}, and smoothing width ε. Right: The smooth plus function.

The smooth saturation function (7.5) will be used for a robust re-design of the observers, and it useful to point out the following properties

P7.1)	$\pi\left(x\right) \equiv x$	$, \forall x \in [x_{\mathrm{lb}} + \varepsilon, x_{\mathrm{ub}} - \varepsilon]$
P7.2)	$\pi\left(x\right) \in [x_{\mathrm{lb}}, x_{\mathrm{ub}}]$	$, \forall x \in \mathbb{R}.$

In Lyapunov-based designs of controller and identifier, we may use quadratic control Lyapunov functions (CLF). However, when introducing saturation of states,

it is convenient to use a modified CLF, as proposed in [95]. We illustrate this for the case with a scalar quadratic function

$$V_0\left(\tilde{\theta}\right) = \frac{1}{2}\tilde{\theta}^2, \tag{7.7}$$

where $\tilde{\theta}(t) \triangleq \theta - \hat{\theta}(t)$ is the estimation error, with $\hat{\theta}(t)$ being a time-varying estimate of the constant parameter θ. When the estimate $\hat{\theta}$ is replaced by the saturated estimate $\hat{\theta}_\pi \triangleq \pi(\hat{\theta})$, the estimation error becomes $\tilde{\theta}_\pi \triangleq \theta - \pi(\hat{\theta})$. We may then use a modified function in the form

$$V\left(\tilde{\theta}\right) \triangleq \int_0^{\tilde{\theta}} \left[\theta - \pi\left(\sigma - \theta\right)\right] d\sigma, \tag{7.8}$$

which has the gradient

$$\frac{\partial V}{\partial \tilde{\theta}}\left(\tilde{\theta}\right) = \theta - \pi\left(\hat{\theta}\right) = \tilde{\theta}_\pi. \tag{7.9}$$

The integral function (7.8) is positive definite and radially unbounded, like its quadratic counterpart (7.7), however, its gradient is bounded, unlike the gradient of the quadratic function (7.7). In Figure 7.2, both the quadratic and the integrated scalar functions and their gradients are plotted to illustrate their main properties.

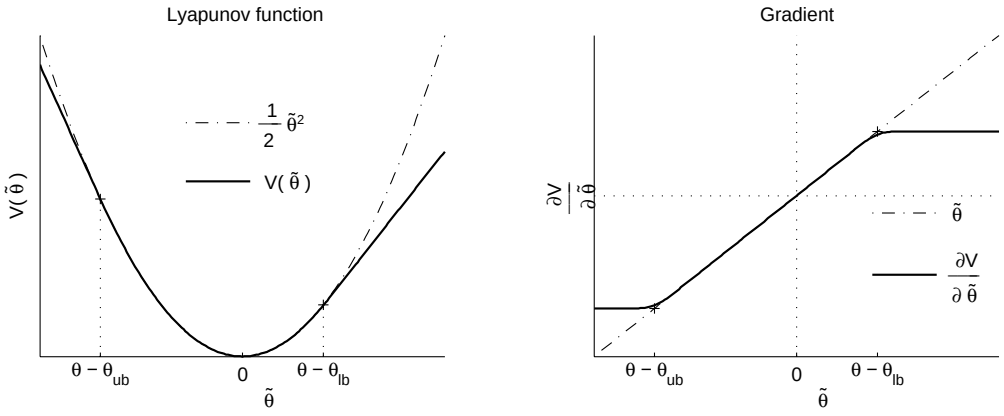

Figure 7.2: Comparison between the quadratic function (7.7) and the integrated function (7.8) utilizing smooth saturation.

7.3 Discontinuous projection

Discontinuous projection will be used to obtain improved transient performance of the observers in Chapters 8 and 9. In this section, we recapitulate the scalar discontinuous projection operator, whose multi-variable extension is commonly used as a robustification tool in the adaptive control literature. See *e.g.* [33, 52, 107] and the references therein.

For the scalar system $\dot{x} = f(x,t)$ with lower and upper bounds x_{lb} and x_{ub} on x, respectively, we define the scalar discontinuous projection

$$\mathcal{P}(f,x) = \mathcal{P}(f,x,x_{\text{lb}},x_{\text{ub}}) \triangleq \begin{cases} 0, & x \geq x_{\text{ub}} \quad \wedge \quad f > 0 \\ 0, & x \leq x_{\text{lb}} \quad \wedge \quad f > 0 \\ f, & \text{otherwise} \end{cases} . \tag{7.10}$$

Note that the projection (7.10) satisfies

$$\textbf{P7.3)} \qquad \mathcal{P}(f,x) \equiv f \quad , \forall x \in \langle x_{\text{lb}}, x_{\text{ub}} \rangle ,$$

and with initial state $x(0) \in [x_{\text{lb}}, x_{\text{ub}}]$, and constant bounds x_{lb} and x_{ub}, the projection $\dot{x} = \mathcal{P}(f,x)$ guarantees that the state is bounded, *i.e.*,

$$\textbf{P7.4)} \qquad x \in [x_{\text{lb}}, x_{\text{ub}}] \quad , \forall t \geq 0.$$

For proof of P7.4), see *e.g.* [106].

Remark 29 *For time-varying bounds, $x_{\text{lb}}(t)$ and $x_{\text{ub}}(t)$, the projection can only guarantee boundedness according to*

$$x \in \left[\inf_{\forall t \geq 0} x_{\text{lb}}(t) , \sup_{\forall t \geq 0} x_{\text{ub}}(t) \right] .$$

7.4 Characterization of uncertainties

A state-space model of a physical system will always contain uncertainties in some form, due to errors in parameters, model inaccuracies, simplifications, or unknown exogenous disturbances. These uncertainties can usually be lumped together and be represented by additive time-varying terms in the model, such as

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x},\mathbf{u},t) + \boldsymbol{\delta}(t) , \tag{7.11}$$

where $\mathbf{f}(\mathbf{x},\mathbf{u},t)$ represents the precisely known nominal dynamic model of the system with the states $\mathbf{x}$, the inputs $\mathbf{u}$, and $\boldsymbol{\delta}$ a lumped uncertainty term. In the following, we make an attempt to clarify some commonly used characterizations of uncertainties.

Uncertainties for a given state-space model may be characterized by its structural properties. An uncertainty may be characterized as an *unstructured uncertainty* if it may be represented as a purely time-varying quantity $\delta = \delta(t)$, *e.g.* as a result of unmodeled dynamics, or exogenous disturbances. Unstructured uncertainties, like $\delta(t)$, are also referred to as *uncertain dynamics*, or simply *disturbances*. When the uncertainty is also a function of the modeled states or inputs to the system, like $\delta = \delta(\mathbf{x},\mathbf{u},t)$, it may be characterized as a *structured uncertainty*, or an *uncertain nonlinearity*. In the special case when the structured uncertainty is a pure static relation of the states or other inputs, such as $\delta = \delta(\mathbf{x},\mathbf{u})$, it is usually referred to as an *uncertain static nonlinearity*. In the general case when a structured uncertainty

is also time-varying ($\delta = \delta(\mathbf{x}, \mathbf{u}, t)$), it may be referred to as an *uncertain dynamic nonlinearity*.

A class of structured uncertainties are so-called *parametric uncertainties*, which are uncertainties due to uncertain parameters. For example, for the nonlinear function $f(\mathbf{x}, \boldsymbol{\theta})$, where the estimate $\hat{\boldsymbol{\theta}}$ is used for the actual parameter vector $\boldsymbol{\theta}$, the parametric uncertainty is given as

$$\delta(\mathbf{x}) = \delta\left(\mathbf{x}, \boldsymbol{\theta}, \hat{\boldsymbol{\theta}}\right) = f(\mathbf{x}, \boldsymbol{\theta}) - f\left(\mathbf{x}, \hat{\boldsymbol{\theta}}\right). \tag{7.12}$$

An important class of parametric uncertainties is obtained for nonlinearities which can be written in parameter-affine form, $f(\mathbf{x}, \boldsymbol{\theta}) = \boldsymbol{\phi}(\mathbf{x})^T \boldsymbol{\theta}$, so that the parametric uncertainty can be expressed as

$$\delta(\mathbf{x}) = \delta\left(\mathbf{x}, \tilde{\boldsymbol{\theta}}\right) = \boldsymbol{\phi}(\mathbf{x})^T \tilde{\boldsymbol{\theta}}, \tag{7.13}$$

affine in its parameter uncertainty $\tilde{\boldsymbol{\theta}} \triangleq \boldsymbol{\theta} - \hat{\boldsymbol{\theta}}$.

Using robust control techniques, high gain is usually required in order to account for uncertainties in the model. In order to reduce the amount of uncertainty that needs to be accounted for by robust control, a structured uncertainty can often be partitioned into a parametric part and a non-parametric part according to

$$\delta(\mathbf{x}, \mathbf{u}, t) = \boldsymbol{\phi}(\mathbf{x}, \mathbf{u})^T \tilde{\boldsymbol{\theta}} + \Delta(\mathbf{x}, \mathbf{u}, t). \tag{7.14}$$

In this way, adaptive techniques can be used to account for the parametric uncertainty $\boldsymbol{\phi}(\mathbf{x}, \mathbf{u})^T \tilde{\boldsymbol{\theta}}$, and robust techniques can be used to deal with the remaining uncertainty $\Delta(\mathbf{x}, \mathbf{u}, t)$, resulting in less need for high gain control. Furthermore, it may usually be advantageous to exploit the structure of the remaining uncertainty $\Delta(\mathbf{x}, \mathbf{u}, t)$ in order to reduce unnecessary high gain. For this objective, we may want to express the structured uncertainty $\Delta(\mathbf{x}, \mathbf{u}, t)$ in the affine form

$$\Delta(\mathbf{x}, \mathbf{u}, t) = \mathbf{q}(\mathbf{x}, \mathbf{u})^T \cdot \mathbf{w}(\mathbf{x}, \mathbf{u}, t), \tag{7.15}$$

where $\mathbf{q}(\mathbf{x}, \mathbf{u})$ is a known function of the system states and inputs, and $\mathbf{w}$ is an unknown, bounded uncertainty, or disturbance. Then, the structural properties of the uncertainty, given by $\mathbf{q}(\mathbf{x}, \mathbf{u})$, can be exploited by so-called *nonlinear damping* terms in the controller, using the knowledge of $\mathbf{q}(\mathbf{x}, \mathbf{u})^T$ to increase the gain only when required.

Uncertain dynamics typically arises from simplifications of the model, or are caused by exogenous disturbances, or a combination of both. A common simplification in fluid power systems, is to disregard fast dynamics in the electromagnetic valve actuator, or model it by a simplified, low-order model, which causes what is referred to as unmodeled input dynamics (or actuator dynamics). Another common simplification is to disregard the fast dynamics of the sensor, resulting in unmodeled output dynamics (or sensor dynamics). Further examples of unmodeled dynamics, are the unmodeled supply dynamics due to treating the supply pressure as constant,

and our use of the reduced-order pressure dynamics instead of the full-order air dynamics, resulting in unmodeled temperature dynamics in our system. An example of an *exogenous disturbance*, is the effect of heat transfer on the pressure dynamics of the pneumatic chamber, which we have disregarded. Note that these uncertainties become structured uncertainties when they enter the nominal state-space model through some nonlinear function of the modeled states — even though they are purely unstructured by themselves.

Uncertain static *nonlinearities* are due to inaccurate static terms in the model. Typical examples of uncertain nonlinearities in the model of the electro-pneumatic clutch actuator, are the parametric uncertainty due to uncertain parameters in the clutch load characteristic, and inaccuracies in the model of the static friction characteristic.

Chapter 8

Nonlinear Observer Design

In this chapter we address the design of nonlinear observers for the estimation of the unmeasured states of the electro-pneumatic clutch actuation system.

In Section 8.1, we review the design model and discuss some of its properties relevant for the subsequent observer design. In Section 8.2, we show that the system model itself can be used as an open-loop observer for the unmeasured states of the electro-pneumatic actuator. In Section 8.3, we present a full-order observer, and introduces a redesign for improved robustness and performance. In Section 8.4, we present a simplified reduced-order observer based on a change of coordinates, and its robust redesign. We analyze some results from an experimental implementation of the observers in Section 8.6, and summarizes the chapter in Section 8.7.

8.1 Design model

The 6th-order smooth design model of the electro-pneumatic clutch actuator and some of its properties which are relevant for an observer design, are reviewed in this section. Since we never intentionally attempt to drive the actuator into end stroke, we neglect the hardstop term $f_h(y, v)$ in the dynamics (6.1). The full 6th-order model used for nonlinear observer design in this chapter, is given in state-space form as

$$\begin{aligned}
\dot{y} &= v \\
\dot{v} &= \frac{A_0 P_0}{M} + \frac{A_A}{M} p_A - \frac{A_B}{M} p_B - \frac{1}{M} f_l(y) - \frac{1}{M} f_f(v, z) \\
\dot{p}_A &= -A_A \frac{1}{V_A(y)} v p_A + \rho_0 T_0 R C_v \frac{1}{V_A(y)} \psi_v(p_A, x_v) \\
\dot{x}_v &= -\frac{1}{\tau_v} x_v + \frac{1}{\tau_v} \pi_u(u + U_0) \\
\dot{p}_B &= A_B \frac{1}{V_B(y)} v \cdot p_B + \rho_0 T_0 R C_r \frac{1}{V_B(y)} \psi_r(p_B) \\
\dot{z} &= v - \frac{K_z}{F_C} |v|_s z.
\end{aligned} \tag{8.1}$$

Here, the states are y, v, p_A, x_v, p_B, z, the scaled control input is $u \triangleq K_v u_v$ and scaled offset $U_0 \triangleq K_v U_{v0}$, and the measured output is the position y. The functions $f_l(y)$ and $f_f(v,z)$ are the nonlinear clutch load characteristic and the friction force, given by (3.2) and (3.16); $V_A(y)$ and $V_B(y)$ are the chamber volumes, which are positive linear functions of y given by (4.3) and (4.7); $\psi_v(p_A, x_v)$ and $\psi_r(p_B)$ are the valve and restriction flow functions given by (5.62) and (5.25), respectively; and $|\cdot|_s$ is a smooth approximation to the absolute value, defined as

$$|v|_s \triangleq \sqrt{v^2 + \varepsilon_0^2}.$$

For the subsequent stability analysis, it is useful to note that the valve flow function $\psi_v(p_A, x_v)$ is monotonically decreasing in p_A, increasing in x_v, and bounded, as can be seen from the plots in Figure 6.1 (page 97). More precisely, $\psi_v(p_A, x_v)$ satisfies

$$\begin{aligned}
\frac{\partial \psi_v}{\partial p_A}(p_A, x_v) &\leq 0, \quad \forall p_A \in [P_0, P_S], \forall x_v \in [-1, 1] \\
\frac{\partial \psi_v}{\partial x_v}(p_A, x_v) &\geq 0, \quad \forall p_A \in [P_0, P_S], \forall x_v \in [-1, 1] \\
|\psi_v(p_A, x_v)| &\leq P_S, \quad \forall p_A \in [P_0, P_S], \forall x_v \in [-1, 1].
\end{aligned}$$

In a similar manner, the restriction flow function $\psi_r(p_B)$, plotted in Figure 6.2 (page 97), is decreasing in p_B, and bounded from above, *i.e.*,

$$\begin{aligned}
\frac{\partial \psi_r}{\partial p_B}(p_B) &\leq 0, \quad \forall p_B \in [0, \infty\rangle \\
\psi_r(p_B) &\leq P_0, \quad \forall p_B \in [0, \infty\rangle.
\end{aligned}$$

Region of validity

The *region of validity*, or *region of feasibility*, of (8.1) is the set $\mathcal{X}_0 \subset \mathbb{R}^6$ defined by

$$\mathcal{X}_0 \triangleq \{\forall \mathbf{x} : \mathbf{x}_{\min} \leq \mathbf{x} \leq \mathbf{x}_{\max}\}, \tag{8.2}$$

where $\mathbf{x} \triangleq [y, v, p_A, x_v, p_B, z]^T$ is the full state vector, and $\mathbf{x}_{\min} \triangleq [y_{\min},\ v_{\min},\ p_{A\min},\ x_{v\min},\ p_{B\min},\ z_{\min}]^T$ and $\mathbf{x}_{\max} \triangleq [y_{\max},\ v_{\max},\ p_{A\max},\ x_{v\max},\ p_{B\max},\ z_{\max}]^T$, are the physically feasible ranges of the state vector $\mathbf{x}$, which are identified in Chapter 6.

Region of normal operation

In normal operation of the actuator, the states will stay within some compact region in state-space, *i.e.*, each state will be bounded. In the following, we assume knowledge about some lower and upper bounds on each state in normal operation, which we later utilize for a robust re-design to achieve global stability and improved transient performance of the observer. We denote these lower and upper bounds by $\mathbf{x}_{\text{lb}} \triangleq [y_{\text{lb}}, v_{\text{lb}}, p_{A\,\text{lb}}, x_{v\,\text{lb}}, p_{B\,\text{lb}}, z_{\text{lb}}]^T$ and $\mathbf{x}_{\text{ub}} \triangleq [y_{\text{ub}}, v_{\text{ub}}, p_{A\,\text{ub}}, x_{v\,\text{ub}}, p_{B\,\text{ub}}, z_{\text{ub}}]^T$, and introduce a *region of normal operation*, defined as

$$\mathcal{X} \triangleq \{\forall \mathbf{x} \in \mathcal{X}_0 \colon \mathbf{x}_{\text{lb}} \leq \mathbf{x} \leq \mathbf{x}_{\text{ub}}\}. \tag{8.3}$$

Scaling of states

In order to minimize numerical errors with an implementation, we introduce a practical scaling of states, utilizing the bounds in the definition (8.3) of the region of normal operation.

Define a vector of scaling factors $\mathbf{s} = [s_1, \cdots, s_6]^T$ as

$$\mathbf{s} \triangleq \mathbf{x}_{\text{ub}} - \mathbf{x}_{\text{lb}},$$

and introduce the scaling of states

$$\bar{\mathbf{x}} \triangleq \mathbf{s}^{-1} \times \mathbf{x},$$

where $\bar{\mathbf{x}} = [\bar{x}_1, \cdots, \bar{x}_6]^T$ is the scaled state vector.

A practical result of this scaling, is that the estimation error in scaled states becomes *normalized* with respect to the set $\mathcal{X}$, made precise in the following:

Proposition 30 (Scaling of States) *Define a vector* $\bar{\mathbf{e}} = [\bar{e}_1, \cdots, \bar{e}_6]^T$ *of scaled estimation errors as*

$$\bar{\mathbf{e}} \triangleq \mathbf{s}^{-1} \times (\mathbf{x} - \hat{\mathbf{x}}),$$

where $\hat{\mathbf{x}}$ *is the estimate of* $\mathbf{x}$*, and the scaling vector* $\mathbf{s} \triangleq \mathbf{x}_{\text{ub}} - \mathbf{x}_{\text{lb}}$*. With* $\mathbf{x}$ *and* $\hat{\mathbf{x}}$ *contained in* $\mathcal{X}$*, defined by (8.3), the magnitude of the scaled observer errors,* $\bar{e}_i = s_i^{-1} \cdot (x_i - \hat{x}_i)$*,* $i = 1, \cdots, 6$*, is* normalized, i.e.*,*

$$(\mathbf{x}, \hat{\mathbf{x}}) \in \mathcal{X} \Longrightarrow |\bar{e}_i| \leq 1, \qquad i = 1, \cdots, 6.$$

Remark 31 *We employ this scaling of states when implementing the observer, both for simulations and experimentally in the test rig. For clarity, however, we omit scaling in our presentation of the observer and controller design.*

Observability

In Section 8.3, we propose an observer with output-injection terms for the three upper states, y, v and p_A. We require that these states are observable from the output y, which is straightforward to establish. A simple definition of observability of the state $\xi = [\xi_1, \xi_2, \xi_3]^T \triangleq [y, v, p_A]$ from the output y, is that with y known, it is possible to identify the corresponding state $\boldsymbol{\xi}$. Note that the assumption that y is known is a stronger concept than being measured, it means that not only y may be considered known, but also its higher-order derivatives. Hence, observability of $\xi_1 = y$ and $\xi_2 = \dot{y}$ is trivial. To establish observability of $\boldsymbol{\xi}$ from y, in general, we may assume that p_B and z (and their derivatives) are known. Denoting the known terms by $h(y, p_B, z, \dot{z}) \triangleq -\frac{A_B}{M} p_B - \frac{K_z}{M} z - \frac{D_{\dot{z}}}{M} \dot{z} + \frac{A_0 P_0}{M} - \frac{1}{M} f_l(y)$, from (8.1) we see that the observability mapping for the three first states may be expressed as

$$\mathbf{y}' \triangleq \begin{bmatrix} y \\ \dot{y} \\ \ddot{y} \end{bmatrix} \triangleq \mathcal{H}(\boldsymbol{\xi}) = \begin{bmatrix} \xi_1 \\ \xi_2 \\ -\frac{D_v}{M}\xi_2 + \frac{A_A}{M}\xi_3 + h(y, p_B, z, \dot{z}) \end{bmatrix}.$$

Clearly, the mapping $\mathcal{H}(\boldsymbol{\xi})$ is invertible and we can write $\boldsymbol{\xi} = \mathcal{H}^{-1}(\mathbf{y}')$, which establishes that $\boldsymbol{\xi}$ is observable from y.

8.2 Open-loop observer

In this section we analyze the use of the model (8.1) as an open-loop observer for the unmeasured states of the system, in the case that only the position y is measured. We show that this observer is asymptotically stable provided both the system and observer states remains in the region of validity $\mathcal{X}_0$ of the model (8.1), and the convergence properties of the observer is demonstrated by simulations.

An open-loop observer for the system (8.1) is given as

$$\begin{aligned}
\dot{\hat{v}} &= \frac{A_0 P_0}{M} + \frac{A_A}{M}\hat{p}_A - \frac{A_B}{M}\hat{p}_B - \frac{1}{M}f_l(y) - \frac{1}{M}f_f(\hat{v},\hat{z}) \\
\dot{\hat{p}}_A &= -A_A\frac{1}{V_A(y)}\hat{v}\hat{p}_A + \rho_0 T_0 R C_v \frac{1}{V_A(y)}\psi_v(\hat{p}_A,\hat{x}_v) \\
\dot{\hat{x}}_v &= -\frac{1}{\tau_v}\hat{x}_v + \frac{1}{\tau_v}\pi_u(u+U_0) \\
\dot{\hat{p}}_B &= A_B\frac{1}{V_B(y)}\hat{v}\hat{p}_B + \rho_0 T_0 R C_r \frac{1}{V_B(y)}\psi_r(\hat{p}_B) \\
\dot{\hat{z}} &= \hat{v} - \frac{K_z}{F_C}|\hat{v}|_s\,\hat{z},
\end{aligned} \tag{8.4}$$

where the states $\hat{\mathbf{x}}_u \triangleq [\hat{v},\hat{p}_A,\hat{x}_v,\hat{p}_B,\hat{z}]^T$ are the estimates of the unmeasured states $\mathbf{x}_u \triangleq [v,p_A,x_v,p_B,z]^T$, and y is the measured output.

8.2.1 Stability and convergence

In order to be able to use an observer in an actual application, *e.g.* for clutch actuation in an heavy-duty truck, we must be able to ascertain that the observer is stable under all possible conditions, and preferably, that its estimates converges to the actual states, with negligible error, during normal operation of the actuator.

An analysis of the stability and convergence properties of the observer error dynamics is in general useful of two main reasons: it provides improved insight into the dynamics of the observer, and it may reveal potentially destabilizing nonlinearities that would make the observer algorithm prone to *blow up* or *drift* unboundedly when subjected to particular operating conditions.

The following analysis establishes that the open-loop observer is asymptotically stable in the entire region of validity of the model. This is a particular useful result with respect to nonlinear observer design for electro-pneumatic actuators, because it shows that the pneumatic actuators have some inherent stability properties which enables the design of simple nonlinear observers.

Observer error dynamics

The stability and convergence properties of the observer is established by analyzing the dynamics of the resulting observer errors

$$\begin{aligned}\tilde{v} &\triangleq v - \hat{v}\\ \tilde{z} &\triangleq z - \hat{z}\\ \tilde{p}_A &\triangleq p_A - \hat{p}_A\\ \tilde{p}_B &\triangleq p_B - \hat{p}_B\\ \tilde{x}_v &\triangleq x_v - \hat{x}_v.\end{aligned}$$

The first step of the analysis is to express the dynamics of the observer error in a suitable form, which is done in detail in this section.

First consider the velocity dynamics. The dynamics of the estimation error

$$\tilde{v} \triangleq v - \hat{v}, \tag{8.5}$$

is obtained by differentiating $\tilde{v}$:

$$\begin{aligned}\dot{\tilde{v}} &= \dot{v} - \dot{\hat{v}}\\ &= \frac{A_0 P_0}{M} + \frac{A_A}{M}p_A - \frac{A_B}{M}p_B - \frac{1}{M}f_l(y) - \frac{1}{M}f_f(v,z)\\ &\quad - \left(\frac{A_0 P_0}{M} + \frac{A_A}{M}\hat{p}_A - \frac{A_B}{M}\hat{p}_B - \frac{1}{M}f_l(y) - \frac{1}{M}f_f(\hat{v},\hat{z})\right)\\ &= \frac{A_A}{M}\tilde{p}_A - \frac{A_B}{M}\tilde{p}_B - \frac{1}{M}\left[f_f(v,z) - f_f(\hat{v},\hat{z})\right].\end{aligned}$$

The only nonlinearity entering the $\tilde{v}$–dynamics, is the resulting error in the friction force f_f, given by (3.16). The friction error can be written as

$$f_f(v,z) - f_f(\hat{v},\hat{z}) = D_v\tilde{v} + K_z\tilde{z} + D_{\dot{z}}\dot{\tilde{z}},$$

where the last term is the time-derivative of $\tilde{z}$, which will be discussed in the next section. The resulting error dynamics can be written in the form

$$M\dot{\tilde{v}} = -D_v\tilde{v} - K_z\tilde{z} - D_{\dot{z}}\dot{\tilde{z}} + A_A\tilde{p}_A - A_B\tilde{p}_B. \tag{8.6}$$

Next, consider the pre-sliding friction dynamics governing the estimation error

$$\tilde{z} \triangleq z - \hat{z}. \tag{8.7}$$

Denoting the right-hand side of the pre-sliding dynamics

$$\dot{z} = v - \frac{K_z}{F_C}|v|_s\, z$$

by the function

$$g_z(v,z) \triangleq v - \frac{K_z}{F_C}|v|_s z, \tag{8.8}$$

the resulting error dynamics can be expressed

$$\begin{aligned}\dot{\tilde{z}} &= \dot{z} - \dot{\hat{z}} \\ &= v - \frac{K_z}{F_C}|v|_s z - \left(\hat{v} - \frac{K_z}{F_C}|\hat{v}|_s \hat{z}\right) \\ &= g_z(v,z) - g_z(\hat{v},\hat{z}). \end{aligned} \tag{8.9}$$

Since $g_z(v,z)$ is smooth in v and z, we may use the Mean Value Theorem (Lemma 27) to rewrite the error $g_z(v,z) - g_z(\hat{v},\hat{z})$ as linear in $\tilde{v}$ and $\tilde{z}$. The friction error dynamics can thus be written as

$$\begin{aligned}\dot{\tilde{z}} &= g_z(v,z) - g_z(\hat{v},\hat{z}) \\ &= \frac{\partial g_z(\bar{v},\bar{z})}{\partial v}(v-\hat{v}) + \frac{\partial g_z(\bar{v},\bar{z})}{\partial z}(z-\hat{z}), \end{aligned} \tag{8.10}$$

with gradients evaluated at a time-varying point $(\bar{v}(t), \bar{z}(t))$ constrained to

$$\bar{v} \in \mathcal{S}_v \triangleq \{\min(v,\hat{v}), \max(v,\hat{v})\} \tag{8.11}$$

$$\bar{z} \in \mathcal{S}_z \triangleq \{\min(z,\hat{z}), \max(z,\hat{z})\}. \tag{8.12}$$

Computing the gradients of $g_z(v,z)$, we get

$$\frac{\partial g_z(v,z)}{\partial v} = 1 - \frac{K_z}{F_C}\mathrm{sgn}_s(v) z \tag{8.13}$$

$$\frac{\partial g_z(v,z)}{\partial z} = -\frac{K_z}{F_C}|v|_s, \tag{8.14}$$

where the derivative of the smooth absolute operator $|\cdot|_s$, is the smooth signum function

$$\mathrm{sgn}_s(v) \triangleq \frac{d|v|_s}{dv} = \frac{v}{\sqrt{v^2+\varepsilon_0^2}}.$$

As the time-varying evaluation point $(\bar{v},\bar{z})$ is constrained by $\mathcal{S}_v$ and $\mathcal{S}_z$, defined in (8.11)–(8.12), we show in the following that the gradients must satisfy

$$\frac{\partial g_z(\bar{v},\bar{z})}{\partial v} \in \langle 0, 2\rangle \tag{8.15}$$

$$\left(-\frac{\partial g_z(\bar{v},\bar{z})}{\partial z}\right) \in \left[\frac{K_z}{F_C}\varepsilon_0, \frac{K_z}{F_C}v_{\mathrm{ub}}\right], \tag{8.16}$$

where v_{ub} denotes an upper bound on $|v|_s$ and $|\hat{v}|_s$. The upper and lower bounds on $\partial g_z(\bar{v},\bar{z})/\partial v$ follows from the inherent boundedness of the LuGre friction model, as established by Proposition 11 (page 51 in the chapter on friction modeling[1]).

[1]Note that for the original non-smooth LuGre model, the inequalities are not strict, *i.e.*, the friction state is constrained to $|z(t)| \leq F_C/K_z$, compared to $|z(t)| < F_C/K_z$, for the smooth LuGre model.

Consequently, the friction states $z(t)$ and $\hat{z}(t)$, whose dynamics are governed by the friction dynamics (3.15), will remain bounded according to

$$\begin{aligned} z(0) &< \frac{F_C}{K_z} \Rightarrow |z(t)| < \frac{F_C}{K_z} \quad \forall t \geq 0 \\ \hat{z}(0) &< \frac{F_C}{K_z} \Rightarrow |\hat{z}(t)| < \frac{F_C}{K_z} \quad \forall t \geq 0, \end{aligned}$$

which again constrains the evaluation point $\bar{z}$ to

$$|\bar{z}(t)| < \frac{F_C}{K_z} \quad \forall t \geq 0,$$

This gives

$$\begin{aligned} \frac{\partial g_z(\bar{v},\bar{z})}{\partial v} &= 1 - \frac{K_z}{F_C}\operatorname{sgn}_s(\bar{v})\,\bar{z} \\ &\Downarrow \\ 1 - \frac{K_z}{F_C}|\bar{z}| \leq & \frac{\partial g_z(\bar{v},\bar{z})}{\partial v} \leq 1 + \frac{K_z}{F_C}|\bar{z}| \\ &\Downarrow \\ 1 - 1 < & \frac{\partial g_z(\bar{v},\bar{z})}{\partial v} < 1 + 1. \end{aligned}$$

The bounds on $\partial g_z(\bar{v},\bar{z})/\partial z$ follows straightforwardly from $\bar{v} \in \mathcal{S}_v$ and $|v|_s \geq \varepsilon_0$. *q.e.d.*

Summarizing, we may write the dynamics of the $\tilde{z}$–subsystem in the linear time-varying form

$$\begin{aligned} \dot{\tilde{z}} &= \left(1 - \frac{K_z}{F_C}\operatorname{sgn}_s(\bar{v})\,\bar{z}\right)\tilde{v} - \frac{K_z}{F_C}|\bar{v}|_s\,\tilde{z} && (8.17) \\ &\triangleq \sigma_1(t)\,\tilde{v} - \sigma_2(t)\,\tilde{z}, && (8.18) \end{aligned}$$

where σ_1 and σ_2 are strictly positive, time-varying coefficients that satisfy

$$\begin{aligned} \sigma_1(t) &\in \langle 0, 2\rangle && (8.19) \\ \sigma_2(t) &\in \left[\frac{K_z}{F_C}\varepsilon_0, \frac{K_z}{F_C}v_{\text{ub}}\right], && (8.20) \end{aligned}$$

where $\bar{v} \in \mathcal{S}_v$ and $\bar{z} \in \mathcal{S}_z$, and v_{ub} is an upper bound on $|v|_s$ and $|\hat{v}|_s$.

To analyze the pressure dynamics, we introduce the following new coordinates

$$\begin{aligned} q_A &\triangleq V_A(y)\,p_A && (8.21) \\ q_B &\triangleq V_B(y)\,p_B, && (8.22) \end{aligned}$$

which can be viewed as measures of the mass of air in each cylinder chamber. In these coordinates, the dynamics of pressure p_A in the front chamber is given indirectly by

$$\begin{aligned} \dot{q}_A &= \frac{\partial V_A}{\partial y}\dot{y}p_A + V_A(y)\,\dot{p}_A \\ &= A_A p_A v - A_A v p_A + \rho_0 T_0 R C_v \psi_v(p_A, x_v) \\ &= \rho_0 T_0 R C_v \psi_v(p_A, x_v), && (8.23) \end{aligned}$$

and equivalently for p_B by

$$\dot{q}_B = \rho_0 T_0 R C_r \psi_r (p_B) . \tag{8.24}$$

Introducing the same change of coordinates for the observer states,

$$\begin{aligned} \hat{q}_A &\triangleq V_A (y) \hat{p}_A & (8.25) \\ \hat{q}_B &\triangleq V_B (y) \hat{p}_B, & (8.26) \end{aligned}$$

we obtain the dynamics

$$\begin{aligned} \dot{\hat{q}}_A &= \frac{\partial V_A}{\partial y} \dot{y} \hat{p}_A + V_A (y) \dot{\hat{p}}_A \\ &= A_A v \hat{p}_A - A_A \hat{v} \hat{p}_A + \rho_0 T_0 R C_v \psi_v (\hat{p}_A, \hat{x}_v) \\ &= A_A \hat{p}_A (v - \hat{v}) + \rho_0 T_0 R C_v \psi_v (\hat{p}_A, \hat{x}_v) , \end{aligned} \tag{8.27}$$

and

$$\dot{\hat{q}}_B = -A_B \hat{p}_B (v - \hat{v}) + \rho_0 T_0 R C_r \psi_r (\hat{p}_B) . \tag{8.28}$$

In the new coordinates, the errors in the estimated pressures are given indirectly by the error variables

$$\begin{aligned} \tilde{q}_A &\triangleq q_A - \hat{q}_A & (8.29) \\ \tilde{q}_B &\triangleq q_B - \hat{q}_B, & (8.30) \end{aligned}$$

as

$$\begin{aligned} \tilde{p}_A &= \frac{1}{V_A (y)} \tilde{q}_A & (8.31) \\ \tilde{p}_B &= \frac{1}{V_B (y)} \tilde{q}_B. & (8.32) \end{aligned}$$

Now consider the pressure dynamics in the new coordinates. Differentiating $\tilde{q}_A$, this gives

$$\begin{aligned} \dot{\tilde{q}}_A &= \rho_0 T_0 R C_v \psi_v (p_A, x_v) - \{ A_A \hat{p}_A (v - \hat{v}) + \rho_0 T_0 R C_v \psi_v (\hat{p}_A, \hat{x}_v) \} \\ &= -A_A \hat{p}_A \tilde{v} + \rho_0 T_0 R C_v [\psi_v (p_A, x_v) - \psi_v (\hat{p}_A, \hat{x}_v)] \end{aligned} \tag{8.33}$$

and likewise for $\tilde{q}_B$

$$\begin{aligned} \dot{\tilde{q}}_B &= \rho_0 T_0 R C_r \psi_r (p_B) - \{ -A_B \hat{p}_B (v - \hat{v}) + \rho_0 T_0 R C_r \psi_r (\hat{p}_B) \} \\ &= A_B \hat{p}_B \tilde{v} + \rho_0 T_0 R C_r [\psi_r (p_B) - \psi_r (\hat{p}_B)] . \end{aligned} \tag{8.34}$$

Since the nonlinear flow functions $\psi_v (p_A, x_v)$ and $\psi_r (p_A)$ are smooth in the arguments, we also here use the Mean Value Theorem (Lemma 27) to rewrite the corresponding errors as linear in the error variables $\tilde{p}_A$, $\tilde{x}_v$ and $\tilde{p}_B$ according to

$$\begin{aligned} \psi_v (p_A, x_v) - \psi_v (\hat{p}_A, \hat{x}_v) &= \frac{\partial \psi_v (\bar{p}_A, \bar{x}_v)}{\partial p_A} (p_A - \hat{p}_A) + \frac{\partial \psi_v (\bar{p}_A, \bar{x}_v)}{\partial x_v} (x_v - \hat{x}_v) \\ \psi_r (p_B) - \psi_r (\hat{p}_B) &= \frac{\partial \psi_r (\bar{p}_B)}{\partial p_B} (p_B - \hat{p}_B) , \end{aligned} \tag{8.35}$$

where the gradients are evaluated at the time-varying points $(\bar{p}_A, \bar{x}_v)$ and $\bar{p}_B$, constrained to

$$\begin{aligned} \bar{p}_A &\in \mathcal{S}_{p_A} \triangleq \{\min(p_A, \hat{p}_A), \max(p_A, \hat{p}_A)\} && (8.36)\\ \bar{p}_B &\in \mathcal{S}_{p_B} \triangleq \{\min(p_B, \hat{p}_B), \max(p_B, \hat{p}_B)\} && (8.37)\\ \bar{x}_v &\in \mathcal{S}_{x_v} \triangleq \{\min(x_v, \hat{x}_v), \max(x_v, \hat{x}_v)\}. && (8.38) \end{aligned}$$

Recalling that

$$\begin{aligned} p_A - \hat{p}_A &= \tilde{p}_A = \frac{1}{V_A(y)}\tilde{q}_A \\ p_B - \hat{p}_B &= \tilde{p}_B = \frac{1}{V_B(y)}\tilde{q}_B, \end{aligned}$$

we may write the $(\tilde{q}_A, \tilde{q}_B)$–dynamics in the following form, linear in the error variables $\tilde{v}$, $\tilde{q}_A$, $\tilde{q}_B$ and $\tilde{x}_v$:

$$\begin{aligned} \dot{\tilde{q}}_A &= -A_A\hat{p}_A\tilde{v} + \rho_0 T_0 R C_v \frac{\partial\psi_v(\bar{p}_A, \bar{x}_v)}{\partial p_A}\frac{1}{V_A(y)}\tilde{q}_A + \rho_0 T_0 R C_v \frac{\partial\psi_v(\bar{p}_A, \bar{x}_v)}{\partial x_v}\tilde{x}_v \\ \dot{\tilde{q}}_B &= A_B\hat{p}_B\tilde{v} + \rho_0 T_0 R C_r \frac{\partial\psi_r(\bar{p}_B)}{\partial p_B}\frac{1}{V_B(y)}\tilde{q}_B. \qquad (8.39) \end{aligned}$$

As the time-varying evaluation points $(\bar{p}_A, \bar{x}_v)$ and $\bar{p}_B$ are constrained to the sets $\mathcal{S}_{p_A}$, $\mathcal{S}_{p_B}$, and $\mathcal{S}_{x_v}$ as defined in (8.36)–(8.38), the gradients satisfy

$$\begin{aligned} \frac{\partial\psi_v(\bar{p}_A, \bar{x}_v)}{\partial p_A} &\leq 0 && (8.40)\\ \frac{\partial\psi_r(\bar{p}_B)}{\partial p_B} &\leq 0 && (8.41)\\ \frac{\partial\psi_v(\bar{p}_A, \bar{x}_v)}{\partial x_v} &\geq 0, && (8.42) \end{aligned}$$

in the entire region of validity $\mathcal{X}_0$ of the model. To verify this, recall from Subsection 8.1 (page 113) that the pressure–flow gradients, $\partial\psi_v(p_A, x_v)/\partial p_A$ and $\partial\psi_v(p_B)/\partial p_B$ are either negative or zero, and that the input–flow gradient $\partial\psi_v(p_A, x_v)/\partial x_v$ is positive or zero, for all valid pressures p_A and p_B, and valve openings x_v, *i.e.*, $[p_A, p_B, x_v]^T \in \mathcal{X}_0$. Furthermore, if both the actual and the estimated states are contained in $\mathcal{X}_0$, then also are the evaluation points $(\bar{p}_A, \bar{x}_v)$ and $\bar{p}_B$ for the gradients.

Now consider the $\tilde{x}_v$–dynamics. Differentiating the estimation error

$$\tilde{x}_v \triangleq x_v - \hat{x}_v, \qquad (8.43)$$

we obtain the simple error dynamics

$$\begin{aligned} \dot{\tilde{x}}_v &= \dot{x}_v - \dot{\hat{x}}_v \\ &= -\frac{1}{\tau_v}x_v + \frac{1}{\tau_v}\pi_u(u + U_0) - \left(-\frac{1}{\tau_v}\hat{x}_v + \frac{1}{\tau_v}\pi_u(u + U_0)\right) \\ &= -\frac{1}{\tau_v}\tilde{x}_v. \qquad (8.44) \end{aligned}$$

Recapitulating so far, the complete error dynamics can be expressed in the linear, time-varying form

$$\begin{aligned}
M\dot{\tilde{v}} &= -\left[D_v + D_{\dot{z}}\sigma_1(t)\right]\tilde{v} - \left[K_z - D_{\dot{z}}\sigma_2(t)\right]\tilde{z} + \frac{A_A}{V_A(y)}\tilde{q}_A - \frac{A_B}{V_B(y)}\tilde{q}_B \\
\dot{\tilde{z}} &= \sigma_1(t)\,\tilde{v} - \sigma_2(t)\,\tilde{z} \\
\dot{\tilde{q}}_A &= -A_A\hat{p}_A\tilde{v} + \rho_0 T_0 R C_v \frac{\partial\psi_v(\bar{p}_A,\bar{x}_v)}{\partial p_A}\frac{1}{V_A(y)}\tilde{q}_A + \rho_0 T_0 R C_v \frac{\partial\psi_v(\bar{p}_A,\bar{x}_v)}{\partial x_v}\tilde{x}_v \\
\dot{\tilde{q}}_B &= A_B\hat{p}_B\tilde{v} + \rho_0 T_0 R C_r \frac{\partial\psi_r(\bar{p}_B)}{\partial p_B}\frac{1}{V_B(y)}\tilde{q}_B \\
\tau_v\dot{\tilde{x}}_v &= -\tilde{x}_v,
\end{aligned} \tag{8.45}$$

where the gradients of the friction dynamics,

$$\sigma_1(t) \triangleq \frac{\partial g_z(\bar{v},\bar{z})}{\partial v} = 1 - \frac{K_z}{F_C}\operatorname{sgn}_s(\bar{v})\,\bar{z} \tag{8.46}$$

$$\sigma_2(t) \triangleq -\frac{\partial g_z(\bar{v},\bar{z})}{\partial z} = \frac{K_z}{F_C}\left|\bar{v}\right|_s, \tag{8.47}$$

are strictly positive, time-varying coefficients that satisfy

$$\sigma_1(t) \in \langle 0, 2\rangle \tag{8.48}$$

$$\sigma_2(t) \in \left[\frac{K_z}{F_C}\varepsilon_0, \frac{K_z}{F_C}v_{\mathrm{ub}}\right]. \tag{8.49}$$

Provided that both the actual and estimated states remain in the region of validity of the model, the pressure–flow gradients, $\partial\psi_v(p_A,x_v)/\partial p_A$ and $\partial\psi_v(p_B)/\partial p_B$ are either negative or zero, and the input–flow gradient $\partial\psi_v(p_A,x_v)/\partial x_v$ positive or zero.

Stabilizing mechanisms

The remaining part of the analysis is simply to identify the stabilizing (or destabilizing) effects of each term in the error dynamics. Denoting

$$\sigma_1(t) \triangleq 1 - \frac{K_z}{F_C}\operatorname{sgn}_s(\bar{v})\,\bar{z} \tag{8.50}$$

$$\sigma_2(t) \triangleq \frac{K_z}{F_C}\left|\bar{v}\right|_s \tag{8.51}$$

$$\alpha_A(t) \triangleq -\rho_0 T_0 R C_v \frac{\partial\psi_v(\bar{p}_A,\bar{x}_v)}{\partial p_A}\frac{1}{V_A(y)} \tag{8.52}$$

$$\beta_A(t) \triangleq \rho_0 T_0 R C_v \frac{\partial\psi_v(\bar{p}_A,\bar{x}_v)}{\partial x_v} \tag{8.53}$$

$$\alpha_B(t) \triangleq -\rho_0 T_0 R C_r \frac{\partial\psi_r(\bar{p}_B)}{\partial p_B}\frac{1}{V_B(y)} \tag{8.54}$$

the error dynamics can be expressed in the form

$$\begin{aligned} M\dot{\tilde{v}} &= -\left[D_v + D_{\dot{z}}\sigma_1(t)\right]\tilde{v} - \left[K_z - D_{\dot{z}}\sigma_2(t)\right]\tilde{z} + \frac{A_A}{V_A(y)}\tilde{q}_A - \frac{A_B}{V_B(y)}\tilde{q}_B \\ \dot{\tilde{z}} &= \sigma_1(t)\,\tilde{v} - \sigma_2(t)\,\tilde{z} \\ \dot{\tilde{q}}_A &= -A_A\hat{p}_A\tilde{v} - \alpha_A(t)\,\tilde{q}_A + \beta_A(t)\,\tilde{x}_v \\ \dot{\tilde{q}}_B &= A_B\hat{p}_B\tilde{v} - \alpha_B(t)\,\tilde{q}_B \\ \tau_v\dot{\tilde{x}}_v &= -\tilde{x}_v. \end{aligned} \tag{8.55}$$

The structure of the error dynamics becomes even more apparent by expressing the system in the matrix form

$$\begin{bmatrix} M\dot{\tilde{v}} \\ \dot{\tilde{z}} \\ \dot{\tilde{q}}_A \\ \dot{\tilde{q}}_B \\ \tau_v\dot{\tilde{x}}_v \end{bmatrix} = \begin{bmatrix} -\left[D_v + D_{\dot{z}}\sigma_1(t)\right] & -\left[K_z - D_{\dot{z}}\sigma_2(t)\right] & \frac{A_A}{V_A(y)} & -\frac{A_B}{V_B(y)} & 0 \\ \sigma_1(t) & -\sigma_2(t) & 0 & 0 & 0 \\ -A_A\hat{p}_A & 0 & -\alpha_A(t) & 0 & \beta_A(t) \\ A_B\hat{p}_B & 0 & 0 & -\alpha_B(t) & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \tilde{v} \\ \tilde{z} \\ \tilde{q}_A \\ \tilde{q}_B \\ \tilde{x}_v \end{bmatrix}. \tag{8.56}$$

Knowing that all the time-varying coefficients are non-negative, the structure of the error dynamics reveals the main stability properties of the system. Because finding a Lyapunov function which explicitely establishes the stability of the complete error dynamics is difficult, we settle with a qualitative analysis where we identify and discuss the stabilizing mechanisms in the system based on the structure of its error dynamics. The two stabilizing mechanisms in the pneumatic system, we refer to as *negative diagonal stabilization* and *skew-symmetrical-like* stabilization.

Negative diagonal feedback terms, or feedback damping terms, are the most obvious stabilizing terms in the system. If the negative diagonal terms appears at all diagonal places, and they are strong enough that they in sum dominate the remaining cross-diagonal terms, they are the main mechanism which stablizes the system. In these cases, stability may be established via simple, diagonal Lyapunov functions. A trivial case of dominating negative diagonal, is the $\tilde{x}_v$–dynamics

$$\tau_v\dot{\tilde{x}}_v = -\tilde{x}_v,$$

which is completely decoupled from the remaining system states, with the damping term $-\tilde{x}_v$. In the absense of other (cross-diagonal) terms, the negative diagonal term ensures that this subsystem is exponentially stable (ES). This is straightforwardly established with the simple quadratic Lyapunov function $W(\tilde{x}_v) = \tau_v\tilde{x}_v^2$, whose derivative becomes $\dot{W} = -2\tilde{x}_v^2$, thus, ES is proven. Because $\tilde{x}_v$ converges exponentially to zero, the requirement on the remaining $(\tilde{v}, \tilde{z}, \tilde{q}_A, \tilde{q}_B)$–subsystem for the complete system to be stable, is that it is input-to-state stable with respect to $\tilde{x}_v$ as input.

The remaining $(\tilde{v}, \tilde{z}, \tilde{q}_A, \tilde{q}_B)$–subsystem has the feedback damping terms: $-\left[D_v + D_{\dot{z}}\sigma_1(t)\right]\tilde{v}$, $-\sigma_2(t)\,\tilde{z}$, $-\alpha_A(t)\,\tilde{q}_A$, and $-\alpha_B(t)\,\tilde{q}_B$. When the time-varying

damping coefficients $[D_v + D_{\dot{z}}\sigma_1(t)]$, $\sigma_2(t)$, $\alpha_A(t)$, and $\alpha_B(t)$ are sufficiently large, they provide the main stabilization of the system. To demonstrate this, we neglect the friction and valve dynamics by assuming $\tilde{z} \equiv 0 \Rightarrow \dot{\tilde{z}} = 0$ and $\tilde{x}_v \equiv 0 \Rightarrow \dot{\tilde{x}}_v$, resulting in the simplified error dynamics

$$\begin{aligned} M\dot{\tilde{v}} &= -D_v\tilde{v} + \frac{A_A}{V_A(y)}\tilde{q}_A - \frac{A_B}{V_B(y)}\tilde{q}_B \\ \dot{\tilde{q}}_A &= -A_A\hat{p}_A\tilde{v} - \alpha_A(t)\,\tilde{q}_A \\ \dot{\tilde{q}}_B &= A_B\hat{p}_B\tilde{v} - \alpha_B(t)\,\tilde{q}_B. \end{aligned} \tag{8.57}$$

When the negative diagonal terms are sufficiently large, the simple Lyapunov function candidate

$$U(\tilde{v}, \tilde{q}_A, \tilde{q}_B) = \frac{M}{2}\tilde{v}^2 + \frac{1}{2}\tilde{q}_A + \frac{1}{2}\tilde{q}_B$$

can be use to establish stability for this system. The time-derivative of U becomes

$$\begin{aligned} \dot{U} &= \tilde{v}\left(-D_v\tilde{v} + \frac{A_A}{V_A}\tilde{q}_A - \frac{A_B}{V_B}\tilde{q}_B\right) + \tilde{q}_A\left(-A_A\hat{p}_A\tilde{v} - \alpha_A\tilde{q}_A\right) + \tilde{q}_B\left(A_B\hat{p}_B\tilde{v} - \alpha_B\tilde{q}_B\right) \\ &= -D_v\tilde{v}^2 - \alpha_A\tilde{q}_A^2 - \alpha_B\tilde{q}_B^2 + \left(\frac{A_A}{V_A} - A_A\hat{p}_A\right)\tilde{v}\tilde{q}_A - \left(\frac{A_B}{V_B} - A_B\hat{p}_B\right)\tilde{v}\tilde{q}_B, \end{aligned}$$

which can be expressed in the matrix form

$$\dot{U} = -\tilde{\mathbf{x}}^T\mathbf{Q}(t)\,\tilde{\mathbf{x}}, \quad \tilde{\mathbf{x}} \triangleq [\tilde{v}, \tilde{q}_A, \tilde{q}_B]^T, \tag{8.58}$$

with

$$\mathbf{Q}(t) \triangleq \begin{bmatrix} D_v & \frac{1}{2}\left(\frac{A_A}{V_A(y)} - A_A\hat{p}_A\right) & -\frac{1}{2}\left(\frac{A_B}{V_B(y)} - A_B\hat{p}_B\right) \\ \frac{1}{2}\left(\frac{A_A}{V_A(y)} - A_A\hat{p}_A\right) & \alpha_A(t) & 0 \\ -\frac{1}{2}\left(\frac{A_B}{V_B(y)} - A_B\hat{p}_B\right) & 0 & \alpha_B(t) \end{bmatrix}. \tag{8.59}$$

Applying Sylvester's Theorem (see *e.g.* [47]) we find the requirements for $\mathbf{Q}$ to be positive definite so that $\dot{U}$ becomes negative definite:

$$D_v\alpha_A(t) - \frac{1}{4}\left(\frac{A_A}{V_A(y)} - A_A\hat{p}_A\right)^2 > 0 \tag{8.60}$$

$$D_v\alpha_A(t)\,\alpha_B(t) - \frac{1}{4}\left(\frac{A_A}{V_A(y)} - A_A\hat{p}_A\right)^2 - \frac{1}{4}\left(\frac{A_B}{V_B(y)} - A_B\hat{p}_B\right) > 0. \tag{8.61}$$

Hence, the simplified system is asymptotically stable (AS) provided the negative diagonal gains D_v, α_A, and α_B are sufficiently large to dominate the coefficients of the cross-terms, thus making $\dot{U}$ negative definite. However, the feedback damping provided by α_A, and α_B (resulting from the stablizing effect of the flow functions ψ_v and ψ_r), and σ_2 (resulting from the LuGre friction dynamics), is not always the dominating stablizing mechanism. For example when the velocity is small, the

internal damping of the LuGre friction dynamics is very small, resulting in a too small σ_2 to provide convergence of $\tilde{z}$. Furthermore, when pressure and valve states are such that choked flow occurs in the valve, the gradient of the valve's flow function ψ_v becomes small, resulting in a to small α_A to ensure convergence of $\tilde{p}_A$. In these cases, the main stabilization of the pneumatic system is due to a skew-symmetrical-like connection between the states.

Skew-symmetrical-like terms are not as obvious stabilizing terms, and are alone not sufficient to provide stabilization. The skew-symmetrical-like terms depend on a negative damping term to provide stabilization (without, they simply result in poles on the imaginary axis, in the linear case). For example, the stability of a (Luenberger) observer design with feedback output-injection terms, is based on skew-symmetrical-like stabilization in combination with a diagonal output-injection term in the first equation. Likewise, is a cancelling backstepping design based on skew-symmetrical cancellation of cross-terms in combination with negative diagonal damping terms. In the pneumatic system, the main stability mechanism alternates between a negative diagonal and skew-symmetrical-like stabilization. Establishing stability in the case of skew-symmetrical-like stabilization, using a Lyapunov function, is far more difficult than in the case with dominating diagonal terms.

What we mean by skew-symmetrical-like stabilization is better explained with an example. Take the simplified error dynamics

$$\begin{aligned} M\dot{\tilde{v}} &= -D_v\tilde{v} + a\left(t\right)\tilde{q}_A \\ \dot{\tilde{q}}_A &= -b\left(t\right)\tilde{v}, \end{aligned}$$

where there is no diagonal stabilizing term for the $\tilde{q}_A$–dynamics ($\alpha_A = 0$). In this case, the stabilization of $\tilde{q}_A$ is due to the terms $a\left(t\right)\tilde{q}_A$ and $-b\left(t\right)\tilde{v}$, which have the skew-symmetrical-like properties that they are non-zero and has different sign. Even in this simple case, it is complicated to find a Lyapunov function which establiches stability, primarily due to the time-variance of the cross-terms. Consider the Lyapunov function candidate

$$V\left(\tilde{v},\tilde{q}_A\right) = \frac{M}{2}\tilde{v}^2 + \frac{C_1}{2}\tilde{q}_A^2 - C_2M\tilde{v}\tilde{q}_A,$$

where C_1 and C_2 are positive constants which satisfies $C_1 - MC_2^2 > 0$, such that V is positive definite. The time-derivative of V becomes

$$\dot{V} = -\left[D_v - C_2Mb\left(t\right)\right]\tilde{v}^2 - C_2a\left(t\right)\tilde{q}_A^2 + \left[a\left(t\right) + C_2D_v - C_1b\left(t\right)\right]\tilde{v}\tilde{q}_A.$$

We see that if a and b had been constants, we could use C_1 and C_2 to cancel the cross-term $\tilde{v}\tilde{q}_A$ to make$\dot{V}$ negative definite. This is possible because the skew-symmetrical terms have different sign.

To show that also the time-varying, skew-symmetrical-like terms provide stability, we can express $\dot{V}$ as

$$\dot{V} = -\tilde{\mathbf{x}}^T\mathbf{Q}\left(t\right)\tilde{\mathbf{x}}, \quad \tilde{\mathbf{x}} \triangleq \left[\tilde{v},\tilde{q}_A\right]^T,$$

with

$$\mathbf{Q}(t) \triangleq \begin{bmatrix} D_v - C_2 M b(t) & a(t) + C_2 D_v - C_1 b(t) \\ a(t) + C_2 D_v - C_1 b(t) & C_2 a(t) \end{bmatrix}. \tag{8.62}$$

Applying Sylvester's Theorem we find that in order to show asymptotic stability with the proposed V, there must exists constants C_1 and C_2 such that

$$\begin{aligned} D_v - C_2 M b(t) &> 0 \\ [D_v - C_2 M b(t)] C_2 a(t) - \frac{1}{4}[a(t) + C_2 D_v - C_1 b(t)]^2 &> 0, \end{aligned}$$

are satisfied for all $a(t)$ and $b(t)$. Provided the time-variance of $a(t)$ and $b(t)$ are limited, we see that because they have different signs, they partly cancel eachother. That is, we can use C_1 and C_2 to make the square term $[a(t) + C_2 D_v - C_1 b(t)]^2$ small, relative to the term $[D_v - C_2 M b(t)] C_2 a(t)$, thereby satisfying the inequality and prove asymptotic stability. On the other hand, with the time-variance of $a(t)$ and $b(t)$ being large, stability will no longer be possible to establish using the suggested Lyapunov function candidate V.

In the $(\tilde{v}, \tilde{z}, \tilde{q}_A, \tilde{q}_B)$–subsystem, the skew-symmetrical-like terms works through the $\tilde{v}$–dynamics, dependent on the negative damping $-[D_v + D_{\dot{z}}\sigma_1(t)]\tilde{v}$ to provide stabilization. The skew-symmetrical-like terms in the $\tilde{v}$–dynamics are: $\sigma_1(t)$, $-A_A\hat{p}_A$, and $A_B\hat{p}_B$, which corresponds to $-[K_z - D_{\dot{z}}\sigma_2(t)]$, $\frac{A_A}{V_A(y)}$, and $-\frac{A_B}{V_B(y)}$, in the $\tilde{z}$, $\tilde{q}_A$, and $\tilde{q}_B$–dynamics, respectively. When the diagonal damping gains $\sigma_2(t)$, $\alpha_A(t)$, and $\alpha_B(t)$ are too small to ensure stability, the $\tilde{z}$, $\tilde{q}_A$, and $\tilde{q}_B$–dynamics, are stabilized by the corresponding skew-symmetrical-like terms and the negative damping term $-[D_v + D_{\dot{z}}\sigma_1(t)]\tilde{v}$.

Summarizing, the structure of the the pneumatic system implies that the system is asymptotically stable, where the type of stabilization alternates between stabilization due to dominating diagonal terms, and stabilization via skew-symmetrical-like terms.

8.2.2 Simulation results

The strong convergence properties of the open-loop nonlinear observer, is here illustrated by simulations. The observer (8.4) and system model (8.1) were implemented in continuous-time in Simulink/Matlab using the model parameters summarized in Table 6.1 (page 103), and the parameters $\boldsymbol{\theta}_l$ of the clutch load characteristic printed below the plot in Figure 3.1 (page 45).

The performance of the observer is illustrated on a simulated response of the system (8.1) subjected to the square wave valve control input

$$u_v = U_0 + U_1 \operatorname{sgn}\left(\sin\left(\frac{2\pi}{T}t\right)\right),$$

with bias $U_0 = -1.5\,\mathrm{V}$, amplitude $U_1 = 5\,\mathrm{V}$, and period $T = 0.15\,\mathrm{s}$. The simulated valve input u_v and corresponding output y, is plotted in Figure 8.1. In order to

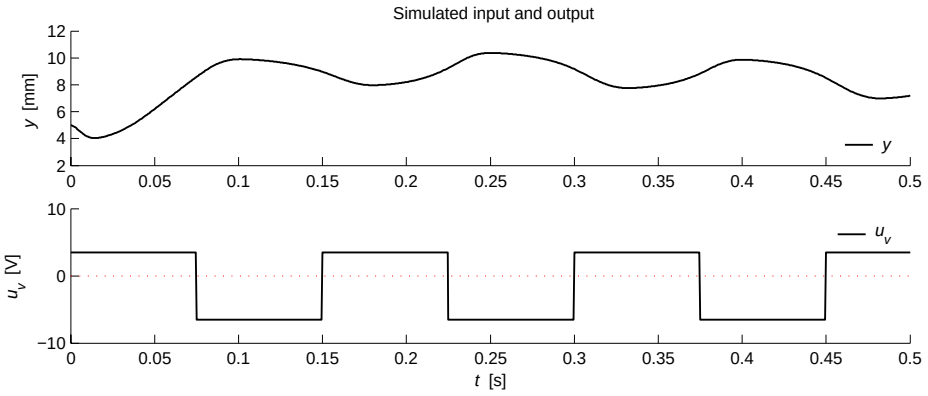

Figure 8.1: Simulated valve input u_v and the corresponding response in the output, actuator position y.

obtain comparable results, we use the same simulated system response (as plotted in Figure 8.1) throughout this chapter for analysis of different observer properties.

The initial convergence properties of the observer is illustrated in Figure 8.2, with the initial observer errors

$$\begin{bmatrix} \tilde{v}\,(0) \\ \tilde{p}_A\,(0) \\ \tilde{x}_v\,(0) \\ \tilde{p}_B\,(0) \\ \tilde{z}\,(0) \end{bmatrix} = \begin{bmatrix} 0\,\mathrm{mm/\,s} \\ 250\,\mathrm{kPa} \\ 2 \\ 0\,\mathrm{kPa} \\ 0\,\mathrm{mm} \end{bmatrix}.$$

The simulation illustrates how a moderately large initial error in the estimate of the chamber pressure p_A, may cause considerable transient errors in the estimates of v, p_B and z, which thereafter, converge to the actual states within approximately 0.3 s. A property of the observer error dynamics worth noting, which is evident from the observer dynamics (8.4), but not apparent from the plotted response, is that the error dynamics of the spool position error $\tilde{x}_v$ is completely decoupled from the remaining observer error dynamics.

8.2.3 Remarks on robustness

From the plot in Figure 8.2, the initial transients in the observer estimates remain within the feasible region $\mathcal{X}_0$ of the system (8.1), while the actual states remain within the region of normal operation $\mathcal{X}$, thus, the observer errors converge asymptotically to zero. However, the simulations also indicates the possibility that certain estimates may move outside of the region of validity $\mathcal{X}_0$ during initial transients. Hence, a modification of the nominal observer is necessary in order to guarantee stability of the observer. A redesign of the nominal observer dynamics, by which we obtain global stability properties and improved transient performance, is addressed in Section 8.5.

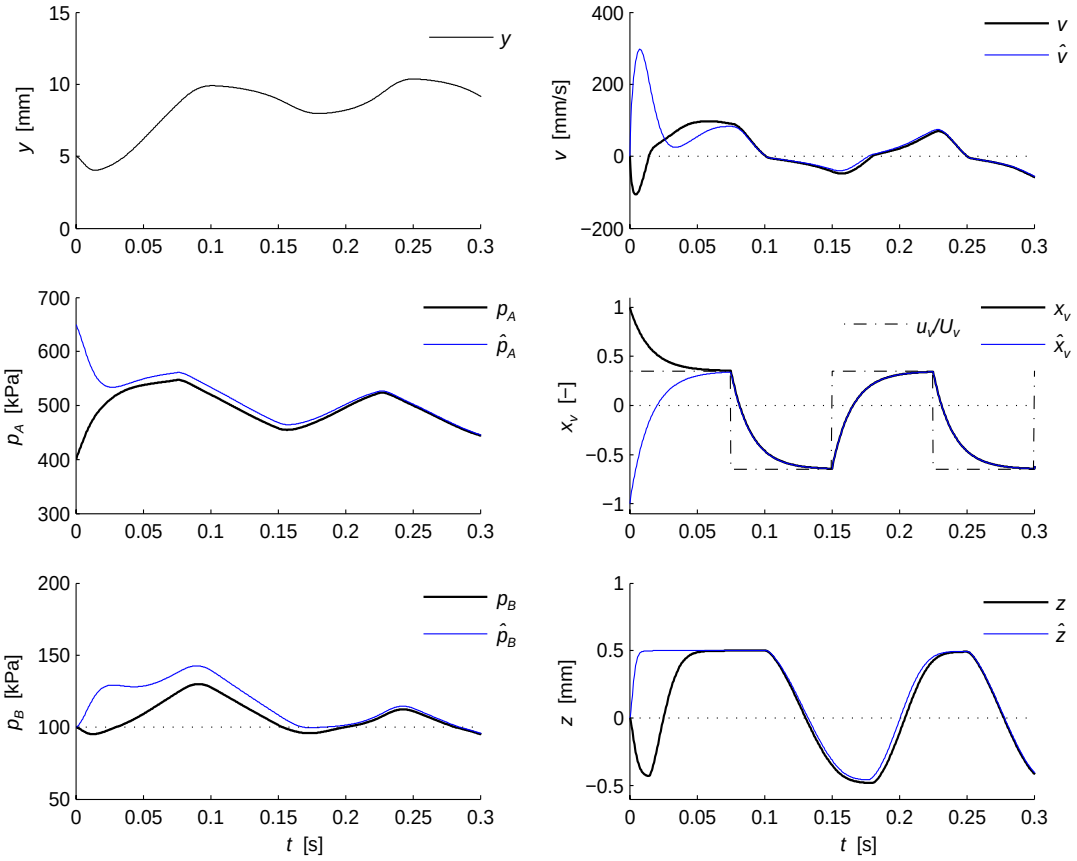

Figure 8.2: Illustration of the convergence properties of the open-loop observer.

8.3 Full-order observer

In this section we propose a simple full-order observer for the electro-pneumatic actuator, which combines estimation using output-injection correction terms for the for main states y, v, and p_A, and open-loop estimation of the remaining states x_v, p_B, and z. The observer inherits the convergence properties from the open-loop observer, but the convergence is strengthened as the output-injection terms introduce the possibility to improve the convergence rate and strengthen the disturbance attenuation properties of the observer. To guarantee stability, we propose a robust modification of the observer, compatible with backstepping, where smooth saturation is utilized to constrain the observer dynamics to the feasible region of the state-space, thus, achieving global stability properties. We further improve the transient performance of the observer by applying discontinuous projection of the observer states, which, in combination with smooth saturation, preserves smooth observer estimates. The convergence properties of the observer is illustrated by simulations.

A full-order observer for the system (8.1) with linear output-injection terms, is given as

$$\begin{aligned}
\dot{\hat{y}} &= k_1 (y - \hat{y}) + \hat{v} \\
\dot{\hat{v}} &= k_2 (y - \hat{y}) + \frac{A_0 P_0}{M} + \frac{A_A}{M}\hat{p}_A - \frac{A_B}{M}\hat{p}_B - \frac{1}{M} f_l (y) - \frac{1}{M} f_f (\hat{v}, \hat{z}) \\
\dot{\hat{p}}_A &= k_3 (y - \hat{y}) - A_A \frac{1}{V_A (y)} \hat{v}\hat{p}_A + \rho_0 T_0 R C_v \frac{1}{V_A (y)} \psi_v (\hat{p}_A, \hat{x}_v) \\
\dot{\hat{x}}_v &= -\frac{1}{\tau_v}\hat{x}_v + \frac{1}{\tau_v}\pi_u (u + U_0) \\
\dot{\hat{p}}_B &= A_B \frac{1}{V_B (y)} \hat{v}\hat{p}_B + \rho_0 T_0 R C_r \frac{1}{V_B (y)} \psi_r (\hat{p}_B) \\
\dot{\hat{z}} &= \hat{v} - \frac{K_z}{F_C} |\hat{v}|_s \hat{z},
\end{aligned} \tag{8.63}$$

where the observer states $\hat{\mathbf{x}} \triangleq [\hat{y}, \hat{v}, \hat{p}_A, \hat{x}_v, \hat{p}_B, \hat{z}]^T$ are the estimates of all of the system states $\mathbf{x} \triangleq [y, v, p_A, x_v, p_B, z]^T$, the control input is u, and y is the measured output.

The observer can be said to be of partially Luenberger-type: It is of Luenberger-type due to the output-error injection terms for the estimation of the main states y, v, and p_A, similar to the Luenberger observer for linear systems, but only partially, because the estimation of the remaining states x_v, p_B, and z is open-loop estimation without output injection terms.

Partitioning the estimated states according to $\hat{\mathbf{x}} = [\hat{\boldsymbol{\xi}}, \hat{\boldsymbol{\zeta}}]^T$, where $\hat{\boldsymbol{\xi}} \triangleq [\hat{y}, \hat{v}, \hat{p}_A]^T$ denotes the estimates with output-injection, and $\hat{\boldsymbol{\zeta}} \triangleq [\hat{x}_v, \hat{p}_B, \hat{z}]^T$ are the open-loop estimated states, the observer (8.63) can be compactly expressed in the form

$$\begin{aligned}
\dot{\hat{\boldsymbol{\xi}}} &= \mathbf{k}\tilde{y} + \mathbf{f}_\xi (\hat{\mathbf{x}}_u, y) \\
\dot{\hat{\boldsymbol{\zeta}}} &= \mathbf{f}_\zeta (\hat{\mathbf{x}}_u, y, u),
\end{aligned} \tag{8.64}$$

where $\tilde{y} \triangleq y - \hat{y}$ denotes the error in the estimated output, $\mathbf{k} \triangleq [k_1, k_2, k_3]^T$ the observer feedback gains, $\hat{\mathbf{x}}_u \triangleq [\hat{v}, \hat{p}_A, \hat{\boldsymbol{\zeta}}]^T$ the estimated unmeasured states, and where the vectors describing the system dynamics are given by

$$\mathbf{f}_\xi (\hat{\mathbf{x}}_u, y) = \begin{bmatrix} f_{\xi 1} \\ f_{\xi 2} \\ f_{\xi 3} \end{bmatrix} \triangleq \begin{bmatrix} \hat{v} \\ \frac{A_0 P_0}{M} + \frac{A_A}{M}\hat{p}_A - \frac{A_B}{M}\hat{p}_B - \frac{1}{M} f_l (y) - \frac{1}{M} f_f (\hat{v}, \hat{z}) \\ -A_A \frac{1}{V_A (y)} \hat{v}\hat{p}_A + \rho_0 T_0 R C_v \frac{1}{V_A (y)} \psi_v (\hat{p}_A, \hat{x}_v) \end{bmatrix} \tag{8.65}$$

$$\mathbf{f}_\zeta (\hat{\mathbf{x}}_u, y, u_v) = \begin{bmatrix} f_{\zeta 1} \\ f_{\zeta 2} \\ f_{\zeta 3} \end{bmatrix} \triangleq \begin{bmatrix} -\frac{1}{\tau_v}\hat{x}_v + \frac{1}{\tau_v}\pi_u (u + U_0) \\ A_B \frac{1}{V_B (y)} \hat{v}\hat{p}_B + \rho_0 T_0 R C_r \frac{1}{V_B (y)} \psi_r (\hat{p}_B) \\ \hat{v} - \frac{K_z}{F_C} |\hat{v}|_s \hat{z} \end{bmatrix}. \tag{8.66}$$

In Figure 8.3, the full-order observer is visualized in the compact form (8.64) by a block diagram.

8.3.1 Stability and convergence

The proposed Luenberger-type observer (8.63) inherits the stability properties of the system (8.1). Due to the output-injection terms $k_1\tilde{y}$, $k_2\tilde{y}$ and $k_3\tilde{y}$, the observer

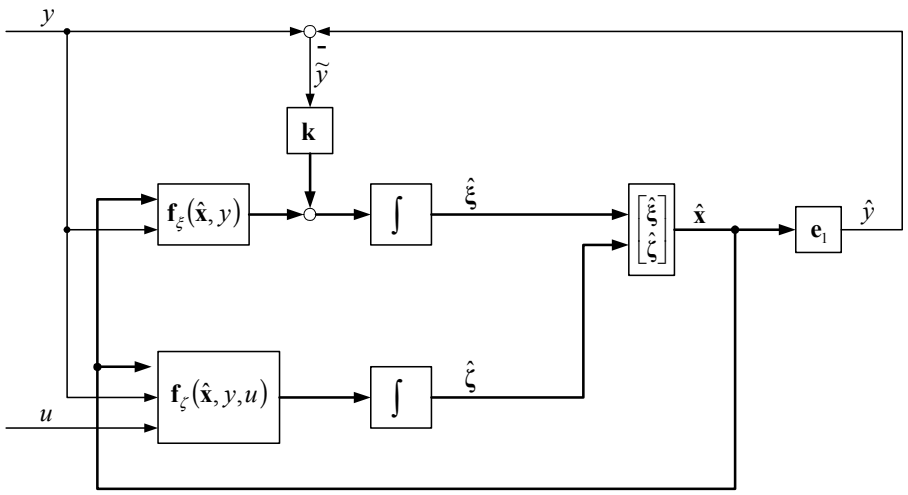

Figure 8.3: Block diagram of the combined open-loop and Luenberger-type full-order observer.

has improved convergence rate and disturbance attenuation properties compared to the open-loop observer (8.4).

8.3.2 Observer gains

Consider the dynamics of the observer error $(\tilde{y}, \tilde{v}, \tilde{p}_A)$–subsystem, which can be written

$$\begin{aligned}
\dot{\tilde{y}} &= -k_1\tilde{y} + \tilde{v} \\
\dot{\tilde{v}} &= -k_2\tilde{y} - \frac{D_v}{M}\tilde{v} + \frac{A_A}{M}\tilde{p}_A - \frac{K_z}{M}\tilde{z} - \frac{D_z}{M}\dot{\tilde{z}} \\
\dot{\tilde{p}}_A &= -k_3\tilde{y} - A_A\frac{1}{V_A(y)}\left[p_A v - \hat{p}_A\hat{v}\right] + \rho_0 T_0 R C_v \frac{1}{V_A(y)}\left[\psi_v(p_A, x_v) - \psi_v(\hat{p}_A, \hat{x}_v)\right].
\end{aligned}$$

Assuming that $p_A, \hat{p}_A \geq P_0$, and neglecting the dynamic friction and the stabilizing flow function terms, we extract the linear time-invariant subsystem

$$\begin{bmatrix}\dot{\tilde{y}}\\ \dot{\tilde{v}}\\ \dot{\tilde{p}}_A\end{bmatrix} = \begin{bmatrix}-k_1 & 1 & 0\\ -k_2 & -\frac{D_v}{M} & \frac{A_A}{M}\\ -k_3 & -\frac{A_A P_0}{V_A(y_{ub})} & 0\end{bmatrix}\begin{bmatrix}\tilde{y}\\ \tilde{v}\\ \tilde{p}_A\end{bmatrix},$$

which we use to facilitate the tuning of observer gains $\mathbf{k} \triangleq [k_1, k_2, k_3]^T$ for the nonlinear observer (8.63). The linear system is compactly written as

$$\dot{\tilde{\boldsymbol{\xi}}} = \mathbf{A}_o\tilde{\boldsymbol{\xi}}, \tag{8.67}$$

where $\tilde{\boldsymbol{\xi}} \triangleq \boldsymbol{\xi} - \hat{\boldsymbol{\xi}} = [\tilde{y}, \tilde{v}, \tilde{p}_A]^T$, and has the characteristic polynomial

$$|s\mathbf{I} - \mathbf{A}_o| = s^3 + \left(k_1 + \frac{D_v}{M}\right) s^2 + \left(k_1 \frac{D_v}{M} + \frac{A_A^2 P_0}{V_A(y_{ub})\, M} + k_2\right) s + (k_1 + k_3) \frac{A_A}{M}. \tag{8.68}$$

The observer gains are chosen such that the constant observer matrix is Hurwitz, *i.e.*, so that $\mathbf{A}_o$ satisfies the Lyapunov equation $\mathbf{A}_o\mathbf{P}_o + \mathbf{P}_o\mathbf{A}_o^T = -\mathbf{Q}$ for some $\mathbf{P}_o = \mathbf{P}_o^T > 0$ and $\mathbf{Q} > 0$. We can show that a sufficient condition for $\mathbf{A}_o$ to be Hurwitz, is that $k_1 > 0$ and $k_2, k_3 \geq 0$.

One way to choose observer gains k_1, k_2, and k_3, is to place multiple poles at $s = -\lambda_o < 0$, which gives a critically damped response in the output $\hat{\xi}_1 = \tilde{y}$. The dynamics of the linear system (8.67) is then governed by the characteristic equation

$$|s\mathbf{I} - \mathbf{A}_o| = (s + \lambda_o)^3 = s^3 + 3\lambda_o s^2 + 3\lambda_o^2 s + \lambda_o^3 = 0. \tag{8.69}$$

By comparing the coefficients of the two polynomials (8.68) and (8.69), the observer gains can be expressed as functions of λ_o:

$$\begin{aligned} k_1 &= 3\lambda_o - \tfrac{D_v}{M} \\ k_2 &= 3\lambda_o^2 - 3\lambda_o \tfrac{D_v}{M} + \left(\tfrac{D_v}{M}\right)^2 - \tfrac{A_A^2 P_0}{V_A(y_{ub})M} \\ k_3 &= \lambda_o^3 \tfrac{M}{A_A} - 3\lambda_o + \tfrac{D_v}{M}. \end{aligned} \tag{8.70}$$

With the observer gains determined according to (8.70), the poles of the linear dynamics (8.67) are placed at $s = -\lambda_o$. Due to the unstable integrator for the position, we must choose $k_1 > 0$ in order to obtain an asymptotically stable observer, and we would like to avoid negative feedback gains k_i for small λ_o. Consequently, we determine the observer gains according to

$$\begin{aligned} k_1 &= \max\left(3\lambda_o - \tfrac{D_v}{M}, 10\right) \\ k_2 &= \max\left(3\lambda_o^2 - 3\lambda_o \tfrac{D_v}{M} + \left(\tfrac{D_v}{M}\right)^2 - a\tfrac{A_A}{M}, 0\right) \\ k_3 &= \max\left(\lambda_o^3 \tfrac{M}{A_A} - 3\lambda_o + \tfrac{D_v}{M}, 0\right), \end{aligned} \tag{8.71}$$

to ensure that $k_1 \geq 10$ and $k_2, k_3 > 0$. In the following, we refer to the parameter λ_o as the *design bandwidth* of the nonlinear observer (8.63).

8.3.3 Simulation results

Figure 8.4 illustrates the convergence properties of the full-order observer (8.63) for increasing λ_o, with the observer gains k_1, k_2, k_3 determined according to (8.71).

With $\lambda_o = 0$, the feedback gains become $k_1 = 10$, and $k_2 = k_3 = 0$, such that the convergence rate is approximately equal (or slightly less than) the convergence rate of the open-loop observer (8.4). In addition, the figure clearly illustrates a desirable property of the observer for the estimated states $\hat{y}$, $\hat{v}$ and $\hat{p}_A$, namely, that the convergence rate can be systematically increased by increasing the observer feedback gains, or more precisely, the design bandwidth λ_o. Notice also that the improved convergence of $\hat{y}$, $\hat{v}$, $\hat{p}_A$, improves the convergence of the open-loop estimates $\hat{p}_B$ and $\hat{z}$, while the estimate $\hat{x}_v$ remains unaffected by changes in the observer gains.

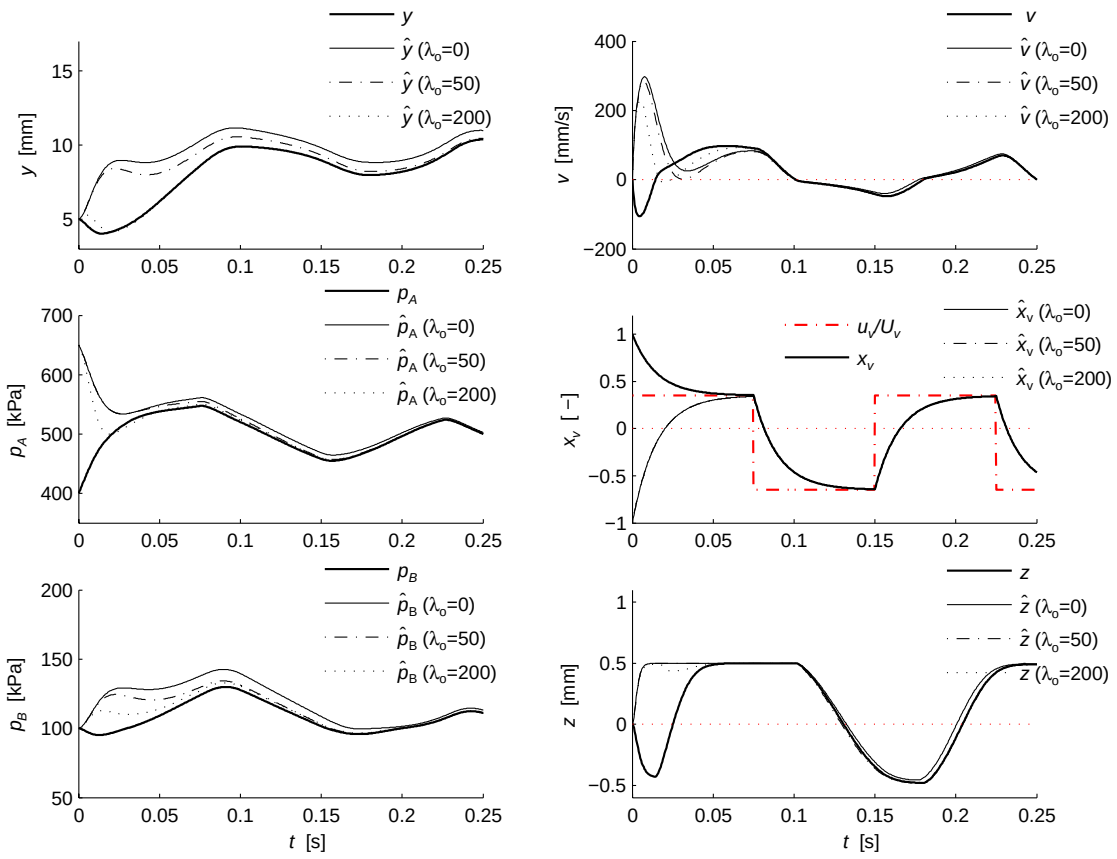

Figure 8.4: Illustration of the convergence properties of the full-order observer for increasing feedback gains.

8.4 Reduced-order observer

In this section we propose a simple reduced-order observer for the electro-pneumatic actuator. The observer is of 4th-order, and estimates only the unmeasured states based on a simplified design model where the back-chamber pressure p_B is assumed to be constant.

8.4.1 Simplified design model

Denote $p \triangleq p_A$, $A \triangleq A_A$, $V(y) = V_0 + Ay \triangleq V_A(y)$, and take $p_B = P_0$ and $A_B = A$, and assume that the dynamics of the electro-pneumatic actuator can be given by the model

$$\begin{aligned}
\dot{y} &= v \\
\dot{v} &= \frac{A}{M}p - \frac{A}{M}P_0 - \frac{1}{M}f_l(y) - \frac{1}{M}f_f(v,z) \\
\dot{p} &= -A\frac{1}{V(y)}vp + \rho_0 T_0 R C_v \frac{1}{V(y)}\psi_v(p,x_v) \\
\dot{x}_v &= -\frac{1}{\tau_v}x_v + \frac{1}{\tau_v}u \\
\dot{z} &= v - \frac{K_z}{F_C}|v|_s z,
\end{aligned} \tag{8.72}$$

where y, v, p, x_v, and z are the actuator position, velocity, chamber pressure, valve spool opening, and friction deflection state, respectively, while $f_l(y)$, $f_f(v,z)$, $V(y)$, and $\psi_v(p,x_v)$, are known functions of the clutch load force, friction force, chamber volume, and the valve flow function, respectively.

The *region of validity* of model (8.1) is the set $\mathcal{X}_0 \subset \mathbb{R}^5$, defined by

$$\mathcal{X}_0 \triangleq \{\forall \mathbf{x} : \mathbf{x}_{\min} \leq \mathbf{x} \leq \mathbf{x}_{\max}\}, \tag{8.73}$$

where the minimum and maximum feasible actuator states are given by $\mathbf{x}_{\min} = [y_{\min}, v_{\min}, p_{\min}, x_{v,\min}, z_{\min}]$ and $\mathbf{x}_{\max} = [y_{\max}, v_{\max}, p_{\max}, x_{v,\max}, z_{\max}]$, as described in Chapter 6. Moreover, the *region of normal operation,* in which the actuator states will remain under normal operation, is defined by the lower and upper bounds $\mathbf{x}_{\text{lb}} \triangleq [y_{\text{lb}}, v_{\text{lb}}, p_{\text{lb}}, x_{v\,\text{lb}}, z_{\text{lb}}]^T$ and $\mathbf{x}_{\text{ub}} \triangleq [y_{\text{ub}}, v_{\text{ub}}, p_{\text{ub}}, x_{v\,\text{ub}}, z_{\text{ub}}]^T$ as

$$\mathcal{X} \triangleq \{\forall \mathbf{x} \in \mathcal{X}_0 \colon \mathbf{x}_{\text{lb}} \leq \mathbf{x} \leq \mathbf{x}_{\text{ub}}\}. \tag{8.74}$$

In the following section, we illustrate the development of a reduced-order observer for the system (8.72) which only estimates the unmeasured states v, p, x_v, and z.

8.4.2 Observer development

We introduce the change of coordinates

$$\begin{aligned}
\xi_1 &\triangleq v - k_1 y \\
\xi_2 &\triangleq V(y)\, p - k_2 y,
\end{aligned} \tag{8.75}$$

where $k_1, k_2 \geq 0$ are design parameters, whose meaning will become clear in the following. Noting that

$$\begin{aligned}
\dot{\xi}_1 &= \dot{v} - k_1 \dot{y} \\
&= \frac{A}{M}p - \frac{A}{M}P_0 - \frac{1}{M}f_l(y) - \frac{1}{M}f_f(v,z) - k_1 v
\end{aligned}$$

and

$$\begin{aligned}
\dot{\xi}_2 &= \frac{\partial V(y)}{\partial y}\dot{y}p + V(y)\,\dot{p} - k_2\dot{y} \\
&= Avp + V(y)\left(-A\frac{1}{V(y)}vp + \rho_0 T_0 R C_v \frac{1}{V(y)}\psi_v(p,x_v)\right) - k_2 v \\
&= \rho_0 T_0 R C_v \psi_v(p,x_v) - k_2 v,
\end{aligned}$$

we can rewrite the dynamics of the (v,p)–subsystem in the new coordinates as

$$\begin{aligned}\dot{\xi}_1 &= \frac{A}{M}p\left(y,\xi_2\right)-\frac{A}{M}P_0-\frac{1}{M}f_l\left(y\right)-\frac{1}{M}f_f\left(v\left(y,\xi_1\right),z\right)-k_1v\left(y,\xi_1\right)\\ \dot{\xi}_2 &= \rho_0T_0RC_v\psi_v\left(p\left(y,\xi_2\right),x_v\right)-k_2v\left(y,\xi_1\right)\end{aligned} \tag{8.76}$$

where v and p are now functions of y, and the new states ξ_1 and ξ_2, given by (8.75) according to

$$\begin{aligned}v\left(y,\xi_1\right) &= \xi_1+k_1y\\ p\left(y,\xi_2\right) &= \frac{1}{V\left(y\right)}\left(\xi_2+k_2y\right).\end{aligned} \tag{8.77}$$

Writing $-k_1v$ and $-k_2v$ in (8.76) in terms of ξ_1 and ξ_2, gives

$$\begin{aligned}\dot{\xi}_1 &= -k_1\xi_1-k_1^2y+\frac{A}{M}p\left(y,\xi_2\right)-\frac{A}{M}P_0-\frac{1}{M}f_l\left(y\right)-\frac{1}{M}f_f\left(v\left(y,\xi_1\right),z\right)\\ \dot{\xi}_2 &= -k_2\xi_1-k_1k_2y+\rho_0T_0RC_v\psi_v\left(p\left(y,\xi_2\right),x_v\right),\end{aligned} \tag{8.78}$$

where we notice that the change of coordinates (8.75) has introduced the stabilizing feedback terms $-k_1\xi_1$ and $k_2\xi_1$ in the dynamics of the (ξ_1,ξ_2)–subsystem. Estimate ξ_1 and ξ_2 using the observer

$$\begin{aligned}\dot{\hat{\xi}}_1 &= -k_1\hat{\xi}_1-k_1^2y+\frac{A}{M}p\left(y,\hat{\xi}_2\right)-\frac{A}{M}P_0-\frac{1}{M}f_l\left(y\right)-\frac{1}{M}f_f\left(v\left(y,\hat{\xi}_1\right),z\right)\\ \dot{\hat{\xi}}_2 &= -k_2\hat{\xi}_1-k_1k_2y+\rho_0T_0RC_v\psi_v\left(p\left(y,\hat{\xi}_2\right),x_v\right),\end{aligned} \tag{8.79}$$

and let estimates of v and p be given by

$$\begin{aligned}\hat{v}\left(y,\hat{\xi}_1\right) &= \hat{\xi}_1+k_1y\\ \hat{p}\left(y,\hat{\xi}_2\right) &= \frac{1}{V\left(y\right)}\left(\hat{\xi}_2+k_2y\right).\end{aligned} \tag{8.80}$$

First note that the estimation errors $\tilde{v}\triangleq v-\hat{v}$ and $\tilde{p}\triangleq p-\hat{p}$ are given by the observer errors $\tilde{\xi}_1\triangleq\xi_1-\hat{\xi}_1$ and $\tilde{\xi}_2\triangleq\xi_2-\hat{\xi}_2$ according to

$$\begin{aligned}\tilde{v} &\triangleq v-\hat{v}=\tilde{\xi}_1\\ \tilde{p} &\triangleq p-\hat{p}=\frac{1}{V\left(y\right)}\tilde{\xi}_2.\end{aligned}$$

Now assume that, in addition to the measured output y, that also the friction f_f and the flow function ψ_v are known. Then, f_f and ψ_v are cancelled, and it is straightforward to show that the estimation error is governed by the time-varying linear dynamics

$$\begin{bmatrix}\dot{\tilde{\xi}}_1\\ \dot{\tilde{\xi}}_2\end{bmatrix}=\begin{bmatrix}-k_1 & \frac{A}{M}\frac{1}{V(y)}\\ -k_2 & 0\end{bmatrix}\begin{bmatrix}\tilde{\xi}_1\\ \tilde{\xi}_2\end{bmatrix}, \tag{8.81}$$

which can be shown to be exponentially stable for $y\in[y_{\mathrm{lb}},y_{\mathrm{ub}}]$ with $k_1,k_2>0$. Motivated by this result, we propose the following reduced-order observer.

Reduced-order observer

A reduced-order observer for the system (8.72) is given by

$$\begin{aligned}
\dot{\hat{\xi}}_1 &= \frac{A}{M}\hat{p}\left(y,\hat{\xi}_2\right) - \frac{A}{M}P_0 - \frac{1}{M}f_l\left(y\right) - \frac{1}{M}f_f\left(\hat{v}(y,\hat{\xi}_1),\hat{z}\right) - k_1\cdot\hat{v}(y,\hat{\xi}_1)\\
\dot{\hat{\xi}}_2 &= \rho_0 T_0 R C_v\cdot\psi_v\left(\hat{p}(y,\hat{\xi}_2),\hat{x}_v\right) - k_2\cdot\hat{v}(y,\hat{\xi}_1)\\
\dot{\hat{x}}_v &= -\frac{1}{\tau_v}\hat{x}_v + \frac{1}{\tau_v}u\\
\dot{\hat{z}} &= \hat{v} - \frac{K_z}{F_C}\left|\hat{v}\right|_s\hat{z},
\end{aligned} \tag{8.82}$$

where $\hat{v}$ and $\hat{p}$ are functions of y, $\hat{\xi}_1$, and $\hat{\xi}_2$, given by

$$\begin{aligned}
\hat{v}(y,\hat{\xi}_1) &= \hat{\xi}_1 + k_1 y\\
\hat{p}(y,\hat{\xi}_2) &= \frac{1}{V\left(y\right)}\left(\hat{\xi}_2 + k_2 y\right).
\end{aligned} \tag{8.83}$$

Here, the observer states are denoted by $\hat{\boldsymbol{\rho}} \triangleq [\hat{\xi}_1,\hat{\xi}_2,\hat{x}_v,\hat{z}]^T$, the corresponding estimates of the unmeasured states by $\hat{\mathbf{x}}_u \triangleq [\hat{v},\hat{p},\hat{x}_v,\hat{z}]^T$, and the actuator states by $\mathbf{x} \triangleq [y,\mathbf{x}_u]^T = [y,v,p,x_v,z]^T$.

Like the full-order observer, the reduced-order observer (8.82)–(8.83) combines output-injection based estimation of the main states v and p, with open-loop estimation of the remaining states x_v and z. Partitioning the observer states according to $\hat{\boldsymbol{\rho}} = [\hat{\boldsymbol{\xi}},\hat{\boldsymbol{\zeta}}]^T$, where $\hat{\boldsymbol{\xi}} \triangleq [\hat{\xi}_1,\hat{\xi}_2]^T$ denotes the output-corrected states, and $\hat{\boldsymbol{\zeta}} \triangleq [\hat{x}_v,\hat{z}]^T$ denotes the open-loop estimated states, the observer (8.63) can be compactly expressed in the form

$$\begin{aligned}
\dot{\hat{\boldsymbol{\xi}}} &= \mathbf{f}_\xi\left(\hat{\mathbf{x}}_u,y\right) - \mathbf{k}\hat{v}\\
\dot{\hat{\boldsymbol{\zeta}}} &= \mathbf{f}_\zeta\left(\hat{\mathbf{x}}_u,y,u\right),
\end{aligned} \tag{8.84}$$

where $\mathbf{k} \triangleq [k_1,k_2]^T$ denotes the observer feedback gains, $\hat{\mathbf{x}}_u \triangleq [\hat{v},\hat{p}_A,\hat{x}_v,\hat{z}]^T$ the estimated unmeasured states, and where the vectors describing the system dynamics are given by

$$\mathbf{f}_\xi\left(\hat{\mathbf{x}}_u,y\right) = \begin{bmatrix} f_{\xi 1}\\ f_{\xi 2}\end{bmatrix} \triangleq \begin{bmatrix} \frac{A}{M}\hat{p} - \frac{A}{M}P_0 - \frac{1}{M}f_l\left(y\right) - \frac{1}{M}f_f\left(\hat{v},\hat{z}\right)\\ \rho_0 T_0 R C_v \psi_v\left(\hat{p},\hat{x}_v\right)\end{bmatrix} \tag{8.85}$$

$$\mathbf{f}_\zeta\left(\hat{\mathbf{x}}_u,y,u_v\right) = \begin{bmatrix} f_{\zeta 1}\\ f_{\zeta 2}\end{bmatrix} \triangleq \begin{bmatrix} -\frac{1}{\tau_v}\hat{x}_v + \frac{1}{\tau_v}u\\ \hat{v} - \frac{K_z}{F_C}\left|\hat{v}\right|_s\hat{z}\end{bmatrix}. \tag{8.86}$$

8.4.3 Stability and convergence properties

The stability and convergence properties of the reduced-order observer (8.82) are qualitatively identical to that of the full-order observer (8.63).

With the introduced change of coordinates (8.75), the dynamics of the system (8.72) can be represented by

$$\begin{aligned}
\dot{y} &= \xi_1 + k_1 y \\
\dot{\xi}_1 &= \frac{A}{M}p - \frac{A}{M}P_0 - \frac{1}{M}f_l(y) - \frac{1}{M}f_f(v,z) - k_1 v \\
\dot{\xi}_2 &= \rho_0 T_0 R C_v \psi_v(p, x_v) - k_2 v \\
\dot{x}_v &= -\frac{1}{\tau_v}x_v + \frac{1}{\tau_v}u \\
\dot{z} &= v - \frac{K_z}{F_C}|v|_s z,
\end{aligned} \tag{8.87}$$

where

$$v(y, \xi_1) = \xi_1 + k_1 y \tag{8.88}$$

$$p(y, \xi_2) = \frac{1}{V(y)}(\xi_2 + k_2 y). \tag{8.89}$$

Observer errors dynamics

The stability and convergence properties of the observer (8.82) is determined by analyzing the dynamics of the resulting observer errors

$$\begin{aligned}
\tilde{\xi}_1 &\triangleq \xi_1 - \hat{\xi}_1 \\
\tilde{\xi}_2 &\triangleq \xi_2 - \hat{\xi}_2 \\
\tilde{x}_v &\triangleq x_v - \hat{x}_v \\
\tilde{z} &\triangleq z - \hat{z},
\end{aligned} \tag{8.90}$$

where the resulting velocity and pressure errors are given by the observer errors as

$$\tilde{v} = \tilde{\xi}_1 \tag{8.91}$$

$$\tilde{p} = \frac{1}{V(y)}\tilde{\xi}_2. \tag{8.92}$$

The first step of the analysis is to express the dynamics of the observer error in a suitable form. Since the details of these steps were illustrated for the open-loop observer, we shorten the presentation here, illustrating only the main steps.

By differentiation, we can show that the observer errors are governed by the dynamics

$$\begin{aligned}
M\dot{\tilde{\xi}}_1 &= -Mk_1\tilde{v} + A\tilde{p} - [f_f(v,z) - f_f(\hat{v},\hat{z})] \\
\dot{\tilde{\xi}}_2 &= -k_2\tilde{v} + \rho_0 T_0 R C_v [\psi_v(p, x_v) - \psi_v(\hat{p}, \hat{x}_v)] \\
\dot{\tilde{z}} &= g_z(v,z) - g_z(\hat{v},\hat{z}) \\
\tau_v\dot{\tilde{x}}_v &= -\tilde{x}_v,
\end{aligned} \tag{8.93}$$

where the friction error can be written

$$f_f(v,z) - f_f(\hat{v},\hat{z}) = D_v\tilde{v} + K_z\tilde{z} + D_{\dot{z}}\dot{\tilde{z}},$$

and the pre-sliding friction dynamics is given by the function

$$g_z(v,z) = v - \frac{K_z}{F_C}|v|_s z.$$

Using the Mean Value Theorem (Lemma 27), we recall that the errors in the pre-sliding dynamics, and the nonlinear flow function, can be written as linear in its error variables according to

$$\dot{\tilde{z}} = g_z(v,z) - g_z(\hat{v},\hat{z}) = \frac{\partial g_z(\bar{v},\bar{z})}{\partial v}\tilde{v} + \frac{\partial g_z(\bar{v},\bar{z})}{\partial z}\tilde{z} \tag{8.94}$$

$$\psi_v(p,x_v) - \psi_v(\hat{p},\hat{x}_v) = \frac{\partial \psi_v(\bar{p},\bar{x}_v)}{\partial p}\tilde{p} + \frac{\partial \psi_v(\bar{p},\bar{x}_v)}{\partial x_v}\tilde{x}_v \tag{8.95}$$

with the gradients evaluated at time-varying points constrained to

$$\begin{aligned} \bar{v} &\in \mathcal{S}_v \triangleq \{\min(v,\hat{v}), \max(v,\hat{v})\} \\ \bar{p} &\in \mathcal{S}_p \triangleq \{\min(p,\hat{p}), \max(p,\hat{p})\} \\ \bar{z} &\in \mathcal{S}_z \triangleq \{\min(z,\hat{z}), \max(z,\hat{z})\} \\ \bar{x}_v &\in \mathcal{S}_{x_v} \triangleq \{\min(x_v,\hat{x}_v), \max(x_v,\hat{x}_v)\}. \end{aligned} \tag{8.96}$$

Using these results, and substituting

$$\begin{aligned} \tilde{v} &= \tilde{\xi}_1 \\ \tilde{p} &= \frac{1}{V(y)}\tilde{\xi}_2, \end{aligned}$$

the complete error dynamics can be expressed in the linear, time-varying form

$$\begin{aligned} M\dot{\tilde{\xi}}_1 &= -\left[Mk_1 + D_v + D_{\dot{z}}\frac{\partial g_z(\bar{v},\bar{z})}{\partial v}\right]\tilde{\xi}_1 + \frac{A}{V(y)}\tilde{\xi}_2 - \left[K_z + D_{\dot{z}}\frac{\partial g_z(\bar{v},\bar{z})}{\partial z}\right]\tilde{z} \\ \dot{\tilde{\xi}}_2 &= -k_2\tilde{\xi}_1 + \rho_0 T_0 R C_v \frac{\partial \psi_v(\bar{p},\bar{x}_v)}{\partial p}\frac{1}{V(y)}\tilde{\xi}_2 + \rho_0 T_0 R C_v \frac{\partial \psi_v(\bar{p},\bar{x}_v)}{\partial x_v}\tilde{x}_v \\ \dot{\tilde{z}} &= \frac{\partial g_z(\bar{v},\bar{z})}{\partial v}\tilde{v} + \frac{\partial g_z(\bar{v},\bar{z})}{\partial z}\tilde{z} \\ \tau_v\dot{\tilde{x}}_v &= -\tilde{x}_v. \end{aligned} \tag{8.97}$$

Denoting

$$\sigma_1(t) \triangleq \frac{\partial g_z(\bar{v},\bar{z})}{\partial v} \tag{8.98}$$

$$\sigma_2(t) \triangleq -\frac{\partial g_z(\bar{v},\bar{z})}{\partial z} \tag{8.99}$$

$$\alpha(t) \triangleq -\rho_0 T_0 R C_v \frac{\partial \psi_v(\bar{p},\bar{x}_v)}{\partial p}\frac{1}{V(y)} \tag{8.100}$$

$$\beta(t) \triangleq \rho_0 T_0 R C_v \frac{\partial \psi_v(\bar{p},\bar{x}_v)}{\partial x_v}, \tag{8.101}$$

the error dynamics can be expressed as

$$\begin{aligned} M\dot{\tilde{\xi}}_1 &= -\left[Mk_1 + D_v + D_{\dot{z}}\sigma_1(t)\right]\tilde{\xi}_1 + \frac{A}{V(y)}\tilde{\xi}_2 - \left[K_z + D_{\dot{z}}\sigma_2(t)\right]\tilde{z} \\ \dot{\tilde{\xi}}_2 &= -k_2\tilde{\xi}_1 - \alpha(t)\frac{1}{V(y)}\tilde{\xi}_2 + \beta(t)\tilde{x}_v \\ \dot{\tilde{z}} &= \sigma_1(t)\tilde{v} - \sigma_2(t)\tilde{z} \\ \tau_v\dot{\tilde{x}}_v &= -\tilde{x}_v, \end{aligned} \tag{8.102}$$

where the gradients of the friction dynamics satisfy,

$$\begin{aligned} \sigma_1(t) &\triangleq \frac{\partial g_z(\bar{v},\bar{z})}{\partial v} = 1 - \frac{K_z}{F_C}\operatorname{sgn}_s(\bar{v})\bar{z} \\ &\in \langle 0, 2\rangle \end{aligned} \tag{8.103}$$

$$\begin{aligned} \sigma_2(t) &\triangleq -\frac{\partial g_z(\bar{v},\bar{z})}{\partial z} = \frac{K_z}{F_C}|\bar{v}|_s \\ &\in \left[\frac{K_z}{F_C}\varepsilon_0, \frac{K_z}{F_C}v_{\text{ub}}\right], \end{aligned} \tag{8.104}$$

and the flow gradients

$$\alpha(t) \triangleq -\rho_0 T_0 R C_v \frac{\partial \psi_v(\bar{p},\bar{x}_v)}{\partial p}\frac{1}{V(y)} \geq 0 \tag{8.105}$$

$$\beta(t) \triangleq \rho_0 T_0 R C_v \frac{\partial \psi_v(\bar{p},\bar{x}_v)}{\partial x_v} \geq 0, \tag{8.106}$$

provided both the actual and estimated states remain in the region of validity of the model.

The structure of the error dynamics becomes more apparent by expressing the system in the matrix form

$$\begin{bmatrix} M\dot{\tilde{\xi}}_1 \\ \dot{\tilde{\xi}}_2 \\ \dot{\tilde{z}} \\ \tau_v\dot{\tilde{x}}_v \end{bmatrix} = \begin{bmatrix} -\left[Mk_1 + D_v + D_{\dot{z}}\sigma_1(t)\right] & \frac{A}{V(y)} & -\left[K_z - D_{\dot{z}}\sigma_2(t)\right] & 0 \\ -k_2 & -\alpha(t) & 0 & \beta(t) \\ \sigma_1(t) & 0 & -\sigma_2(t) & 0 \\ 0 & 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \tilde{\xi}_1 \\ \tilde{\xi}_2 \\ \tilde{z} \\ \tilde{x}_v \end{bmatrix}. \tag{8.107}$$

Knowing that all the time-varying coefficients are non-negative, the structure of the error dynamics reveals the stability properties of the system: Briefly states, when the time-varying coefficients $-\alpha(t)$, and $-\sigma_2(t)$ are sufficiently large, the error system will have a dominating negative diagonal that ensures convergence of the complete error dynamics to zero, otherwise, the error system is stabilized by the diagonal damping term $-\left[Mk_1 + D_v + D_{\dot{z}}\sigma_1(t)\right]\tilde{\xi}_1$, with the $\tilde{\xi}_2$ error stabilized by the skew-symmetrical-like connection between the cross-terms $-k_2\tilde{\xi}_1$ and $A/V(y)\,\tilde{\xi}_2$, and the $\tilde{z}$ error by $\sigma_1(t)\tilde{\xi}_1$ and $-\left[K_z - D_{\dot{z}}\sigma_2(t)\right]\tilde{z}$. Since $\beta(t)$ is bounded, the resulting $\left(\tilde{\xi}_1, \tilde{\xi}_2, \tilde{z}\right)$-subsystem will be ISS with respect to the $\tilde{x}_v$-subsystem, which is exponentially stable. Consequently, the complete $\left(\tilde{\xi}_1, \tilde{\xi}_2, \tilde{z}, \tilde{x}_v\right)$-system must be asymptotically stable.

8.4.4 Observer gains

Consider the simplified time-varying linear dynamics (8.81) of the $\left(\tilde{\xi}_1, \tilde{\xi}_2\right)$-subsystem from the development of the reduced-order observer. We extract the time-invariant dynamics

$$\begin{bmatrix} \dot{\tilde{v}} \\ \dot{\tilde{p}} \end{bmatrix} = \begin{bmatrix} -k_1 & \frac{A}{M}\frac{1}{V(y_{\text{ub}})} \\ -k_2 & 0 \end{bmatrix} \begin{bmatrix} \tilde{v} \\ \tilde{p} \end{bmatrix},$$

which we use to facilitate the tuning of observer gains $\mathbf{k} \triangleq [k_1, k_2]^T$ for the nonlinear observer (8.82). This linear system is compactly written as

$$\dot{\tilde{\boldsymbol{\xi}}} = \mathbf{A}_o \tilde{\boldsymbol{\xi}}, \tag{8.108}$$

and has the characteristic polynomial

$$|s\mathbf{I} - \mathbf{A}_o| = s^2 + k_1 s + k_2 \frac{A}{MV\left(y_{\text{ub}}\right)}. \tag{8.109}$$

The observer gains are chosen such that the constant observer matrix is Hurwitz, *i.e.*, so that $\mathbf{A}_o$ satisfies the Lyapunov equation $\mathbf{A}_o\mathbf{P}_o + \mathbf{P}_o\mathbf{A}_o^T = -\mathbf{Q}$ for some $\mathbf{P}_o = \mathbf{P}_o^T > 0$ and $\mathbf{Q} > 0$, which is satisfied for $\forall\ k_1, k_2 > 0$.

We choose the observer gains k_1 and k_2 by placing multiple poles at $s = -\lambda_o < 0$, which gives a critically damped response in the output $\tilde{\xi}_1 = \tilde{v}$. The dynamics of the linear system (8.67) is then governed by the characteristic equation

$$|s\mathbf{I} - \mathbf{A}_o| = (s + \lambda_o)^2 = s^2 + 2\lambda_o s + \lambda_o^2 = 0. \tag{8.110}$$

Comparing the coefficients of the two polynomials (8.109) and (8.110), the observer gains can be expressed as functions of λ_o:

$$\begin{aligned} k_1 &= 2\lambda_o \\ k_2 &= \tfrac{MV(y_{\text{ub}})}{A}\lambda_o^2. \end{aligned} \tag{8.111}$$

With the observer feedback gains determined according to (8.111), we refer to the parameter $\lambda_o > 0$ as the *design bandwidth* of the nonlinear observer (8.82).

8.4.5 Simulation results

The simulation results are qualitatively same as the results obtained using the full-order observer, plotted in Figures 8.4 (page 132), and is therefore omitted.

8.5 Robust re-design

In this section, we introduce a simple robust modification of the observers by utilizing smooth saturation to constrain the observer states to the feasible region $\mathcal{X}_0$. Moreover, we show that we can improve the initial transient performance of the observer by using discontinuous projection of the unsaturated observer states. The

robust re-design is developed for the full-order observer (8.63), and then applied also for the reduced-order observer (8.82).

We first introduce the smooth saturation operator, followed by presenting the modified observer with saturated observer states. Next, we introduce the discontinuous projection operator, which we subsequently apply for discontinuous projection of observer states. Finally, we illustrate by simulations some robustness properties resulting from the introduced modifications.

8.5.1 Smooth saturation of observer estimates

In normal operation of the actuator, the states will stay within some compact region in state-space, which means that each state will be bounded. With the knowledge about some upper and lower bounds on each system state in normal operation, given by

$$\begin{aligned} \mathbf{x}_{\text{ub}} &\triangleq [y_{\text{ub}}, v_{\text{ub}}, p_{A\,\text{ub}}, x_{v\,\text{ub}}, p_{B\,\text{ub}}, z_{\text{ub}}]^T \\ \mathbf{x}_{\text{lb}} &\triangleq [y_{\text{lb}}, v_{\text{lb}}, p_{A\,\text{lb}}, x_{v\,\text{lb}}, p_{B\,\text{lb}}, z_{\text{lb}}]^T, \end{aligned}$$

an estimate of the *region of normal operation* is given by

$$\mathcal{X} \triangleq \{\forall \mathbf{x} \in \mathcal{X}_0 \colon \mathbf{x}_{\text{lb}} \leq \mathbf{x} \leq \mathbf{x}_{\text{ub}}\}, \tag{8.112}$$

where $\mathcal{X}_0$ is the set containing all physically feasible states of the system (8.1).

Since we are primarily interested in tracking control of the actuator in normal operation, it is sufficient to require that the observer provides asymptotically converging estimates as long as the actuator states remain within $\mathcal{X}$. In addition, it is required that the observer states remain bounded (or constrained to some set containing $\mathcal{X}$) if the actuator in particular situations exceed its normal operating range.

We may achieve this by projecting the observer states to the set $\mathcal{X}$ using projection (see *e.g.* [33, 52]), however, this makes the resulting observer dynamics incompatible with an observer backstepping design because the projection operator introduces a non-differentiability in the observer (see *e.g.* [32]). Motivated by the use of smooth parameter saturation in nonlinear adaptive control (see *e.g.* [55,95,105]), we use smooth saturation of the observer estimates to constrain the observer dynamics to the set $\mathcal{X}$, which achieves global stability properties of the observer.

We define the smooth saturation operator[2]

$$\boldsymbol{\pi}(\hat{\mathbf{x}}) \triangleq [\pi_1(\hat{x}_1), \pi_2(\hat{x}_2), \cdots, \pi_6(\hat{x}_6)]^T, \tag{8.113}$$

where

$$\pi_i(\hat{x}_i) \triangleq \pi(\hat{x}_i, x_{i,\text{lb}}, x_{i,\text{ub}}, \varepsilon_{\pi,i}), \qquad i = 1, \cdots, 6, \tag{8.114}$$

utilizing the scalar saturation function $\pi(\cdot)$ defined by (7.5), page 108, where $x_{i,\text{lb}}$, $x_{i,\text{ub}}$, $\varepsilon_{\pi,i}$, $i = 1, \cdots, 6$, are the lower and upper bounds and smoothing widths,

[2] Defined here for our 6th-order system at hand, however, the generalization to nth-order systems is trivial.

respectively. The saturation function $\boldsymbol{\pi}(\hat{\mathbf{x}})$ defined in this way, is decentralized[3], where each $\pi_i(\hat{x}_i)$ is smooth, nondecreasing, and satisfies the following properties, as stated in Chapter 7:

P7.1) $\pi_i(\hat{x}_i) \equiv \hat{x}_i$ $, \forall \hat{x}_i \in [x_{i,\mathrm{lb}} + \varepsilon_{\pi,i}, x_{i,\mathrm{ub}} - \varepsilon_{\pi,i}]$

P7.2) $\pi_i(\hat{x}_i) \in [x_{i,\mathrm{lb}}, x_{i,\mathrm{ub}}]$ $, \forall \hat{x}_i \in \mathbb{R}$.

For a compact notation in the following, we denote the saturated estimates by $\hat{\mathbf{x}}_\pi = [\hat{y}_\pi, \hat{\mathbf{x}}_{u,\pi}] \triangleq \boldsymbol{\pi}(\hat{\mathbf{x}})$, componentwise by

$$\hat{\mathbf{x}}_\pi = \begin{bmatrix} \hat{x}_{1,\pi} \\ \hat{x}_{2,\pi} \\ \hat{x}_{3,\pi} \\ \hat{x}_{4,\pi} \\ \hat{x}_{5,\pi} \\ \hat{x}_{6,\pi} \end{bmatrix} = \begin{bmatrix} \hat{y}_\pi \\ \hat{v}_\pi \\ \hat{p}_{A,\pi} \\ \hat{x}_{v,\pi} \\ \hat{p}_{B,\pi} \\ \hat{z}_\pi \end{bmatrix} \triangleq \begin{bmatrix} \pi_1(\hat{y}) \\ \pi_2(\hat{v}) \\ \pi_3(\hat{p}_A) \\ \pi_4(\hat{x}_v) \\ \pi_5(\hat{p}_B) \\ \pi_6(\hat{z}) \end{bmatrix}.$$

A robust redesign of the full-order observer (8.64) is obtained simply by saturating the observer estimates in the observer dynamics. Utilizing the above introduced smooth saturation, a robust observer for the electro-pneumatic actuator is given by

$$\begin{aligned} \dot{\hat{\boldsymbol{\xi}}} &= \mathbf{k}\tilde{y}_\pi + \mathbf{f}_\xi(\hat{\mathbf{x}}_\pi, y) \\ \dot{\hat{\boldsymbol{\zeta}}} &= \mathbf{f}_\zeta(\hat{\mathbf{x}}_\pi, y, u), \end{aligned} \tag{8.115}$$

where $\tilde{y}_\pi \triangleq y - \hat{y}_\pi = y - \pi_1(\hat{y})$ is the saturated output estimation error, and the system dynamics is given by

$$\mathbf{f}_\xi(\hat{\mathbf{x}}_\pi, y) = \begin{bmatrix} f_{\xi 1} \\ f_{\xi 2} \\ f_{\xi 3} \end{bmatrix} \triangleq \begin{bmatrix} \hat{v}_\pi \\ \frac{A_0 P_0}{M} + \frac{A_A}{M}\hat{p}_{A,\pi} - \frac{A_B}{M}\hat{p}_{B,\pi} - \frac{1}{M}f_l(y) - \frac{1}{M}f_f(\hat{v}_\pi, \hat{x}_{v,\pi}) \\ -A_A \frac{1}{V_A(y)}\hat{v}_\pi \hat{p}_{A,\pi} + \rho_0 T_0 R C_v \frac{1}{V_A(y)}\psi_v(\hat{p}_{A,\pi}, \hat{x}_{v,\pi}) \end{bmatrix} \tag{8.116}$$

$$\mathbf{f}_\zeta(\hat{\mathbf{x}}_\pi, y, u_v) = \begin{bmatrix} f_{\zeta 1} \\ f_{\zeta 2} \\ f_{\zeta 3} \end{bmatrix} \triangleq \begin{bmatrix} -\frac{1}{\tau_v}\hat{x}_{v,\pi} + \frac{1}{\tau_v}u \\ A_B \frac{1}{V_B(y)}\hat{v}_\pi \hat{p}_{B,\pi} + \rho_0 T_0 R C_r \frac{1}{V_B(y)}\psi_r(\hat{p}_{B,\pi}) \\ -\frac{K_z}{F_C}|\hat{v}_\pi|_s \hat{z}_\pi + \hat{v}_\pi \end{bmatrix}. \tag{8.117}$$

Figure 8.5, the re-designed full-order observer (8.115) is visualized by a block diagram.

Remark 32 *Note that the introduced saturation of the estimated friction deflection state $\hat{z}$ is superfluous when $\hat{z}(0) \in [z_{\mathrm{lb}}, z_{\mathrm{ub}}] \subset [-Z_{\max}, Z_{\max}]$ because the finite deflection property of the friction model (see proof of Proposition 11, page 54), guarantees that $\hat{z}(t) \in [-Z_{\max}, Z_{\max}]$ for $\forall t \geq 0$.*

[3]The ith component of $\boldsymbol{\pi}(\hat{\mathbf{x}})$ depends only on the ith component of $\hat{\mathbf{x}}$.

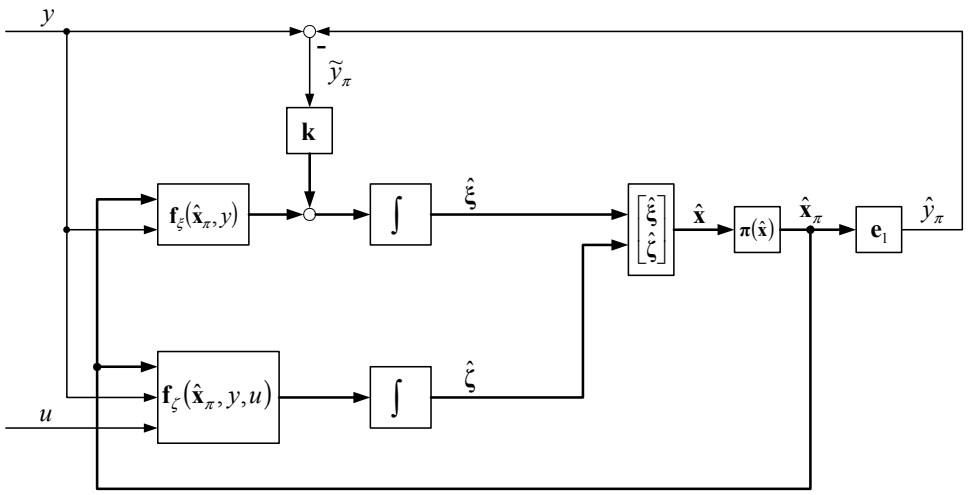

Figure 8.5: Block diagram of the full-order observer utilizing smooth saturation of state estimates

8.5.2 Discontinuous projection of observer states

Though the redesigned observer (8.115) ensures global stability properties of the observer, the actual (non-saturated) observer states $\hat{\boldsymbol{\xi}}$ and $\hat{\boldsymbol{\zeta}}$ (see the block diagram in Figure 8.5) are unconstrained, and may move outside the region of normal operation $\mathcal{X}$ during initial transients of the observer, while the saturated estimates $\hat{\boldsymbol{\xi}}_\pi$ and $\hat{\boldsymbol{\zeta}}_\pi$ are kept at the boundary of $\mathcal{X}$. By using projection to stop integration of $\hat{\boldsymbol{\xi}}$ and $\hat{\boldsymbol{\zeta}}$ for corresponding estimates $\hat{\boldsymbol{\xi}}_\pi$ and $\hat{\boldsymbol{\zeta}}_\pi$ which are fully saturated, the observer states $\hat{\boldsymbol{\xi}}$ and $\hat{\boldsymbol{\zeta}}$ are constrained to some small boundary layer around $\mathcal{X}$, and the transient performance of the observer can be significantly improved. Moreover, since the projection is then only active for estimates which are fully saturated, the smoothness of the saturated estimates $\hat{\boldsymbol{\xi}}_\pi$ and $\hat{\boldsymbol{\zeta}}_\pi$ is preserved. Hence, a control law can be designed by a backstepping approach using the observer with saturated estimates, and can be implemented using the observer with saturation and projection of the unsaturated states, without introducing discontinuities in the control input, thus, preserving the stability properties of the closed-loop system. See Chapter 9.

To introduce the projection operator, first note that we may write the observer dynamics in the general form

$$\dot{\hat{\mathbf{x}}} = \mathbf{f}\left(\hat{\mathbf{x}}, y, u\right), \tag{8.118}$$

where $\mathbf{f}\left(\hat{\mathbf{x}}, y, u\right) \triangleq [f_1, f_2, \cdots, f_6]^T \in \mathbb{R}^6$. We define the discontinuous projection operator as

$$\mathcal{P}_{\mathbf{x}}\left(\mathbf{f}, \hat{\mathbf{x}}\right) \triangleq \left[\mathcal{P}_{x_1}\left(f_1, \hat{x}_1\right), \mathcal{P}_{x_2}\left(f_2, \hat{x}_2\right), \cdots, \mathcal{P}_{x_6}\left(f_6, \hat{x}_6\right)\right]^T, \tag{8.119}$$

where

$$\mathcal{P}_{x_i}\left(f_i, \hat{x}_i\right) \triangleq \mathcal{P}\left(f_i, \hat{x}_i, x_{i,\text{lb}} - \varepsilon_{\pi,i}, x_{i,\text{ub}} + \varepsilon_{\pi,i}\right), \qquad i = 1, \cdots, 6, \tag{8.120}$$

utilizing the scalar projection $\mathcal{P}\left(f, x, x_{\rm lb}, x_{\rm ub}\right)$ defined by (7.10), page 110. Defined in this way, the discontinuous projection (8.119) constrains the estimates to the slightly enlarged estimate of the region of normal operation, given by

$$\mathcal{X}_{\varepsilon_\pi^+} \triangleq \left\{\forall \mathbf{x} \in \mathcal{X}_0 \colon \mathbf{x}_{\rm lb} - \boldsymbol{\varepsilon}_\pi \leq \mathbf{x} \leq \mathbf{x}_{\rm ub} + \boldsymbol{\varepsilon}_\pi\right\}, \tag{8.121}$$

which is the union of $\mathcal{X}$ and a small boundary layer around it, determined by the smoothing widths $\boldsymbol{\varepsilon}_\pi \triangleq [\varepsilon_{\pi 1}, \cdots, \varepsilon_{\pi 6}]^T$ of the smooth saturation function (8.113).

Like the saturation function $\boldsymbol{\pi}\left(\hat{\mathbf{x}}\right)$, the vectorized projection operator $\mathcal{P}_{\mathbf{x}}(\mathbf{f}, \hat{\mathbf{x}})$ is decentralized, and each scalar projection $\mathcal{P}_i\left(f_i, \hat{x}_i\right)$, $i = 1, \cdots, 6$ satisfies

$$\textbf{P7.3)} \qquad \mathcal{P}_i\left(f_i, \hat{x}_i\right) \equiv f_i \quad , \forall \hat{x}_i \in \left\langle x_{i,\rm lb} - \varepsilon_{\pi,i}, x_{i,\rm ub} + \varepsilon_{\pi,i}\right\rangle.$$

Letting $\mathring{\mathcal{X}}_{\varepsilon_\pi^+}$ denote the interior of $\mathcal{X}_{\varepsilon_\pi^+}$, Property P7.3 means that for $\forall \hat{\mathbf{x}} \in \mathring{\mathcal{X}}_{\varepsilon_\pi^+}$ the projection remains inactive, *i.e.*, $\mathcal{P}_{\mathbf{x}}\left(\mathbf{f}, \hat{\mathbf{x}}\right) \equiv \mathbf{f}$. Furthermore, with initial estimate $\hat{\mathbf{x}}\left(0\right) \in \mathcal{X}_{\varepsilon_\pi^+}$, the projection $\dot{\hat{\mathbf{x}}} = \mathcal{P}_{\mathbf{x}}(\mathbf{f}, \hat{\mathbf{x}})$ guarantees

$$\textbf{P7.4)} \qquad \hat{\mathbf{x}} \in \mathcal{X}_{\varepsilon_\pi^+} = \left\{\mathbf{x}_{\rm lb} - \boldsymbol{\varepsilon}_\pi, \mathbf{x}_{\rm ub} + \boldsymbol{\varepsilon}_\pi\right\} \quad , \forall t \geq 0.$$

8.5.3 Robust full-order observer

The redesigned observer utilizing smooth saturation combined with discontinuous projection can compactly be expressed as

$$\begin{aligned} \dot{\hat{\boldsymbol{\xi}}} &= \mathcal{P}_{\boldsymbol{\xi}}\left(\mathbf{k}\tilde{y}_\pi + \mathbf{f}_\xi\left(\hat{\mathbf{x}}_\pi, y\right), \hat{\boldsymbol{\xi}}\right) \\ \dot{\hat{\boldsymbol{\zeta}}} &= \mathcal{P}_{\boldsymbol{\zeta}}\left(\mathbf{f}_\zeta\left(\hat{\mathbf{x}}_\pi, y, u\right), \hat{\boldsymbol{\zeta}}\right), \end{aligned} \tag{8.122}$$

where $\hat{\mathbf{x}}_\pi = \boldsymbol{\pi}\left(\hat{\mathbf{x}}\right) = \boldsymbol{\pi}([\hat{\boldsymbol{\xi}}, \hat{\boldsymbol{\zeta}}]^T)$ is the saturated observer state vector with $\boldsymbol{\pi}\left(\cdot\right)$ defined by (8.113, and where the discontinuous projection operator, defined by (8.119), is partitioned according to $\mathcal{P}_{\mathbf{x}}\left(\mathbf{f}, \hat{\mathbf{x}}\right) \triangleq [\mathcal{P}_{\boldsymbol{\xi}}(\mathbf{k}\tilde{y}_\pi + \mathbf{f}_\xi, \hat{\boldsymbol{\xi}}), \mathcal{P}_{\boldsymbol{\xi}}(\mathbf{f}_\zeta, \hat{\boldsymbol{\xi}})]^T$.

Figure 8.6 illustrates by a block diagram, the re-designed full-order observer (8.122).

Simulation results

In this section we illustrate the improved robustness and convergence properties of the redesigned full-order observer with smooth saturation (8.115), and the observer with combined saturation and projection (8.122), compared to that of the nominal Luenberger-type observer (8.63).

The parameters of the electro-pneumatic system (8.1) are the same as in the preceding simulations. The observer is implemented with upper and lower bounds, $\mathbf{x}_{\rm ub}$ and $\mathbf{x}_{\rm lb}$, and smoothing width $\boldsymbol{\varepsilon}_\pi$, set according to

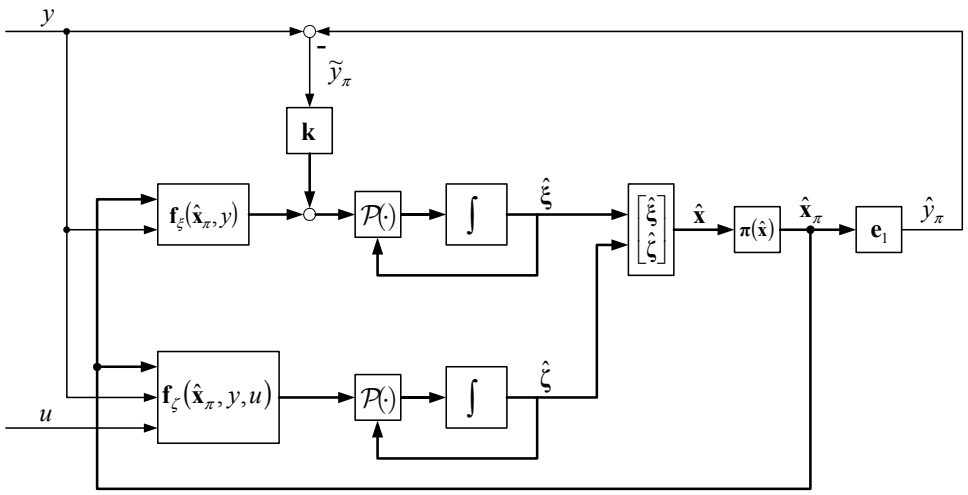

Figure 8.6: Block diagram of the re-designed full-order observer utilizing smooth saturation and discontinuous projection of the unsaturated observer states.

$$
\begin{aligned}
\mathbf{x}_{\text{ub}} &= \begin{bmatrix} 15\,\text{mm} & 200\,\text{mm/s} & 995\,\text{kPa} & 1 & 150\,\text{kPa} & 0.5\,\text{mm} \end{bmatrix}^T \\
\mathbf{x}_{\text{lb}} &= \begin{bmatrix} -5\,\text{mm} & -200\,\text{mm/s} & 105\,\text{kPa} & -1 & 50\,\text{kPa} & -0.5\,\text{mm} \end{bmatrix}^T \\
\boldsymbol{\varepsilon}_\pi &= \begin{bmatrix} 0.1\,\text{mm} & 1\,\text{mm/s} & 1\,\text{kPa} & 0.01 & 1\,\text{kPa} & 0.001\,\text{mm} \end{bmatrix}^T .
\end{aligned}
$$

We compute the feedback gains of the observer according to (8.71) using design bandwidth $\lambda_o = 100$:

$$
\lambda_o = 100 \Longrightarrow \mathbf{k} = \begin{bmatrix} k_1 \\ k_2 \\ k_3 \end{bmatrix} = \begin{bmatrix} 50 \\ 16 \cdot 10^3 \\ 1.2 \cdot 10^3 \end{bmatrix} \begin{matrix} \text{mm/mm} \\ \text{mm/(mm s)} \\ \text{kPa/mm.} \end{matrix}
$$

The performance of the observer is illustrated for the simulation plotted in Figure 8.1, where the actuator is subjected to an open-loop square wave control input. To demonstrate the global stability of the modified observer, we simply set all observer estimates initially equal to zero, giving

$$
\begin{aligned}
\hat{\mathbf{x}}(0) &= \mathbf{0} \\
&\Downarrow \\
\tilde{\mathbf{x}}(0) &= \begin{bmatrix} 5\,\text{mm} & 0\,\text{mm/s} & 400\,\text{kPa} & 1 & 100\,\text{kPa} & 0\,\text{mm} \end{bmatrix}^T \\
\tilde{\mathbf{x}}_\pi(0) &= \begin{bmatrix} 5\,\text{mm} & 0\,\text{mm/s} & 295\,\text{kPa} & 1 & 50\,\text{kPa} & 0\,\text{mm} \end{bmatrix}^T .
\end{aligned}
$$

Notice that this places the initial estimate of pressure p_A outside the region of validity $\mathcal{X}_0$ of the model, since $p_{A,\min} = P_0 = 100\,\text{kPa}$, while the initial estimate of pressure p_B lies outside the region of normal operation $\mathcal{X}$.

Figure 8.7 plots the performance of the observer with saturation (8.115). In the figure, also the unsaturated observer states are plotted with dash-dot lines ($-\cdot-$),

without legend. When the unsaturated estimates in $\hat{\mathbf{x}}$ remain well within $\mathcal{X}$, they are identical to the saturated estimates in $\hat{\mathbf{x}}_\pi$, however, during the initial transient period, the observer states $\hat{v}$, $\hat{p}_A$ and $\hat{p}_B$ also operates outside of $\mathcal{X}$, while their saturated counterparts $\hat{v}_\pi$, $\hat{p}_{A,\pi}$ and $\hat{p}_{B,\pi}$ remain inside $\mathcal{X}$ due to the saturation. We see that the velocity estimate $\hat{v}$ experiences large transient peaks before eventually converging to the actual velocity v. Notice that both the estimated pressures $\hat{p}_A$ and $\hat{p}_B$, operates outside of the feasibility region $\mathcal{X}_0$ of the model (8.1) during the initial transients. This illustrates the global stability of the redesigned observer (8.115), which is due to the saturation.

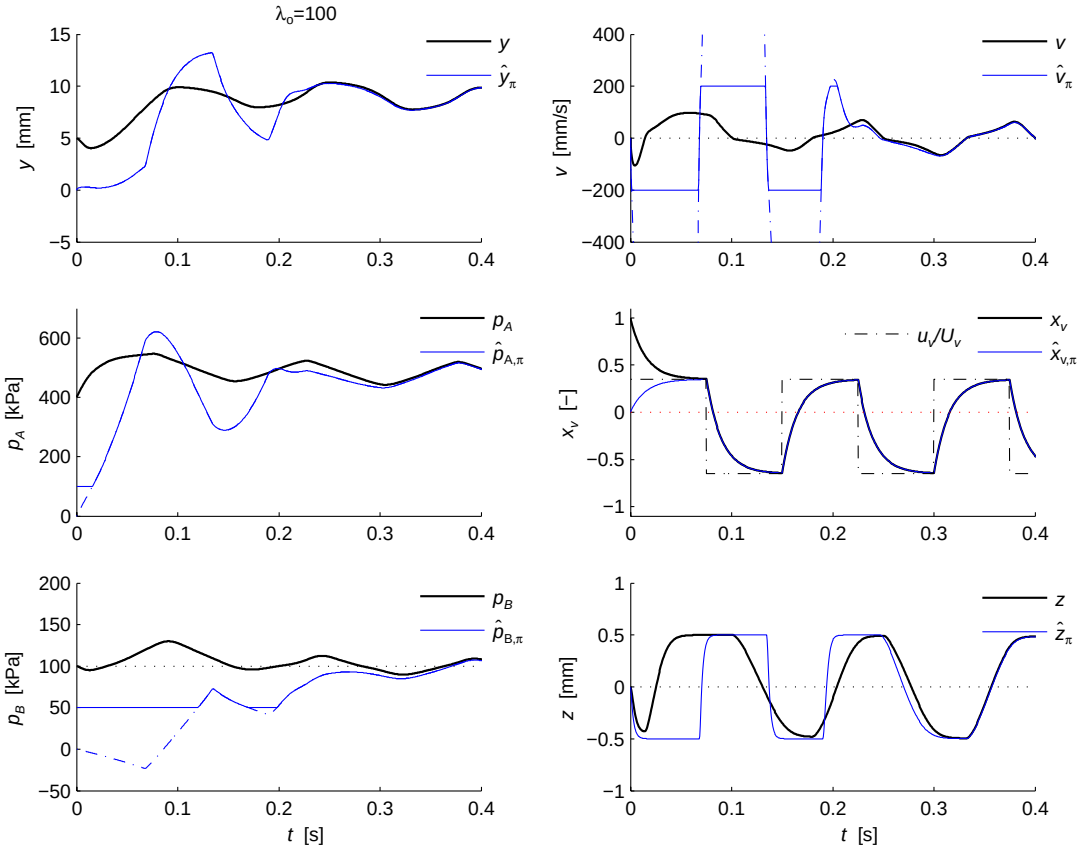

Figure 8.7: Illustration of the global stability properties of the full-order observer with saturation of estimates, but without discontinuous projection. The unsaturated estimates are plotted with dash-dot line $(-.-)$.

Figure 8.8 plots the performance of the observer (8.122) with combined saturation and projection. To illustrate the effect of projection, we simulated the observer with the same conditions as in the simulation plotted in Figure 8.7. The simulation illustrates a dramatical improvement in initial transient performance of the observer utilizing projection of observer states.

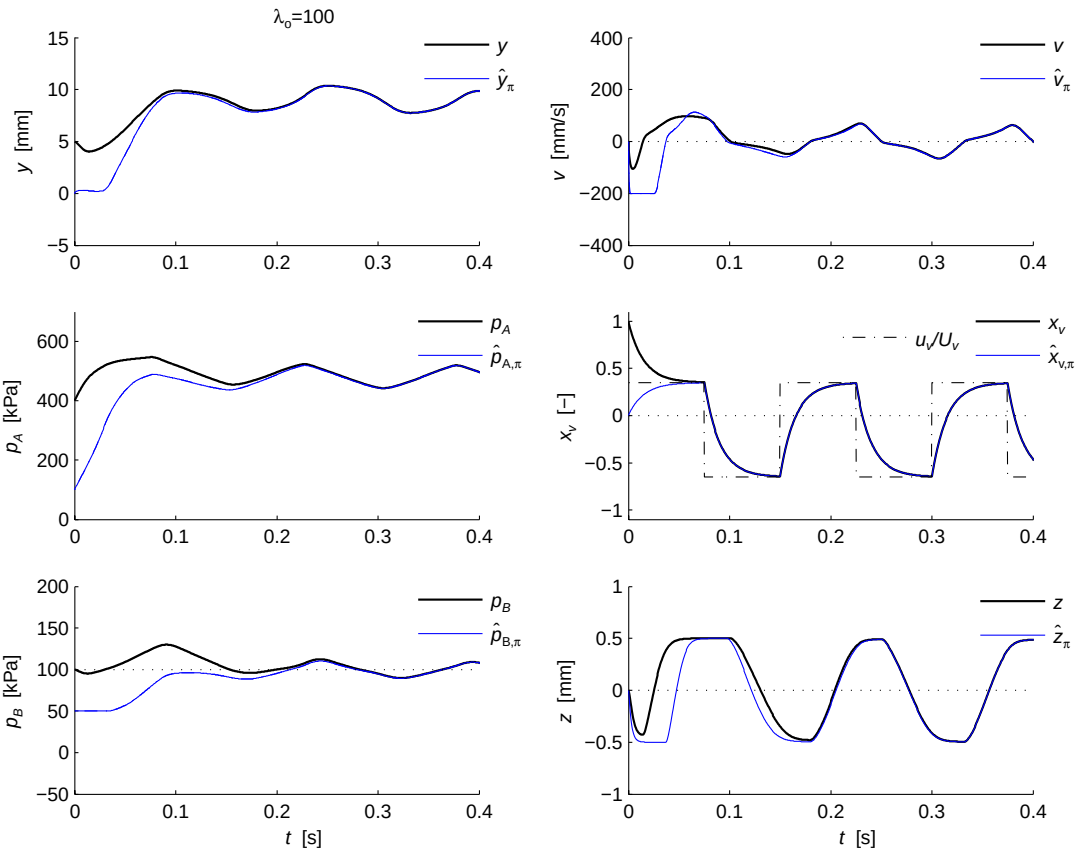

Figure 8.8: Transient performance improvement with discontinuous projection of observer states.

8.5.4 Robust reduced-order observer

In this section we describe the robust redesign of the reduced-order observer (8.82) using smooth saturation and discontinuous projection, as described for the full-order observer in previous section.

Since the friction deflection estimate $\hat{z}$ is bounded according to $\hat{z}(t) \in [-Z_{\max}, Z_{\max}]$ for $\forall \hat{z}(0) \in [-Z_{\max}, Z_{\max}]$, saturation (and projection) of $\hat{z}$ becomes superfluous. Furthermore, by constraining the control input according to $u(t) \in [-1, 1]$, the estimate $\hat{x}_v$ of the valve opening, becomes bounded according to $\hat{x}_v(t) \in [-1, 1]$ for $\forall \hat{x}_v(0) \in [-1, 1]$. We utilize this knowledge to simplify the redesign of the reduced-order observer, by employing saturation and projection only for the estimation of the main states v and p.

The reduced-order observer (8.82), redesigned with saturation of estimates and projection of states, can be expressed as

$$\begin{aligned}
\dot{\hat{\xi}}_1^P &= \mathcal{P}\left(\dot{\hat{\xi}}_1, \hat{v}_\pi^P(y, \hat{\xi}_1^P), \hat{v}_{\text{lb}} + \varepsilon_{\pi,2}, \hat{v}_{\text{ub}} - \varepsilon_{\pi,2}\right) \\
\dot{\hat{\xi}}_2^P &= \mathcal{P}\left(\dot{\hat{\xi}}_2, \hat{p}_\pi^P(y, \hat{\xi}_2^P), \hat{p}_{\text{lb}} + \varepsilon_{\pi,3}, \hat{p}_{\text{ub}} - \varepsilon_{\pi,3}\right) \\
\dot{\hat{x}}_v &= -\frac{1}{\tau_v}\hat{x}_v + \frac{1}{\tau_v}\pi_u(u) \\
\dot{\hat{z}} &= -\frac{K_z}{F_C}\left|\hat{v}_\pi^P\right|_s \hat{z} + \hat{v}_\pi^P,
\end{aligned} \tag{8.123}$$

with

$$\begin{aligned}
\dot{\hat{\xi}}_1 &= \frac{A}{M}\hat{p}_\pi^P - \frac{A}{M}P_0 - \frac{1}{M}f_l(y) - \frac{1}{M}f_f\left(\hat{v}_\pi^P, \hat{z}\right) - k_1\hat{v}_\pi^P \\
\dot{\hat{\xi}}_2 &= \rho_0 T_0 R C_v \cdot \psi_v\left(\hat{p}_\pi^P, \hat{x}_v\right) - k_2\hat{v}_\pi^P,
\end{aligned} \tag{8.124}$$

and where the estimates $\hat{v}_\pi^P$ and $\hat{p}_\pi^P$ are saturated functions of y, and the projected observer states $\hat{\xi}_1^P$ and $\hat{\xi}_2^P$, given according to

$$\begin{aligned}
\hat{v}_\pi^P(y, \hat{\xi}_1^P) &\triangleq \pi_v\left(\hat{\xi}_1^P + k_1 y\right) \\
\hat{p}_\pi^P(y, \hat{\xi}_2^P) &\triangleq \pi_p\left(\frac{1}{V(y)}\left(\hat{\xi}_2^P + k_2 y\right)\right).
\end{aligned} \tag{8.125}$$

In the above equations, the saturation functions $\pi_u(\cdot)$, $\pi_v(\cdot)$ and $\pi_p(\cdot)$ are defined from (7.5) according to

$$\begin{aligned}
\pi_u(u) &\triangleq \pi(u, -1, 1, \varepsilon_{\pi,u}) \\
\pi_v(\hat{v}) &\triangleq \pi(\hat{v}, \hat{v}_{\text{lb}}, \hat{v}_{\text{ub}}, \varepsilon_{\pi,2}) \\
\pi_p(\hat{p}) &\triangleq \pi(\hat{p}, \hat{p}_{\text{lb}}, \hat{p}_{\text{ub}}, \varepsilon_{\pi,3}),
\end{aligned}$$

with the estimated lower and upper bounds on the velocity v and pressure p given by $\hat{v}_{\text{lb}}$, $\hat{v}_{\text{ub}}$, $\hat{p}_{\text{lb}}$ and $\hat{p}_{\text{ub}}$, respectively, and the parameters $\varepsilon_{\pi,u}$, $\varepsilon_{\pi,2}$ and $\varepsilon_{\pi,3}$ are the

smoothing widths of the saturation functions. The scalar projection $\mathcal{P}\left(f, x, x_{\text{lb}}, x_{\text{ub}}\right)$ is defined by (7.10), page 110.

For the redesigned reduced-order observer given by (8.123)–(8.125), the estimates $\mathbf{x}_{u,\pi}^{P} \triangleq [\hat{v}_{\pi}^{P}, \hat{p}_{\pi}^{P}, \hat{x}_{v}, \hat{z}]^{T}$ are constrained to the slightly enlarged region of normal operation,

$$\mathcal{X}_{\varepsilon_{\pi}^{+}} \triangleq \left\{\forall \mathbf{x} \in \mathcal{X}_0 \colon \mathbf{x}_{\text{lb}} - \boldsymbol{\varepsilon}_{\pi} \leq \mathbf{x} \leq \mathbf{x}_{\text{ub}} + \boldsymbol{\varepsilon}_{\pi}\right\},$$

which means that since $\mathcal{X}_{\varepsilon_{\pi}^{+}}$ is a subset of the region of feasibility $\mathcal{X}_0$, the stability properties hold globally.

8.6 Experimental results

The redesigned robust full-order observer with saturation and projection of observer states, given by (8.122), was implemented and tested experimentally on the test rig described in Chapter 2. Results for the reduced-order observer are qualitatively similar, and the reduced-order observer is validated experimentally in Chapter 9 in closed-loop.

The observer gains are computed according to (8.71) on page 131, and the following parameter sets were used to illustrate the performance of the observer:

$$\begin{array}{llll}
\lambda_o = 0: & k_1 = 10\,\text{mm}/\,\text{mm} & k_2 = 0\,\text{mm}/\,(\,\text{mm}\,\text{s}) & k_3 = 0\,\text{kPa}/\,\text{mm} \\
\lambda_o = 100: & k_1 = 50\,\text{mm}/\,\text{mm} & k_2 = 15 \cdot 10^3\,\text{mm}/\,(\,\text{mm}\,\text{s}) & k_3 = 1.0 \cdot 10^3\,\text{kPa}/\,\text{mm} \\
\lambda_o = 300: & k_1 = 650\,\text{mm}/\,\text{mm} & k_2 = 100 \cdot 10^3\,\text{mm}/\,(\,\text{mm}\,\text{s}) & k_3 = 30 \cdot 10^3\,\text{kPa}/\,\text{mm} \\
\lambda_o = 400: & k_1 = 950\,\text{mm}/\,\text{mm} & k_2 = 240 \cdot 10^3\,\text{mm}/\,(\,\text{mm}\,\text{s}) & k_3 = 75 \cdot 10^3\,\text{kPa}/\,\text{mm}.
\end{array}$$

For the given numerical solver algorithm and sample time, the practical range of observer gains are $\lambda_o \in [0, 300]^T$, while for $\lambda_o \geq 400$ the observer becomes sensitive to measurement noise. In all plotted responses, we illustrate the initial convergence of the observer using the initial estimates

$$\hat{\mathbf{x}}\,(0) = \begin{bmatrix} 0\,\text{mm} & 0\,\text{mm}/\,\text{s} & 100\,\text{kPa} & 0 & 100\,\text{kPa} & 0\,\text{mm} \end{bmatrix}^{T}.$$

Like in the preceding simulations, the observer was implemented with the model parameters given by Table 6.1 on page 103, while the parameters $\boldsymbol{\theta}_l$ of the clutch load characteristic were tuned manually to approximately fit the clutch spring on the current test rig, giving

$$\begin{aligned}
\boldsymbol{\theta}_l \;=\; \big[\; & -4.0 \quad -4.0 \quad -4.0 \quad -4.0 \quad -4.0 \quad -4.0 \quad -4.0 \quad -4.0 \quad -4.0 \quad -4.0 \\
& 4.0 \quad 6.1 \quad 6.3 \quad 6.3 \quad 6.2 \quad 6.1 \quad 6.1 \quad 6.1 \quad 6.2 \quad 6.4 \;\big]^{T} \cdot 10^3\,\text{N}.
\end{aligned}$$

Remark 33 *To make the observer design realistic with respect to implementation in an actual application, we have implemented the observer with parameter estimates (except for the load characteristic), based purely on* a priori *knowledge of the actuator and clutch. That is, we have relative accurate estimates of the geometric parameters of the actuator (A_0, A_A, V_{A0}, A_B, V_{B0}), fairly good estimates of the parameters of the valve (C_v, b_v, C_r, b_r, τ_v, K_v, U_{v0}) and the physical pressures (P_0, P_S), but*

only rough estimates of the friction parameters (D_v, F_C, K_z, $D_{\dot{z}}$) and the lumped mass of the actuator, release bearing and clutch (M). The parameters of the load characteristic were tuned manually to fit the current application. This was required in order to obtain good estimates of the unmeasured states because of the strong influence of the load characteristic on the actuator dynamics.

8.6.1 General observer performance

We illustrate the general estimation performance of the observer using the sine wave input

$$u_v = U_{v0} + U_{v1} \sin\left(T^{-1}t\right),$$

with bias $U_{v0} = -1.5\,\mathrm{V}$, amplitude $U_{v1} = 3.5\,\mathrm{V}$, and period $T = 1.0\,\mathrm{s}$. This produces a steady periodic response of the actuator in the range $y\,(t) \in [2, 7]\,\mathrm{mm}$, which encompasses the grip point of the clutch and also the region with the strongest nonlinearity in the load characteristic (see Figure 3.1, page 45). Since the plotted experiment includes operation in the grip point of the clutch, it illustrates the performance of the observer in the region where good observer performance is the most important. Figures 8.9–8.11, illustrate the performance of the observer with increasing observer gains.

In Figure 8.9, the observer is implemented with $\lambda_o = 0$, where practically no feedback correction is used in the observer, only a low correction in the output $\hat{y}$ to avoid divergence due to the unstable integrator of the position. Hence, the figure can be taken as the performance of a pure simulation of the model (8.1), which also can be taken as a good indication of the performance of the open-loop observer (8.4). Like the simulations, the experiment demonstrates the strong stability of the actuator, seen from the fast convergence of the estimates. The figure also reveals model errors for small actuator positions and low velocities, which manifest themselves as an offset in the estimated position and poor velocity estimates. These errors are most likely due to poorly tuned parameters of the load characteristic in this region[4], and poor estimates of the parameters of the friction model.

Figure 8.10 demonstrates the disturbance attenuating effect of the output-injection terms in the observer, where we see that with a moderate observer bandwidth of $\lambda_o = 100$, the observer provides estimates with good accuracy.

Figure 8.11, demonstrates the observer performance with a relatively high observer design bandwidth of $\lambda_o = 300$. The observer provides highly accurate estimates of the position and velocity, and also the pressure of chamber B, while there is a slight offset in the estimated pressure in chamber A. This fortifies the assumption of parameter errors in the load characteristic and friction model.

With too high observer gains, the observer is prone to amplify measurement noise. This undesirable property is demonstrated in Figure 8.12 with observer design bandwidth $\lambda_o = 400$.

[4]The parameters of the load characteristic were tuned manually, using the measured pressure force $f_p = A_0 P_0 + A_A p_A - A_B p_B$ as an indication of goodness of fit. Consequently, the estimates are only approximate, and can not rival the online estimation of the adaptive controller.

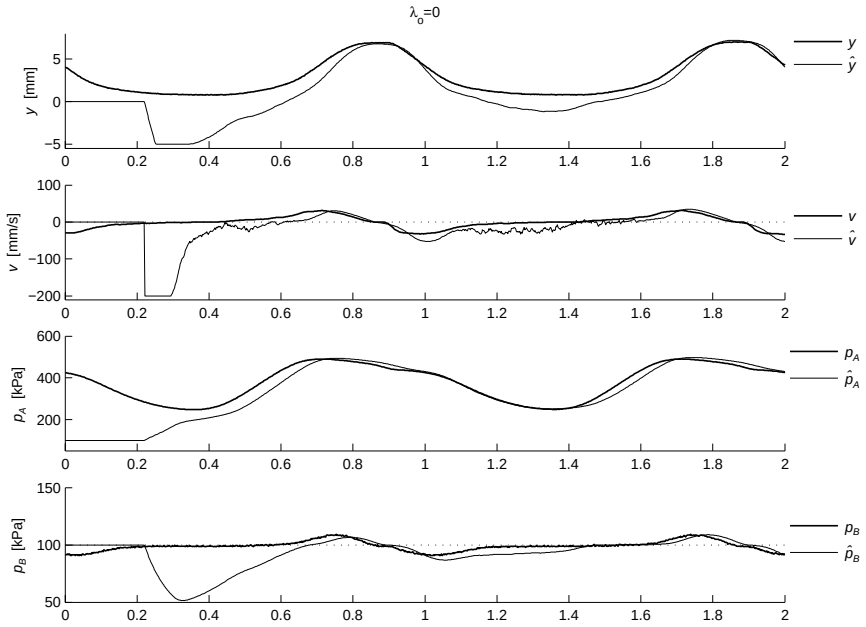

Figure 8.9: Measured and estimated observer states with practically no feedback correction in the observer ($\lambda_o = 0$).

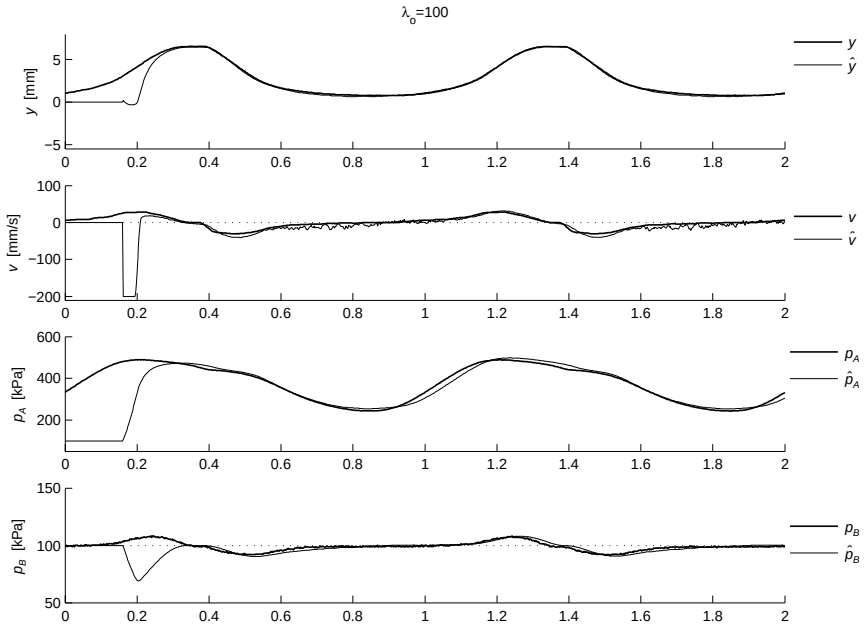

Figure 8.10: Measured and estimated observer states with observer feedback gains according to an average design bandwidth $\lambda_o = 100$.

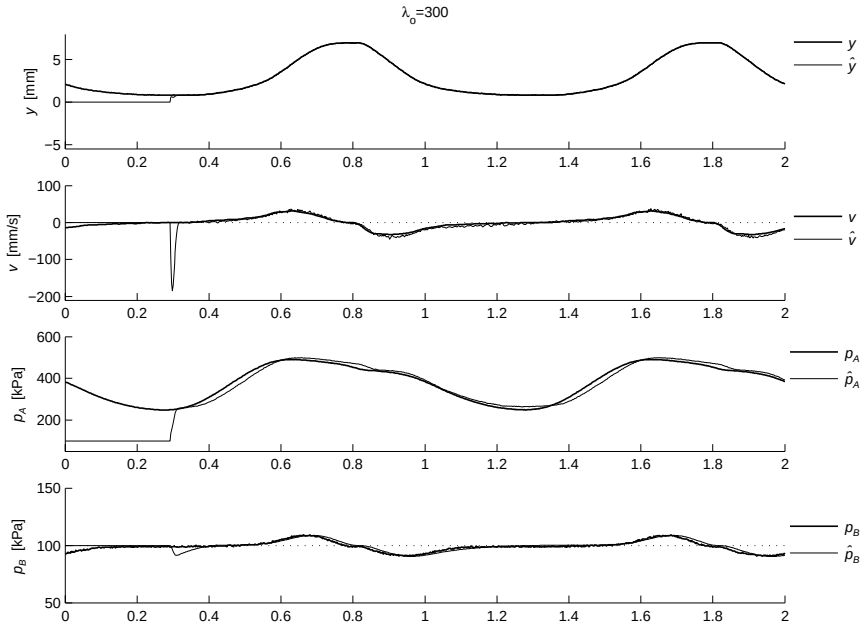

Figure 8.11: Measured and estimated observer states with observer feedback gains according to a relatively high design bandwidth $\lambda_o = 300$.

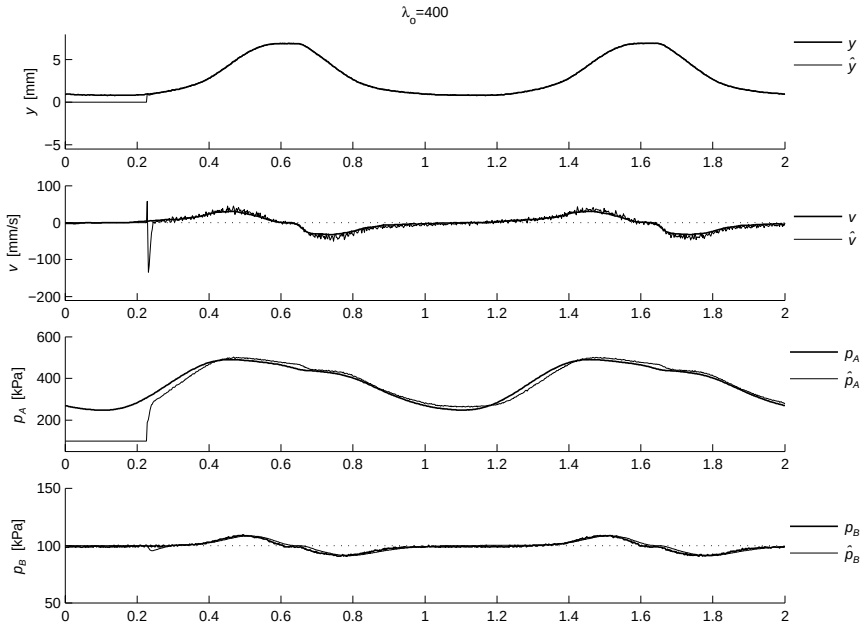

Figure 8.12: Noise amplification in the estimates for high observer gain ($\lambda_o = 400$).

8.6.2 Disturbance attenuation

In the following two subsections illustrate the performance of the observer when subjected to disturbances. We consider two types of disturbances: parametric errors in the modeled clutch load characteristic, and the excitation of unmodeled dynamics in the proportional valve and supply pressure.

Detuned parameters of the load characteristic

Figures 8.13 and 8.14 illustrate the performance of the observer with the parameters of the load characteristic corresponding to the load characteristic of a worn clutch.

Figure 8.13 illustrates the significance of an accurate load characteristic for good performance of the observer. With detuned parameters of the load characteristic, the observer without feedback correction ($\lambda_o = 0$) provides highly inaccurate estimates of all states. The poor response of the detuned observer, emphasizes the usefulness of adaptation of the load characteristic in order to obtain good performance of the observer.

Figure 8.14, comparing with Figure 8.13, illustrates that the performance of the observer is improved by increasing the observer gain. In the figure, the performance for $\lambda_o = 100$ is plotted. The output is approximately recovered, while the estimated velocity is improved, but still with significant estimation errors. Further increasing the observer bandwidth (λ_o), increases the accuracy of the estimated position and velocity, and also the pressure of chamber B. However, due to the errors in the load characteristic, the estimated pressure of chamber A is impossible to improve with increasing gain.

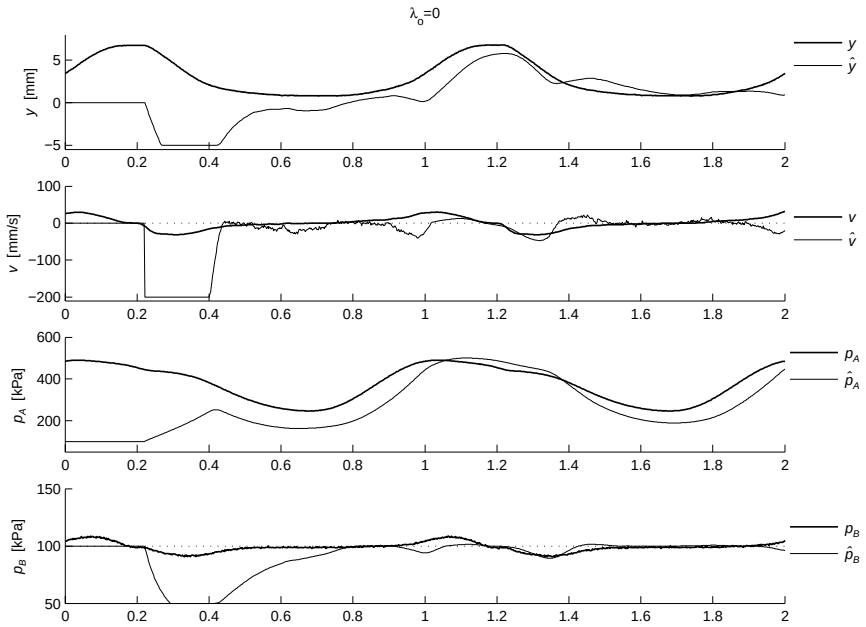

Figure 8.13: Observer performance for $\lambda_o = 0$ with a detuned load characteristic corresponding to a worn clutch.

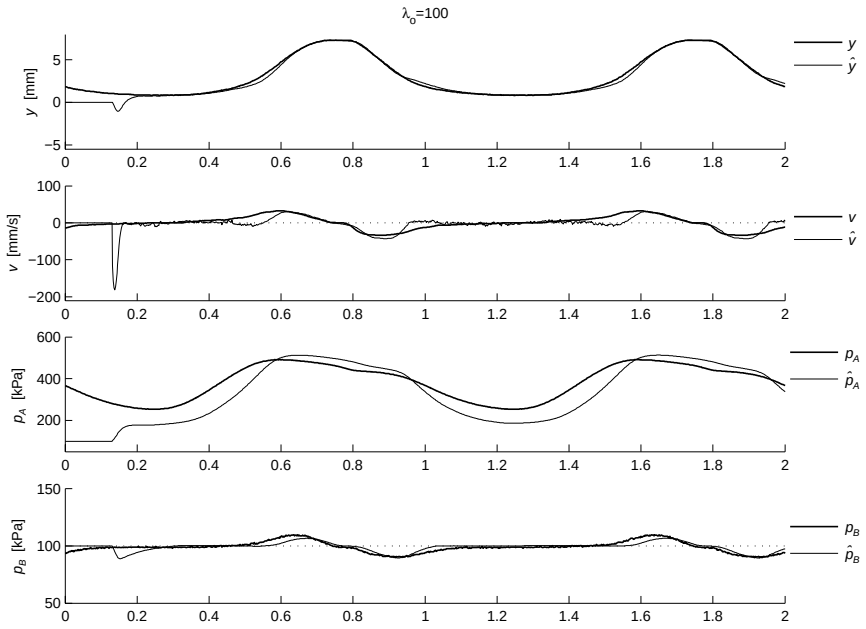

Figure 8.14: Observer performance for $\lambda_o = 100$ with a detuned load characteristic.

Unmodeled dynamics

High-frequency control inputs with high amplitudes tend to excite unmodeled dynamics of the proportional valve. In addition, high-amplitude control inputs result in significant pressure drops in the supply pressure (modeled as a constant pressure, given by P_S), which also can be viewed as unmodeled dynamics. This undesirable phenomena is clearly exhibited for the square wave valve input

$$u_v = U_{v0} + U_{v1} \operatorname{sgn}\left(\sin\left(T^{-1}t\right)\right),$$

with bias $U_{v0} = -2.5\,\mathrm{V}$, amplitude $U_{v1} = 4.5\,\mathrm{V}$, and period $T = 1.0\,\mathrm{s}$. This high-amplitude input produces a steady periodic response of the actuator in the range $y(t) \in [1, 14]\,\mathrm{mm}$, covering the entire region of the actuator in which tracking control is required. Principally, the step changes in the square wave represent control inputs with infinitely high frequency. The square wave input and the corresponding pressure drop in the supply pressure is plotted in Figure 8.15.

The response of the observer for $\lambda_o = 0$ is plotted in Figure 8.16. The excitation of the unmodeled dynamics is apparent by viewing the pressure response immediately after a step change in the input. The resulting pressure peak is caused by an overshoot in the positioning of the valve spool, which occur for high-amplitude step inputs. This overshooting behavior of the valve assumably due to a high proportional (and possibly derivative) feedback in the positioning loop of the spool, combined with slow integral action.

Figure 8.17 illustrates the disturbance attenuating effect of feedback correction in the observer for moderate gains ($\lambda_o = 100$). Increasing the gain, further improves es-

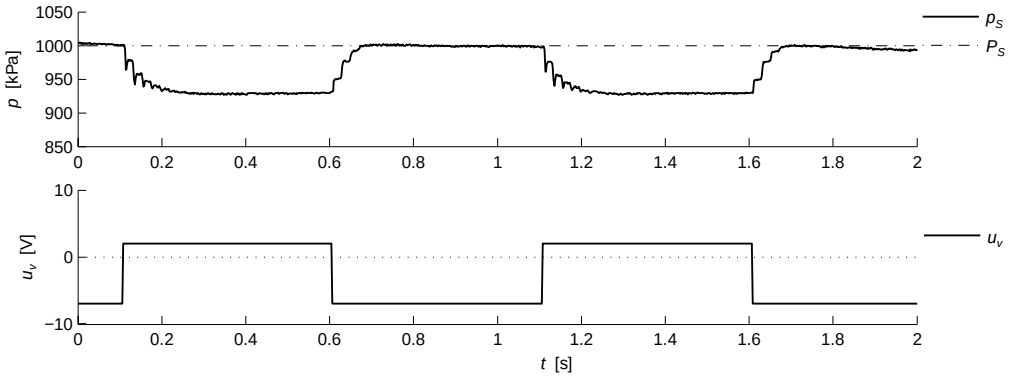

Figure 8.15: The square wave control input and the corresponding drop in the supply pressure due to high air flow.

timates of the position, velocity and pressure of chamber B. However, the estimated pressure of chamber A can not be further improved by increased observer gain because the error still remaining in Figure 8.17 is caused by errors in the open-loop estimate of the spool position x_v.

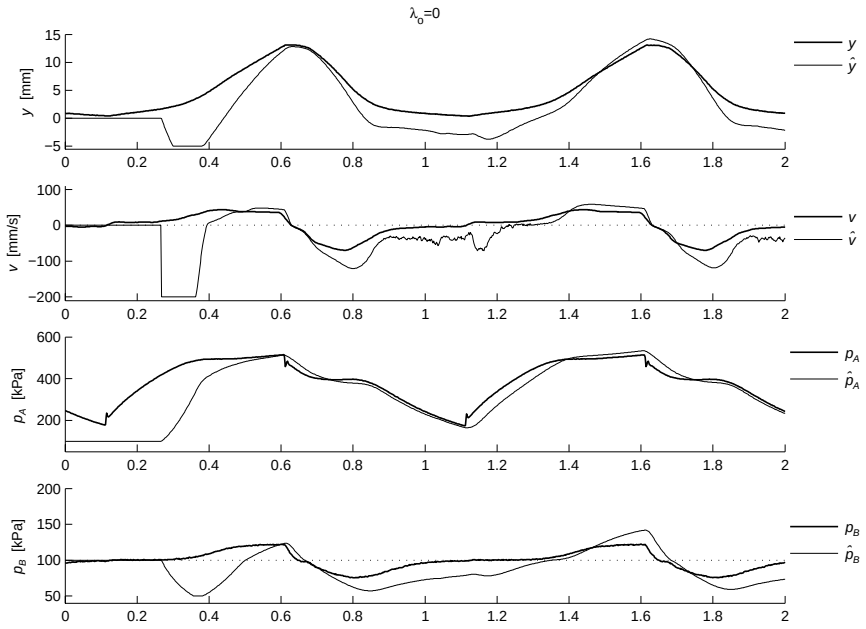

Figure 8.16: Observer performance with $\lambda_o = 0$, illustrating the deterioriating effect of excitation of unmodelled dynamics with a high-amplitude square wave control input.

8.7 Summary

In this chapter we have analyzed the problem of nonlinear observer design for estimating the unmeasured states of the electro-pneumatic actuator, in the case when only the position y is measured. We first analyze the use of the model as an open-loop nonlinear observer for the unmeasured states. The stability properties of this observer, thus also the dynamics of the unmeasured states, is shown to be asymptotically stable. This again establishes a detectability property of the system, and shows that the pneumatic actuator has some inherent stability properties which enables the design of simple nonlinear observers.

Based on these results, two simple nonlinear observers for the electro-pneumatic clutch actuator are proposed, both compatible with output-feedback control by an observer-based backstepping approach: a full-order observer and a simplified reduced-order observer. Both observers combine closed-loop estimation using linear output-injection of the main states, with open-loop estimation of the remaining states. The observers are asymptotically stable and robust to bounded disturbances, as long as the estimated actuator states remain within the region of validity of the model.

As a robust redesign of the observers, smooth saturation of the state estimates is introduced to constrain the observer dynamics to the feasible region of the state-space, by which we are able to guarantee global uniform stability properties even if the unsaturated observer states enter the non-feasible region in state-space during initial transients. The differentiability of the introduced smooth saturation ensures

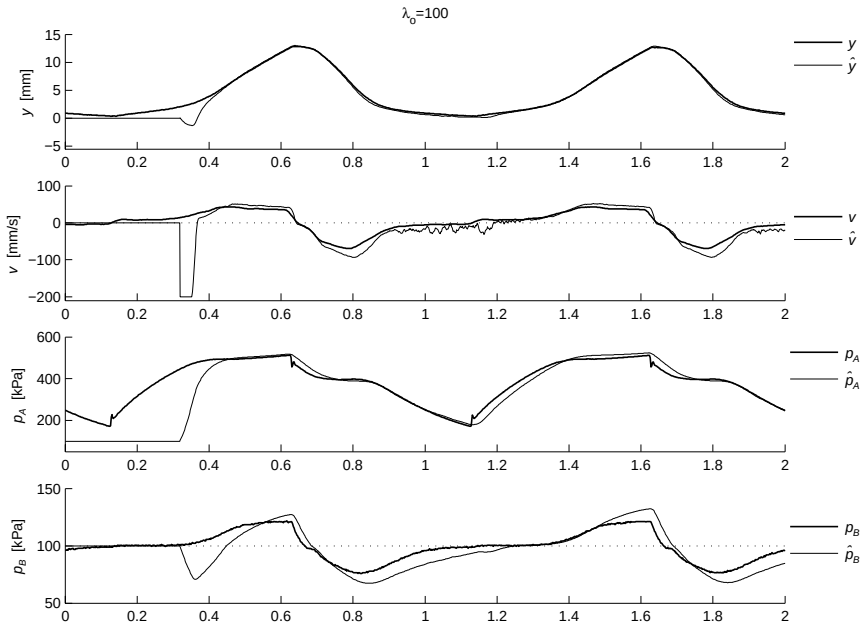

Figure 8.17: Observer performance for $\lambda_o = 100$, illustrating the disturbance attenuating effect of the observer when subjected to excitation of unmodeled valve and supply pressure dynamics.

compatibility with observer backstepping. We further improve initial transients of the observer by projecting its non-saturated observer states by discontinuous projection such that the discontinuous projection is only active for estimates which are fully saturated, thus, preserving the smoothness of the saturated estimates. Hence, a control law can be designed by a backstepping approach using the observer with saturated estimates, and then implemented using the observer with combined saturation and projection, without introducing discontinuities in the control input, thus, preserving the stability properties of the closed-loop system.

The performance of the observers are validated by simulations, and experimentally on the test rig.

Chapter 9

Nonlinear Output-feedback Control

In this chapter, we address the design of an output-feedback tracking controller for the electro-pneumatic clutch actuation system. Using the nonlinear reduced-order observer introduced in Chapter 8.4, we present a robust output-feedback design by an observer backstepping approach, which is a recursive procedure performed in four steps. We further present an approximate backstepping design, where we simplify the last two steps of the design by using high-gain observers to estimate, rather than calculate analytically, the derivative of the stabilizing function designed in the previous steps. The performance of the controller is analyzed by simulations and experimentally on the test rig.

In Section 9.1, we describe the reference model which is used to generate a smooth reference trajectory from an arbitrary reference input. Next, we present the observer backstepping design in detail in Section 9.2, and give an overview of the implemented output-feedback controller in Section 9.3. We present the simulation results in Section 9.4, and the experimental results in Section 9.5. Finally, the chapter is summarized in Section 9.6.

9.1 Reference model

For tracking control applications, the objective is to track a reference input r by the output y. It is convenient to design a controller for the alternative objective of tracking the output y_r of a linear reference model in the form

$$\dot{\mathbf{z}}_r = \begin{bmatrix} 0 & & \\ \vdots & & \mathbf{I}_{n-1} \\ 0 & & \\ -m_0 & \cdots & -m_{n-1} \end{bmatrix} \mathbf{z}_r + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ m_0 \end{bmatrix} r, \tag{9.1}$$

where r is the actual reference, $\mathbf{z}_r = [y_r, \dot{y}_r, \cdots, y_r^{(n-1)}]^T$ is the state vector of the filter, and $\mathbf{I}_{n-1} \in \mathbb{R}^{(n-1)\times(n-1)}$ is the identity matrix. The characteristic polynomial

$s^n + m_{n-1}s^{n-1} + \cdots + m_1 s + m_0$ of the filter is chosen to be Hurwitz, so that the filter dynamics is exponentially stable. For most tracking tasks, a good choice of the parameters of the reference model is obtained by placing the multiple (n) poles at $s = -\lambda_r$, resulting in a critically damped dynamics with time constant $\tau_r = n/\lambda_r$. This dynamics may be regarded as the "best conditioned" among the linear dynamics for a given bandwidth.

Using this type of reference model to generate the tracking trajectory has several advantages, outlined below:

- The tracking reference trajectory y_r becomes smooth.
- Noise and discontinuities in the reference r is filtered out.
- The derivatives of y_r, which is required by a Lyapunov-based tracking design, becomes available as states in $\mathbf{z}_r$.
- The reference model can be used to reduce initial transients by *trajectory initialization* (see [52, Section 4.3.2]).
- The use of a reference model facilitates the incorporation of a desired performance specification for the closed-loop controller. That is, for a properly designed tracking controller, which achieves tracking of the reference $y_r(t)$, the closed-loop dynamical tracking properties are in essence determined by the properties of the reference model.

Reference model for the electro-pneumatic actuator

For a system with relative degree ρ, a backstepping tracking design requires the first $(\rho + 1)$ derivatives of the tracking trajectory to be available for control. For the electro-pneumatic actuator, the relative degree is $\rho = 4$, thus, it is convenient to use a 5th-order reference model such that the state vector contains all the required derivatives, *i.e.*,

$$\mathbf{z}_r = \begin{bmatrix} z_{r1} \\ z_{r2} \\ z_{r3} \\ z_{r4} \\ z_{r5} \end{bmatrix} \triangleq \begin{bmatrix} y_r \\ \dot{y}_r \\ \ddot{y}_r \\ y_r^{(3)} \\ y_r^{(4)} \end{bmatrix}.$$

The reference model (9.1) is conveniently expressed in the compact form

$$\dot{\mathbf{z}}_r = \mathbf{A}_r \mathbf{z}_r + \mathbf{c}_r r, \tag{9.2}$$

where in the 5th-order case, the matrix $\mathbf{A}_r$ and the vector $\mathbf{c}_r$ are

$$\mathbf{A}_r \triangleq \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ -m_0 & -m_1 & -m_2 & -m_3 & -m_4 \end{bmatrix}, \qquad \mathbf{c}_r \triangleq \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ m_0 \end{bmatrix}. \tag{9.3}$$

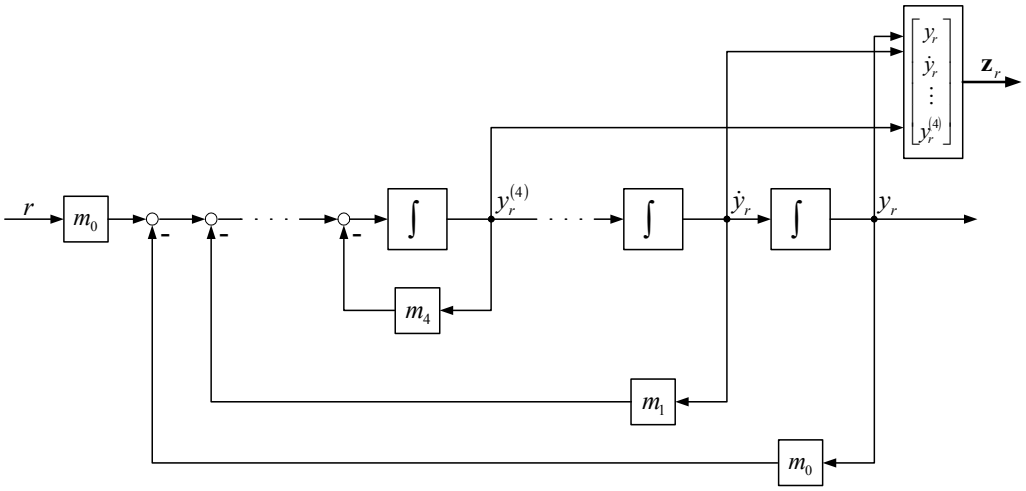

Figure 9.1: Block diagram of the reference model.

A block diagram of this 5th-order reference model (9.1), is given in Figure 9.1.

The characteristic polynomial of the filter becomes

$$|s\mathbf{I} - \mathbf{A}_r| = s^5 + m_4 s^4 + m_3 s^3 + m_2 s^2 + m_1 s + m_0. \tag{9.4}$$

Placing all the poles of the reference model at $s = -\lambda_r$ amounts to choosing the filter coefficients which results in the characteristic polynomial

$$(s + \lambda_r)^5 = s^5 + 5\lambda_r s^4 + 10\lambda_r^2 s^3 + 10\lambda_r^3 s^2 + 5\lambda_r^4 s + \lambda_r^5. \tag{9.5}$$

Comparing (9.4) with (9.5), we can express the filter coefficients as functions of λ_r, giving

$$\begin{aligned} m_0 &= \lambda_r^5 \\ m_1 &= 5\lambda_r^4 \\ m_2 &= 10\lambda_r^3 \\ m_3 &= 10\lambda_r^2 \\ m_4 &= 5\lambda_r. \end{aligned}$$

We refer to λ_r as the bandwidth of the reference model, where in this 5th-order case, the corresponding time-constant is

$$\tau_r = \frac{5}{\lambda_r}.$$

9.2 Observer backstepping design

System

We consider the electro-pneumatic clutch actuator given by the model

$$\begin{aligned}
\dot{y} &= v \\
\dot{v} &= \tfrac{A}{M}p - \tfrac{A}{M}P_0 - \tfrac{1}{M}f_l(y) - \tfrac{1}{M}f_f(v,z) \\
\dot{p} &= -A\tfrac{1}{V(y)}vp + \rho_0 T_0 R C_v \tfrac{1}{V(y)}\psi_v(p,x_v) \\
\dot{x}_v &= -\tfrac{1}{\tau_v}x_v + \tfrac{1}{\tau_v}u \\
\dot{z} &= v - \tfrac{K_z}{F_C}|v|_s z,
\end{aligned} \tag{9.6}$$

where y, v, p, x_v, and z are the actuator position, velocity, chamber pressure, valve spool opening and friction deflection state, respectively; while $f_l(y)$, $f_f(v,z)$, $V(y)$ and $\psi_v(p,x_v)$, are known functions of the clutch load spring force, friction force, chamber volume and the valve flow function.

The *region of validity* (or *region of feasibility*), of (9.6) is the set $\mathcal{X}_0 \subset \mathbb{R}^5$ defined by

$$\mathcal{X}_0 \triangleq \{\forall \mathbf{x} : \mathbf{x}_{\min} \leq \mathbf{x} \leq \mathbf{x}_{\max}\}, \tag{9.7}$$

where the minimum and maximum feasible actuator states are given by

$$\begin{aligned}
\mathbf{x}_{\min} &= [y_{\min}, v_{\min}, p_{\min}, x_{v,\min}, z_{\min}]^T \\
\mathbf{x}_{\max} &= [y_{\max}, v_{\max}, p_{\max}, x_{v,\max}, z_{\max}]^T,
\end{aligned}$$

as described in Chapter 6. Moreover, the *region of normal operation,* in which the actuator states will remain under normal operation, is defined by the lower and upper bounds

$$\begin{aligned}
\mathbf{x}_{\text{lb}} &\triangleq [y_{\text{lb}}, v_{\text{lb}}, p_{\text{lb}}, x_{v\,\text{lb}}, z_{\text{lb}}]^T \\
\mathbf{x}_{\text{ub}} &\triangleq [y_{\text{ub}}, v_{\text{ub}}, p_{\text{ub}}, x_{v\,\text{ub}}, z_{\text{ub}}]^T,
\end{aligned}$$

as

$$\mathcal{X} \triangleq \{\forall \mathbf{x} \in \mathcal{X}_0 \colon \mathbf{x}_{\text{lb}} \leq \mathbf{x} \leq \mathbf{x}_{\text{ub}}\}. \tag{9.8}$$

Problem formulation

The objective is to design an output-feedback controller that achieves asymptotic (practical) tracking of the reference y_r by the output y, while keeping all the states of the closed-loop system bounded.

Observer

The basis for the output-feedback control design is the reduced-order observer, given by (8.82), presented in Chapter 8. The observer is given as

$$\begin{aligned}
\dot{\hat{\xi}}_1 &= \tfrac{A}{M}\hat{p}\left(y,\hat{\xi}_2\right) - \tfrac{A}{M}P_0 - \tfrac{1}{M}f_l(y) - \tfrac{1}{M}f_f\left(\hat{v}(y,\hat{\xi}_1),\hat{z}\right) - k_1 \cdot \hat{v}(y,\hat{\xi}_1) \\
\dot{\hat{\xi}}_2 &= \rho_0 T_0 R C_v \cdot \psi_v\left(\hat{p}(y,\hat{\xi}_2),\hat{x}_v\right) - k_2 \cdot \hat{v}(y,\hat{\xi}_1) \\
\dot{\hat{x}}_v &= -\tfrac{1}{\tau_v}\hat{x}_v + \tfrac{1}{\tau_v}u \\
\dot{\hat{z}} &= \hat{v} - \tfrac{K_z}{F_C}|\hat{v}|_s \hat{z},
\end{aligned} \tag{9.9}$$

where $\hat{v}$ and $\hat{p}$ are functions of y, $\hat{\xi}_1$, and $\hat{\xi}_2$ according to

$$\begin{aligned}\hat{v}(y,\hat{\xi}_1) &= \hat{\xi}_1 + k_1 y \\ \hat{p}(y,\hat{\xi}_2) &= \frac{1}{V(y)}\left(\hat{\xi}_2 + k_2 y\right),\end{aligned}$$

and not states in the implemented observer.

In preparation for our backstepping design, we rewrite the actuator dynamics in terms of the observer velocity and the corresponding estimation error, according to

$$\begin{aligned}\dot{y} &= \hat{v} + \tilde{v} \\ &= \hat{\xi}_1 + k_1 y + \tilde{v}.\end{aligned}$$

Utilizing the observer (9.9), we apply backstepping to the system

$$\begin{aligned}&\dot{y} = \hat{\xi}_1 + k_1 y + \tilde{v} \\ &\dot{\hat{\xi}}_1 = \tfrac{A}{M}\tfrac{1}{V(y)}\hat{\xi}_2 + \tfrac{A}{M}\tfrac{1}{V(y)}k_2 y - \tfrac{A}{M}P_0 - \tfrac{1}{M}f_l(y) - \tfrac{1}{M}f_f(\hat{v},\hat{z}) - k_1\cdot\hat{v}(y,\hat{\xi}_1) \\ &\dot{\hat{\xi}}_2 = \rho_0 T_0 R C_v\cdot\psi_v\left(\hat{p}(y,\hat{\xi}_2),\hat{x}_v\right) - k_2\cdot\hat{v}(y,\hat{\xi}_1) \\ &\dot{\hat{x}}_v = -\tfrac{1}{\tau_v}\hat{x}_v + \tfrac{1}{\tau_v}u \\ &\dot{\hat{z}} = \hat{v} - \tfrac{K_z}{F_C}\left|\hat{v}\right|_s\hat{z},\end{aligned} \tag{9.10}$$

with

$$\begin{aligned}\hat{v}(y,\hat{\xi}_1) &= \hat{\xi}_1 + k_1 y \\ \hat{p}(y,\hat{\xi}_2) &= \frac{1}{V(y)}\left(\hat{\xi}_2 + k_2 y\right).\end{aligned}$$

Denoting

$$b_1(y) \triangleq \tfrac{A}{M}\tfrac{1}{V(y)}, \quad b_2 \triangleq \rho_0 T_0 R C_v, \quad b_3 \triangleq \tfrac{1}{\tau_v}$$

and re-defining functions in terms of the observer states according to

$$\psi_2\left(y,\hat{\xi}_2,\hat{x}_v\right) \triangleq \psi_v\left(\hat{p}(y,\hat{\xi}_2),\hat{x}_v\right) \quad \bar{f}_f\left(y,\hat{\xi}_1,\hat{z}\right) \triangleq f_f\left(\hat{v}(y,\hat{\xi}_1),\hat{z}\right),$$

the system can be written in the form

$$\begin{aligned}&\dot{y} = \hat{\xi}_1 + k_1 y + \tilde{\xi}_1 \\ &\dot{\hat{\xi}}_1 = b_1(y)\cdot\hat{\xi}_2 + b_1(y)\cdot k_2 y - \tfrac{A}{M}P_0 - \tfrac{1}{M}f_l(y) - \tfrac{1}{M}\bar{f}_f\left(y,\hat{\xi}_1,\hat{z}\right) - k_1\hat{\xi}_1 - k_1^2 y \\ &\dot{\hat{\xi}}_2 = b_2\cdot\psi_2\left(y,\hat{\xi}_2,\hat{x}_v\right) - k_2\hat{\xi}_1 - k_1 k_2 y \\ &\dot{\hat{x}}_v = b_3 u - b_3\hat{x}_v \\ &\dot{\hat{z}} = \hat{\xi}_1 + k_1 y - \tfrac{K_z}{F_C}\left|\hat{v}(y,\hat{\xi}_1)\right|_s\hat{z}.\end{aligned} \tag{9.11}$$

The system is in *pure feedback form*, hence, suitable for a nonlinear control design based on an integrator backstepping approach.

9.2.1 Exact backstepping

The backstepping design is performed recursively in four steps on the system (9.11). We illustrate the design procedure in detail in the following section, and summarizes the controller and its properties in the section following that.

Controller development

Design – Step 1 First define the tracking error

$$e_1 \triangleq y - y_r. \tag{9.12}$$

The time-derivative of e_1 becomes

$$\begin{aligned} \dot{e}_1 &= \dot{y} - \dot{y}_r \\ &= \hat{\xi}_1 + k_1 y + \tilde{\xi}_1 - \dot{y}_r. \end{aligned}$$

Assuming $u_1 \triangleq \hat{\xi}_1$ is the actual control, we design a control law for u_1 which renders the error dynamics

$$\dot{e}_1 = u_1 - \dot{y}_r + k_1 y + \tilde{\xi}_1 \tag{9.13}$$

exponentially stable for $\tilde{\xi}_1 = 0$, and input-to-state stable with respect to the observer error $\tilde{\xi}_1$. We take a control law in the form

$$u_1 = \dot{y}_r + \alpha_1 (y, y_r), \tag{9.14}$$

where $\dot{y}_r$ is a feedforward tracking term (which compensates for a time-varying reference), and α_1 is a stabilizing function which is designed to stabilize a zero tracking error, $e_1 = 0$.

For conformity with the following design steps, we illustrate the design of the stabilizing function α_1 using a Control Lyapunov Function (CLF). We use the CLF

$$V_1 = \frac{1}{2} e_1^2, \tag{9.15}$$

whose time-derivative becomes

$$\begin{aligned} \dot{V}_1 &= e_1 \dot{e}_1 \\ &= e_1 \left(u_1 - \dot{y}_r + k_1 y + \tilde{\xi}_1 \right) \\ &= e_1 (\alpha_1 + k_1 y) + e_1 \tilde{\xi}_1. \end{aligned}$$

We choose the stabilizing function

$$\alpha_1 = -c_1 e_1 - k_1 y, \tag{9.16}$$

which gives

$$\dot{V}_1 = -c_1 e_1^2 + e_1 \tilde{\xi}_1,$$

where $c_1 > 0$ is a design constant.

Note that for zero observer error, the time-derivative of V_1 becomes

$$\tilde{\xi}_1 = 0 \Longrightarrow \dot{V}_1 = -c_1 e_1^2,$$

and that

$$|e_1| > \frac{1}{c_1}\left|\tilde{\xi}_1\right| \Longrightarrow \dot{V}_1 < 0.$$

This proves that for the system (9.13), the control $u_1 = \dot{y}_r + \alpha_1$ with the stabilizing function α_1 given by (9.16) makes the error state e_1 ISS with respect to the input $\tilde{\xi}_1$, and exponentially stable for zero observer error, $\tilde{\xi}_1 = 0$.

Design – Step 2 Taking into consideration that u_1 is merely a *virtual control* that we are not able to control directly, the control law designed at step 1 is taken as the desired control law

$$u_{1d} = \dot{y}_r + \alpha_1\left(y, y_r\right),$$

and we define the corresponding error

$$e_2 \triangleq u_1 - u_{1d} = u_1 - \dot{y}_r - \alpha_1. \tag{9.17}$$

This error appears in the e_1–dynamics by substituting $u_1 = e_2 + \dot{y}_r + \alpha_1$ into (9.13), giving

$$\begin{aligned} \dot{e}_1 &= e_2 + \alpha_1 + k_1 y + \tilde{\xi}_1 \\ &= -c_1 e_1 + e_2 + \tilde{\xi}_1. \end{aligned}$$

To deal with this error, we backstep it through the first integrator by differentiation, and obtain

$$\begin{aligned} \dot{e}_2 &= \dot{u}_1 - \ddot{y}_r - \dot{\alpha}_1 \\ &= \dot{\hat{\xi}}_1 - \ddot{y}_r - \frac{\partial \alpha_1}{\partial y}\dot{y} - \frac{\partial \alpha_1}{\partial y_r}\dot{y}_r. \end{aligned}$$

Note that the time-derivatives of $\hat{\xi}_1$ and y are given by (9.11), while the partial derivatives of α_1 is obtained upon differentiating (9.16), giving

$$\frac{\partial \alpha_1}{\partial y} = -\left(c_1 + k_1\right), \quad \frac{\partial \alpha_1}{\partial y_r} = c_1.$$

The e_2–dynamics then becomes

$$\begin{aligned} \dot{e}_2 &= b_1\hat{\xi}_2 + b_1 k_2 y - \frac{A}{M}P_0 - \frac{1}{M}f_l - \frac{1}{M}f_f - k_1\hat{\xi}_1 - k_1^2 y \\ &\quad -\ddot{y}_r + \left(c_1 + k_1\right)\left(\hat{\xi}_1 + k_1 y + \tilde{\xi}_1\right) - c_1\dot{y}_r \\ &= b_1\hat{\xi}_2 - \ddot{y}_r + b_1 k_2 y - \frac{A}{M}P_0 - \frac{1}{M}f_l - \frac{1}{M}f_f + c_1\left(\hat{\xi}_1 + k_1 y\right) \\ &\quad -c_1\dot{y}_r + \left(c_1 + k_1\right)\tilde{\xi}_1. \end{aligned}$$

Denoting

$$\begin{aligned} u_2\left(y,\hat{\xi}_2\right) &\triangleq b_1(y)\cdot\hat{\xi}_2 \\ \Sigma_2\left(y,\hat{\xi}_1,\hat{z},\dot{y}_r\right) &\triangleq b_1(y)\cdot k_2 y - \frac{A}{M}P_0 - \frac{1}{M}f_l(y) - \frac{1}{M}\bar{f}_f\left(y,\hat{\xi}_1,\hat{z}\right) \\ &\quad + c_1\left(\hat{\xi}_1 + k_1 y\right) - c_1\dot{y}_r \\ w_2 &\triangleq c_1 + k_1, \end{aligned}$$

the dynamics of the (e_1, e_2)–subsystem can be written

$$\begin{aligned} \dot{e}_1 &= -c_1 e_1 + e_2 + \tilde{\xi}_1 \\ \dot{e}_2 &= u_2 - \ddot{y}_r + \Sigma_2 + w_2\tilde{\xi}_1. \end{aligned} \tag{9.18}$$

Assuming $u_2 \triangleq b_1(y)\cdot\hat{\xi}_2$ is the actual control, we use a control in the form

$$u_2 = \ddot{y}_r + \alpha_2,$$

and design a stabilizing function α_2 which makes the system exponentially stable for $\tilde{\xi}_1 = 0$, and ISS with respect to $\tilde{\xi}_1$. Like in Step 1, the stabilizing function α_2 is obtained through a Lyapunov-based design, using a control Lyapunov function for the system (9.18). We use the CLF

$$V_2 = V_1 + \frac{1}{2}e_2^2, \tag{9.19}$$

whose derivative becomes

$$\begin{aligned} \dot{V}_2 &= e_1\dot{e}_1 + e_2\dot{e}_2 \\ &= e_1\left(-c_1 e_1 + e_2 + \tilde{\xi}_1\right) + e_2\left(u_2 - \ddot{y}_r + \Sigma_2 + w_2\tilde{\xi}_1\right) \\ &= -c_1 e_1^2 + \underbrace{e_1 e_2} + e_2(\alpha_2 + \Sigma_2) + e_1\tilde{\xi}_1 + w_2 e_2\tilde{\xi}_1 \\ &= -c_1 e_1^2 + e_2\left(\underbrace{e_1} + \alpha_2 + \Sigma_2\right) + e_1\tilde{\xi}_1 + w_2 e_2\tilde{\xi}_1. \end{aligned}$$

We choose the stabilizing function

$$\alpha_2 = -e_1 - c_2 e_2 - \Sigma_2, \tag{9.20}$$

where $c_2 > 0$ is a design constant. Notice that we are now able to eliminate the $e_1 e_2$–term from the previous design step with α_2. This gives

$$\dot{V}_2 = -c_1 e_1^2 - c_2 e_2^2 + e_1\tilde{\xi}_1 + w_2 e_2\tilde{\xi}_1,$$

which is negative definite for zero observer error,

$$\tilde{\xi}_1 = 0 \Longrightarrow \dot{V}_2 = -c_1 e_1^2 - c_2 e_2^2.$$

For the 2nd-order system (9.18), the control $u_2 = \ddot{y}_r + \alpha_2$ with α_2 given by (9.20, makes the system exponentially stable for $\tilde{\xi}_1 = 0$, and since w_2 is constant, the linear damping term $-c_2 e_2$ is sufficient to make the closed-loop system ISS with respect to the observer error $\tilde{\xi}_1$.

Design – Step 3 We are not able to control u_2 directly, hence, u_2 is only a *virtual control*, and the control law designed at previous step becomes the desired control law

$$u_{2d} = \ddot{y}_r + \alpha_2\left(y, \hat{\xi}_1, y_r, \dot{y}_r\right).$$

Define the corresponding error

$$e_3 \triangleq u_2 - u_{2d} = u_2 - \ddot{y}_r - \alpha_2, \tag{9.21}$$

and note in passing that e_3 appears in the e_2–dynamics according to

$$\begin{aligned}\dot{e}_2 &= e_3 + \alpha_2 + \Sigma_2 + w_2\tilde{\xi}_1 \\ &= -e_1 - c_2e_2 + e_3 + w_2\tilde{\xi}_1.\end{aligned}$$

Then backstep the error through the last integrator by differentiating e_3:

$$\begin{aligned}\dot{e}_3 &= \dot{u}_2 - y_r^{(3)} - \dot{\alpha}_2 \\ &= \frac{d}{dt}\left(b_1(y)\,\hat{\xi}_2\right) - y_r^{(3)} - \frac{d}{dt}\left(\alpha_2\left(y, \hat{\xi}_1, \hat{z}, y_r, \dot{y}_r\right)\right) \\ &= \frac{\partial b_1}{\partial y}\dot{y}\hat{\xi}_2 + b_1\dot{\hat{\xi}}_2 - y_r^{(3)} - \frac{\partial \alpha_2}{\partial y}\dot{y} - \frac{\partial \alpha_2}{\partial \hat{\xi}_1}\dot{\hat{\xi}}_1 - \frac{\partial \alpha_2}{\partial \hat{z}}\dot{\hat{z}} - \frac{\partial \alpha_2}{\partial y_r}\dot{y}_r - \frac{\partial \alpha_2}{\partial \dot{y}_r}\ddot{y}_r. \\ &= \frac{\partial b_1}{\partial y}\left(\hat{\xi}_1 + k_1y + \tilde{\xi}_1\right)\hat{\xi}_2 + b_1\left(b_2\psi_2 - k_2\hat{\xi}_1 - k_1k_2y\right) \\ &\quad -y_r^{(3)} - \frac{\partial \alpha_2}{\partial y}\left(\hat{\xi}_1 + k_1y + \tilde{\xi}_1\right) - \frac{\partial \alpha_2}{\partial \hat{\xi}_1}\dot{\hat{\xi}}_1 - \frac{\partial \alpha_2}{\partial \hat{z}}\dot{\hat{z}} - \frac{\partial \alpha_2}{\partial y_r}\dot{y}_r - \frac{\partial \alpha_2}{\partial \dot{y}_r}\ddot{y}_r.\end{aligned}$$

Substituting with expressions for $\dot{y}$ and $\dot{\hat{\xi}}_2$from (9.11) and re-ordering, gives

$$\begin{aligned}\dot{e}_3 &= \frac{\partial b_1}{\partial y}\left(\hat{\xi}_1 + k_1y + \tilde{\xi}_1\right)\hat{\xi}_2 + b_1\left(b_2\psi_2 - k_2\hat{\xi}_1 - k_1k_2y\right) \\ &\quad -y_r^{(3)} - \frac{\partial \alpha_2}{\partial y}\left(\hat{\xi}_1 + k_1y + \tilde{\xi}_1\right) - \frac{\partial \alpha_2}{\partial \hat{\xi}_1}\dot{\hat{\xi}}_1 - \frac{\partial \alpha_2}{\partial \hat{z}}\dot{\hat{z}} - \frac{\partial \alpha_2}{\partial y_r}\dot{y}_r - \frac{\partial \alpha_2}{\partial \dot{y}_r}\ddot{y}_r \\ &= b_1b_2\psi_2 - y_r^{(3)} + \left(-b_1k_2 + \frac{\partial b_1}{\partial y}\hat{\xi}_2 - \frac{\partial \alpha_2}{\partial y}\right)\left(\hat{\xi}_1 + k_1y\right) \\ &\quad -\frac{\partial \alpha_2}{\partial \hat{\xi}_1}\dot{\hat{\xi}}_1 - \frac{\partial \alpha_2}{\partial \hat{z}}\dot{\hat{z}} - \frac{\partial \alpha_2}{\partial y_r}\dot{y}_r - \frac{\partial \alpha_2}{\partial \dot{y}_r}\ddot{y}_r + \left(\frac{\partial b_1}{\partial y}\hat{\xi}_2 - \frac{\partial \alpha_2}{\partial y}\right)\tilde{\xi}_1.\end{aligned}$$

Denote

$$u_3\left(y, \hat{\xi}_2, \hat{x}_v\right) \triangleq b_1b_2\psi_2,$$

which will be the third virtual control, and collect all the remaining known terms except $y_r^{(3)}$ in the function

$$\begin{aligned}\Sigma_3\left(y, \hat{\xi}_1, \hat{\xi}_2, \hat{z}, \dot{y}_r, \ddot{y}_r\right) &\triangleq \left(-b_1k_2 + \frac{\partial b_1}{\partial y}\hat{\xi}_2 - \frac{\partial \alpha_2}{\partial y}\right)\left(\hat{\xi}_1 + k_1y\right) \\ &\quad -\frac{\partial \alpha_2}{\partial \hat{\xi}_1}\dot{\hat{\xi}}_1 - \frac{\partial \alpha_2}{\partial \hat{z}}\dot{\hat{z}} - \frac{\partial \alpha_2}{\partial y_r}\dot{y}_r - \frac{\partial \alpha_2}{\partial \dot{y}_r}\ddot{y}_r.\end{aligned}$$

This gives

$$\dot{e}_3 = u_3 - y_r^{(3)} + \Sigma_3 + \left(\frac{\partial b_1}{\partial y}\hat{\xi}_2 - \frac{\partial \alpha_2}{\partial y}\right)\tilde{\xi}_1.$$

Denoting

$$w_3\left(y, \hat{\xi}_2\right) \triangleq \frac{\partial b_1}{\partial y}\hat{\xi}_2 - \frac{\partial \alpha_2}{\partial y},$$

the dynamics of the (e_1, e_2, e_3)–subsystem can be written

$$\begin{aligned} \dot{e}_1 &= -c_1 e_1 + e_2 + \tilde{\xi}_1 \\ \dot{e}_2 &= -e_1 - c_2 e_2 + e_3 + w_2\tilde{\xi}_1 \\ \dot{e}_3 &= u_3 - y_r^{(3)} + \Sigma_3 + w_3\tilde{\xi}_1. \end{aligned} \tag{9.22}$$

Assuming $u_3 \triangleq b_1(y)\, b_2\psi_2(y, \hat{\xi}_2, \hat{x}_v)$ is the actual control, we use a control in the form

$$u_3 = y_r^{(3)} + \alpha_3,$$

and design a stabilizing function α_3 to stabilize the (e_1, e_2, e_3)–subsystem. We use the CLF

$$V_3 = V_2 + \frac{1}{2}e_3^2,$$

which gives the derivative

$$\begin{aligned} \dot{V}_3 &= e_1\dot{e}_1 + e_2\dot{e}_2 + e_3\dot{e}_3 \\ &= e_1\left(-c_1 e_1 + e_2 + \tilde{\xi}_1\right) + e_2\left(-e_1 - c_2 e_2 + e_3 + w_2\tilde{\xi}_1\right) \\ &\quad + e_3\left(u_3 - y_r^{(3)} + \Sigma_3 + w_3\tilde{\xi}_1\right) \\ &= -c_1 e_1^2 - c_2 e_2^2 + e_3\left(e_2 + \alpha_3 + \Sigma_3 + w_3\tilde{\xi}_1\right) + e_1\tilde{\xi}_1 + e_2 w_2\tilde{\xi}_1 + e_3 w_3\tilde{\xi}_1. \end{aligned}$$

To make $\dot{V}_3$ negative definite for $\tilde{\xi}_2 = 0$, we could choose the stabilizing function

$$\alpha_3' = -e_2 - c_3 e_3 - \Sigma_3, \quad c_3 > 0 \tag{9.23}$$

which would give

$$\dot{V}_3' = -c_1 e_1^2 - c_2 e_2^2 - c_3 e_3^2 + e_1\tilde{\xi}_1 + e_2 w_2\tilde{\xi}_1 + e_3 w_3\tilde{\xi}_1.$$

For α_3' to make the closed-loop system ISS with respect to the observer error $\tilde{\xi}_1$, the term $w_3(y, \hat{\xi}_2)$ must be bounded. That is, for bounded w_3, the negative definite term $-c_3 e_3^2$ in $\dot{V}_3'$ due to linear damping, will dominate the disturbance term $e_3 w_3\tilde{\xi}_1$ for sufficiently large e_3, thus, ensuring the boundedness of e_3. However, the term $w_3(y, \hat{\xi}_2)$ is bounded only if y and $\hat{\xi}_2$ are bounded. Hence, the linear damping term $-c_3 e_3$ in α_3' is not sufficient to ensure boundedness of e_3 in the presence of observer errors, unless y and $\hat{\xi}_2$ can be assumed to be bounded.

To enhance performance and guarantee global boundedness in the presence of the observer error $\tilde{\xi}_1$, we strengthen the stabilizing function with a *nonlinear damping* term according to

$$\alpha_3 = -e_2 - c_3 e_3 - d_3 w_3^2 e_3 - \Sigma_3, \quad c_3, d_3 > 0. \tag{9.24}$$

This gives

$$\dot{V}_3 = -c_1 e_1^2 - c_2 e_2^2 - c_3 e_3^2 - d_3 w_3^2 e_3^2 + e_1 \tilde{\xi}_1 + e_2 w_2 \tilde{\xi}_1 + e_3 w_3 \tilde{\xi}_1,$$

and for zero observer error

$$\tilde{\xi}_1 = 0 \Longrightarrow \dot{V}_3 = -c_1 e_1^2 - c_2 e_2^2. - c_3 e_3^2 - d_3 w_3^2 e_3^2.$$

For the 3rd-order system (9.22), the control $u_3 = y_r^{(3)} + \alpha_3$, with α_3 given by (9.24), makes the system exponentially stable for $\tilde{\xi}_1 = 0$, and due to the nonlinear damping term $-d_3 w_3^2 e_3$, it also makes the closed-loop system ISS with respect to the observer error $\tilde{\xi}_1$ even for unbounded states y and $\hat{\xi}_2$.

Design – Final step Since u_3 is only a virtual control, the control law designed in the previous step is not directly implementable, but becomes our a desired control law

$$u_{3d} \triangleq y_r^{(3)} + \alpha_3 \left(y, \hat{\xi}_1, \hat{\xi}_2, \hat{z}, y_r, \dot{y}_r, \ddot{y}_r \right),$$

with the corresponding error

$$e_4 \triangleq u_3 - u_{3d} = u_3 - y_r^{(3)} - \alpha_3. \tag{9.25}$$

The error e_4 appears in the e_3–dynamics according to

$$\begin{aligned} \dot{e}_3 &= e_4 + \alpha_3 + \Sigma_3 + w_3 \tilde{\xi}_1 \\ &= -e_2 - c_3 e_3 - d_3 w_3^2 e_3 + e_4 + w_3 \tilde{\xi}_1. \end{aligned}$$

Proceeding as in previous design steps, we backstep the error e_4 through the last integrator by differentiation, and obtain

$$\begin{aligned} \dot{e}_4 &= \frac{d}{dt}\left(b_1(y)\, b_2 \psi_2 \left(y, \hat{\xi}_2, \hat{x}_v \right) \right) - y_r^{(4)} - \dot{\alpha}_3 \\ &= \frac{\partial b_1}{\partial y}\dot{y} + b_1 b_2 \frac{\partial \psi_2}{\partial y}\dot{y} + b_1 b_2 \frac{\partial \psi_2}{\partial \hat{\xi}_2}\dot{\hat{\xi}}_2 + b_1 b_2 \frac{\partial \psi_2}{\partial \hat{x}_v}\dot{\hat{x}}_v - y_r^{(4)} \\ &\quad - \frac{\partial \alpha_3}{\partial y}\dot{y} - \frac{\partial \alpha_3}{\partial \hat{\xi}_1}\dot{\hat{\xi}}_1 - \frac{\partial \alpha_3}{\partial \hat{\xi}_2}\dot{\hat{\xi}}_2 - \frac{\partial \alpha_3}{\partial \hat{z}}\dot{\hat{z}} - \frac{\partial \alpha_3}{\partial y_r}\dot{y}_r - \frac{\partial \alpha_3}{\partial \dot{y}_r}\ddot{y}_r - \frac{\partial \alpha_3}{\partial \ddot{y}_r} y_r^{(3)}. \end{aligned}$$

Substituting with expressions for $\dot{y}$ and $\dot{x}_v$ from (9.11) and re-ordering, gives

$$\begin{aligned} \dot{e}_4 &= b_1 b_2 b_3 \frac{\partial \psi_2}{\partial \hat{x}_v} u - y_r^{(4)} + \left(\frac{\partial b_1}{\partial y} + b_1 b_2 \frac{\partial \psi_2}{\partial y} - \frac{\partial \alpha_3}{\partial y} \right) \left(\hat{\xi}_1 + k_1 y \right) - \frac{\partial \alpha_3}{\partial \hat{\xi}_1}\dot{\hat{\xi}}_1 \\ &\quad + \left(b_1 b_2 \frac{\partial \psi_2}{\partial \hat{\xi}_2} - \frac{\partial \alpha_3}{\partial \hat{\xi}_2} \right) \dot{\hat{\xi}}_2 - b_1 b_2 b_3 \frac{\partial \psi_2}{\partial \hat{x}_v}\hat{x}_v - \frac{\partial \alpha_3}{\partial \hat{z}}\dot{\hat{z}} - \frac{\partial \alpha_3}{\partial y_r}\dot{y}_r - \frac{\partial \alpha_3}{\partial \dot{y}_r}\ddot{y}_r - \frac{\partial \alpha_3}{\partial \ddot{y}_r} y_r^{(3)} \\ &\quad + \left(\frac{\partial b_1}{\partial y} + b_1 b_2 \frac{\partial \psi_2}{\partial y} - \frac{\partial \alpha_3}{\partial y} \right) \tilde{\xi}_1. \end{aligned}$$

For simplicity of notation, introduce the scaled control variable u_4, which is one-to-one with the actual control u

$$u_4\left(y,\hat{\xi}_2,\hat{x}_v\right) \triangleq b_1b_2b_3\frac{\partial\psi_2}{\partial\hat{x}_v}u,$$

and collect all the remaining known terms except $y_r^{(4)}$ in the function

$$\begin{aligned}\Sigma_4\left(y,\hat{\xi}_1,\hat{\xi}_2,\hat{z},\hat{x}_v,\dot{y}_r,\ddot{y}_r,y_r^{(3)}\right) \triangleq & \left(\frac{\partial b_1}{\partial y}+b_1b_2\frac{\partial\psi_2}{\partial y}-\frac{\partial\alpha_3}{\partial y}\right)\left(\hat{\xi}_1+k_1y\right)-\frac{\partial\alpha_3}{\partial\hat{\xi}_1}\dot{\hat{\xi}}_1\\ &+\left(b_1b_2\frac{\partial\psi_2}{\partial\hat{\xi}_2}-\frac{\partial\alpha_3}{\partial\hat{\xi}_2}\right)\dot{\hat{\xi}}_2-b_1b_2b_3\frac{\partial\psi_2}{\partial\hat{x}_v}\hat{x}_v-\frac{\partial\alpha_3}{\partial\hat{z}}\dot{\hat{z}}\\ &-\frac{\partial\alpha_3}{\partial y_r}\dot{y}_r-\frac{\partial\alpha_3}{\partial\dot{y}_r}\ddot{y}_r-\frac{\partial\alpha_3}{\partial\ddot{y}_r}y_r^{(3)},\end{aligned}$$

so that the e_4–dynamics can be compactly written as

$$\dot{e}_4=u_4-y_r^{(4)}+\Sigma_4+\left(\frac{\partial b_1}{\partial y}+b_1b_2\frac{\partial\psi_2}{\partial y}-\frac{\partial\alpha_3}{\partial y}\right)\tilde{\xi}_1.$$

Denoting

$$w_4\left(y,\hat{\xi}_1,\hat{\xi}_2,\hat{x}_v\right) \triangleq \frac{\partial b_1}{\partial y}+b_1b_2\frac{\partial\psi_2}{\partial y}-\frac{\partial\alpha_3}{\partial y},$$

the dynamics of the full (e_1,e_2,e_3,e_4)–system can be written

$$\begin{aligned}\dot{e}_1&=-c_1e_1+e_2+\tilde{\xi}_1\\ \dot{e}_2&=-e_1-c_2e_2+e_3+w_2\tilde{\xi}_1\\ \dot{e}_3&=-e_2-c_3e_3-d_3w_3^2e_3+e_4+w_3\tilde{\xi}_1\\ \dot{e}_4&=u_4-y_r^{(4)}+\Sigma_4+w_4\tilde{\xi}_1.\end{aligned}\tag{9.26}$$

Taking

$$u_4=y_r^{(4)}+\alpha_4,$$

we design the final stabilizing function α_4 to stabilize the complete (e_1,e_2,e_3,e_4)–system. We use the CLF

$$V_4=V_3+\frac{1}{2}e_4^2,$$

giving the derivative

$$\begin{aligned}\dot{V}_4 &= e_1\left(-c_1e_1+e_2+\tilde{\xi}_1\right)+e_2\left(-e_1-c_2e_2+e_3+w_2\tilde{\xi}_1\right)\\ &\quad +e_3\left(-e_2-c_3e_3-d_3w_3^2e_3+e_4+w_3\tilde{\xi}_1\right)+e_4\left(u_4-y_r^{(4)}+\Sigma_4+w_4\tilde{\xi}_1\right)\\ &= -c_1e_1^2-c_2e_2^2-c_3e_3^2-d_3w_3^2e_3^2+e_4\left(e_3+\alpha_4+\Sigma_4\right)\\ &\quad +\left(e_1+w_2e_2+e_3w_3+e_4w_4\right)\tilde{\xi}_1.\end{aligned}$$

To make $\dot{V}_4$ negative definite for $\tilde{\xi}_2=0$, it would be sufficient to choose a stabilizing function with only linear damping according to

$$\alpha_4^{'}=-e_3-c_4e_4-\Sigma_4,\quad c_4>0,\tag{9.27}$$

which would give

$$\dot{V}_4 = -c_1 e_1^2 - c_2 e_2^2 - c_3 e_3^2 - d_3 w_3^2 e_3^2 - c_4 e_4^2 + (e_1 + w_2 e_2 + e_3 w_3 + e_4 w_4)\,\tilde{\xi}_1.$$

However, the disturbance term $w_4(y, \hat{\xi}_1, \hat{\xi}_2, \hat{x}_v)$ is an unbounded term which has the potential to grow very large, so that the linear damping term $-c_4 e_4$ in α_4' will not be sufficient to guarantee boundedness of the error state e_4. Consequently, to enhance performance and guarantee global boundedness in the presence of observer errors $\tilde{\xi}_1$, we strengthen the stabilizing function with a *nonlinear damping* term which counteracts the potentially destabilizing effect of observer errors multiplied with w_4. That is, we choose the stabilizing function

$$\alpha_4 = -e_3 - c_4 e_4 - d_4 w_4^2 e_4 - \Sigma_4, \quad c_4, d_4 > 0, \tag{9.28}$$

which gives

$$\begin{aligned} \dot{V}_4 \;=\; & -c_1 e_1^2 - c_2 e_2^2 - c_3 e_3^2 - d_3 w_3^2 e_3^2 - c_4 e_4^2 - d_4 w_4^2 e_4^2 - c_4 e_4^2 \\ & + (e_1 + w_2 e_2 + e_3 w_3 + e_4 w_4) \cdot \tilde{\xi}_1. \end{aligned}$$

For the system (9.22), the control $u_4 = y_r^{(4)} + \alpha_4$, with α_4 given by (9.28), makes the system exponentially stable for $\tilde{\xi}_1 = 0$, and due to the nonlinear damping term $-d_3 w_3^2 e_3$, it also makes the closed-loop system ISS with respect to the observer error $\tilde{\xi}_1$ even for unbounded states y and $\hat{\xi}_2$. The actual control input is given as

$$u = \left(b_1 b_2 b_3 \frac{\partial \psi_2}{\partial \hat{x}_v} \right)^{-1} \left[y_r^{(4)} + \alpha_4 \right]. \tag{9.29}$$

The complete set of equations describing the exact backstepping control law is summarized in the following section.

Summary — exact backstepping control law

The exact backstepping control law

$$u\left(y, \hat{\mathbf{x}}_u, \mathbf{z}_r\right) = \left(b_1 b_2 b_3 \frac{\partial \psi_2}{\partial \hat{x}_v} \right)^{-1} \left[y_r^{(4)} + \alpha_4 \right], \tag{9.30}$$

asymptotically stabilizes the system in the error coordinates

$$\begin{aligned} e_1 &= y - y_r \\ e_2 &= u_1 - \dot{y}_r - \alpha_1 \\ e_3 &= u_2 - \ddot{y}_r - \alpha_2 \\ e_4 &= u_3 - y_r^{(3)} - \alpha_3, \end{aligned} \tag{9.31}$$

and thereby solves the problem of asymptotic output-feedback tracking for the electro-pneumatic clutch actuator (9.6). The *virtual control* variables are

$$\begin{aligned} u_1 &= \hat{\xi}_1 \\ u_2 &= b_1 \hat{\xi}_2 \\ u_3 &= b_1 b_2 \psi_2, \end{aligned} \tag{9.32}$$

and the *stabilizing functions* are defined by the following expressions:

$$\begin{aligned}
&\alpha_1\left(y, y_r\right) = -c_1 e_1 - \Sigma_1 \\
&\alpha_2\left(y, \hat{\xi}_1, \hat{z}, y_r, \dot{y}_r\right) = -e_1 - c_2 e_2 - \Sigma_2 \\
&\alpha_3\left(y, \hat{\xi}_1, \hat{\xi}_2, \hat{z}, y_r, \dot{y}_r, \ddot{y}_r\right) = -e_2 - c_3 e_3 - d_3 w_3^2 e_3 - \Sigma_3 \\
&\alpha_4\left(y, \hat{\xi}_1, \hat{\xi}_2, \hat{x}_v, \hat{z}, y_r, \dot{y}_r, \ddot{y}_r, y_r^{(3)}\right) = -e_3 - c_4 e_4 - d_4 w_4^2 e_4 - \Sigma_4,
\end{aligned} \tag{9.33}$$

where $c_1, \cdots, c_4 > 0$ and $d_3, d_4 > 0$ are design parameters. The cancelled observer dynamics Σ_i and the disturbance gains w_i, $i = 1, \cdots, 4$, respectively, are given by

$$\begin{aligned}
&\Sigma_1\left(y\right) = k_1 y \\
&\Sigma_2\left(y, \hat{\xi}_1, \hat{z}, \dot{y}_r\right) = b_1 k_2 y - \tfrac{A}{M} P_0 - \tfrac{1}{M} f_l - \tfrac{1}{M} f_f + c_1 \hat{v} - c_1 \dot{y}_r \\
&\Sigma_3\left(y, \hat{\xi}_1, \hat{\xi}_2, \hat{z}, \dot{y}_r, \ddot{y}_r\right) = \left(-b_1 k_2 + \tfrac{\partial b_1}{\partial y}\hat{\xi}_2 - \tfrac{\partial \alpha_2}{\partial y}\right)\hat{v} - \tfrac{\partial \alpha_2}{\partial \hat{\xi}_1}\dot{\hat{\xi}}_1 - \tfrac{\partial \alpha_2}{\partial \hat{z}}\dot{\hat{z}} - \tfrac{\partial \alpha_2}{\partial y_r}\dot{y}_r - \tfrac{\partial \alpha_2}{\partial \dot{y}_r}\ddot{y}_r \\
&\Sigma_4\left(y, \hat{\xi}_1, \hat{\xi}_2, \hat{x}_v, \hat{z}, \dot{y}_r, \ddot{y}_r, y_r^{(3)}\right) = \left(\tfrac{\partial b_1}{\partial y} + b_1 b_2 \tfrac{\partial \psi_2}{\partial y} - \tfrac{\partial \alpha_3}{\partial y}\right)\hat{v} - \tfrac{\partial \alpha_3}{\partial \hat{\xi}_1}\dot{\hat{\xi}}_1 \\
&+ \left(b_1 b_2 \tfrac{\partial \psi_2}{\partial \hat{\xi}_2} - \tfrac{\partial \alpha_3}{\partial \hat{\xi}_2}\right)\dot{\hat{\xi}}_2 \\
&- b_1 b_2 b_3 \tfrac{\partial \psi_2}{\partial \hat{x}_v}\hat{x}_v - \tfrac{\partial \alpha_3}{\partial \hat{z}}\dot{\hat{z}} - \tfrac{\partial \alpha_3}{\partial y_r}\dot{y}_r - \tfrac{\partial \alpha_3}{\partial \dot{y}_r}\ddot{y}_r - \tfrac{\partial \alpha_3}{\partial \ddot{y}_r}y_r^{(3)},
\end{aligned} \tag{9.34}$$

and

$$\begin{aligned}
&w_1 = 1 \\
&w_2 = c_1 + k_1 \\
&w_3\left(y, \hat{\xi}_2\right) = \tfrac{\partial b_1}{\partial y}\hat{\xi}_2 - \tfrac{\partial \alpha_2}{\partial y} \\
&w_4\left(y, \hat{\xi}_1, \hat{\xi}_2, \hat{x}_v\right) = \tfrac{\partial b_1}{\partial y} + b_1 b_2 \tfrac{\partial \psi_2}{\partial y} - \tfrac{\partial \alpha_3}{\partial y}.
\end{aligned} \tag{9.35}$$

The main properties of the exact backstepping controller is summarized in the following Proposition:

Proposition 34 (Exact Backstepping Controller) *Consider the output-feedback controller consisting of the observer (9.9) and the control law (9.30)–(9.35) applied to position tracking of the electro-pneumatic clutch actuator (9.6). For bounded initial conditions, and any sufficiently smooth reference trajectory $y_r(t)$, assuming the observer is exponentially stable in the feasible region $\mathcal{X}_0$ of the model (9.6), the following properties hold for the closed-loop system:*

i) Boundedness: All signals of the closed-loop control system are bounded.

ii) Exponential tracking: The closed-loop system has an exponentially stable (ES) equilibrium at $(\mathbf{e}, \tilde{\mathbf{x}}_u) = \mathbf{0}$, which means that exponential tracking is achieved:

$$\lim_{t\to\infty}\left[y\left(t\right) - y_r\left(t\right)\right] = 0.$$

iii) Robustness: The closed-loop system is robust with respect to bounded disturbances entering additively in the system dynamics (9.6). These disturbances can be exogenous, or caused by model mismatches due to simplifications, parameter

errors or unmodeled dynamics. To be precise, the system is exponentially input-to-state stable (ISS) with respect to these disturbances as inputs. This means, in particular, that in the presence of bounded disturbances, the controller achieves tracking within a certain precision Δ_0 which depends on the upper bound of the disturbances:

$$\lim_{t\to\infty} |y(t) - y_r(t)| \leq \Delta_0.$$

Proof. The closed-loop system consists of the actuator states $\mathbf{x} = [y, v, p, x_v, z]^T$, which can be partitioned into the measured output y and the unmeasured states $\mathbf{x}_u = [v, p, x_v, z]^T$, and the states of the reduced-order observer $\hat{\mathbf{x}}_u = [\hat{\xi}_1, \hat{\xi}_2, \hat{x}_v, \hat{z}]^T$. Since the reference trajectory $\mathbf{z}_r$ is smooth, its components are bounded. The control law (9.30)–(9.35) stabilizes the system in the new error coordinates $\mathbf{e} = [e_1, e_2, e_3, e_4]^T$, where the change of coordinates (9.31), which we compactly write as

$$\mathbf{e} = \mathbf{\Phi}(y, \hat{\mathbf{x}}_u, \mathbf{z}_r),$$

is smooth in y, $\hat{\mathbf{x}}_u$ and $\mathbf{z}_r$, and the inverse transformation

$$\hat{\mathbf{x}}_u = \mathbf{\Phi}^{-1}(y, \mathbf{e}, \mathbf{z}_r),$$

is smooth in y, $\mathbf{e}$ and $\mathbf{z}_r$. Consequently, the boundedness of y and $\hat{\mathbf{x}}_u$ will follow from the boundedness of the error variables $\mathbf{e}$, and the boundedness of $\mathbf{x}_u$ from the boundedness of the observer error $\tilde{\mathbf{x}}_u$.

First, it is assumed that observer error $\tilde{\mathbf{x}}_u \triangleq [\tilde{\xi}_1, \tilde{\xi}_2, \tilde{x}_v, \tilde{z}]^T = [\tilde{v}, V(y)\tilde{p}, \tilde{x}_v, \tilde{z}]^T$ is exponentially stable, such that there exists a Lyapunov function

$$\underline{c}\,|\tilde{\mathbf{x}}_u|^2 \leq V_o(\tilde{\mathbf{x}}_u) \leq \overline{c}\,|\tilde{\mathbf{x}}_u|^2, \quad \overline{c} \geq \underline{c} > 0,$$

whose derivative satisfy

$$\dot{V}_o \leq -2\alpha_o V_o + \gamma_o |\boldsymbol{\delta}|^2, \quad \alpha_o, \gamma_o > 0,$$

where $\boldsymbol{\delta}(t) \in \mathbb{R}^{n-1}$ represents an additive disturbance in the actuator dynamics, which can be exogenous, or a result of model mismatches caused by modeling simplifications, parameter errors or unmodeled dynamics. This means that the observer error $\tilde{\mathbf{x}}_u$ is input-to-state stable with respect to $\boldsymbol{\delta}(t)$ as input, that is, exponentially stable for $\boldsymbol{\delta}(t) \equiv \mathbf{0}$, and bounded when subjected to bounded disturbances $\boldsymbol{\delta}(t)$.

Next, we establish that the dynamics of the closed-loop system in the error coordinates $e_1, \cdots, e_4$, is ISS with respect to the observer error $\tilde{\xi}_1$ as input. Partition the first two damping constants into $c_1 = \bar{c}_1 + \bar{d}_1$ and $c_2 = \bar{c}_2 + \bar{d}_2 w_2^2$, for some $\bar{c}_1, \bar{d}_1, \bar{c}_2, \bar{d}_2 > 0$ (which is possible for c_2 because w_2 is constant). The dynamics of the $\mathbf{e}$–system can be compactly written as

$$\dot{\mathbf{e}} = \mathbf{A}_e(t)\,\mathbf{e} + \mathbf{w}_e(t)\,\tilde{\xi}_1, \tag{9.36}$$

where the time-varying system matrix $\mathbf{A}_e(t)$, and disturbance gain vector $\mathbf{w}_e(t)$,

are given by

$$\mathbf{A}_e(t) \triangleq \begin{bmatrix} -\bar{c}_1 - \bar{d}_1 & 1 & 0 & 0 \\ -1 & -\bar{c}_2 - \bar{d}_2 w_2^2 & 1 & 0 \\ 0 & -1 & -c_3 - d_3 w_3\left(y, \hat{\xi}_2\right)^2 & 1 \\ 0 & 0 & -1 & -c_4 - d_4 w_4\left(y, \hat{\xi}_2, \hat{x}_v\right)^2 \end{bmatrix}, \tag{9.37}$$

$$\mathbf{w}_e(t) \triangleq \begin{bmatrix} 1 \\ w_2 \\ w_3\left(y, \hat{\xi}_2\right) \\ w_4\left(y, \hat{\xi}_2, \hat{x}_v\right) \end{bmatrix}. \tag{9.38}$$

To analyze the stability of this error dynamics, consider the CLF (V_4) from the final design step:

$$\begin{aligned} V_e &= \frac{1}{2}e_1^2 + \frac{1}{2}e_2^2 + \frac{1}{2}e_3^2 + \frac{1}{2}e_4^2 \\ &= \frac{1}{2}|\mathbf{e}|^2. \end{aligned}$$

The time-derivative of V_e becomes

$$\begin{aligned} \dot{V}_e &= -\bar{c}_1 e_1^2 - \bar{c}_2 e_2^2 - c_3 e_3^2 - c_4 e_4^2 \\ &\quad -\bar{d}_1 e_1^2 - \bar{d}_2 e_2^2 w_2^2 - d_3 e_3^2 w_3^2 - d_4 e_4^2 w_4^2 \\ &\quad +e_1\tilde{\xi}_1 + e_2 w_2 \tilde{\xi}_1 + e_3 w_3 \tilde{\xi}_1 + e_4 w_4 \tilde{\xi}_1. \end{aligned}$$

Using Lemma 26 (completion of squares), we obtain the inequalities

$$\begin{aligned} e_1\tilde{\xi}_1 &\leq \bar{d}_1 e_1^2 + \frac{1}{4\bar{d}_1}\tilde{\xi}_1^2 \\ e_2 w_2 \tilde{\xi}_1 &\leq \bar{d}_2 e_2^2 w_2^2 + \frac{1}{4\bar{d}_2}\tilde{\xi}_1^2 \\ e_i w_i \tilde{\xi}_1 &\leq d_i e_i^2 w_i^2 + \frac{1}{4d_i}\tilde{\xi}_1^2, \quad i = 3, 4, \end{aligned}$$

by which we can show that $\dot{V}_e$ satisfies

$$\begin{aligned} \dot{V}_e &\leq -\bar{c}_1 e_1^2 - \bar{c}_2 e_2^2 - c_3 e_3^2 - c_4 e_4^2 \\ &\quad + \left(\frac{1}{4\bar{d}_1} + \frac{1}{4\bar{d}_2} + \frac{1}{4d_3} + \frac{1}{4d_4}\right)\tilde{\xi}_1^2. \end{aligned}$$

Denoting

$$\begin{aligned} c_0 &\triangleq \min\{\bar{c}_1, \bar{c}_2, c_3, c_4\} \\ d_0 &\triangleq \left(\frac{1}{\bar{d}_1} + \frac{1}{\bar{d}_2} + \frac{1}{d_3} + \frac{1}{d_4}\right)^{-1}, \end{aligned}$$

we obtain

$$\begin{aligned}\dot{V}_e &\leq -c_0 |\mathbf{e}|^2 + \frac{1}{4d_0}\tilde{\xi}_1^2 \\ &= -2c_0 V_e + \frac{1}{4d_0}\tilde{\xi}_1^2,\end{aligned}$$

which proves that the $\mathbf{e}$–system is ISS with respect to the observer error $\tilde{\xi}_1$ as input. This means that the equilibrium $\mathbf{e} = \mathbf{0}$ is exponentially stable for $\tilde{\boldsymbol{\xi}}_1(t) \equiv \mathbf{0}$, and bounded when subjected to bounded observer errors $\tilde{\xi}_1(t)$.

In preparation for the analysis of the complete $(\mathbf{e}, \tilde{\mathbf{x}}_u)$–system, we rewrite $\dot{V}_e$ in terms of V_o. Using the inequality

$$\tilde{\xi}_1^2 \leq |\tilde{\mathbf{x}}_u|^2 \leq \frac{2}{\underline{c}} V_o,$$

we obtain

$$\dot{V}_e \leq -2c_0 V_e + \frac{1}{2d_0 \underline{c}} V_o.$$

An ISS Lyapunov function for the complete $(\mathbf{e}, \tilde{\mathbf{x}}_u)$–system is given by

$$\begin{aligned}V &= V_e + m_o V_o \\ &= \frac{1}{2}|\mathbf{e}|^2 + \frac{m_o}{2} V_o(\tilde{\mathbf{x}}_u)\end{aligned}$$

where $m_o > 0$ is a scaling constant which is determined below. The derivative of V becomes

$$\begin{aligned}\dot{V} &= \dot{V}_e + m_o \dot{V}_o \\ &\leq -2c_0 V_e - \left(2m_o \alpha_o - \frac{1}{2d_0 \underline{c}}\right) V_o + m_o \gamma_o |\boldsymbol{\delta}|^2.\end{aligned}$$

Taking

$$m_o \triangleq \frac{1}{4(\alpha_o - \sigma) d_0 \underline{c}},$$

we get

$$\dot{V} \leq -2c_0 V_e - 2\sigma m_o V_o + m_o \gamma_o |\boldsymbol{\delta}|^2.$$

By choosing σ according to

$$0 < \sigma < \min\{c_0, \alpha_o\},$$

the scaling factor m_o becomes finite and positive, and $\dot{V}$ satisfies

$$\begin{aligned}\dot{V} &\leq -2\sigma(V_e - m_o V_o) + m_o \gamma_o |\boldsymbol{\delta}|^2 \\ &= -2\sigma V + m_o \gamma_o |\boldsymbol{\delta}|^2.\end{aligned}$$

This proves that the complete $(\mathbf{e}, \tilde{\mathbf{x}}_u)$–system is ISS with respect to the disturbance $\boldsymbol{\delta}(t)$. This means that for $\boldsymbol{\delta}(t) \equiv \mathbf{0}$, the equilibrium $(\mathbf{e}, \tilde{\mathbf{x}}_u) = \mathbf{0}$ is exponentially

stable, *i.e.*, both the controller error $\mathbf{e}$ and observer error $\tilde{\mathbf{x}}_u$ converges exponentially to zero. Furthermore, for a bounded disturbance $\boldsymbol{\delta}(t)$, both $\mathbf{e}(t)$ and $\tilde{\mathbf{x}}_u(t)$ are guaranteed to be bounded.

We proceed to obtain an ISS bound on the tracking error $e_1(t) = y(t) - y_r(t)$. Using Lemma 28, with $c = 2\sigma$ and $d = m_o\gamma_o$, it follows that

$$V(t) \leq V(0)\,\mathrm{e}^{-2\sigma\cdot t} + \frac{m_o\gamma_o}{2\sigma}\left\|\boldsymbol{\delta}(t)\right\|_\infty^2,$$

where we use the simplified notation $V(t) = V(\mathbf{e}(t), \tilde{\mathbf{x}}_u(t))$. Noting that

$$|\mathbf{e}|^2 = 2V,$$

we can write

$$\begin{array}{rcl} |\mathbf{e}(t)|^2 & \leq & 2V(0)\,\mathrm{e}^{-2\sigma\cdot t} + \dfrac{m_o\gamma_o}{\sigma}\left\|\boldsymbol{\delta}(t)\right\|_\infty^2 \\ & \Downarrow & \\ |\mathbf{e}(t)| & \leq & \sqrt{2V(0)}\,\mathrm{e}^{-\sigma\cdot t} + \sqrt{\dfrac{m_o\gamma_o}{\sigma}}\left\|\boldsymbol{\delta}(t)\right\|_\infty, \end{array}$$

From

$$|e_1| \leq |\mathbf{e}|,$$

this provides an ISS bound on the tracking error is given by

$$|y(t) - y_r(t)| = |e_1(t)| \leq \sqrt{2V(0)}\,\mathrm{e}^{-\sigma\cdot t} + \sqrt{\frac{m_o\gamma_o}{\sigma}}\left\|\boldsymbol{\delta}(t)\right\|_\infty,$$

where an upper bound on the final tracking precision becomes

$$\begin{array}{rcl} \lim\limits_{t\to\infty}|y(t) - y_r(t)| & = & \lim\limits_{t\to\infty}|e_1(t)| \\ & \leq & \sqrt{\dfrac{m_o\gamma_o}{\sigma}}\left\|\boldsymbol{\delta}(t)\right\|_\infty \triangleq \Delta_0. \end{array}$$

■

Remark 35 *Note that, in contrast to the full state-feedback case, the cancelling control law (9.30)–(9.35) designed by observer backstepping, is robust because the terms of the observer are exactly known, and because the controller render the closed-loop system ISS with respect to observer errors at each step of the design. However, the control law may not be optimal, as it may cancel stabilizing nonlinearities, hence, wasting control effort.*

9.2.2 Approximate backstepping

In the recursive backstepping design, each new design step requires the calculation of the analytical expression of the derivative of a virtual control law designed at the previous step, whose complexity grows discouraging complex with each step. In

the process of simplifying the backstepping design, we may estimate the derivative of the virtual control law rather than calculate its analytical expression. To this end, we may use a high-gain observer to obtain (theoretically) an arbitrary accurate estimate of the 1st-order derivative $\dot{y}(t)$ of a smooth signal $y(t)$, using high gain, and the fact that any smooth signal has a 1st-order derivative which is continuous, thus, a 2nd-order derivative which is bounded.

First, we introduce the high-gain observer which will be used for derivative estimation. Next, we illustrate the design procedure using an estimate obtained by the high-gain observer, rather than calculating analytically, the derivative of the stabilizing functions.

Reduced-order high-gain observer for derivative estimation

The smooth time-varying signal $y(t)$ can be represented by the state-space model

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \ddot{y}, \end{aligned} \tag{9.39}$$

where the states are $x_1 = y$, and $x_2 = \dot{y}$, and the input is the 2nd-order derivative $\ddot{y}$. A high-gain observer can be designed to estimate the states x_1 and x_2 using feedback from $y = x_1$. Since $x_1 = y$ is already known, it may be preferable to design a reduced-order observer which only estimates x_2.

To design a reduced-order observer, we introduce the change of coordinate

$$\xi \triangleq x_2 - ky, \tag{9.40}$$

where $k > 0$ is a design constant yet to be determined. In the new coordinate ξ, we can rewrite the system (9.39) as

$$\begin{aligned} \dot{x}_1 &= \xi + ky \\ \dot{\xi} &= \ddot{y} - k(\xi + ky). \end{aligned}$$

This can be verified by direct calculation of the derivative of x_2 and substituting (9.40) for ξ:

$$\begin{aligned} \dot{x}_2 &= \dot{\xi} + k\dot{y} \\ &= \ddot{y} - k(\xi + ky) + kx_2 \\ &= \ddot{y} - k(x_2 - ky + ky) + kx_2 \\ &= \ddot{y}. \end{aligned}$$

An observer for estimation of the derivative $\dot{y}(t)$ of the smooth signal $y(t)$ is now obtained by defining the filter

$$\dot{\hat{\xi}} = -k\left(\hat{\xi} + ky\right), \tag{9.41}$$

with the estimate of $x_2 = \dot{y}$ taken as

$$\hat{x}_2 = \hat{\xi} + ky. \tag{9.42}$$

The estimate $\hat{x}_2$ obtained by the reduced-order observer, given by (9.41)–(9.42), can be made arbitrary accurate by increasing the filter gain k. It is instructive to realize that this high-gain observer is in essence a filtered differentiator, as can be clearly seen from its transfer function from y to $\hat{x}_2$, which is

$$\hat{x}_2 = \frac{s}{\frac{1}{k}s + 1} y. \tag{9.43}$$

We make the properties of the high-gain observer precise in the following Proposition.

Proposition 36 (Derivative Observer) *The estimate $\hat{x}_2(t)$ of the time-derivative $\dot{y}(t)$ of $y(t)$ obtained using the reduced-order observer (9.41)–(9.42), is exponentially input-to-state stable (ISS) with respect to the 2nd-order derivative $\ddot{y}(t)$. In particular, for any smooth signal $y(t) \in C^1$, its 2nd-order derivative $\ddot{y}(t)$ will be bounded, which means that for any prescribed accuracy $\varepsilon_0 > 0$, there exist a sufficiently large observer gain $k > 0$, which makes the estimation error $\tilde{x}_2 \triangleq \dot{y} - \hat{x}_2$ converge exponentially to within the prescribed accuracy. That is, the estimation error satisfies*

$$|\tilde{x}_2(t)| \leq |\tilde{x}_2(0)| e^{-k \cdot t} + \varepsilon_0 \tag{9.44}$$

where $\varepsilon_0 = \Delta / k$, where $\Delta \triangleq \|\ddot{y}(t)\|_\infty$ is the upper bound on $\ddot{y}(t)$.

Proof. Denote the filter state error $\tilde{\xi} \triangleq \xi - \hat{\xi}$, and note from (9.39) and (9.42) that $\tilde{\xi}$ equals the estimation error

$$\tilde{x}_2 \triangleq x_2 - \hat{x}_2 = \xi - \hat{\xi} = \tilde{\xi}.$$

The estimation error is governed by the dynamics

$$\dot{\tilde{\xi}} = -k\tilde{\xi} + \ddot{y}(t),$$

which is ISS with respect to $\ddot{y}(t)$. This can be established by the Lyapunov function

$$V = \frac{1}{2}\tilde{\xi}^2,$$

which has the derivative

$$\dot{V} = \tilde{\xi}(\ddot{y} - k\tilde{\xi}).$$

Clearly, $\left|\tilde{\xi}\right| > |\ddot{y}| / k \Longrightarrow \dot{V} < 0$, which proves the ISS property. We prove the ISS bound (9.44) with application of the simple convergence Lemma 28. Noting that

$$\begin{aligned} \frac{d}{dt}\left|\tilde{\xi}\right| &= \operatorname{sgn}\tilde{\xi}\dot{\tilde{\xi}} = -k \operatorname{sgn}\tilde{\xi}\tilde{\xi} + \operatorname{sgn}\tilde{\xi}\ddot{y}(t) \\ &\Downarrow \\ \frac{d}{dt}\left|\tilde{\xi}\right| &\leq -k\left|\tilde{\xi}\right| + |\ddot{y}(t)|, \end{aligned}$$

(taking $v = |\tilde{\xi}|$, $c = k$, $d = 1$, and $w^2 = |\ddot{y}|$), the lemma gives

$$\left|\tilde{\xi}(t)\right| \leq \left|\tilde{\xi}(0)\right| e^{-k \cdot t} + \frac{1}{k}\|\ddot{y}(t)\|_\infty.$$

■

Controller development

In this section, we illustrate the design procedure using approximate backstepping in the last two steps of the design in order to reduce the complexity of the resulting control law. That is, rather than computing the analytic expressions for the derivatives of the stabilizing functions $\alpha_2(y, \hat{\xi}_1, \hat{\xi}_2, \hat{z}, \mathbf{z}_r)$ and $\alpha_3(y, \hat{\xi}_1, \hat{\xi}_2, \hat{x}_v, \hat{z}, \mathbf{z}_r)$ in Steps 3 and 4 of the design procedure, respectively, we employ the high-gain observer introduced in the previous section to estimate these derivatives. This approximation introduces errors in the cancellation of nonlinearities, thus causing tracking errors, which again requires high feedback gain in the controller in order to achieve a high tracking precision. High gain in the controller can be disadvantageous because it makes the controller prone to amplify measurement noise, excite unmodeled dynamics, and introduce chattering in the control input. Therefore, this approximation is only applied in last steps of the design process, where the complexity of the controller due to differentiation of the stabilizing function is significant.

Design – Step 3 Recall that the virtual control from the previous step was

$$u_2\left(y, \hat{\xi}_2\right) \triangleq b_1 \hat{\xi}_2,$$

where the desired control law designed by an exact backstepping approach, was given by

$$u_{2d} = \ddot{y}_r + \alpha_2\left(y, \hat{\xi}_1, y_r, \dot{y}_r\right).$$

with the corresponding error

$$e_3 \triangleq u_2 - u_{2d} = u_2 - \ddot{y}_r - \alpha_2.$$

Denote $\beta_2 \triangleq \dot{\alpha}_2$, its estimate $\hat{\beta}_2$, and the corresponding error $\tilde{\beta}_2 \triangleq \beta_2 - \hat{\beta}_2$. Defining the filter

$$\dot{\hat{\zeta}}_2 = -l_2 \cdot \left(\hat{\zeta}_2 + l_2 \alpha_2\right), \tag{9.45}$$

an estimate of β_2 is given by

$$\hat{\beta}_2 = \hat{\zeta}_2 + l_2 \alpha_2. \tag{9.46}$$

Now, backstep the error through the last integrator using the estimate $\hat{\beta}_2$:

$$\begin{aligned} \dot{e}_3 &= \dot{u}_2 - y_r^{(3)} - \dot{\alpha}_2 \\ &= \frac{d}{dt}\left(b_1(y)\,\hat{\xi}_2\right) - y_r^{(3)} - \left(\hat{\beta}_2 + \tilde{\beta}_2\right) \\ &= \frac{\partial b_1}{\partial y}\dot{y}\hat{\xi}_2 + b_1\dot{\hat{\xi}}_2 - y_r^{(3)} - \left(\hat{\beta}_2 + \tilde{\beta}_2\right). \end{aligned}$$

Substituting with expressions for $\dot{y}$ and $\dot{\hat{\xi}}_2$ from (9.11) and re-ordering gives

$$\dot{e}_3 = b_1 b_2 \psi_2 - y_r^{(3)} + \left(\frac{\partial b_1}{\partial y}\hat{\xi}_2 - b_1 k_2\right)\left(\hat{\xi}_1 + k_1 y\right) - \hat{\beta}_2 + \frac{\partial b_1}{\partial y}\hat{\xi}_2\tilde{\xi}_1 - \tilde{\beta}_2.$$

The virtual control is the same:

$$u_3\left(y,\hat{\xi}_2,\hat{x}_v\right) \triangleq b_1 b_2 \psi_2.$$

A significant difference from exact backstepping is the cancellation term

$$\Sigma_3\left(y,\hat{\xi}_1,\hat{\xi}_2,\hat{\beta}_2\right) \triangleq \left(\frac{\partial b_1}{\partial y}\hat{\xi}_2 - b_1 k_2\right)\left(\hat{\xi}_1 + k_1 y\right) - \hat{\beta}_2,$$

which is significantly simplified, and the function

$$w_3\left(y,\hat{\xi}_2\right) \triangleq \frac{\partial b_1}{\partial y}\hat{\xi}_2,$$

multiplying the observer error. With the redefined Σ_3 and w_3, the last difference in the dynamics of the (e_1,e_2,e_3)–subsystem

$$\begin{aligned}
\dot{e}_1 &= -c_1 e_1 + e_2 + \tilde{\xi}_1 \\
\dot{e}_2 &= -e_1 - c_2 e_2 + e_3 + w_2\tilde{\xi}_1 \\
\dot{e}_3 &= u_3 - y_r^{(3)} + \Sigma_3 + w_3\tilde{\xi}_1 - \tilde{\beta}_2,
\end{aligned}$$

compared to exact backstepping, is the appearance of the estimation error $\tilde{\beta}_2$.

Taking

$$u_3 = y_r^{(3)} + \alpha_3,$$

we design a stabilizing function α_3 to stabilize the (e_1,e_2,e_3)–subsystem using the CLF

$$\begin{aligned}
V_3 &= V_2 + \frac{1}{2\nu_3}e_3^2 \\
&= \frac{1}{2}e_1^2 + \frac{1}{2}e_2^2 + \frac{1}{2\nu_3}e_3^2.
\end{aligned}$$

Here, the constant $\nu_3 > 0$ is introduced as an additional design parameter in order to more effectively compensate for the error $\tilde{\beta}_2$. The derivative of V_3 becomes

$$\begin{aligned}
\dot{V}_3 &= e_1\left(-c_1 e_1 + e_2 + \tilde{\xi}_1\right) + e_2\left(-e_1 - c_2 e_2 + e_3 + w_2\tilde{\xi}_1\right) \\
&\quad + \frac{1}{\nu_3}e_3\left(\alpha_3 + \Sigma_3 + w_3\tilde{\xi}_1 - \tilde{\beta}_2\right) \\
&= -c_1 e_1^2 - c_2 e_2^2 + \frac{1}{\nu_3}e_3\left(\nu_3 e_2 + \alpha_3 + \Sigma_3\right) \\
&\quad + e_1\tilde{\xi}_1 + e_2 w_2\tilde{\xi}_1 + \frac{1}{\nu_3}e_3 w_3\tilde{\xi}_1 - \frac{1}{\nu_3}e_3\tilde{\beta}_2.
\end{aligned}$$

We choose the stabilizing function

$$\alpha_3 = -\nu_3 e_2 - c_3 e_3 - d_3 w_3^2 e_3 - \Sigma_3, \quad c_3, d_3 > 0,$$

which gives

$$\begin{aligned}\dot{V}_3 &= -c_1e_1^2 - c_2e_2^2 - \frac{c_3}{\nu_3}e_3^2 - \frac{d_3}{\nu_3}w_3^2e_3^2 \\ &\quad + e_1\tilde{\xi}_1 + e_2w_2\tilde{\xi}_1 + \frac{1}{\nu_3}e_3w_3\tilde{\xi}_1 - \frac{1}{\nu_3}e_3\tilde{\beta}_2.\end{aligned}$$

Note that each term in the stabilizing function α_3 has a particular task: The first term $-\nu_3 e_2$ cancels the effect of the error e_2 caused by the virtual control variable u_1 being different from the desired control law u_{1d}. The last term Σ_3 implements a feedforward cancellation of known nonlinear dynamics, which allows us to replace it with a desired dynamics by adding feedback damping terms: The linear damping c_3–term determines the local convergence properties of the e_3–dynamics, and ensures boundedness with respect to the bounded disturbance $\tilde{\beta}_3$. The nonlinear damping d_3–term is introduced to counteract the potentially destabilizing effect of the observer error $\tilde{\xi}_1$ multiplied with the unbounded w_3–term. Moreover, the design parameter ν_3, introduces the possibility to compensate for the disturbances appearing in the e_3–dynamics. This is apparent from the scaling of the last two terms in the expression for $\dot{V}_3$ above, which shows that by increasing ν_3, the effect of $\tilde{\xi}_1$ and $\tilde{\beta}_2$ reduces.

Design – Final step With u_3 being a virtual control, the control law designed at the previous step becomes our desired control law

$$u_{3d} \triangleq y_r^{(3)} + \alpha_3\left(y, \hat{\xi}_1, \hat{\xi}_2, y_r, \dot{y}_r, \ddot{y}_r, \hat{\beta}_2\right),$$

with the corresponding error

$$e_4 \triangleq u_3 - u_{3d} = u_3 - y_r^{(3)} - \alpha_3. \tag{9.47}$$

As in Step 3, denote $\hat{\beta}_3$ the estimate of $\beta_3 \triangleq \dot{\alpha}_3$, and the estimation error $\tilde{\beta}_3 \triangleq \beta_3 - \hat{\beta}_3$. Introduce the filter

$$\dot{\hat{\zeta}}_3 = -l_3 \cdot \left(\hat{\zeta}_3 + l_3\alpha_3\right), \tag{9.48}$$

and take the estimate of β_3 as

$$\hat{\beta}_3 = \hat{\zeta}_3 + l_3\alpha_3. \tag{9.49}$$

Backstepping the error e_4 through the last integrator using approximate differentiation of α_3, we obtain

$$\begin{aligned}\dot{e}_4 &= \frac{d}{dt}\left(b_1(y)\, b_2\psi_2\left(y, \hat{\xi}_2, \hat{x}_v\right)\right) - y_r^{(4)} - \dot{\alpha}_3 \\ &= \frac{\partial b_1}{\partial y}\dot{y} + b_1b_2\frac{\partial \psi_2}{\partial y}\dot{y} + b_1b_2\frac{\partial \psi_2}{\partial \hat{\xi}_2}\dot{\hat{\xi}}_2 + b_1b_2\frac{\partial \psi_2}{\partial \hat{x}_v}\dot{\hat{x}}_v - y_r^{(4)} - \left(\hat{\beta}_3 + \tilde{\beta}_3\right).\end{aligned}$$

Substituting with expressions for $\dot{y}$ and $\dot{x}_v$ from (9.11) and re-ordering, gives

$$\begin{aligned}\dot{e}_4 &= b_1b_2b_3\frac{\partial\psi_2}{\partial\hat{x}_v}u - y_r^{(4)} + \left(\frac{\partial b_1}{\partial y} + b_1b_2\frac{\partial\psi_2}{\partial y}\right)\left(\hat{\xi}_1 + k_1y\right) + b_1b_2\frac{\partial\psi_2}{\partial\hat{\xi}_2}\dot{\hat{\xi}}_2 \\ &\quad -b_1b_2b_3\frac{\partial\psi_2}{\partial\hat{x}_v}\hat{x}_v - \hat{\beta}_3 + \left(\frac{\partial b_1}{\partial y} + b_1b_2\frac{\partial\psi_2}{\partial y}\right)\tilde{\xi}_1 - \tilde{\beta}_3.\end{aligned}$$

Denoting

$$u_4\left(y,\hat{\xi}_2,\hat{x}_v\right) \triangleq b_1b_2b_3\frac{\partial\psi_2}{\partial\hat{x}_v}u,$$

$$\Sigma_4\left(y,\hat{\xi}_1,\hat{\xi}_2,\hat{z},\hat{x}_v,\hat{\beta}_3\right) \triangleq \left(\frac{\partial b_1}{\partial y} + b_1b_2\frac{\partial\psi_2}{\partial y}\right)\left(\hat{\xi}_1 + k_1y\right) + b_1b_2\frac{\partial\psi_2}{\partial\hat{\xi}_2}\dot{\hat{\xi}}_2 - b_1b_2b_3\frac{\partial\psi_2}{\partial\hat{x}_v}\hat{x}_v - \hat{\beta}_3,$$

$$w_4\left(y,\hat{\xi}_2,\hat{x}_v\right) \triangleq \frac{\partial b_1}{\partial y} + b_1b_2\frac{\partial\psi_2}{\partial y},$$

the dynamics of the complete (e_1,e_2,e_3,e_4)–system can be written in the same form as before

$$\begin{aligned}\dot{e}_1 &= -c_1e_1 + e_2 + \tilde{\xi}_1 \\ \dot{e}_2 &= -e_1 - c_2e_2 + e_3 + w_2\tilde{\xi}_1 \\ \dot{e}_3 &= -\nu_3e_2 - c_3e_3 - d_3w_3^2e_3 + e_4 + w_3\tilde{\xi}_1 - \tilde{\beta}_2 \\ \dot{e}_4 &= u_4 - y_r^{(4)} + \Sigma_4 + w_4\tilde{\xi}_1 - \tilde{\beta}_3,\end{aligned} \tag{9.50}$$

where the differences compared to exact backstepping, are the appearance of the estimation errors $\tilde{\beta}_2$ and $\tilde{\beta}_3$, and the redefined functions, Σ_3, Σ_4, w_3 and w_4.

Taking

$$u_4 = y_r^{(4)} + \alpha_4,$$

we design the stabilizing function α_4 to stabilize the (e_1,e_2,e_3,e_4)–system using the CLF

$$V_4 = V_3 + \frac{1}{2\nu_4}e_4^2,$$

where $\nu_4 > 0$ is introduced as an additional design parameter. The derivative of V_4 becomes

$$\begin{aligned}\dot{V}_4 &= e_1\left(-c_1e_1 + e_2 + \tilde{\xi}_1\right) + e_2\left(-e_1 - c_2e_2 + e_3 + w_2\tilde{\xi}_1\right) \\ &\quad + \frac{1}{\nu_3}e_3\left(-\nu_3e_2 - c_3e_3 - d_3w_3^2e_3 + e_4 + w_3\tilde{\xi}_1 - \tilde{\beta}_2\right) \\ &\quad + \frac{1}{\nu_4}e_4\left(\alpha_4 + \Sigma_4 + w_4\tilde{\xi}_1 - \tilde{\beta}_3\right) \\ &= -c_1e_1^2 - c_2e_2^2 - \frac{c_3}{\nu_3}e_3^2 - \frac{d_3}{\nu_3}w_3^2e_3^2 + \frac{1}{\nu_4}e_4\left(\frac{\nu_4}{\nu_3}e_3 + \alpha_4 + \Sigma_4\right) \\ &\quad + e_1\tilde{\xi}_1 + e_2w_2\tilde{\xi}_1 + \frac{1}{\nu_3}e_3w_3\tilde{\xi}_1 - \frac{1}{\nu_3}e_3\tilde{\beta}_2 + \frac{1}{\nu_4}e_4w_4\tilde{\xi}_1 - \frac{1}{\nu_4}e_4\tilde{\beta}_3.\end{aligned}$$

We choose the stabilizing function

$$\alpha_4 = -\frac{\nu_4}{\nu_3}e_3 - c_4e_4 - d_4w_4^2e_4 - \Sigma_4, \quad c_4, d_4 > 0,$$

which gives

$$\begin{aligned}\dot{V}_4 &= -c_1e_1^2 - c_2e_2^2 - \frac{c_3}{\nu_3}e_3^2 - \frac{d_3}{\nu_3}w_3^2e_3^2 - -\frac{c_4}{\nu_4}e_4^2 - \frac{d_4}{\nu_4}w_4^2e_4^2 \\ &\quad +e_1\tilde{\xi}_1 + e_2w_2\tilde{\xi}_1 + \frac{1}{\nu_3}e_3w_3\tilde{\xi}_1 - \frac{1}{\nu_3}e_3\tilde{\beta}_2 + \frac{1}{\nu_4}e_4w_4\tilde{\xi}_1 - \frac{1}{\nu_4}e_4\tilde{\beta}_3.\end{aligned}$$

Note again that each term in the stabilizing function α_4 has similar tasks as in the previous design step: The first term $-\frac{\nu_4}{\nu_3}e_3$ cancels the effect of the error between the virtual control variable u_2 and the desired control law u_{2d}, while Σ_4 cancels known nonlinear dynamics in order to replace it with a desired dynamics by adding feedback damping terms: The linear damping c_4–term determines the local convergence properties of the e_4–dynamics, and ensures boundedness with respect to $\tilde{\beta}_4$. The nonlinear damping d_4–term counteracts the potentially destabilizing effect of the observer error $\tilde{\xi}_1$ multiplied with the unbounded w_4–term. Moreover, the design parameter ν_4, is introduced to make possible improved compensation of the disturbances appearing in the e_4–dynamics.

Summary — approximate backstepping control law

The main difference in the approximate backstepping control law compared to exact backstepping, is due to new definitions of the cancellation terms Σ_3 and Σ_4, where the analytic expressions for $\beta_2 \triangleq \dot{\alpha}_2$ and $\beta_3 \triangleq \dot{\alpha}_3$ are replaced with estimates $\hat{\beta}_2$ and $\hat{\beta}_3$, respectively. Furthermore, some additional design parameters ν_3 and ν_4 are introduced in the stabilizing functions α_3 and α_4 for improved compensation of the estimation errors $\tilde{\beta}_2$ and $\tilde{\beta}_3$. The complete set of equations describing the approximate backstepping control law is summarized below:

The approximate backstepping control law

$$u\left(y, \hat{\mathbf{x}}_u, \mathbf{z}_r\right) = \left(b_1b_2b_3\frac{\partial\psi_2}{\partial\hat{x}_v}\right)^{-1}\left[y_r^{(4)} + \alpha_4\right], \tag{9.51}$$

stabilizes the system in the error coordinates

$$\begin{aligned}e_1 &= y - y_r \\ e_2 &= u_1 - \dot{y}_r - \alpha_1 \\ e_3 &= u_2 - \ddot{y}_r - \alpha_2 \\ e_4 &= u_3 - y_r^{(3)} - \alpha_3,\end{aligned} \tag{9.52}$$

and thereby solves the problem of practical asymptotic output-feedback tracking for the electro-pneumatic clutch actuator (9.6). The *virtual control* variables are

$$\begin{aligned}u_1 &= \hat{\xi}_1 \\ u_2 &= b_1\left(y\right)\hat{\xi}_2 \\ u_3 &= b_1\left(y\right)b_2\psi_2\left(y, \hat{\xi}_2, \hat{x}_v\right),\end{aligned} \tag{9.53}$$

and the *stabilizing functions* are defined by the following expressions:

$$\begin{aligned}
&\alpha_1\left(y, y_r\right) = -c_1 e_1 - \Sigma_1 \\
&\alpha_2\left(y, \hat{\xi}_1, \hat{z}, y_r, \dot{y}_r\right) = -e_1 - c_2 e_2 - \Sigma_2 \\
&\alpha_3\left(y, \hat{\xi}_1, \hat{\xi}_2, y_r, \dot{y}_r, \ddot{y}_r, \hat{\beta}_2\right) = -\nu_3 e_2 - c_3 e_3 - d_3 w_3^2 e_3 - \Sigma_3 \\
&\alpha_4\left(y, \hat{\xi}_1, \hat{\xi}_2, \hat{x}_v, \hat{z}, y_r, \dot{y}_r, \ddot{y}_r, y_r^{(3)}, \hat{\beta}_3\right) = -\tfrac{\nu_4}{\nu_3} e_3 - c_4 e_4 - d_4 w_4^2 e_4 - \Sigma_4,
\end{aligned} \tag{9.54}$$

where $c_1, \cdots, c_4 > 0$, $d_3, d_4 > 0$ and $\nu_3, \nu_4 > 0$ are design parameters. The time-derivatives of α_2 and α_3 are estimated with the reduced-order observers

$$\begin{aligned}
\dot{\hat{\zeta}}_2 &= -l_2 \cdot \left(\hat{\zeta}_2 + l_2 \alpha_2\right), \quad \hat{\beta}_2 = \hat{\zeta}_2 + l_2 \alpha_2 \\
\dot{\hat{\zeta}}_3 &= -l_3 \cdot \left(\hat{\zeta}_3 + l_3 \alpha_3\right), \quad \hat{\beta}_3 = \hat{\zeta}_3 + l_3 \alpha_3,
\end{aligned} \tag{9.55}$$

where $l_2, l_3 > 0$ are design parameters which determines the time-constants of the filters, given as $\tau_2 = 1/l_2$ and $\tau_3 = 1/l_3$, respectively. The cancelled observer dynamics Σ_i and the nonlinear damping disturbance gains w_i, $i = 1, \cdots, 4$, respectively, are given by

$$\begin{aligned}
&\Sigma_1\left(y\right) = k_1 y \\
&\Sigma_2\left(y, \hat{\xi}_1, \hat{z}, \dot{y}_r\right) = b_1 k_2 y - \tfrac{A}{M} P_0 - \tfrac{1}{M} f_l - \tfrac{1}{M} f_f + c_1 \hat{v} - c_1 \dot{y}_r \\
&\Sigma_3\left(y, \hat{\xi}_1, \hat{\xi}_2, \hat{\beta}_2\right) = \left(\tfrac{\partial b_1}{\partial y} \hat{\xi}_2 - b_1 k_2\right)\left(\hat{\xi}_1 + k_1 y\right) - \hat{\beta}_2 \\
&\Sigma_4\left(y, \hat{\xi}_1, \hat{\xi}_2, \hat{z}, \hat{x}_v, \hat{\beta}_3\right) = \left(\tfrac{\partial b_1}{\partial y} + b_1 b_2 \tfrac{\partial \psi_2}{\partial y}\right)\left(\hat{\xi}_1 + k_1 y\right) + b_1 b_2 \tfrac{\partial \psi_2}{\partial \hat{\xi}_2} \dot{\hat{\xi}}_2 - b_1 b_2 b_3 \tfrac{\partial \psi_2}{\partial \hat{x}_v} \hat{x}_v - \hat{\beta}_3
\end{aligned} \tag{9.56}$$

and

$$\begin{aligned}
&w_1 = 1 \\
&w_2 = c_1 + k_1 \\
&w_3\left(y, \hat{\xi}_2\right) = \tfrac{\partial b_1}{\partial y} \hat{\xi}_2 \\
&w_4\left(y, \hat{\xi}_2, \hat{x}_v\right) = \tfrac{\partial b_1}{\partial y} + b_1 b_2 \tfrac{\partial \psi_2}{\partial y}.
\end{aligned} \tag{9.57}$$

The main properties of the approximate backstepping controller are summarized in the following Proposition:

Proposition 37 (Approximate Backstepping Controller) *Consider the output-feedback controller consisting of the observer (9.9) and the control law (9.51)–(9.57) applied to position tracking of the electro-pneumatic clutch actuator (9.6). For bounded initial conditions, and any sufficiently smooth reference trajectory $y_r(t)$, assuming the observer is exponentially stable in the feasible region $\mathcal{X}_0$ of the model (9.6), the following properties hold for the closed-loop system in the feasible region $\mathcal{X}_0$ of the model (9.6):*

i) Boundedness: All signals of the closed-loop control system are bounded.

ii) Practical exponential tracking: The closed-loop system has an exponentially practically stable equilibrium at $(\mathbf{e}, \tilde{\mathbf{x}}_u, \tilde{\boldsymbol{\beta}}) = \mathbf{0}$, which means that the controller

achieves exponential tracking within a certain precision e_0 which depends on the upper bound of the estimation errors $\tilde{\beta}_2$ and $\tilde{\beta}_3$:

$$\lim_{t\to\infty} |y(t) - y_r(t)| \leq e_0.$$

iii) Robustness: The closed-loop system is robust to bounded disturbances entering additively in the system dynamics (9.6). These disturbances can be exogenous, or caused by model mismatches due to simplifications, parameter errors or unmodeled dynamics. More precisely, the system is exponentially input-to-state practically stable (ISpS) with the disturbances as inputs. This means, in particular, that in the presence of bounded disturbances, the controller exponentially achieves tracking within a certain precision $e_0 + \Delta_0$, which depends on the upper bound of the disturbance $\boldsymbol{\delta}$ (Δ_0) and the estimation errors $\tilde{\beta}_2$ and $\tilde{\beta}_3$ (e_0):

$$\lim_{t\to\infty} |y(t) - y_r(t)| \leq e_0 + \Delta_0.$$

Proof. The main part of the proof is identical to the proof of Proposition 34, and we only outline the differences.

We first establish the properties of the estimation errors $\tilde{\beta}_2$ and $\tilde{\beta}_3$. In the region of validity $\mathcal{X}_0$, the model (9.6) and the observer (9.9) are smooth in the actuator states $\mathbf{x} = [y, v, p, x_v, z]^T$, the estimates $\hat{\mathbf{x}}_u = [\hat{\xi}_1, \hat{\xi}_2, \hat{x}_v, \hat{z}]^T$ and the control input u. Since the reference trajectory $\mathbf{z}_r$ and the estimates $\hat{\beta}_2$ and $\hat{\beta}_3$ (for continuous α_2 and α_3) are smooth, the approximate backstepping controller (9.51)–(9.57) produces a control law u and stabilizing functions α_i, $i = 1, \cdots, 4$, that are smooth in t. Consequently, by Proposition 36, the estimation errors $\tilde{\beta}_2$ and $\tilde{\beta}_3$ are bounded according to

$$\begin{aligned}
\left|\tilde{\beta}_2(t)\right| &\leq \left|\tilde{\beta}_2(0)\right| \mathrm{e}^{-l_2 t} + \frac{1}{l_2} \left\|\ddot{\alpha}_2(t)\right\|_\infty \\
\left|\tilde{\beta}_3(t)\right| &\leq \left|\tilde{\beta}_3(0)\right| \mathrm{e}^{-l_3 t} + \frac{1}{l_3} \left\|\ddot{\alpha}_3(t)\right\|_\infty .
\end{aligned}$$

From the proof of Proposition 36, we have that the error dynamics can be written as

$$\begin{aligned}
\dot{\tilde{\beta}}_2 &= -l_2 \tilde{\beta}_2 + \ddot{\alpha}_2(t) \\
\dot{\tilde{\beta}}_3 &= -l_3 \tilde{\beta}_3 + \ddot{\alpha}_3(t)
\end{aligned}$$

A Lyapunov function which establishes the ISS property of this error dynamics is given by

$$\begin{aligned}
V_\beta &= \frac{1}{2}\tilde{\beta}_2 + \frac{1}{2}\tilde{\beta}_3 \\
&= \frac{1}{2}\left|\tilde{\boldsymbol{\beta}}\right|^2 .
\end{aligned}$$

whose derivative can be written

$$\dot{V}_\beta = -l_2 \tilde{\beta}_2^2 - l_3 \tilde{\beta}_3^2 + \tilde{\beta}_2 \ddot{\alpha}_2 + \tilde{\beta}_3 \ddot{\alpha}_3.$$

Using completion of squares (Lemma 26), we obtain

$$\tilde{\beta}_i \ddot{a}_i \leq \frac{l_i}{2}\tilde{\beta}_i^2 + \frac{1}{2l_i}\ddot{\alpha}_i^2, \quad i = 1, 2,$$

which is used to show that $\dot{V}_\beta$ satisfies

$$\dot{V}_\beta \leq -\frac{l_2}{2}\tilde{\beta}_2^2 - \frac{l_3}{2}\tilde{\beta}_3^2 + \frac{1}{2l_2}\ddot{\alpha}_2(t)^2 + \frac{1}{2l_3}\ddot{\alpha}_3(t)^2.$$

Denoting $l_0 \triangleq \frac{1}{2}\min\{l_2, l_3\}$, with $\tilde{\boldsymbol{\beta}} \triangleq [\tilde{\beta}_2, \tilde{\beta}_3]^T$ and $\ddot{\boldsymbol{\alpha}} \triangleq [\ddot{\alpha}_2, \ddot{\alpha}_3]^T$, we can write

$$\begin{aligned} \dot{V}_\beta &\leq -l_0\left|\tilde{\boldsymbol{\beta}}\right|^2 + \frac{1}{2l_0}\left|\ddot{\boldsymbol{\alpha}}\right|^2 \\ &= -2l_0 V_\beta + \frac{1}{2l_0}\left|\ddot{\boldsymbol{\alpha}}\right|^2, \end{aligned}$$

which establishes that the error $\tilde{\boldsymbol{\beta}}$ is exponentially input-to-state stable (ISS) with respect to $\ddot{\boldsymbol{\alpha}}(t)$ as input.

Next, we establish that the dynamics of the system in the error coordinates $e_1, \cdots, e_4$, is ISS with respect to the observer errors $\tilde{\xi}_1$, $\tilde{\beta}_2$ and $\tilde{\beta}_3$ as inputs. Consider the CLF (V_4) from the final design step

$$\begin{aligned} V_e &= \frac{1}{2}e_1^2 + \frac{1}{2}e_2^2 + \frac{1}{2\nu_3}e_3^2 + \frac{1}{2\nu_4}e_4^2 \\ &= \frac{1}{2}\mathbf{e}^T\mathbf{P}_e\mathbf{e}, \end{aligned}$$

where $\mathbf{P}_e \triangleq \mathrm{diag}\{1, 1, \nu_3^{-1}, \nu_4^{-1}\}$ and $\mathbf{e} \triangleq [e_1, e_2, e_3, e_4]^T$. Substitute $c_1 = \bar{c}_1 + \bar{d}_1$, $c_2 = \bar{c}_2 + \bar{d}_2 w_2^2$, $c_3 = \bar{c}_3 + \kappa_3$ and $c_4 = \bar{c}_4 + \kappa_4$, such that the time-derivative can be written

$$\begin{aligned} \dot{V}_e &= -\bar{c}_1 e_1^2 - \bar{c}_2 e_2^2 - \frac{\bar{c}_3}{\nu_3}e_3^2 - \frac{\bar{c}_4}{\nu_4}e_4^2 \\ &\quad -\bar{d}_1 e_1^2 - \bar{d}_2 e_2^2 w_2^2 - \frac{d_3}{\nu_3}e_3^2 w_3^2 - \frac{d_4}{\nu_4}e_4^2 w_4^2 \\ &\quad +e_1\tilde{\xi}_1 + e_2 w_2\tilde{\xi}_1 + \frac{1}{\nu_3}e_3 w_3\tilde{\xi}_1 + \frac{1}{\nu_4}e_4 w_4\tilde{\xi}_1 \\ &\quad -\frac{\kappa_3}{\nu_3}e_3^2 - \frac{\kappa_4}{\nu_4}e_4^2 - \frac{1}{\nu_3}e_3\tilde{\beta}_2 - \frac{1}{\nu_4}e_4\tilde{\beta}_3. \end{aligned}$$

Using Lemma 26 (completion of squares) to obtain the inequalities

$$\begin{aligned} e_1\tilde{\xi}_1 &\leq \bar{d}_1 e_1^2 + \frac{1}{4\bar{d}_1}\tilde{\xi}_1^2 \\ e_2 w_2\tilde{\xi}_1 &\leq \bar{d}_2 e_2^2 w_2^2 + \frac{1}{4\bar{d}_2}\tilde{\xi}_1^2 \\ e_i w_i\tilde{\xi}_1 &\leq d_i e_i^2 w_i^2 + \frac{1}{4d_i}\tilde{\xi}_1^2, \quad i = 3, 4 \\ e_i\beta_{i-1} &\leq \kappa_i e_i^2 + \frac{1}{4\kappa_i}\tilde{\beta}_{i-1}^2, \quad i = 3, 4, \end{aligned}$$

we can show that $\dot{V}_e$ satisfies

$$\begin{aligned}\dot{V}_e &= -\bar{c}_1 e_1^2 - \bar{c}_2 e_2^2 - \frac{\bar{c}_3}{\nu_3} e_3^2 - \frac{\bar{c}_4}{\nu_4} e_4^2 \\ &\quad + \left(\frac{1}{4\bar{d}_1} + + \frac{1}{4\bar{d}_2} + \frac{1}{4\nu_3 d_3} + \frac{1}{4\nu_4 d_4}\right)\tilde{\xi}_1^2 \\ &\quad + \frac{1}{4\nu_3\kappa_3}\tilde{\beta}_2^2 + \frac{1}{4\nu_4\kappa_4}\tilde{\beta}_3^2.\end{aligned}$$

Denoting

$$\begin{aligned}c_0 &\triangleq \min\left\{\bar{c}_1, \bar{c}_2, \bar{c}_3, \bar{c}_4\right\}, \\ d_0 &\triangleq \left(\frac{1}{\bar{d}_1} + \frac{1}{\bar{d}_2} + \frac{1}{\nu_3 d_3} + \frac{1}{\nu_4 d_4}\right)^{-1} \\ \nu_0\kappa_0 &\triangleq \min\{\nu_3\kappa_3, \nu_4\kappa_4\},\end{aligned}$$

we get

$$\dot{V}_e \leq -c_0 \mathbf{e}^T \mathbf{P}_e \mathbf{e} + \frac{1}{4d_0}\tilde{\xi}_1^2 + \frac{1}{4\nu_0\kappa_0}\left|\tilde{\boldsymbol{\beta}}\right|^2.$$

This proves that the $\mathbf{e}$–system is ISS with respect to the observer errors $\tilde{\xi}_1$, $\tilde{\beta}_2$ and $\tilde{\beta}_3$ as inputs.

Preparing for analysis of the complete $(\mathbf{e}, \tilde{\mathbf{x}}_u, \tilde{\boldsymbol{\beta}})$–system, we show, using the inequality $\tilde{\xi}_1^2 \leq |\tilde{\mathbf{x}}_u|^2 \leq \frac{1}{\underline{c}} V_o$, that the derivative of V_e in terms of V_e, V_o and V_β, satisfies

$$\dot{V}_e \leq -2c_0 V_e + \frac{1}{4d_0\underline{c}} V_o + \frac{1}{2\nu_0\kappa_0} V_\beta.$$

A Lyapunov function for the complete $(\mathbf{e}, \tilde{\mathbf{x}}_u, \tilde{\boldsymbol{\beta}})$–system is given by

$$\begin{aligned}V &= V_e + m_o V_o + m_\beta V_\beta \\ &= \frac{1}{2}\mathbf{e}^T \mathbf{P}_e \mathbf{e} + m_o V_o + \frac{m_\beta}{2}\left|\tilde{\boldsymbol{\beta}}\right|^2, \end{aligned} \tag{9.58}$$

where $m_o, m_\beta > 0$ are positive constants which is determined below. In order to obtain an ISS bound on the tracking error $e_1(t) \triangleq y(t) - y_r(t)$, we need to rewrite $\dot{V}$ in terms of V. With this in mind, we show that the derivative of V in terms of V_e, V_o and V_β satisfies

$$\begin{aligned}\dot{V} &= \dot{V}_e + m_o \dot{V}_o + m_\beta \dot{V}_\beta \\ &\leq -2c_0 V_e + \frac{1}{2d_0\underline{c}} V_o + \frac{1}{2\nu_0\kappa_0} V_\beta + m_o\left(-2\alpha_o V_o + \gamma_o |\boldsymbol{\delta}|^2\right) + m_\beta\left(-2l_0 V_\beta + \frac{1}{2l_0}|\ddot{\boldsymbol{\alpha}}|^2\right) \\ &= -2c_0 V_e - \left(2m_o\alpha_o - \frac{1}{4d_0\underline{c}}\right) V_o - \left(2m_\beta l_0 - \frac{1}{2\nu_0\kappa_0}\right) V_\beta + m_o\gamma_o |\boldsymbol{\delta}|^2 + \frac{m_\beta}{2l_0}|\ddot{\boldsymbol{\alpha}}|^2.\end{aligned}$$

Taking

$$\begin{aligned}m_o &\triangleq \frac{1}{2(\alpha_o - \sigma) d_0\underline{c}} \\ m_\beta &\triangleq \frac{1}{4(l_o - \sigma)\nu_0\kappa_0},\end{aligned}$$

we get

$$\dot{V} \leq -2c_0 V_e - 2\sigma m_o V_o - 2\sigma m_\beta V_\beta + m_o \gamma_o \left|\boldsymbol{\delta}\right|^2 + \frac{m_\beta}{2l_0}\left|\ddot{\boldsymbol{\alpha}}\right|^2.$$

By choosing σ so that it satisfies

$$0 < \sigma < \min\left\{c_0, \alpha_o, l_o\right\},$$

we ensure that both scaling factors m_o and m_β are finite and positive, and that $\dot{V}$ satisfies

$$\begin{aligned}\dot{V} &\leq -2\sigma\left(V_e - m_o V_o - m_\beta V_\beta\right) + m_o\gamma_o\left|\boldsymbol{\delta}\right|^2 + \frac{m_\beta}{2l_0}\left|\ddot{\boldsymbol{\alpha}}\right|^2 \\ &= -2\sigma V + m_o\gamma_o\left|\boldsymbol{\delta}\right|^2 + \frac{m_\beta}{2l_0}\left|\ddot{\boldsymbol{\alpha}}\right|^2.\end{aligned}$$

This proves that the complete $(\mathbf{e}, \tilde{\mathbf{x}}_u, \tilde{\boldsymbol{\beta}})$–system is ISS with respect to the disturbance $\boldsymbol{\delta}(t)$ and $\ddot{\boldsymbol{\alpha}}(t)$ as inputs.

We proceed to obtain an ISS bound on the tracking error $e_1(t) = y(t) - y_r(t)$. Using Lemma 28 for both inputs $\boldsymbol{\delta}(t)$ and $\ddot{\boldsymbol{\alpha}}(t)$, we obtain

$$V(t) \leq V(0)\,\mathrm{e}^{-2\sigma\cdot t} + \frac{m_o\gamma_o}{2\sigma}\left\|\boldsymbol{\delta}(t)\right\|_\infty^2 + \frac{m_\beta}{4\sigma l_0}\left\|\ddot{\boldsymbol{\alpha}}(t)\right\|_\infty^2,$$

where we use the simplified notation $V(t) = V\left(\mathbf{e}(t), \tilde{\mathbf{x}}_u(t), \tilde{\boldsymbol{\beta}}(t)\right)$. From (9.58), we note that

$$\left|e_1\right|^2 \leq \mathbf{e}^T\mathbf{P}_\nu\mathbf{e} \leq 2V,$$

and obtain an ISS bound on the tracking error $e_1(t) \triangleq y(t) - y_r(t)$ according to

$$\begin{aligned}e_1(t)^2 &\leq 2V(0)\,\mathrm{e}^{-2\sigma\cdot t} + \frac{m_o\gamma_o}{\sigma}\left\|\boldsymbol{\delta}(t)\right\|_\infty^2 + \frac{m_\beta}{2\sigma l_0}\left\|\ddot{\boldsymbol{\alpha}}(t)\right\|_\infty^2 \\ &\Downarrow \\ \left|e_1(t)\right| &\leq \sqrt{2V(0)}\,\mathrm{e}^{-\sigma\cdot t} + \sqrt{\frac{m_o\gamma_o}{\sigma}}\left\|\boldsymbol{\delta}(t)\right\|_\infty^2 + \sqrt{\frac{m_\beta}{2\sigma l_0}}\left\|\ddot{\boldsymbol{\alpha}}(t)\right\|_\infty,\end{aligned}$$

where both $\boldsymbol{\delta}(t)$ and $\ddot{\boldsymbol{\alpha}}(t)$ are viewed as inputs. Here, we denote the bounds corresponding to $\boldsymbol{\delta}(t)$ and $\ddot{\boldsymbol{\alpha}}(t)$, respectively, by

$$\begin{aligned}\Delta_0 &\triangleq \sqrt{\frac{m_o\gamma_o}{\sigma}}\left\|\boldsymbol{\delta}(t)\right\|_\infty \\ e_0 &\triangleq \sqrt{\frac{m_\beta}{2\sigma l_0}}\left\|\ddot{\boldsymbol{\alpha}}(t)\right\|_\infty.\end{aligned}$$

An upper bound on the final tracking precision $\Delta_0 + e_0$, is then given by

$$\lim_{t\to\infty}\left|y(t) - y_r(t)\right| = \lim_{t\to\infty} e_1(t) = \Delta_0 + e_0,$$

which reduces to e_0 without disturbances, *i.e.*, for $\boldsymbol{\delta}(t) \equiv \mathbf{0}$.

The boundedness of the uncertain derivatives of the stabilizing functions, is a property which is inherent of the controller design. Therefore, we may view $\ddot{\boldsymbol{\alpha}}(t)$ as an internal controller error rather than an input. In this case, the $(\mathbf{e}, \tilde{\mathbf{x}}_u, \tilde{\boldsymbol{\beta}})$–system is said to be exponentially input-to-state practically stable (ISpS) with respect to $\boldsymbol{\delta}(t)$ as a disturbance input. Without the disturbance, *i.e.*, for $\boldsymbol{\delta}(t) \equiv \mathbf{0}$, the equilibrium $(\mathbf{e}, \tilde{\mathbf{x}}_u, \tilde{\boldsymbol{\beta}}) = \mathbf{0}$, is then said to be exponentially practically stable. ■

9.3 Robust output-feedback control system

The complete tracking control system for the electro-pneumatic actuator (9.6) consists of three main components, illustrated in the block diagram in Figure 9.2:

Reference Model For arbitrary reference inputs r, the reference model generates a smooth reference trajectory y_r and its time-derivatives, given by $\mathbf{z}_r \triangleq [y_r, \dot{y}_r, \ddot{y}_r, y_r^{(3)}, y_r^{(4)}]^T$.

State Estimator Based on a nonlinear model of the actuator, and using the measured position y and the control input u, the observer (9.9) provides estimates of the unmeasured states $\hat{\mathbf{x}}_u \triangleq [v, p, x_v, z]^T$ of the actuator.

Control Law Based on a nonlinear model of the actuator, the controller use the measured position y and the estimate $\hat{\mathbf{x}}_u$ of the unmeasured states to compute a control input u which makes the actuator position y track the reference trajectory y_r.

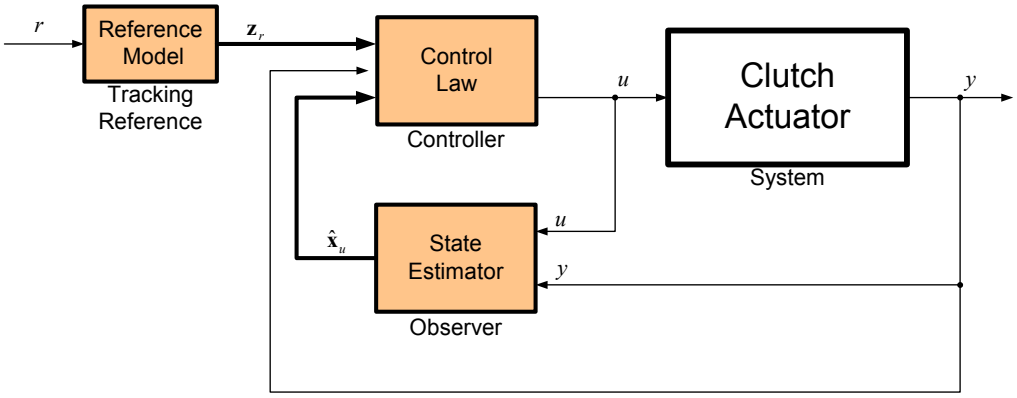

Figure 9.2: Block diagram illustrating the structure of the output-feedback tracking controller.

9.3.1 Reference model

The reference model is given by (9.2)–(9.3), described in Section 9.1. We determine the properties of the reference model by the single design parameter λ_r, referred to as the design bandwidth of the reference model. Since λ_r determines the time-constant of the reference model, it is also viewed as the design bandwidth of the closed-loop control system with respect to tracking the reference input r with the actuator position y.

9.3.2 State estimator

To obtain a robust solution, we use the re-designed robust observer given by (8.123), page 147, rather than the nominal observer (9.9) which was used in the backstepping design of the control law in the previous sections. In the robust observer, smooth saturation and projection of the observer states $\hat{\xi}_1$ and $\hat{\xi}_2$, and smooth saturation of the control input u, is used to constrain the estimates $\hat{v}$, $\hat{p}$ and $\hat{x}_v$ to the feasible region $\boldsymbol{\chi}_0$ of the model (9.6). Hence, global exponential stability of the observer $\mathbf{x}_u$ is obtained for $\forall \mathbf{x}_u \in \mathcal{X}_0$. For simplicity of notation, we denote the estimate obtained from the robust observer (8.123) simply by $\hat{\mathbf{x}}_u$, *i.e.*, dropping 'π' and 'P' in $\hat{\mathbf{x}}_{u,\pi}^{P}$.

The estimation properties of the observer is determined by the design bandwidth λ_o, however, the upper and lower bounds, $\hat{\mathbf{x}}_{u,\mathrm{ub}}$ and $\hat{\mathbf{x}}_{u,\mathrm{lb}}$, have a strong influence on the initial convergence properties, and should be tightened as much as possible in order to improve the initial convergence of the observer.

9.3.3 Control law

The control law is given either by the exact backstepping design (9.30)–(9.35), or the approximate backstepping design (9.51)–(9.57), where the observer estimates $\hat{\mathbf{x}}_u$ are replaced with the estimates obtained from the robust observer (8.123)–(8.125). The actual control input u_v is taken as

$$u_v \triangleq \frac{1}{K_v} \pi_u \left(u \right),$$

where $\pi_u \left(u \right)$ is the smooth saturated control input used in (8.123). Since the projected estimates from the robust observer are smooth, the smoothness of the backstepping control law is preserved. Furthermore, since the estimates from the robust observer are constrained to the feasible region $\mathcal{X}_0$ of the model, the properties established by either Proposition 34 (page 170) or Proposition 37 (page 182), are now valid, globally. Hence, the controller achieves global robust output-feedback (practical) tracking.

In the following, we will implement and analyze only the approximate backstepping controller, given by (9.51)–(9.57).

9.4 Simulation results

In this section, we present some characteristic simulation results when the approximate backstepping controller is applied to output-feedback tracking of the electropneumatic clutch actuator (9.6). We analyze the performance of the controller with a sinusoidal reference trajectory, which results in a relatively simple tracking task, but which demonstrates well the properties of the controller. The reference trajectory $y_r \left(t \right)$ is generated from the reference model using the sine wave input

$$r \left(t \right) = R_0 + R_1 \sin \left(\frac{2\pi}{T} \cdot t \right),$$

with bias $R_0 = 8\,\text{mm}$, amplitude $R_1 = 4\,\text{mm}$, and period $T = 1.0\,\text{s}$.

The reference model is implemented with poles placed at $s = -\lambda_r$, with $\lambda_r = 50$. This corresponds to a time-constant $\tau_r = 5/\lambda_r = 0.100\,\text{s}$ for the critically damped 5th-order reference model. The observer gains k_1 and k_2 of the reduced-order observer (8.123), is placed according to (8.111) with the single design parameter λ_o, referred to as the design bandwidth of the observer. We illustrate the performance using the following observer settings:

$$\begin{array}{lcl} \lambda_o = 25 & \Rightarrow & k_1 = 50 \quad k_2 = 0.5 \cdot 10^3 \\ \lambda_o = 50 & \Rightarrow & k_1 = 100 \quad k_2 = 2.2 \cdot 10^3 \\ \lambda_o = 100 & \Rightarrow & k_1 = 200 \quad k_2 = 8.9 \cdot 10^3. \end{array}$$

The observer is implemented with upper bound $\hat{\mathbf{x}}_{u,\text{ub}}$, lower bound $\hat{\mathbf{x}}_{u,\text{lb}}$, and smoothing width $\boldsymbol{\varepsilon}_\pi$, set according to

$$\begin{aligned} \hat{\mathbf{x}}_{u,\text{ub}} &= \begin{bmatrix} 200\,\text{mm/s} & 995\,\text{kPa} & 1 & 0.5\,\text{mm} \end{bmatrix}^T \\ \hat{\mathbf{x}}_{u,\text{lb}} &= \begin{bmatrix} -200\,\text{mm/s} & 105\,\text{kPa} & -1 & -0.5\,\text{mm} \end{bmatrix}^T \\ \boldsymbol{\varepsilon}_\pi &= \begin{bmatrix} 1\,\text{mm/s} & 1\,\text{kPa} & 0.01 & 0.001\,\text{mm} \end{bmatrix}^T. \end{aligned}$$

Furthermore, the feedback gains l_2 and l_3 of the derivative observers (9.55) are implemented with $l_2 = l_3 \triangleq l_\beta = 200$, which corresponds to low-pass filtering with time-constants $\tau_2 = \tau_3 = 1/l_\beta = 0.005\,\text{s}$.

Controller scaling factors ν_3 and ν_4

In this subsection, we illustrate the effectiveness of using ν_3 and ν_4 to attenuate the disturbing effect of the estimation errors $\tilde{\beta}_2$ and $\tilde{\beta}_3$, introduced by the simplification of the exact backstepping controller. The tuning of ν_3 and ν_4 is simplified by introducing the single design parameter $\nu_c > 0$, and take ν_3 and ν_4 according to the rule

$$\begin{aligned} \nu_3 &= \nu_c \\ \nu_4 &= \nu_c \nu_3. \end{aligned} \tag{9.59}$$

In order to isolate the effect of $\tilde{\beta}_2$ and $\tilde{\beta}_3$, the closed-loop controller is simulated without initial transients, and with zero feedback gain in the controller, *i.e.*, $c_c = 0$. We initialize the observer to produce zero initial error, $\tilde{\mathbf{x}}_u(0) = \mathbf{0}$, which is equivalent to simulating the controller with full state-feedback since $\hat{\mathbf{x}}_u(t) \equiv \mathbf{x}_u(t)$ for $\forall t \geq 0$. Furthermore, we chose the initial value of the state vector of the reference model $\mathbf{z}_r(0)$ to produce $\mathbf{e}(0) = \mathbf{0}$. This is referred to as trajectory initialization, which together with full state-feedback (*i.e.*, zero observer errors), eliminates the initial transients of the closed-loop system.

Figure 9.3 plots three simulations for increasing values of ν_c. The estimation errors $\tilde{\beta}_2$ and $\tilde{\beta}_3$ cause errors which destroy the asymptotic tracking properties of the closed-loop system. However, as illustrated in the figure, the effect of the errors $\tilde{\beta}_2$ and $\tilde{\beta}_3$ reduces for increasing ν_c, and the asymptotic tracking performance of the closed-loop system is approximately recovered for high ν_c.

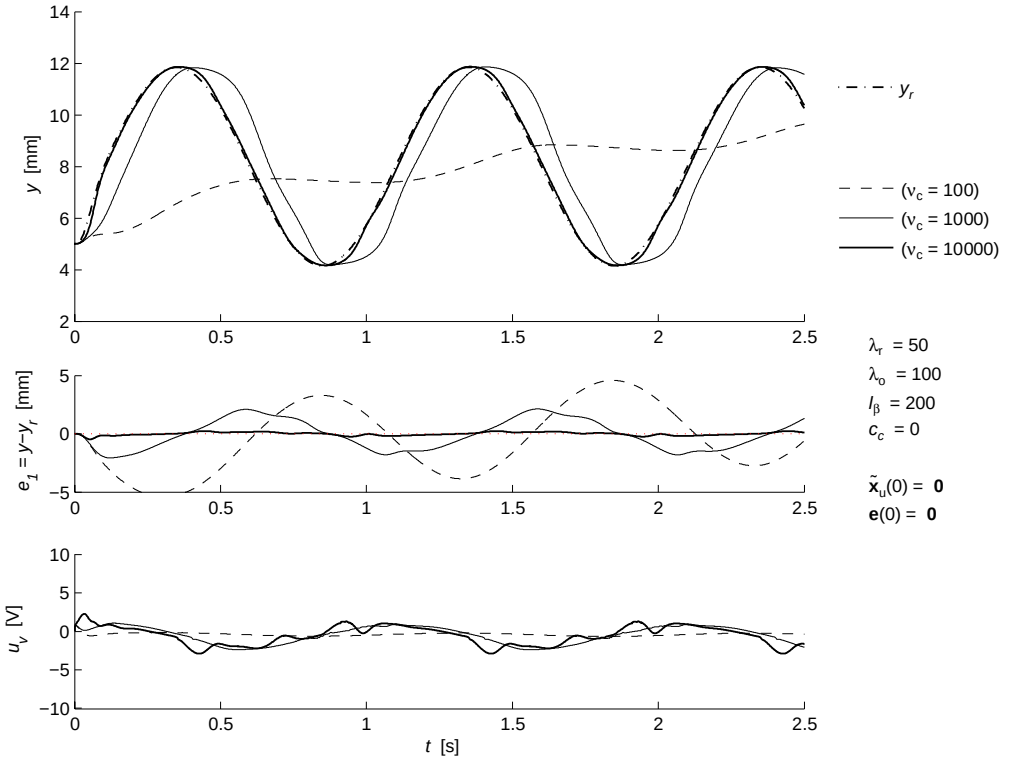

Figure 9.3: Simulations illustrating the effect of the scaling factors $\nu_3 = \nu_c$ and $\nu_4 = \nu_c\nu_3$.

Analyzing the stabilizing functions (9.54), we see that high ν_3, in effect, implements high gain compensation of the error e_2 due the the cancellation term $-\nu_3 e_2$ in the the third stabilizing function α_3. Likewise, will high ν_4 implement high gain compensation of the error e_3 due to $-\nu_4/\nu_3 \cdot e_3$ in α_4. However, too high values of ν_3 or ν_4 is undesirable, as it may make the controller prone to introduce chattering — high-frequency switching — in the control input, and sensitive to noise in the measured output y.

Remark 38 *Both the simulations and experimental results show that a good compromise between high compensation of $\tilde{\beta}_2$ and $\tilde{\beta}_3$ and a reasonably low gain which do not introduce too high control effort, is to take v_3 and ν_4 using (9.59) with $\nu_c = 10^4$. This is, however, a coarse choice, and fine tuning of ν_c should provide a more optimal value.*

Controller feedback gains c_1, c_2, c_3, and c_4

In this subsection, we demonstrate the effect of the controller feedback gains c_1, c_2, c_3, and c_4 on the convergence properties of the output-feedback tracking controller. We choose the controller gains with a single design parameter c_c according to the simple rule

$$c_1 = c_2 = c_3 = c_4 = c_c \tag{9.60}$$

When c_1, c_2, c_3 and c_4 are given by the above tuning law, we refer to c_c as the controller gain. Like for the ν–parameters, this is a very course choice, such that fine tuning of each parameter individually for a given application, should provide a more optimal controller gain setting.

We use observer gains k_1 and k_2 set according to a bandwidth of $\lambda_o = 50$, and for the controller scaling factors, we take $\nu_c = 10^4$ with ν_3 and ν_4 determined according to (9.59). In order to illustrate the strong convergence properties of the closed-loop control system, the observer is initialized with the simple choice

$$\hat{\mathbf{x}}_u(0) = [\hat{v}(0), \hat{p}(0), \hat{x}_v(0), \hat{z}(0)]^T = [0, 0, 0, 0]^T.$$

Since we replace $\hat{\mathbf{x}}_u$ with its saturated and projected version $\hat{\mathbf{x}}^P_{u,\pi}$ in the implemented controller, the actual initial values of the observer becomes

$$\hat{\mathbf{x}}^P_{u,\pi}(0) = [\hat{v}(0), \hat{p}(0), \hat{x}_v(0), \hat{z}(0)]^T = [0, P_0, 0, 0]^T,$$

due to the lower bound on the pressure estimate, $\hat{p}_{\text{lb}} = P_0$. Also for illustration of convergence properties, the state of the reference model $\mathbf{z}_r$ is initialized with the simple choice $\mathbf{z}_r(0) = \mathbf{0}$, rather than using trajectory initialization.

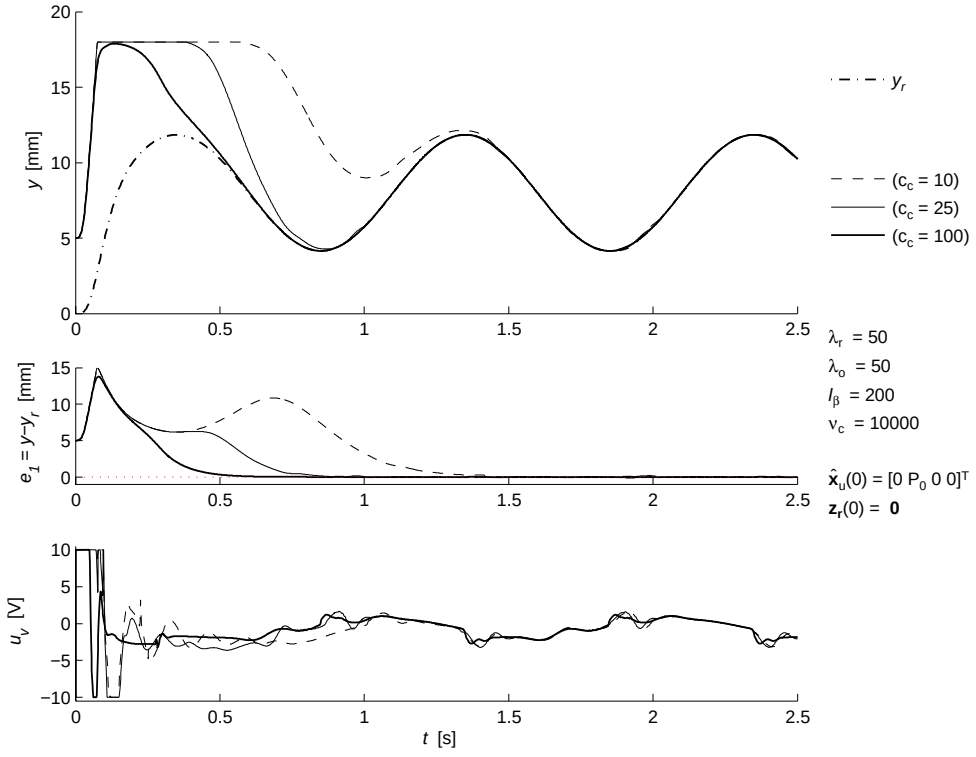

Figure 9.4: Simulations illustrating the effect of controller feedback gains c_i, $i = 1, \cdots, 4$, set according to $c_i = c_c$.

Figure 9.4 plots the simulated tracking performance for increasing values of c_c. For $c_c < 100$, the output y of the actuator hits its physical constraint ($y_{\text{ub}} = 18\,\text{mm}$)

during the initial transients of the controller convergence. The initial transients in the controller are caused by initial observer errors, $\tilde{\mathbf{x}}_u \neq \mathbf{0}$, and initial errors in the $\mathbf{e}$–system, $\mathbf{e} \neq \mathbf{0}$. For all $c_c = \{10, 25, 100\}$, the tracking error $e_1 = y - y_r$, converges to approximately zero, *i.e.*, practical tracking is recovered after the exponentially converging transients. We see from the figure that the control input saturates during the initial transients, and that the control effort increases with increasing gain. However, once the tracking task is met, the control effort lies well within the available control, $u_v \in [-10, 10]\,\mathrm{V}$.

Observer feedback gains — bandwidth λ_o

In this subsection, we demonstrate the effect of the observer feedback gains k_1 and k_2 on the convergence properties of the output-feedback tracking controller. As proposed in Subsection 8.4.4 (page 139), we determine the observer gains according to (8.111) using the single parameter λ_o, referred to as the design bandwidth of the observer.

The parameters and initial conditions are in accordance with the simulations presented in the previous subsection, except the controller gains, which are set according to $c_c = 50$, and the observer gains, which are set according to different values of the design bandwidth λ_o.

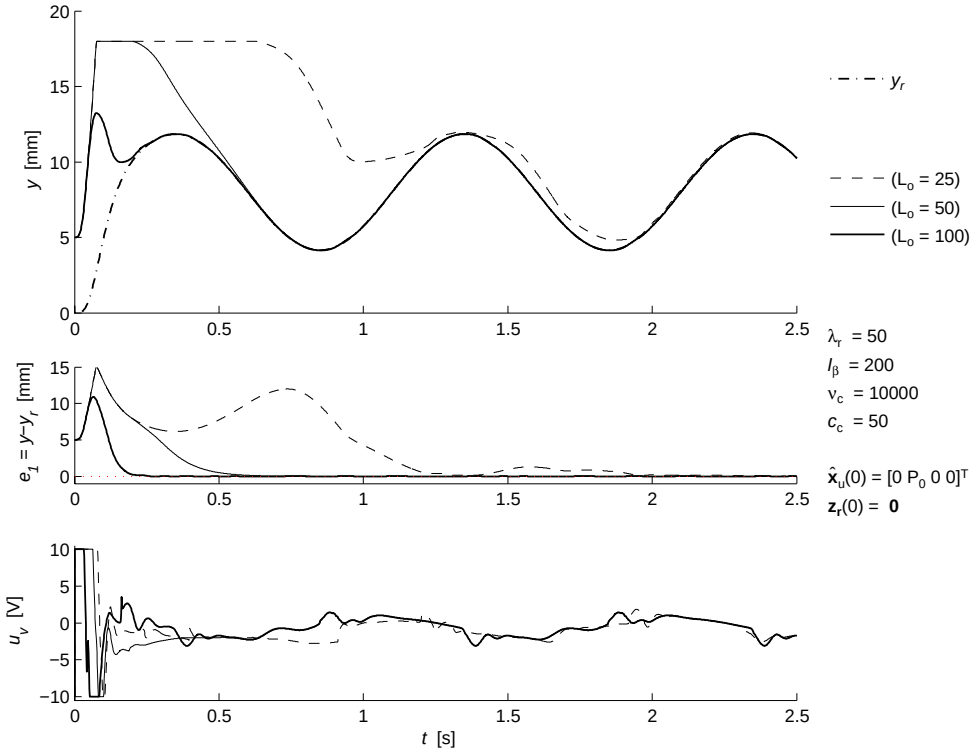

Figure 9.5: Simulations illustrating the effect of observer gains k_1 and k_2, which are given by the observer bandwidth, λ_o.

Figure 9.5 plots the simulated tracking performance for increasing values of λ_o.

With respect to initial controller transients, the effect of increasing λ_o is very similar to the effect obtained when increasing controller gain c_c. With a perfect model, however, the final tracking precision is obviously determined only by c_c and not λ_o (since the observer error converges to zero, $\tilde{\mathbf{x}}_u = 0$). For sufficiently high λ_o (for a given c_c), the controller avoids that the output y of the actuator hits its physical constraint during the initial transients. High λ_o also reduces the undesirable oscillations appearing in the control input during initial transients when the initial observer error $\tilde{\mathbf{x}}_u(0)$ is large.

Dynamic friction compensation

In this subsection, we illustrate the effect that the dynamic friction compensation of the backstepping controller has on the tracking performance. Because of the dynamic friction model

$$\begin{aligned} f_f(v,z) &= D_v v + K_z z + D_z \dot{z} \\ &= D_v v + K_z z + D_z \left(v - \frac{K_z}{F_C} |v|_s z \right) \end{aligned}$$

in the observer, the backstepping controller has an inherent dry friction[1] compensation by design. To illustrate the effect of this dry friction compensation, we implemented a simplified backstepping controller design, replacing the dynamic friction model with the simple static model

$$f_f(v) = D_v v,$$

including viscous friction only.

We use a controller gain $c_c = 50$ and observer bandwidth $\lambda_o = 50$. Otherwise, the parameters and initial conditions are identical with the simulations presented in the two previous subsections.

Figure 9.6 plots the tracking performance with and without dynamic dry friction compensation in the observer used in the backstepping design. The backstepping controller whose design is based on a static viscous friction model, and therefore do not have dry friction compensation, naturally, is not able to track the reference trajectory during a change of direction in $y_r(t)$. In order to be able to track the reference during a change of direction in $y_r(t)$ in the presence of dry (Coulomb) friction, the controller need to anticipate the change in friction by a corresponding pulse in the control input. This is seen from the behavior of the control input for the approximate backstepping controller with a dynamic friction model.

Remark 39 *By increasing the feedback gain c_c and the observer bandwidth λ_o, the backstepping controller design without a dynamic friction model, will to a certain extent attenuate the lack of dry friction compensation. Furthermore, if changes in direction of $y_r(t)$ are slow, then the resulting tracking error due to dry friction becomes smaller compared to faster changes. Hence, for slow tracking tasks, the dynamic friction model may be omitted from the design, without introducing large tracking errors.*

[1] *Dry friction* is also commonly referred to as *Coulomb friction*.

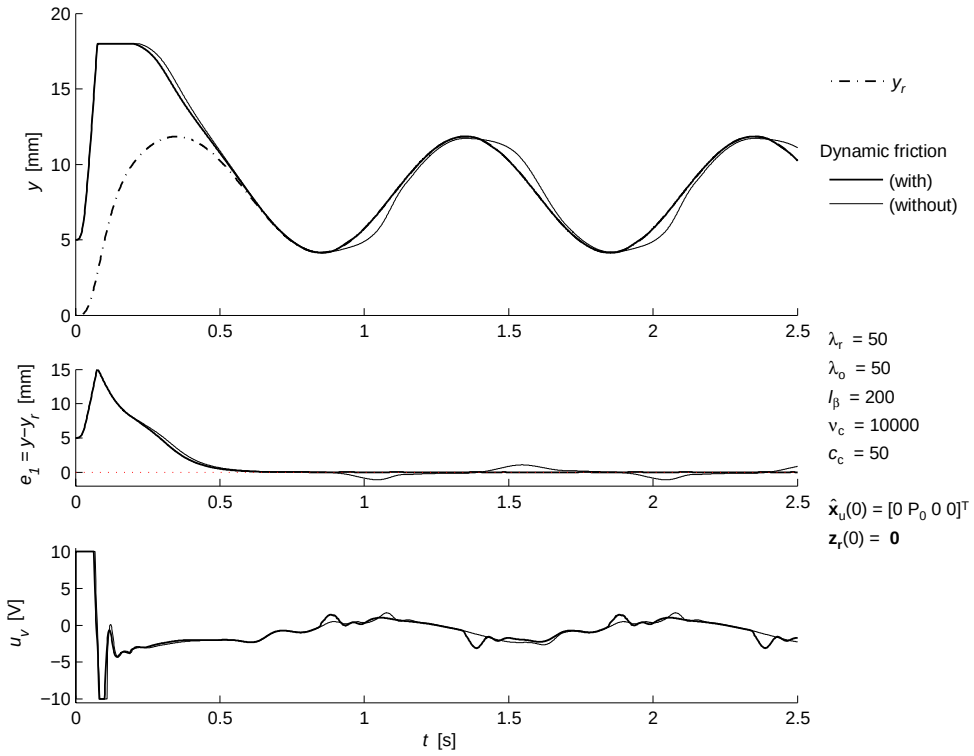

Figure 9.6: Simulations of the controller with and without dynamic dry friction compensation.

9.5 Experimental results

Here we present some characteristic experimental results when the approximate backstepping controller is applied to output-feedback tracking of the electro-pneumatic clutch actuator of the test rig described in Section 2.1.

The observer and controller are implemented with the model parameters summarized in Table 6.1. The design parameters are generally the same as in the simulations, and are also printed in each figure: The reference model is implemented with $\lambda_r = 50$, corresponding to a time-constant $\tau_r = 5/\lambda_r = 0.100\,\mathrm{s}$. The robust reduced-order observer (8.123) is implemented with bounds and smoothing width

$$\begin{aligned} \hat{\mathbf{x}}_{u,\mathrm{ub}} &= \begin{bmatrix} 200\,\mathrm{mm/s} & 995\,\mathrm{kPa} & 1 & 0.5\,\mathrm{mm} \end{bmatrix}^T \\ \hat{\mathbf{x}}_{u,\mathrm{lb}} &= \begin{bmatrix} -200\,\mathrm{mm/s} & 105\,\mathrm{kPa} & -1 & -0.5\,\mathrm{mm} \end{bmatrix}^T \\ \boldsymbol{\varepsilon}_\pi &= \begin{bmatrix} 1\,\mathrm{mm/s} & 1\,\mathrm{kPa} & 0.01 & 0.001\,\mathrm{mm} \end{bmatrix}^T, \end{aligned}$$

and observer gains k_1 and k_2 according to a design bandwidth of $\lambda_o = 100$:

$$\lambda_o = 100 \quad \Rightarrow \quad k_1 = 200 \quad k_2 = 8.9 \cdot 10^3.$$

The feedback gains l_2 and l_3 of the derivative observers (9.55) are implemented with $l_2 = l_3 \triangleq l_\beta = 200$, which corresponds to low-pass filtering with time-constants

$\tau_2 = \tau_3 = 1/l_\beta = 0.005\,\mathrm{s}$. The controller is implemented with the scaling factors $\nu_3 = 10^4$ and $\nu_4 = 10^8$, while we present results using two different gain settings

$$\begin{array}{rlllll} \text{Moderate gain:} & (c_c = 50) & c_1 = 50 & c_2 = 50 & c_3 = 50 & c_4 = 50 \\ \text{High gain:} & & c_1 = 400 & c_2 = 50 & c_3 = 50 & c_4 = 50. \end{array}$$

The logging of an experiment is set up to be triggered after an initialization of the observer, while the reference input $r(t)$ and the state $\mathbf{z}_r(t)$ of the reference model are not initialized at the beginning of a logged experiment. The initial actuator states

$$\mathbf{x}(0) = [y(0), v(0), p(0), x_v(0), z(0)]^T,$$

are different for each experiment, and are printed in the figures in the units [mm], [mm/s], [kPa], [−], [mm], respectively. Like in the simulations, we illustrate the strong convergence properties of the output-feedback controller using the simple choice of initial observer states:

$$\hat{\mathbf{x}}_u(0) = [\hat{v}(0), \hat{p}(0), \hat{x}_v(0), \hat{z}(0)]^T = [0,0,0,0]^T,$$

which due to the lower bound $\hat{p}_{\mathrm{lb}} = P_0$ on the pressure estimate, becomes

$$\hat{\mathbf{x}}^P_{u,\pi}(0) = [\hat{v}(0), \hat{p}(0), \hat{x}_v(0), \hat{z}(0)]^T = [0, P_0, 0, 0]^T,$$

after saturation and projection in the implemented observer.

Sine wave reference input $r(t)$

In this subsection, we validate the performance of the controller with a sinusoidal reference trajectory. The reference trajectory $y_r(t)$ is similar to the simulations in the preceding sections, *i.e.*, it is generated from the reference model using the sine wave input

$$r(t) = R_0 + R_1 \sin\left(\frac{2\pi}{T} \cdot t\right),$$

with bias $R_0 = 8\,\mathrm{mm}$, amplitude $R_1 = 4\,\mathrm{mm}$, and period $T = 1.0\,\mathrm{s}$.

Figure 9.7 illustrates the measured tracking performance of the output-feedback controller for a sinusoidal reference trajectory with the observer and controller gains set according to $\lambda_o = 100$ and $c_c = 50$, respectively. With these observer and controller settings, the control law produces a smooth control input, where the amplification of measurement noise is negligible, and the convergence of the observer and tracking error is fast and monotonic, *i.e.*, without oscillations. Clearly, the tracking performance can be said to be good, with reasonably small tracking errors.

For the given controller setting, the plotted experiment is representative for the tracking performance of the controller when subjected to similar types of reference inputs. That is, with feedback gain $c_c = 50$, tracking is good for reference trajectories with a frequency content in the range up to about $f = T^{-1} = 1\,\mathrm{Hz}$, like the sine wave in Figure 9.7. However, the controller is not able to track faster changes in the reference trajectory without increased tracking errors. In order to ensure

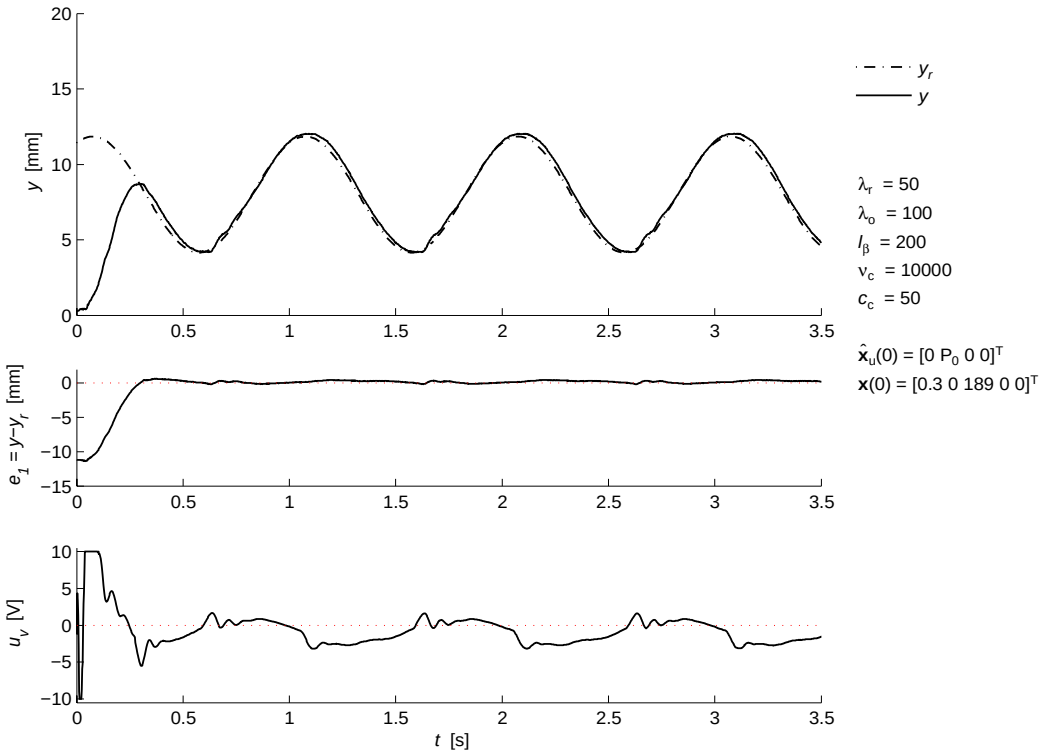

Figure 9.7: Measured tracking performance with feedback gains in the observer and controller gains set according to $\lambda_o = 100$ and $c_c = 50$, respectively.

good tracking with this controller setting for all possible reference inputs $r\,(t)$, the bandwidth of the reference model must be lowered so that it low-pass filters the reference input $r\,(t)$ with a cut-off frequency less than 1 Hz. This is achieved with $\lambda_r = 25$, which gives a time-constant of $\tau_r = 5/25 = 0.200\,\mathrm{s}$, and where the cut-off frequency in Hz is given by

$$f_c = \frac{1}{2\pi\tau_r} = \frac{\lambda_r}{10\pi} = 0.8\,\mathrm{Hz}.$$

In this case, the plotted experiment is representative for the tracking performance of the controller for arbitrary reference inputs $r\,(t)$.

With the reference model implemented with $\lambda_r = 50$, the time-constant is given by $\tau_r = 5/\lambda_r = 0.100\,\mathrm{s}$, giving a cut-off frequency of $f_c = 1.6\,\mathrm{Hz}$. For the controller gain set according to $c_c = 50$, the controller is not able to track the reference $y\,(t)$ for arbitrary inputs in $r\,(t)$ without tracking errors. In the following experiments, rather than reducing the bandwidth of the reference model to $\lambda_r = 25$, we increase the feedback gains in an attempt to achieve acceptable tracking performance for $\lambda_r = 50$. This is done at the expense of higher control effort, which reveals a limitation of the achievable bandwidth of the closed-loop tracking controller caused by unmodeled valve dynamics.

Square wave reference input $r\,(t)$

In this section, we validate the performance of the controller when tracking a filtered square wave reference. That is, with the reference trajectory $y_r\,(t)$ generated from the reference model using the square wave input

$$r\,(t) = R_0 + R_1 \operatorname{sgn}\left[\sin\left(\frac{2\pi}{T}\cdot t\right)\right],$$

with bias $R_0 = 8\,\text{mm}$, amplitude $R_1 = 4\,\text{mm}$, and period $T = 1.0\,\text{s}$.

In the following experiments, we use the controller gains $c_1 = 400$ and $c_2 = c_3 = c_4 = 50$. This controller setting is a compromise between high tracking performance and low control effort.

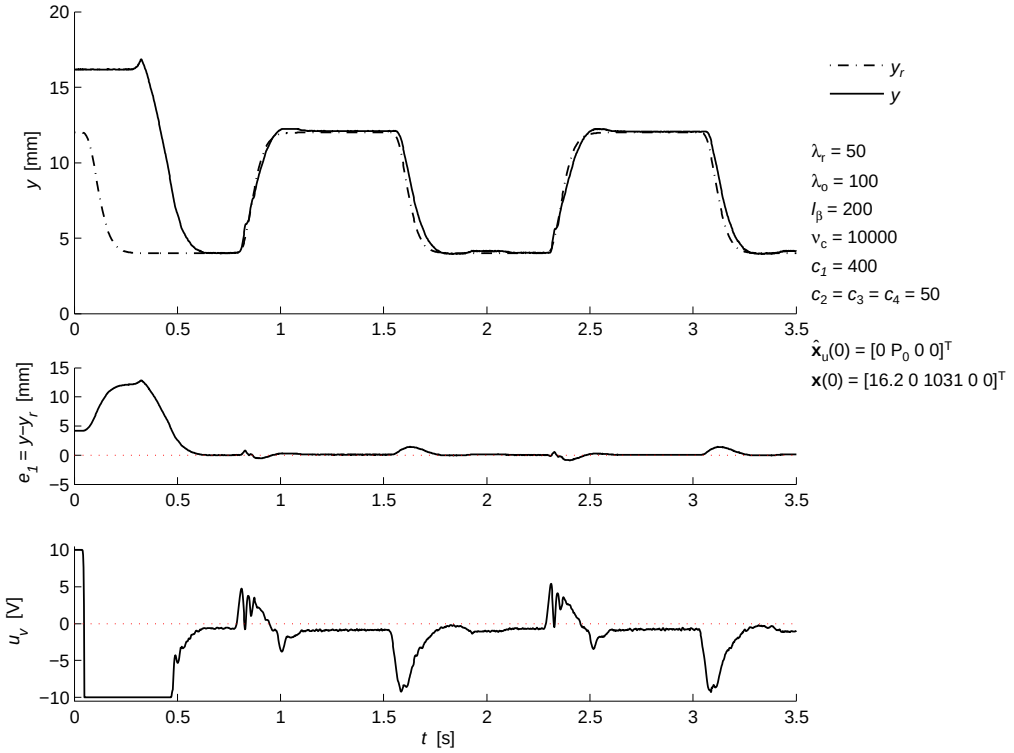

Figure 9.8: Measured tracking performance for a filtered square wave reference trajectory $y\,(t)$.

Figure 9.8 illustrates the performance when attempting to track a filtered square wave trajectory, where tracking errors and high control effort occur during fast changes in $y_r\,(t)$. This experiment is representative for the worst case tracking performance for the approximate backstepping controller with a closed-loop bandwidth set according to $\lambda_r = 50$. The cause for the high control effort is the high feedback gain $c_1 = 400$, which is necessary in order to track $y_r\,(t)$ without unacceptable tracking errors during step changes in $r\,(t)$. Furthermore, by increasing the other gains c_2, c_3, c_4 the tracking performance improves, however, the control effort increases and produces a more aggressive control input.

Notice that the actuator starts initially with a chamber pressure which equals the supply pressure $P_S = 1031\,\text{kPa}$, meaning that the actuator produces the maximum possible actuation force[2]. This again means that the actuator is fully disengaging the clutch, and that the actuator starts initially at its maximum position, $y_{ub} = 16.2\,\text{mm}$. The high initial pressure $p(0) = P_S$ is the reason for the long initial transient period before practical tracking is achieved, which is caused by the large amount of air which have to be evacuated from the actuator chamber.

Step inputs in $r(t)$ produces the fastest change in the tracking reference $y_r(t)$ which is possible for a given reference model setting. The fast changes in $y_r(t)$ is probably a triggering factor for the reduced controller performance during step inputs in $r(t)$, as fast changes in $y_r(t)$ results in fast changes in $\dot{\alpha}_2$ and $\dot{\alpha}_3$ which again results in large errors in their estimates $\hat{\beta}_2$ and $\hat{\beta}_3$. As indicated by Proposition 37 (page 182) and simulations, these errors are attenuated for sufficiently high feedback gains c_1, c_2, c_3 and c_4. However, in a practical implementation, measurement noise, or unmodeled actuator or sensor dynamics, will limit the maximum implementable feedback gain. Upon examination of the observer performance, the limiting factor on maximum achievable bandwidth for this application, appears to be unmodeled valve dynamics which is excited by high-frequency control inputs.

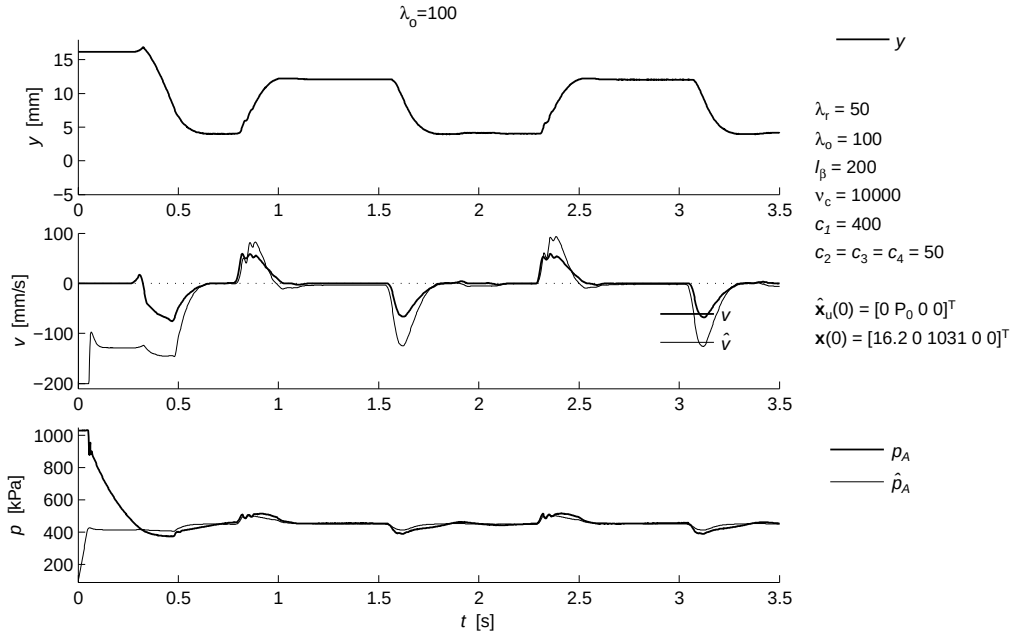

Figure 9.9: Measured closed-loop observer performance with the approximate backstepping controller subjected to a filtered square wave reference trajectory $y(t)$.

Figure 9.9 plots the corresponding observer performance for the experiment plotted in Figure 9.8. During step changes in $r(t)$, where the control law produces fast

[2]The strange peak in the position just before the actuator starts to move is not the actual movement of the actuator, but an error in the measurement due to an elastic deformation of the sensor bracket, which occurs when the actuator push towards the physical constraint of the clutch with excessive force.

changes in the control input u_v, the estimated velocity $\hat{v}$ exhibits an overshooting estimate which is caused by errors in the pressure estimate $\hat{p}$. This error can be traced to unmodeled valve dynamics, as the spool position of the valve experiences a slowly converging overshoot after a fast change in u_v, which obviously, is not captured by the simple linear model of the valve dynamics used by the observer.

Considering the above analysis, we make the following conclusion:

Remark 40 *The maximum bandwidth of the tracking controller with which we are able to achieve acceptable tracking performance for arbitrary reference inputs $r(t)$, is with the poles of the reference model placed at $s = -\lambda_r$, with $\lambda_r = 50$. This corresponds to a time-constant of $\tau_r = 0.100\,\mathrm{s}$ for the closed-loop system. For higher λ_r the observer performance degrades considerably due to the excitation of unmodeled valve dynamics for high-frequency input $u_v(t)$. Since high feedback gains increases the control effort, which again increases the frequency content of the input, thus, the excitation of the unmodeled valve dynamics, the disturbance due to observer errors can not be attenuated using high gain without introducing chattering in the control input u_v.*

In the following sections, we further validate the performance of the approximate backstepping controller implemented with the maximum bandwidth $\lambda_r = 50$, and feedback gains $c_1 = 400$ and $c_2 = c_3 = c_4 = 50$.

Arbitrary reference input $r(t)$

Using a potentiometer as a joystick, we can control the reference input $r(t)$ manually in order to illustrate the tracking performance for the approximate backstepping controller when subjected to arbitrary reference inputs.

Figure 9.10 and 9.11 plots the tracking performance and the corresponding observer performance for an arbitrary reference input $r(t)$, generated manually by hand. Notice the tracking errors which occur for fast changes in $y_r(t)$, which are primarily caused by observer errors due to excitation of the unmodeled valve dynamics, as discussed above.

Typical clutch sequence reference input $r(t)$

A typical clutch sequence during gear-shift, consists of the following sequence:

Full disengagement of the clutch: A gear-shift starts with a full disengagement of the clutch to disconnect the engine from the drive line. This disengagement is usually performed as fast as possible.

Disengagement during gear-shift: The clutch is held in fully disengaged position during the change of transmission.

Engagement to the slip-point: The engagement to reach the point where the clutch starts to transfer torque — the slip-point of the clutch — is usually performed in a precise and relatively fast motion.

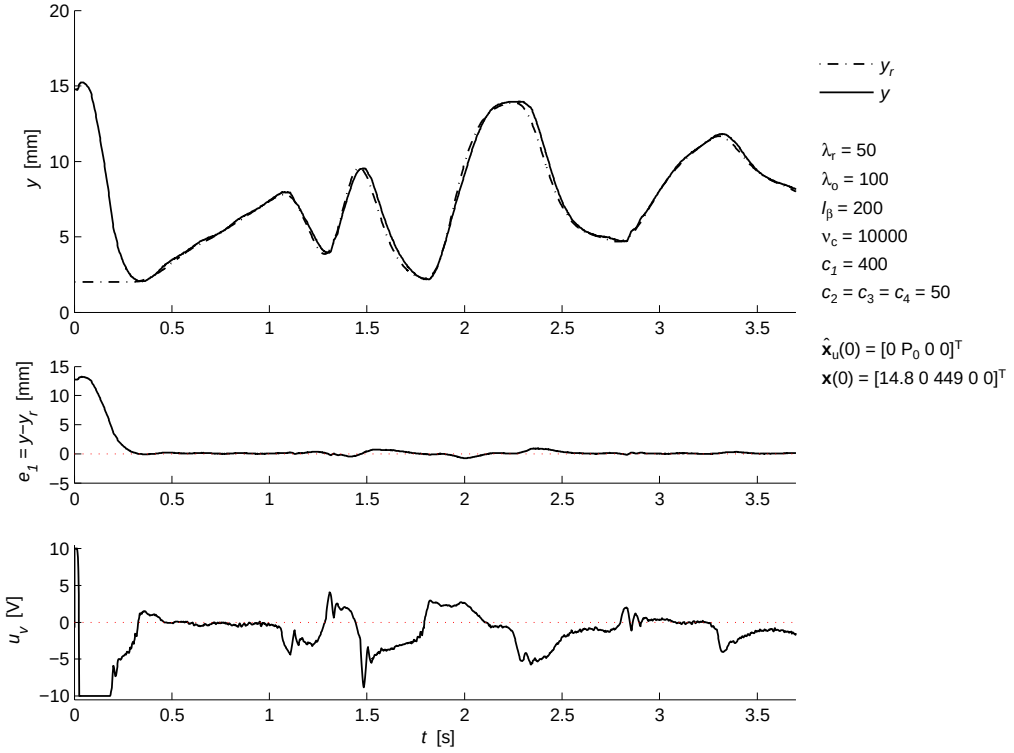

Figure 9.10: Measured tracking performance of the approximate backstepping controller subjected to a reference trajectory $y_r\,(t)$ generated from an arbitrary reference input $r\,(t)$.

Engagement through the slip-point: Once the slip-point is reached, it is followed by a slow engagement to ensure a smooth torque transfer from the engine to the drive line.

Remaining full engagement of the clutch: When the clutch engagement has reached the point where the required torque is transferred by the clutch and the clutch disks locks up, the remaining full engagement of the clutch is performed as fast as possible.

Figure 9.12 and 9.13 illustrate the tracking performance and the corresponding observer performance for a reference input $r\,(t)$ representing a typical clutch sequence during gear-shift.

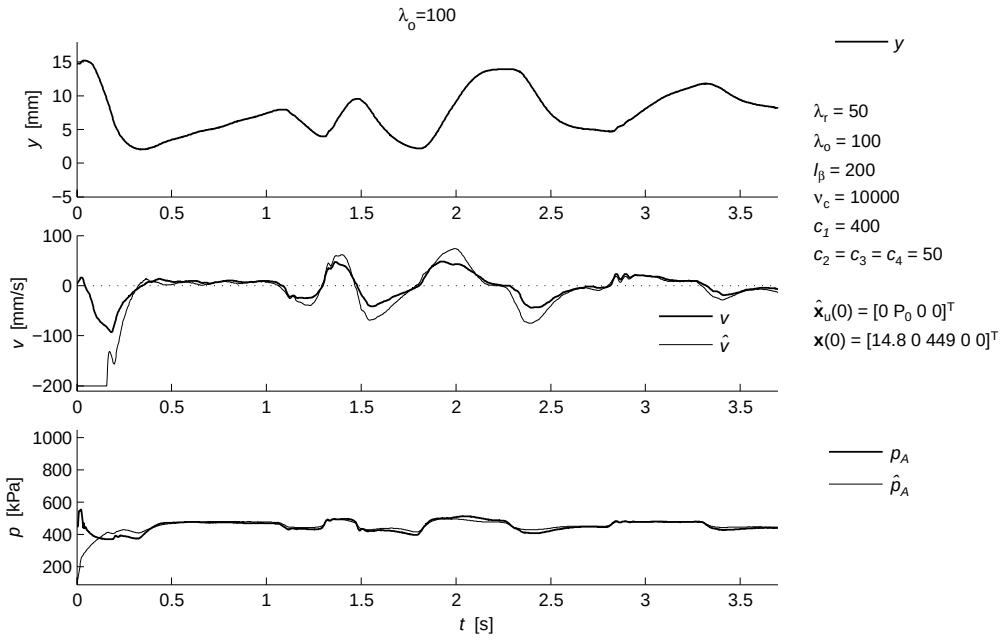

Figure 9.11: Measured observer performance in closed-loop with a reference trajectory $y_r(t)$ generated using a potentiometer as a joystick to produce a manual reference input $r(t)$.

9.6 Summary

Based on a reduced-order observer for the electro-pneumatic actuator, a robust output-feedback tracking controller is designed by a recursive observer-based backstepping procedure in four steps. The backstepping design is simplified for the two last steps by using high-gain observers to estimate, rather than calculate analytically, the derivative of the stabilizing function designed at the previous steps. The approximate backstepping controller achieves exponential practical tracking within a prescribed tracking precision, where an arbitrary small precision is achieved by sufficiently high feedback gain in the observer and controller. The controller is robust to bounded disturbances (*e.g.* modeling errors) appearing additively in the system dynamics. Combined with a robust re-design of the observer using smooth saturation and projection of the estimates to constrain the estimated states to the region of validity of the design model, these strong properties of the controller hold globally.

The controller is in essence tuned according to four main design parameters: The observer gains are set according to the design bandwidth λ_o, and the feedback gains and scaling of control law according to c_c and ν_c. The parameters of the reference model are determined according to the design bandwidth λ_r, which is also viewed as the design bandwidth of the closed-loop tracking controller since it determines the time-constant τ_r of the tracking of the reference input r.

The experimental results show that the maximum achievable bandwidth of the controller is limited by unmodeled valve and pressure dynamics. The controller

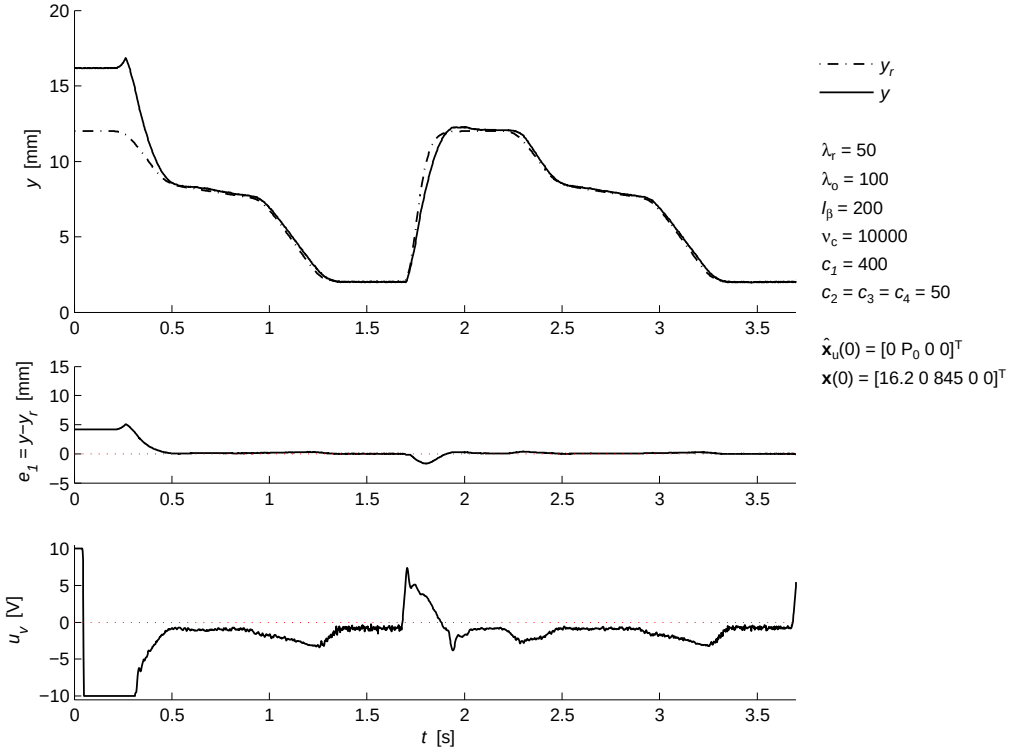

Figure 9.12: Measured tracking performance of the approximate backstepping controller when subjected to a reference trajectory $y\,(t)$ representing a typical clutch sequence for a gear-shift.

achieves accurate tracking for arbitrary reference inputs r for design bandwidths $\lambda_r < 50$, where $\lambda_r = 50$ corresponds to a time-constant of $\tau_r = 0.100\,\mathrm{s}$.

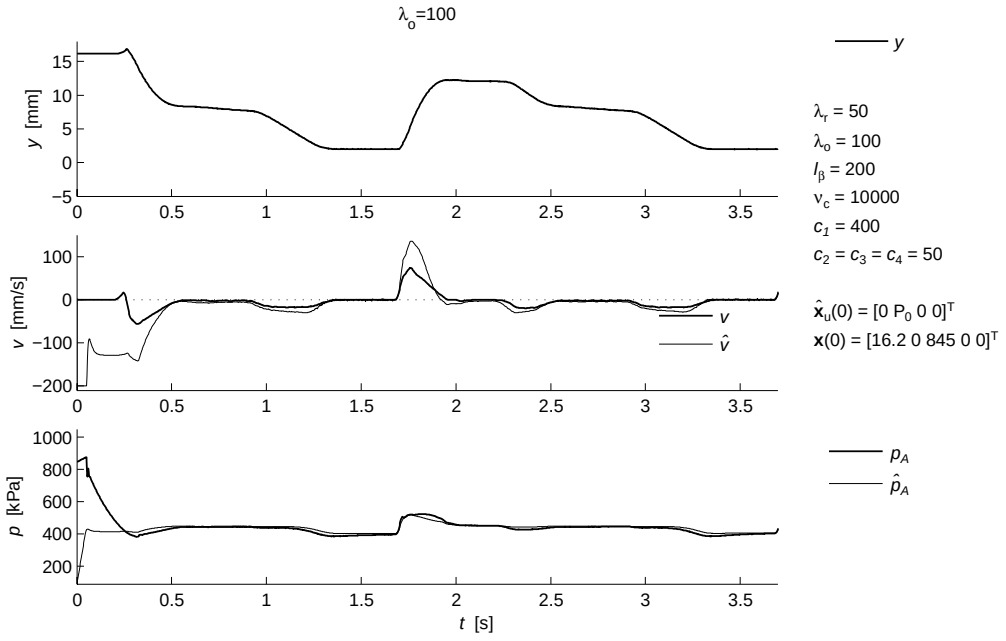

Figure 9.13: Measured observer performance in closed-loop with a reference trajectory $y(t)$ representing a typical clutch sequence for a gear-shift.

Part III

Thesis Conclusions

Chapter 10

Conclusions

This thesis deals with the design of an position tracking controller for a single-acting pneumatic cylinder actuator operated by a three-way proportional valve, using only position feedback. The considered applications are clutch actuation in automated manual transmissions (AMT), and clutch-by-wire (CBW) solutions, in heavy-duty trucks. Particularly, the latter application requires precise and fast position tracking.

For electro-pneumatic actuators in general, the main obstacles with respect to control design, are the high compressibility of air in combination with low damping, and nonlinear flow and friction characteristics. Additionally for the clutch application, the actuator operates against a highly nonlinear clutch compression spring, which constitutes the main load of the actuator. The lack of sensors to provide state measurements for feedback, is another severe contraint on the control system, which further complicate the design.

State of the art

The introduction of Chapter 1, is one of the main contributions of the thesis, providing a survey of the most relevant literature on the modeling and control of electro-pneumatic actuators, and reviews selected nonlinear control theory not yet applied to pneumatic actuators.

The nature of pneumatic actuators is well understood, and accurate models describing their dynamics exists. Still, there are untapped potentials of improvements with respect to aspects regarding parameter estimation and control design. That is, the most accurate models are not particularly suited for control and parameter estimation, *i.e.*, they are in forms which exclude the application of existing solutions for nonlinear output-feedback and adaptive control available in the literature.

The literature related to control of pneumatic actuators applied for clutch actuation is scarce, with results reported mainly by two researchers. On the other hand, the literature on control of pneumatic actuators, in general, is relatively rich, with several research groups doing research on the topic the last two decades.

With a few exceptions, the results from this research on control of pneumatic actuators, address full state-feedback control, *i.e.* , assuming all states being measured. The introduction of this thesis provides an overview of these results on state-feedback control strategies applied to pneumatic actuators, grouped into three

main categories where the essence of the solutions in each category is extracted and described in detail.

Only a few papers address output-feedback control, or observer design for pneumatic actuators, where most results are primarily based on linear control theory. The reason for the few results on output-feedback control of pneumatic actuators, is likely due to the lack of constructive procedures for designing observers for nonlinear systems in general. However, for certain classes of nonlinear systems, solutions for contructive observer design do exist. The largest of these classes being feedback linearizable systems which comprises pneumatic actuators, where a constructive high-gain observer design applies. In the last part of the introduction of this thesis, we provide a review of the main results in the nonlinear output-feedback control literature that is applicable to pneumatic actuators.

Modeling

Chapters 2–6 of Part I of the thesis, address the modeling of the system in the context of nonlinear and adaptive control. That is, besides accuracy, emphasis is laid on deriving models that have particular properties like smoothness and parameter-affinty. Smoothness makes the resulting model fully feedback linearizable, thus, applicable for existing solutions for output-feedback control using high-gain observers, while parameter-affinity, facilitates parameter estimation and the design of adaptive controllers. Main contributions in this regard, is the unified treatment of the modeling of the complete electro-pneumatic clutch actuation system (which until this has been missing in the literature), and improvements of particular models of pneumatic actuators. In particular, Chapter 5 on flow modeling provides a refined review of existing models, and proposes new models with improved, or optimalized, properties with respect to parameter-affinity, applicability to nonlinear control, and accuracy.

Observer design

In Chapter 8, the lack of sensors to provide full state-feedback, is solved by designing nonlinear observers to estimate the unmeasured states based on the measured position. Rather than applying the existing constructive design procedure for high-gain observers that require a transformation to the normal form, the thesis pursues the more novel approach of designing a nonlinear observer in its original states, based on insight to the inherent stability properties of pneumatic actuators. The main contribution of the thesis, is the work presented in this chapter on the design of simple nonlinear observers for pneumatic actuators, where it is shown that by understanding the dynamical properties of pneumatic actuators, simple nonlinear observers can be designed by utilizing the inherent stabilizing nonlinearities of the pneumatic actuator. The main advantages of this approach is that the observer states are the actual physical states rather than the transformed states (*e.g.* pressure rather than acceleration); that the observer does not require high-gain to guarantee stability; and that the observer is compatible with an output-feedback control design by an observer-backstepping approach. Two specific observer designs are proposed. Their

most important properties are demonstrated by simulations, and their performance is validated experimentally, both with open-loop and closed-loop control inputs.

Control design

The high compressibility, combined with the low damping, and the significant nonlinearities in the system, makes high-performance tracking difficult, or even impossible, without model-based nonlinear feedforward compensation of some sort.

In Chapter 9, a model-based nonlinear tracking controller is designed by a recursive observer-based backstepping procedure in four steps, based on the nonlinear observer design of Chapter 8. In each step of the design, the nonlinearites are cancelled, *i.e.*, implementing model-based nonlinear feedforward compensation, and replaced with tunable feedback terms to give the desired closed-loop dynamics. This type of backstepping designs can be referred to as a cancelling backstepping design. In contrast to the state-feedback case, where cancellation of stabilizing nonlinearities may result in a controller which is not robust to model errors, this observer-based cancellation design is robust because it is applied to the exactly known observer dynamics.

In addition to a design based on exact backstepping using the exact derivatives of the stabilizing functions, an alternative approximate backstepping design is also presented, where the backstepping procedure is simplified for the last two steps using high-gain observers to estimate, rather than calculate analytically, the derivative of the stabilizing function designed at the previous step.

The exact and approximate backstepping controllers are shown to possess some strong theoretical stability properties: The (approximate) backstepping controller achieves exponential (practical) tracking according to a prescribed tracking precision, which can be made arbitrary accurate by sufficiently high feedback gain in the observer and controller. The controller is also shown to be robust to bounded modeling errors.

The output-feedback backstepping design was implemented and validated experimentally on the test rig, demonstrating extremely high tracking performance. The parameters of the reference model are determined according to a design bandwidth λ_r, which can be viewed as the design bandwidth of the closed-loop tracking controller since it determines the time-constant τ_r of the tracking of the reference input r. With the bandwidth of the reference filter chosen lower than the maximum achievable bandwidth of the controller, the controller achieves accurate tracking of the filtered reference trajectory y_r for arbitrary reference inputs r, *i.e.*, with tracking errors in the order of magnitude of the sensor accuracy. The maximum achievable bandwidth of the controller, which is limited by unmodeled valve and pressure dynamics, was found experimentally to correspond to a reference filter with bandwidth $\lambda_{r,\max} \approx 50$, corresponding to a time-constant $\tau_r = 0.100\,\mathrm{s}$ for this 5th-order reference model. For comparison, a properly tuned PID controller, requiring the same accuracy, is able to achieve tracking according to a time-constant $\tau_r = 1.500\,\mathrm{s}$, that is, 15 times slower tracking.

10.1 Discussions

This thesis has addressed the main problems related to the design of a control system for electro-pneumatic clutch actuation, where the emphasis has been divided between the modeling of the system dynamics, and the design of observer and controller for clutch actuation. A key factor has been the application of recent advances from research on nonlinear control theory.

Because of this relatively wide scope, there are several aspects of the work that have been covered unsatisfactory in this thesis. To show that we are aware of (some of) them, we briefly summarize important weaknesses with our work, and potential improvements, below:

Model validation and parameter estimation: Much experimental work has been conducted during the doctoral period to provide the basis for Part I of the thesis on mathematical modeling. However, only a fraction of this work has been included in the thesis, primarily because similar work can be found in the litterature on pneumatic actuators.

Comparison with existing output-feedback control strategies: A weak point of the thesis, is the lack of experimental work that compares conventional control strategies for pneumatic actuators with the observer-based backstepping design proposed in this thesis. A main reason for this is that the few output-feedback control strategies that exists for pneumatic actuators, cause severe oscillations when implemented on the clutch actuation system, primarily because they are based on linear techniques which do not handle the clutch spring nonlinearity. An exception is the simple PID controller, which when tuned sufficiently slow, *i.e.*, using sufficiently low gain, avoids oscillations.

Comparison with existing state-feedback control strategies: An interesting analysis would be an experimental comparison of the output-feedback backstepping controller, with the existing full state-feedback control strategies for pneumatic actuators, *i.e.*, the linear PVA controller, the feedback linearizing controller, and the sliding mode controller, described in the literature review of Chapter 1.

Observer validation by existing state-feedback control strategies: By implementing experimentally the proposed observers combined with existing state-feedback controllers, the validity of the proposed observer design, and the contribution of the thesis with respect to nonlinear observer design and output-feedback control for pneumatic actuators, would be better emphasized.

Simplifications of the proposed backstepping controller: A main drawback with the proposed backstepping controller, is its high complexity which makes implementation time-consuming. With emphasis on simplicity, and avoiding cancellation of stabilizing nonlinearities, the complexity of the controller has a significant potential of improvement.

Experimental analysis of robustness: A thorough analysis of the robustness of the proposed control system with respect to different types of clutch characteristics, different types of control valves, engine vibrations, wear, *etc.*, would better demonstrate the validity of the proposed control system.

10.2 Further research

In this section, we briefly outline ongoing and future research that are directly related to the work presented in this thesis.

The research on nonlinear observer design is currently being refined into a journal paper, planned submitted by January 2006. The work on a simplified non-cancelling backstepping design that avoids cancellation of the inherent stabilizing nonlinearities of the pneumatic clutch actuator, will be published in another journal paper, planned submitted by June 2006. Related ongoing research which will be continued in the future, is backstepping design applied for clutch actuation using on-off valves instead of proportional valves, adaptive backstepping and adaptive observer designs for online estimation of the clutch load characteristic to compensate for wear, and online estimation of the valve dynamics of on-off valves to compensate for changes in temperature.

Other more general, but related topics for further research, are improvements of backstepping in directions of simpler control laws, more relaxed (and optimal) control laws for reduced control effort, and handling of rate and magnitude constraints of states. Discrete-time implementation has not been an issue in this thesis, where a continuous-time design was implemented using a standard Runge-Kutta solver with a relatively high sampling frequency of 2 ms. This approach is justified by the assumption that computation is cheap, however, this is not always sufficient. The number of control systems in vehicles increases strongly, hence, in order to avoid an increase in total cost, an important objective is a reduction in the computational load, and required sample frequency of each individual control system on the vehicle. For the electro-pneumatic clutch actuation control system, a reduction of the sample frequency, *e.g.* from 2 ms to 10 ms, dramatically increases the demands on the control system, particularly in the case of output-feedback. Research on output-feedback control, in directions of discrete-time nonlinear observer design, and discrete-time backstepping design, is therefore of particular interest.

Part IV

Appendices

Appendix A

Modeling of Static Nonlinearities

Most existing solutions for adaptive control are restricted to systems with parametric uncertainties which are linearly parametrizable in the uncertain parameters, that is, the uncertain static nonlinearity can be expressed in a parameter-affine form. In general, it is advantageous for parameter identification that the parameters of a model appear in an affine form, because then, the task of identifying the set of parameters which provides the best fit to a nonlinear function, amounts to solving a linear convex optimization problem. Consequently, from a parameter identification point of view, it will often be preferable to model static nonlinearities by parameter-affine models, and even substitute complicated nonlinear functions in the system model with approximations that are parameter-affine.

In general, a smooth nonlinear function can be modeled, or approximated, by a weighted sum of simple basis functions, where increased complexity, simply requires a larger number of basis functions in order to meet a prescribed accuracy. This is for example exploited in neural network (NN) models, which are usually composed of a large number of simple basis functions. With respect to modeling for control, it may be of interest to exploit the known structure of nonlinearities in order to construct customized basis functions where the number of basis functions, hence, number of parameters, are minimal. In this thesis, we employ simple bell-shaped basis functions to obtain parameter-affine models of the nonlinear load characteristic of the clutch as a function of position (see Section 3.1), and the flow conductance characteristic as a function of spool position of each port of the pneumatic proportional valve (see Section 5.4). Furthermore, we develop two customized basis functions to obtain a parameter-affine equation for the flow rate through pneumatic restrictions (see Section 5.2).

This appendix addresses the empirical modeling of smooth static nonlinearities, using bell-shaped Gaussian and B-spline basis functions. In Section A.1, the general formulation of parameter-affine models is briefly addressed, and some notions of *support* and *support width* are made precise. In Section A.3, the properties of the Gaussian and the normalized Gaussian basis functions, and in Section A.4, the properties of the B-Spline basis functions, are explored with respect to their capability to approximate nonlinear functions.

Most of the theory on neural network models are taken from the textbook by

Nelles, [67]. The neural networks community have adopted their own terminology, while in this appendix, we use the terminology from the field of system identification (and statistics). For a reference on the theory of B-splines, see for *e.g.* [19, Chapter 7].

A.1 Parameter-affine models

In general, a scalar multivariable function, $y = f(\mathbf{x})$, can be modeled in the parameter-affine form

$$\hat{y} = \theta_1 \phi_1(\mathbf{x}) + \theta_1 \phi_1(\mathbf{x}) + \cdots + \theta_p \phi_p(\mathbf{x}), \tag{A.1}$$

which can compactly be expressed in vector form as

$$\begin{aligned} \hat{y} &= \begin{bmatrix} \theta_1 & \theta_2 & \cdots & \theta_p \end{bmatrix} \begin{bmatrix} \phi_1(\mathbf{x}) \\ \phi_2(\mathbf{x}) \\ \vdots \\ \phi_p(\mathbf{x}) \end{bmatrix} \\ &\Updownarrow \\ \hat{y} &= \boldsymbol{\theta}^T \cdot \boldsymbol{\phi}(\mathbf{x}), \end{aligned} \tag{A.2}$$

where $\mathbf{x} \in \mathcal{X}_0 \subset \mathbb{R}^m$ is a vector of m input variables, $\hat{y} \in \mathcal{Y} \subset \mathbb{R}$ is the modeled output, and $\boldsymbol{\theta} \in \Omega_\theta \subset \mathbb{R}^p$ is a parameter vector. The regressor elements $\phi_i(\mathbf{x})$, $i = 1, 2, \cdots p$ are called basis functions, which can be of various forms. Let the input set $\mathcal{X}$ be defined so that it contains all possible inputs $\mathbf{x}$ of practical interest, and the output set $\mathcal{Y}$ such that it contains all possible outputs for $\mathbf{x} \in \mathcal{X}$. Furthermore, for a given set of basis functions $\boldsymbol{\phi}(\mathbf{x})$, the parameter vector $\boldsymbol{\theta}$ is constrained to be in the in the set Ω_θ.

A.2 Basis functions

In this section we describe some general properties of basis functions, while in following Sections A.3 and A.4, we explore the modeling properties with two types of bell-shaped basis functions, the Gaussian and B-spline basis functions, respectively.

To better be able to describe the properties of basis functions, we introduce the notion of *support* of a scalar function $\phi(\mathbf{x})$, and for the single-input case, the notion of *support width* for $\phi(x)$. A scalar function $\phi(\mathbf{x})$ is characterized by its *support* on $\mathcal{X}$, which we try to make precise by the following definition:

Definition 41 (Support) *A scalar function $\phi_1(\mathbf{x}) \in \mathbb{R}$, $\mathbf{x} \in \mathcal{X} \subset \mathbb{R}^m$ where $\mathcal{X}$ is the input region of interest, is said to have*

- local support *on $\mathcal{X}$ if it is non-zero only on a compact subset $\mathcal{X}_1 \subset \mathcal{X}$, hence, it is zero on its compliment $\mathcal{X} - \mathcal{X}_1$:*

$$\begin{aligned} \forall \mathbf{x} \in \mathcal{X}_1 &\implies |\phi(\mathbf{x})| > 0 \\ \forall \mathbf{x} \in (\mathcal{X} - \mathcal{X}_1) &\implies \phi(\mathbf{x}) = 0. \end{aligned}$$

- global support *on* $\mathcal{X}$ *if it is non-zero on* $\mathcal{X}$*, except for some singular points* $\mathbf{x}^*$*:*

$$\forall \mathbf{x} \in \mathcal{X} - \{\mathbf{x}^*\} \Longrightarrow |\phi(\mathbf{x})| > 0.$$

- exponentially local support *on* $\mathcal{X}$*, if it tends asymptotically to zero with increasing distance to a finite subset* $\mathcal{X}_1 \subset \mathcal{X}$*:*

$$\operatorname{dist}(\mathbf{x}, \mathcal{X}_1) \to \infty \Longrightarrow \phi(\mathbf{x}) \to 0,$$

where $\operatorname{dist}(\mathbf{z}, \mathcal{X}_1) \triangleq \inf_{\mathbf{x} \in \mathcal{X}_1} |\mathbf{z} - \mathbf{x}|$.

The above definitions applies for scalar functions $\phi(\mathbf{x}) \in \mathbb{R}$ with multiple inputs $\mathbf{x} \in \mathbb{R}^m$. In this appendix, we will mainly consider single-input basis functions, $\phi(x)$ with $x \in \mathbb{R}$.

Activation functions

In neural networks, the basis functions are usually constructed from a simple single-variable function $g(x)$, called an *activation function*. A typical choice of activation function is the Gaussian function

$$g(x) = \mathrm{e}^{-\frac{1}{2}x^2}, \tag{A.3}$$

where the basis functions $\phi_i(u)$ are often defined using a scaling of the input according to $u \triangleq w \cdot x = 1/\sigma^2 \cdot x$, producing a Gaussian distribution with standard deviation σ. Another activation function commonly used in neural networks, is the inverse multi-quadratic function

$$g(x) = \frac{a}{\sqrt{x^2 + a^2}}. \tag{A.4}$$

Both the exponential activation functions are plotted in Figure A.1.

Single-variable input construction

A set of scalar basis functions $\phi_i(x) \in \mathbb{R}$, $i = 1, 2, \cdots, p$ can be obtained using an activation function, like the Gaussian function (A.3), according to

$$\phi_i(x) = g(u_i(x)), \tag{A.5}$$

where each u_i is constructed from the variable x as

$$u_i = w_i \cdot (x - c_i). \tag{A.6}$$

For the ith basis function $\phi_i(x)$, the offset weight c_i determines the center position of $\phi_i(x)$, while the scaling weight w_i determines the width of the function. Following this approach, a set of basis functions $\boldsymbol{\phi}(x) = \left[\phi_1(x), \phi_2(x), \cdots, \phi_p(x)\right]^T$ are completely described by a center vector $\mathbf{c} = [c_1, c_2, \cdots, c_p]^T$, a scaling vector $\mathbf{w} = [w_1, w_2, \cdots, w_p]^T$, and the characteristics of the activation function $g(\cdot)$.

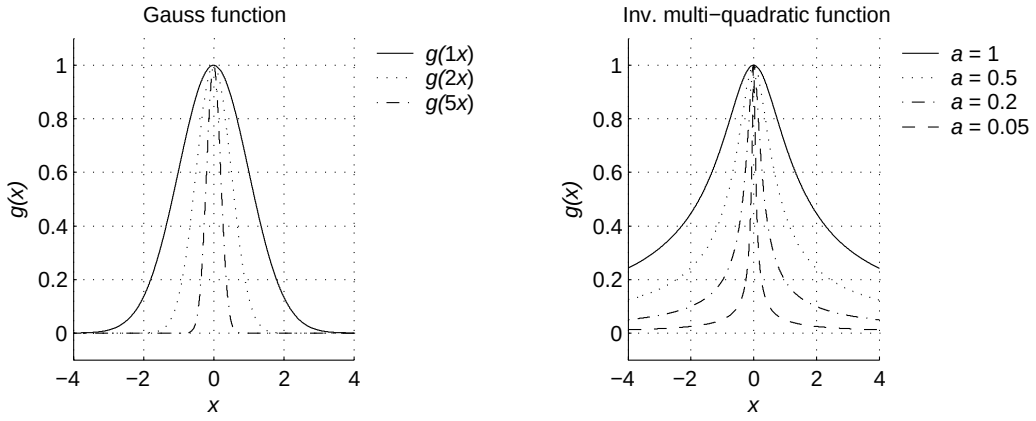

Figure A.1: Left: The Gauss function for different scaling of the input signal x, representing normal distributions with standard deviations $\sigma = \left\{1/\sqrt{1}, 1/\sqrt{2}, 1/\sqrt{5}\right\} \approx \{1, 0.71, 0.45\}$. Right: The inverse multi-quadratic function for $a = \{1.0, 0.5, 0.2, 0.05\}$.

Remark 42 *We may usually regard the parameter vectors* $\mathbf{c}$ *and* $\mathbf{w}$ *as pre-determined and fixed parameters of the basis functions, so that the nonlinear function modeled by (A.2), can be viewed to be affine in its tunable parameters.*

Multidimensional basis functions may be realized either by multivariable input constructions, or by tensor product constructions, briefly outlined in the following subsections.

Multi-variable input-constructions

A multivariable basis function $\phi_i(\mathbf{x})$ may be constructed from a scalar activation function $g(\cdot)$, according to

$$\phi_i(\mathbf{x}) = g(u_i(\mathbf{x})), \tag{A.7}$$

where $\mathbf{x} = [x_1, \cdots, x_m]^T$ is a vector of input variables, and the scalar input function $u_i = u_i(\mathbf{x})$ can be constructed from $\mathbf{x}$ in various ways.

The scalar function input u_i can for example be realized by the *ridge input construction*

$$\begin{aligned} u_i &= \mathbf{w}_i^T \cdot (\mathbf{x} - \mathbf{c}_i) \\ &= w_{i1} \cdot (x_1 - c_{i1}) + \cdots + w_{im} \cdot (x_m - c_{im}), \end{aligned} \tag{A.8}$$

where $\mathbf{x} \in \mathbb{R}^m$ is the input vector, $\mathbf{c}_i \in \mathbb{R}^m$ is the center vector which determines the distance to the origin, and $\mathbf{w}_i$ is the scaling vector which determines the slopes in each direction of the input variable space $\mathbb{R}^m$. For this type of input construction, the scaling vector $\mathbf{w}_i$ points toward the direction of nonlinearity of the basis function, while orthogonal to $\mathbf{w}_i$ the basis function stay constant.

The scalar function input u_i can alternatively be realized by a *radial input construction*

$$u_i = |\mathbf{x} - \mathbf{c}_i|_{\Sigma_i}, \tag{A.9}$$

where the vector norm $|\mathbf{z}|_{\Sigma_i} \triangleq \sqrt{\mathbf{z}^T \boldsymbol{\Sigma}_i \mathbf{z}}$ determines the slopes in each direction of the input parameter space $\mathbb{R}^m$. The scaling matrix $\boldsymbol{\Sigma}_i$ is often taken as a diagonal matrix, where each diagonal element then determines the slope, or scaling in each direction. The resulting basis functions $\phi_i(\mathbf{x})$, are usually referred to as radial basis functions (RBF), often without weighting, *i.e.*, with $u_i = |\mathbf{x} - \mathbf{c}_i|$. An early application of the radial basis functions were in mathematics as multivariable interpolation functions, see *e.g.* [76].

Tensor product constructions

Another alternative for the construction of multivariable basis functions is by forming the tensor product. That is, the ith basis function $\phi_i(\mathbf{x})$, is constructed according to

$$\phi_i(\mathbf{x}) = g_{i1}(u_{i1}) \cdot g_{i2}(u_{i2}) \cdot \cdots \cdot g_{im}(u_{im}), \quad i = 1, 2, \cdots, p, \tag{A.10}$$

where $\mathbf{x} = [x_1, x_2, \cdots, x_m]^T \in \mathbb{R}^m$ is the vector of input variables, and $g_{ij}(\cdot)$ is a scalar function of the jth input variable x_j for the ith basis function $\phi_i(\mathbf{x})$. The scalar inputs $u_{i1}, u_{i2}, \cdots, u_{im}$ for the ith basis function $\phi_i(\mathbf{x})$ can be constructed according to

$$\begin{aligned} u_{i1} &= w_{i1} \cdot (x_1 - c_{i1}) \\ u_{i2} &= w_{i2} \cdot (x_2 - c_{i2}) \\ &\vdots \\ u_{im} &= w_{im} \cdot (x_m - c_{im}), \end{aligned} \tag{A.11}$$

and alternatively, by a radial input construction. In the simplest case, a complete set of basis functions can be constructed using the same activation function $g(\cdot)$, *e.g.*,

$$\phi_i(\mathbf{x}) = g(u_{i1}) \cdot g(u_{i2}) \cdot \cdots \cdot g(u_{im}), \quad i = 1, 2, \cdots, p, \tag{A.12}$$

which is usually the case in neural networks.

A.3 Gaussian basis functions

Any smooth nonlinear function

$$y = f(x), \tag{A.13}$$

can be approximated with the parameter-affine model

$$\hat{y} = \boldsymbol{\theta}^T \cdot \boldsymbol{\phi}(x), \tag{A.14}$$

where the vector $\boldsymbol{\phi}(x) = [\phi_1(x), \phi_2(x), \cdots, \phi_p(x)]^T \in \mathbb{R}^p$ is a set of Gaussian basis functions, defined according to

$$\phi_i(x) = \mathrm{e}^{-\frac{1}{2} w_i^2 (x - c_i)^2}, \tag{A.15}$$

where $\mathbf{w} = [w_1, w_2, \cdots, w_p]^T$ is a vector of scaling parameters, and $\mathbf{c} = [c_1, c_2, \cdots, c_p]^T$ a vector of center offset parameters.

When choosing scaling and center parameters w_i and c_i for a basis function $\phi_i(x)$, it is useful to characterize the region on which it has support. For basis functions which are symmetric, *i.e.*, radial basis functions like the Gaussian function (A.15), it is convenient to introduce the notion of *support width*, which determines the radial distance from its center on which it has support. For functions with exponentially local support, the region with support is somewhat imprecise, hence, we introduce an equivalent notion of approximate support width, which characterizes the region on which symmetrical basis functions have practical support, *i.e.*, where it is greater than some small value $\varepsilon > 0$:

Definition 43 (Support width) *Let c be a positive parameter that determines the center of support, and let $\phi(x)$ be a symmetric, scalar, single-variable basis function with strictly* local support*, defined on the interval $\mathcal{X}$, where $\phi(x)$ is non-zero only on the interval $\mathcal{X}_1 \triangleq [c-\delta, c+\delta] \subset \mathcal{X}$, and zero on its compliment $\mathcal{X} - \mathcal{X}_1$. Then, the positive parameter δ is defined as the* support width *of $\phi(x)$, and satisfies*

$$\begin{aligned} \forall x \quad &\in \quad [c-\delta, c+\delta] \Longrightarrow |\phi(x)| > 0 \\ \forall x \quad &\notin \quad [c-\delta, c+\delta] \Longrightarrow \phi(x) = 0. \end{aligned}$$

Definition 44 (Approximate support width) *Let $\varepsilon > 0$ be a small parameter which determines the level of significance, and c be a parameter that determines the center of support of the function $\phi(x)$. Furthermore, let $\phi(x)$ be a scalar, single-variable basis function $\phi(x)$ with* exponentially local support*, defined on the interval $\mathcal{X}$, where $|\phi(x)| > \varepsilon$ only on the interval $\mathcal{X}_1 \triangleq [c-\delta, c+\delta] \subset \mathcal{X}$, and $|\phi(x)| \leq \varepsilon$ on its compliment $\mathcal{X} - \mathcal{X}_1$. Then, the positive parameter δ is defined as the* approximate support width *of $\phi(x)$, and satisfies*

$$\begin{aligned} \forall x \quad &\in \quad [c-\delta, c+\delta] \Longrightarrow |\phi(x)| \geq \varepsilon \\ \forall x \quad &\notin \quad [c-\delta, c+\delta] \Longrightarrow |\phi(x)| < \varepsilon. \end{aligned}$$

The nonlinear test function

$$y = f(x) = \frac{10x}{(x+10)} + \mathrm{e}^{-(x-2)^2}, \tag{A.16}$$

is used to illustrate some properties with respect to nonlinear function approximation using the Gaussian basis functions. Some general properties utilizing the Gaussian basis functions are briefly summarized below. In Figure A.2, some of these properties are illustrated for the modeling of the above test function (A.16).

- The approximation of increasingly complex nonlinear behavior requires increasing number of basis functions. Too few basis functions, or too small support width δ_i, introduce a non-monotonic, oscillating behavior of the modeled output. This is seen in Figure A.2, where in a), the nonlinear test function is approximated well with $p = 9$ equally spaced basis functions, while in b), with $p = 5$, the output is non-monotonic.

- The number of basis functions needed to model a particular nonlinear function may be reduced by customizing the scaling w_i and center c_i of each basis function $\phi_i(x)$, which is illustrated in case c) in Figure A.2.
- Since the Gaussian basis functions have exponentially local support, their extrapolation behavior tend towards zero, that is, $\boldsymbol{\phi}(x)$ converges to zero with increasing distance from $\mathcal{X}$.
- The Gaussian basis functions are analytic, *i.e.*, they are C^∞ (completely differentiable)

Remark 45 *When using basis functions with (exponential) local support, it may be a problem that the modeled output $\hat{y}$ drop to zero near the boundary of $\mathcal{X}$. This is seen for example in Figure A.2, where $\hat{y}$ dips down near center $c_p = 8$ of the last basis function $\phi_p(x)$ (c_i, $i = 1, 2, \cdots, p$). In order to retain adequate support near the boundary of $\mathcal{X}$, it is often advantageous to extend the set of basis functions somewhat outside of $\mathcal{X}$,* i.e.*, so that $\mathcal{X} \subset [c_1, c_p]$. In the example, the input of interest is defined by the set $\mathcal{X} = \{\forall x : x \in [0, 8]\}$ where the boundary basis functions are placed with centers at $c_1 = 0$ and $c_p = 8$. The drop in $\hat{y}$ would be completely removed by placing the boundary basis functions $\phi_1(x)$ and $\phi_p(x)$,* e.g.*, with centers $c_1 < -1$ and $c_p > 9$. Another common approach, is to redefine the boundary basis functions ϕ_1 and ϕ_p so that they have support outside $\mathcal{X}$.*

Normalization

It is sometimes advantageous to use *normalized* basis functions, defined as

$$\bar{\phi}_i(x) \triangleq \frac{\phi_i(x)}{\Sigma_{j=1}^{p} \phi_j(x)}. \tag{A.17}$$

The set of normalized basis functions $\bar{\phi}_i(x)$ has some desirable properties. First of all, normalization usually improves the undesirable non-monotonic behavior that occurs when the width δ_i is chosen too small. Due to the *unity property*

$$\Sigma_{i=1}^{p} \bar{\phi}_i(x) = 1, \tag{A.18}$$

there is a close relation between the modeled output $\hat{y}$ at $x = c_i$ and the parameter θ_i, thus the interpretation of the weighting parameter θ_i is simpler than for the non-normalized basis function. Due to the normalization, the boundary basis functions $\phi_1(x)$ and $\phi_p(x)$ no longer have local support, but converge to unity outside $\mathcal{X}$. Thus, the modeled output $\hat{y}$ converges to a constant outside $\mathcal{X}$, and the extrapolation behavior is said to be constant. As a result, there is no drops in $\hat{y}$ near the boundaries of $\mathcal{X}$, which was the case with the non-normalized basis functions. In Fig. A.3, the nonlinear test function A.16) is modeled using normalized Gauss functions.

Examples of the use of the normalized Gaussian basis functions, is the modeling of the clutch load characteristic, illustrated in Figure **??**, and the modeling of the conductance function of the valve ports in the smooth flow rate model of the pneumatic proportional valve, plotted in Figure 5.7 (page 85).

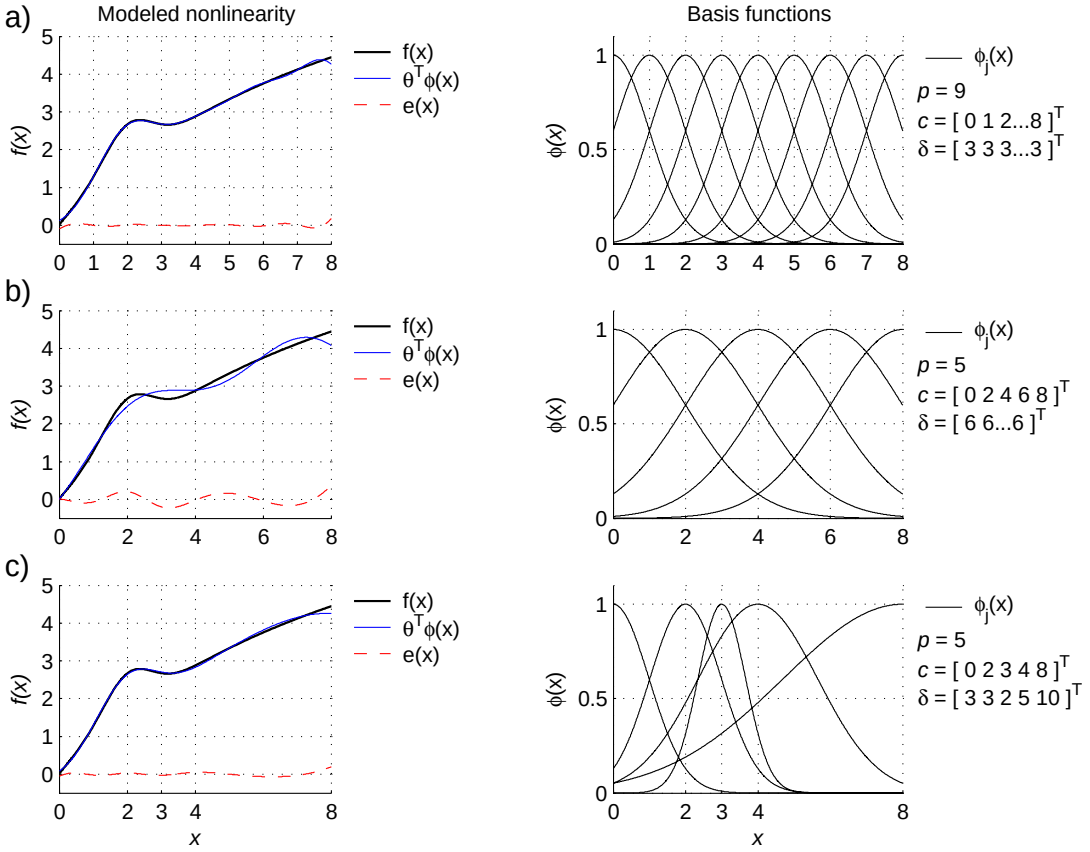

Figure A.2: The left column shows the nonlinear test function $y = f(x)$, given by (A.16), the modeled function $\hat{y} = \boldsymbol{\theta}^T \boldsymbol{\phi}(x)$, and the corresponding error $e(x) \triangleq y - \hat{y}$. The right column shows the corresponding set of basis functions, with input construction given by the center vector $\mathbf{c}$, and the scaling vector $\mathbf{w}$. In a), the set of basis functions are of the same width, and equally spaced. In b), the number of basis functions is reduced. Plot c), illustrates the use of basis functions with customized width and position.

Filtered interpolation

An interesting observation is that normalized Gaussian basis functions — as an implication of the unity property — can be used for filtered interpolation, or smoothing of a set of input and output data, $\{x_k\}$, $\{y_k\}$, $k = 1, 2, \cdots, N$, without the need to fit the parameters to the data. By forming the center vector $\mathbf{c}$ and the parameter vector $\boldsymbol{\theta}$ from the given data as

$$\mathbf{c} = [x_1, x_2, \cdots, x_N]^T \tag{A.19}$$
$$\boldsymbol{\theta} = [y_1, y_2, \cdots, y_N]^T, \tag{A.20}$$

a smooth approximation of the discrete points $\{x_k, y_k\}$ is given by

$$\hat{y}(x) = \boldsymbol{\theta}^T \cdot \overline{\boldsymbol{\phi}}(x), \tag{A.21}$$

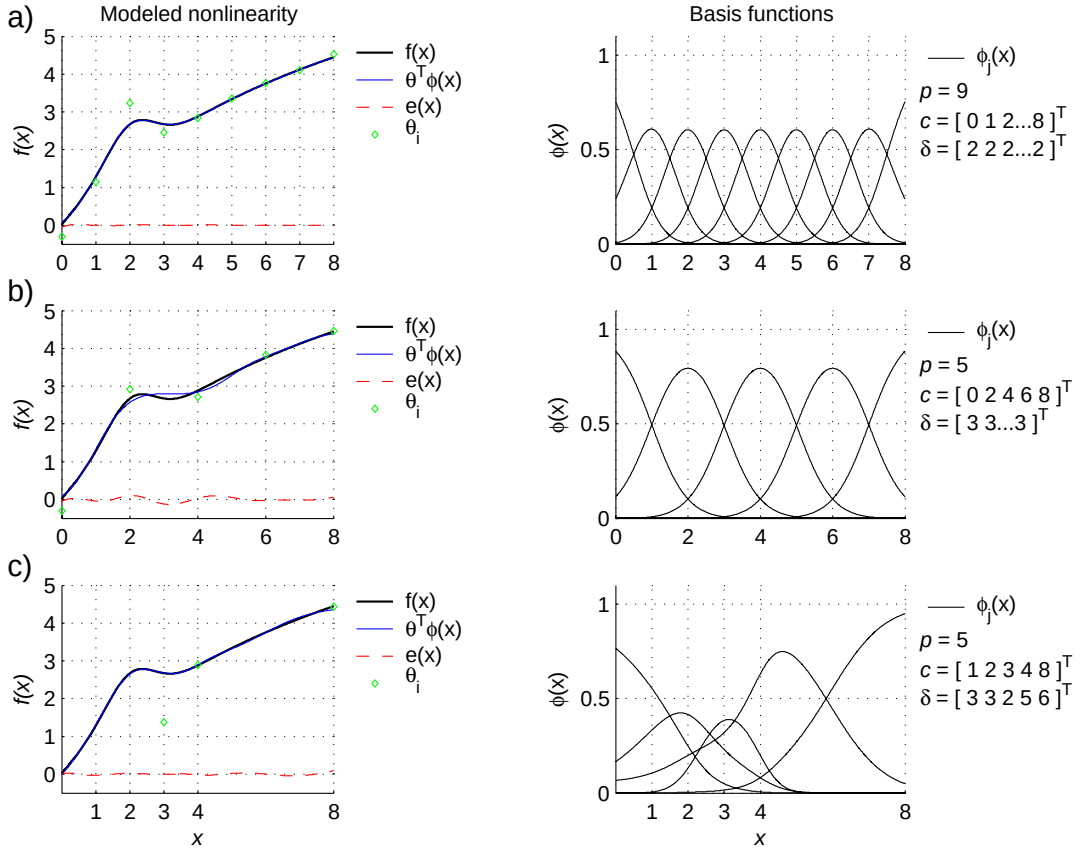

Figure A.3: The left column shows the nonlinear test function (A.16), the modeled output $\hat{y} = \boldsymbol{\theta}^T \overline{\boldsymbol{\phi}}(x)$ using normalized Gauss basis functions, and the corresponding error $e(x) \triangleq y - \hat{y}$. The value of the fitted parameter θ_i are marked with a diamond ($\diamond$) for the corresponding basis function with center at $x = c_i$. The right column shows the corresponding set of basis functions.

where the amount of smoothing is determined by the support width δ_i (scaling w_i) of the basis functions. For sufficiently small support width δ_i, we have that $\hat{y}(x_i) \approx y_i$, with negligible error.

Fig. A.4 illustrates the filtered interpolation of the discrete data set

$$\begin{aligned} \{x_k\} &= \{1, 2, 3, 4, 5, 6, 7, 8\}, \\ \{y_k\} &= \{3.0, 3.1, 3.8, 3.3, 1.9, 0.8, 0.3, 0.1\}, \end{aligned} \tag{A.22}$$

using normalized Gaussian basis functions for different support width δ_i (which is determined by the scaling factor w_i). For extrapolation, *i.e.*, for x outside $\mathcal{X} = \{\forall x : x \in [c_1, c_p] = [0, 8]\}$, the output $\hat{y} = \boldsymbol{\theta}^T \overline{\boldsymbol{\phi}}(x)$ converges to the boundary values $\theta_1 = 3.0$ or $\theta_p = 0.1$.

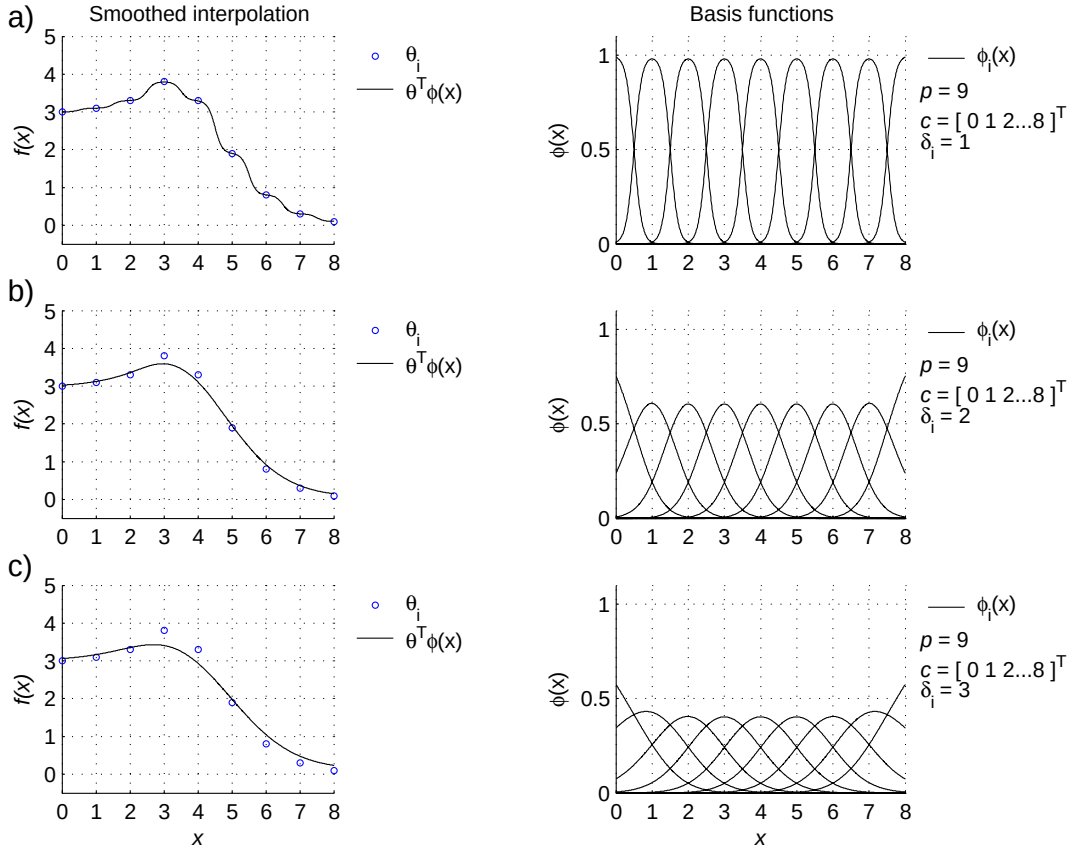

Figure A.4: The left column illustrates the filtered interpolation behaviour of the discrete data set (A.22), marked with circles ($\circ$). The right column shows the corresponding normalized Gaussian basis functions. The filtration, or smooting is illustrated for increasing approximate support width δ_i: a) $\delta_i = 1$ b) $\delta_i = 2$, and c) $\delta_i = 3$.

A.4 B-spline basis functions

While the Gaussian basis functions have exponentially local support, the so-called *B-splines*, or *bell splines*, belongs to a group of functions which have strictly *local support*.

The B-spline of kth-degree is defined recursively by

$$\phi_i^k(x) = \left(\frac{x - c_i}{c_{i+k} - c_i}\right)\phi_i^{k-1}(x) + \left(\frac{c_{i+k+1} - x}{c_{i+k+1} - c_{i+1}}\right)\phi_{i+1}^{k-1}(x), \tag{A.23}$$

where the 0th-degree B-spline is defined as

$$\phi_i^0(x) \triangleq \begin{cases} 1 & , x \in [c_i, c_{i+1}\rangle \\ 0 & , \text{otherwise} \end{cases}. \tag{A.24}$$

The B-splines form the basis for all kth-degree spline functions, and has the following properties:

- The ith B-spline $\phi_i^k(x)$ has support on the interval $[c_i, c_{i+k+1}]$, thus

$$x \notin [c_i, c_{i+k+1}] \Longrightarrow \phi_i^k(x) = 0.$$

  Hence, the complete set of basis functions $\phi_i^k(x)$, $i = 1, 2, \cdots, p$ has support on the interval $[c_1, c_{p+k+1}]$.

- The B-splines has the unity property

$$\sum_{i=-\infty}^{\infty} \phi_i^k(x) \equiv 1.$$

- Since the B-spline functions have strictly local support, their extrapolation behavior tend to zero, *i.e.*, $\boldsymbol{\phi}^k(x) = 0, \forall x \notin [c_1, c_{p+k+1}]$. See Remark 45.

- The linear combination of B-splines $\sum_{i=1}^{p} \theta_i \phi_i^k(x) = \boldsymbol{\theta}^T \boldsymbol{\phi}^k(x)$ is C^{k-1}, *i.e.*, it has continuous derivatives up to order $k-1$.

The recursive definition results in a piecewise defined spline function. For example, with degree $k = 2$ we get

$$\phi_i^2(x) \triangleq \begin{cases} a_{10} + a_{11}x + a_{12}x^2 & , x \in [c_i, c_{i+1}] \\ a_{20} + a_{21}x + a_{22}x^2 & , x \in [c_{i+1}, c_{i+2}\rangle \\ a_{30} + a_{31}x + a_{32}x^2 & , x \in [c_{i+2}, c_{i+3}\rangle \\ 0 & , \text{otherwise} \end{cases}. \tag{A.25}$$

Since the expressions for the polynomial coefficients a_{ji} becomes rather messy for higher order B-splines, it is convenient to use the recursive definition in (A.23).

The 1st-order derivative of the B-spline $\phi_i^k(x)$ can be expressed recursively as

$$\frac{d}{dx}\phi_i^k(x) = \left(\frac{k}{c_{i+k} - c_i}\right)\phi_i^{k-1}(x) - \left(\frac{k}{c_{i+k+1} - c_{i+1}}\right)\phi_{i+1}^{k-1}(x). \tag{A.26}$$

Higher-order derivatives can be obtained recursively by repeated application of (A.26), *e.g.*

$$\frac{d^2}{dx^2}\phi_i^k(x) = \left(\frac{k}{c_{i+k} - c_i}\right)\frac{d\phi_i^{k-1}(x)}{dx} - \left(\frac{k}{c_{i+k+1} - c_{i+1}}\right)\frac{d\phi_{i+1}^{k-1}(x)}{dx}. \tag{A.27}$$

In particular, we have that the 1st-degree B-spline consist of linear line segments, and is thus denoted *linear B-spline*. The linear B-spline $\phi_i^1(x)$ has support on the open interval $\langle c_i, c_{i+2}\rangle$, and the modeled function $\hat{y} = \boldsymbol{\theta}^T \boldsymbol{\phi}(x)$ becomes continuous, but not smooth — it has discontinuous 1st-order derivatives at each knot c_i. The 2nd-degree B-spline consists of 2nd-order polynomial line segments, and is denoted *quadratic B-spline*. The quadratic B-spline $\phi_i^2(x)$ has support on the interval $\langle c_i, c_{i+3}\rangle$, and the modeled function $\hat{y} = \boldsymbol{\theta}^T \boldsymbol{\phi}(x)$ is C^1. The 3rd-degree B-spline consists of 3rd-order polynomial line segments, and is denoted *cubic B-spline*. The cubic B-spline $\phi_i^3(x)$ has support on the interval $\langle c_i, c_{i+4}\rangle$, and the modeled function $\hat{y} = \boldsymbol{\theta}^T \boldsymbol{\phi}(x)$ is C^2. The B-splines of degree $k = 1$, 2, and 3 are plotted in Figure A.5.

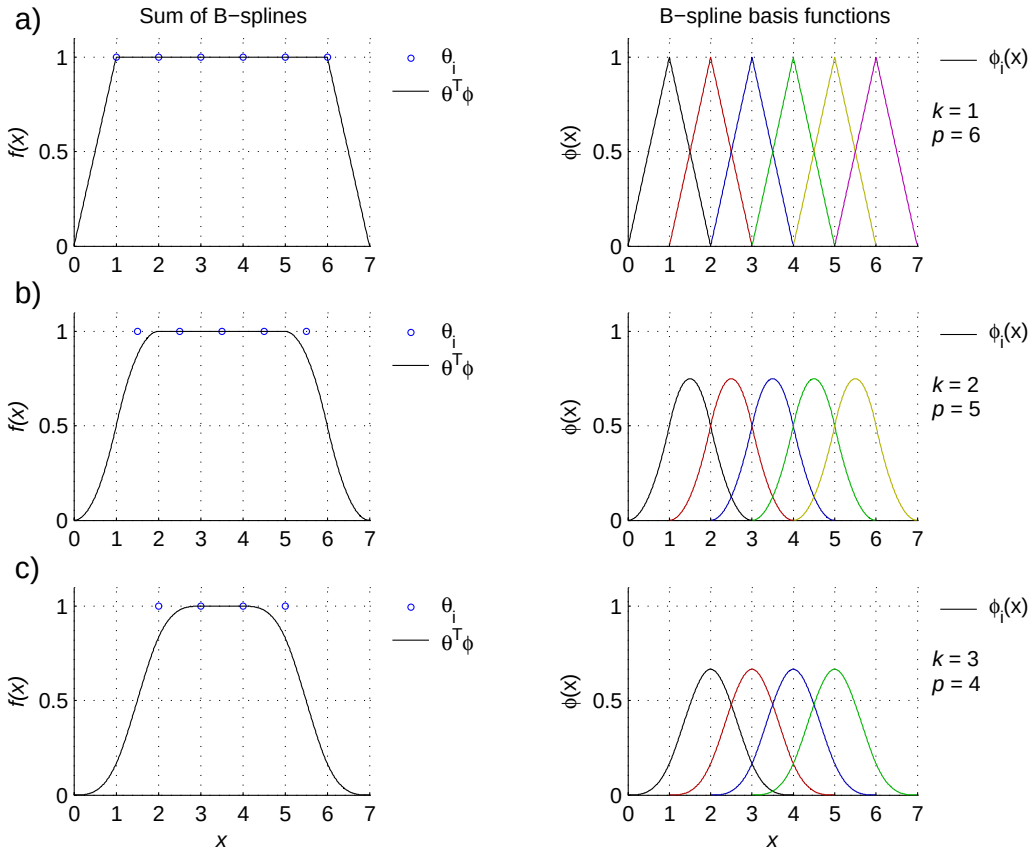

Figure A.5: The right column shows the B-splines of order a) $k = 1$, b) $k = 2$, and c) $k = 3$, for the interval (knot sequence) vector $\mathbf{c} = [0, 1, 2, 3, 4, 5, 6, 7]^T$. The corresponding sum of B-splines, $\hat{y} = \Sigma_{i=1}^{p} \phi_i(x) = \boldsymbol{\theta}^T \boldsymbol{\phi}(x)$, with $\boldsymbol{\theta} = [1, 1, \cdots, 1]^T$, are plotted in left column.

The modeling properties of the B-splines of degree $k > 1$ are very similar to those of the Gaussian basis functions. Notice, however, that the linear B-splines, results in linear interpolation between the knots c_i. In Figure A.6, the nonlinear test function (A.16) is modeled using B-splines of degree $k = \{1, 2, 3\}$. The example uses the same number of basis functions ($p = 9$), and an equidistant knot distribution which can be compared to what was used in a) for the non-normalized (Figure A.2) and normalized (Figure A.3) Gaussian basis functions. Like in the case with the non-normalized Gaussian basis functions, there is a dip in $\hat{y}$ near the upper boundary of $\mathcal{X}$, $x = 8$, which is due to defining the boundary basis function $\phi_p(x)$ with its center at the boundary $x = 8$. The dip would be effectively removed by defining $\phi_p(x)$ with its center somewhat outside of $\mathcal{X}$, *e.g.* at $x = 10$. See Remark 45. Notice, however, that for higher degree ($k > 1$) B-splines with unevenly distributed knots, the interpretation of centers is not straightforward.

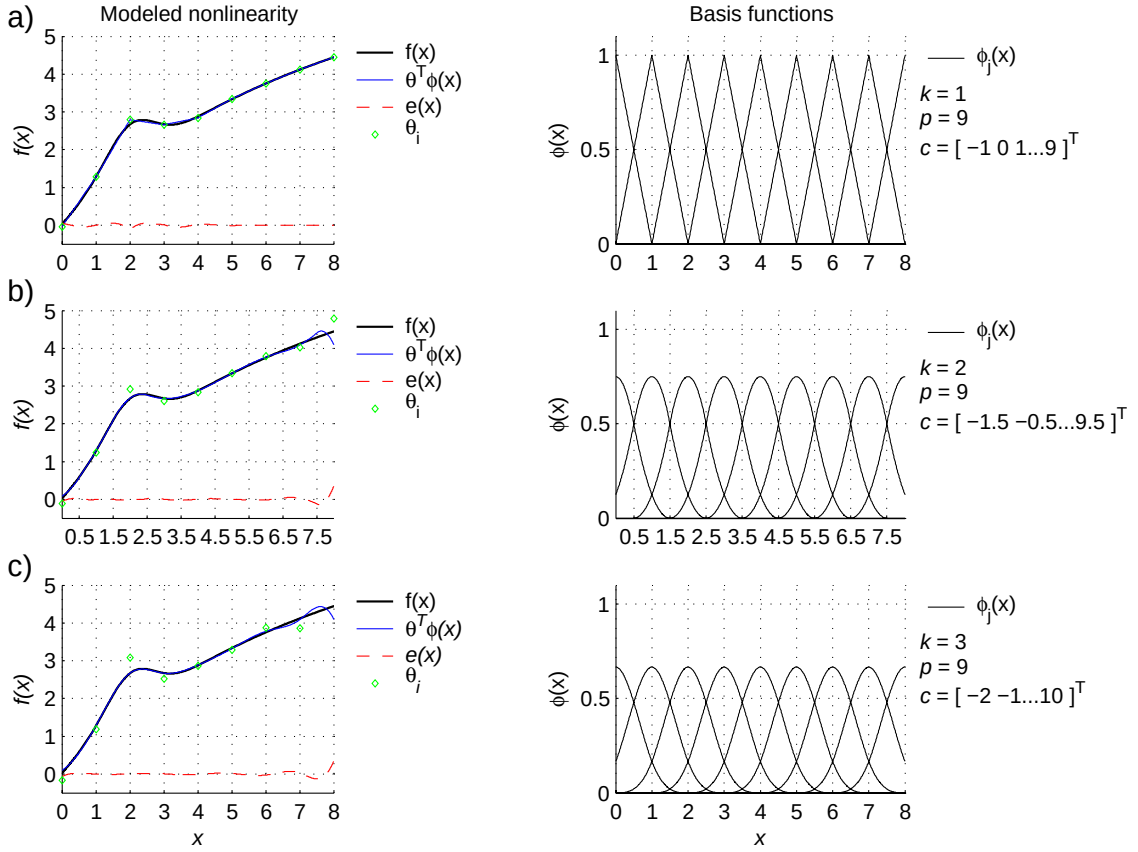

Figure A.6: The left column shows the nonlinear test function (A.16), the modeled output $\hat{y} = \boldsymbol{\theta}^T \boldsymbol{\phi}(x)$ using B-spline basis functions, where linear, quadratic, and cubic B-splines are used in a), b), and c), respectively. Additionally, the corresponding modeling error $e(x) \triangleq y - \hat{y}$ is plotted in each case, and the value of the fitted parameter θ_i are marked with a diamond ($\diamond$) at the center of the ith B-spline. The right column shows the corresponding set of basis functions.

A.4.1 Modeling the clutch load characteristic

In Figure A.7 and A.8, the static load characteristic is modeled using cubic B-spline basis functions. The knot distribution is taken as

$$\mathbf{c} = [-7, -4, -1, 0, 1, 2, 3, 4, 10, 15, 20, 25, 30]^T \text{ mm}, \tag{A.28}$$

where the number of knots is more frequent on the interval $y \in [-1, 4]$ mm in order to allow for stronger nonlinear behavior in this region and its neighborhood. Furthermore, the boundary basis functions $\phi_{l,1}(y)$ and $\phi_{l,p}(y)$, are positioned somewhat outside of $\mathcal{Y}$ so that adequate support is retained on the entire set $\mathcal{Y}$.

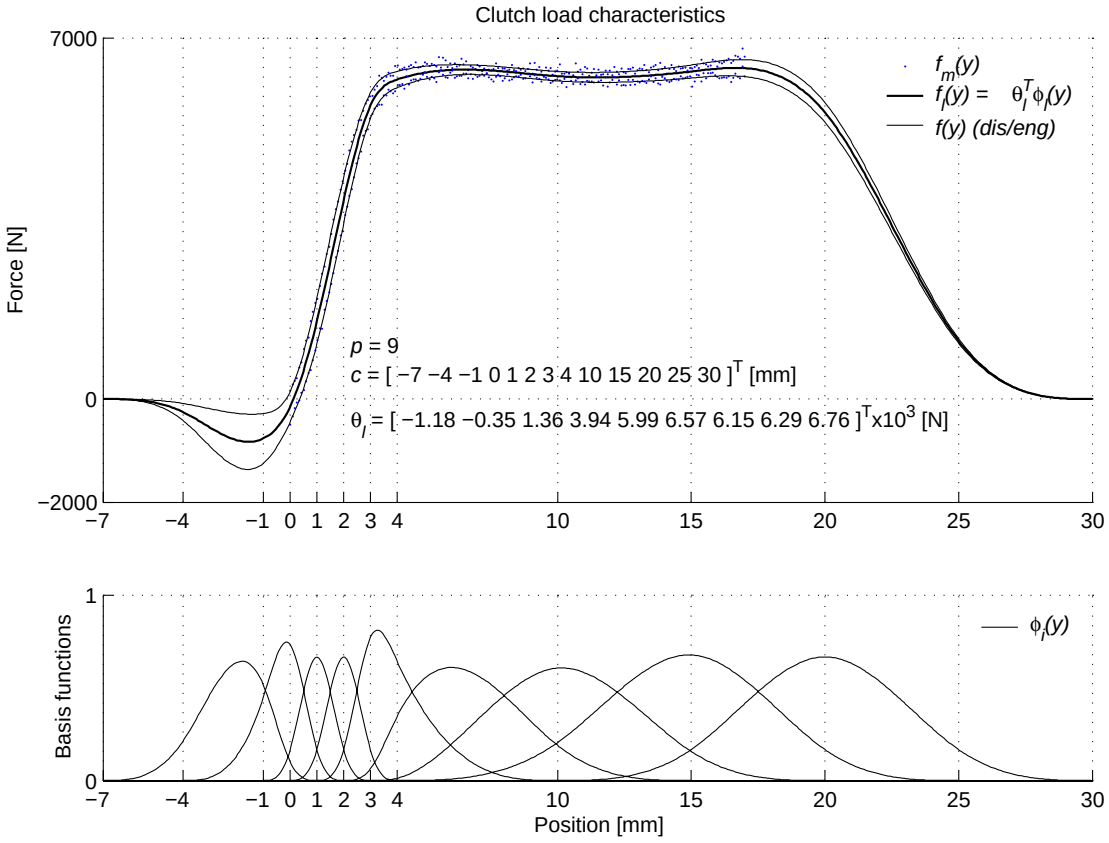

Figure A.7: The modeled clutch characteristics $f_l\,(y)$ fitted to quasi-static measurements, and the corresponding cubic B-spline basis functions.

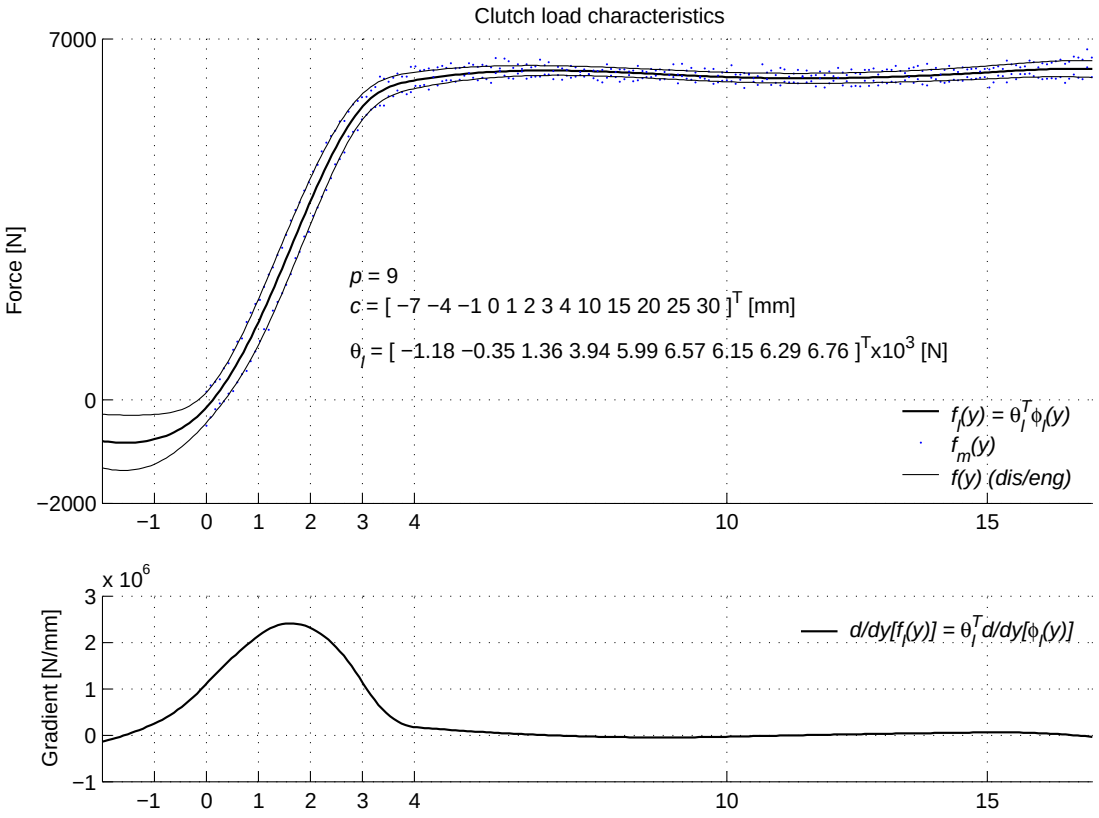

Figure A.8: The clutch characteristics $f_l\,(y)$ and its gradient df_l/dy, modeled using cubic B-splines.

Appendix B

Parameter Estimation

The model of the electro-pneumatic clutch actuator is based on mechanistic modeling, which mean that it should be possible to obtain its parameters from its physical properties by measurements. However, even though the system is relatively simple, the process of obtaining such measurements of all the parameters can be complicated, and in many cases impossible due to economic considerations. Furthermore, for in-vehicle testing, several of the measurements that are possible to obtain on the test rig, are not practically implementable on an actuator mounted on a vehicle. An alternative to measure the physical properties of the system, is to estimate the model parameters from measured input-output data, by finding the parameters of the model which provides the best fit to the measurements. This is often the superior method with respect to efficient modeling. This appendix briefly review some methods which can be used for off-line estimation of the model parameters.

We consider two methods: In Section B.1, the *linear least squares* method for parameter-affine models, and in Section B.2, a *nonlinear least squares* optimization algorithm for the general case when the parameters appear nonlinearly in the model.

B.1 Linear least squares parameter fitting

A numerically advantageous case, is when the model can be expressed in the parameter-affine form

$$\hat{y}_k = \boldsymbol{\theta}^T \cdot \boldsymbol{\phi}\left(\mathbf{x}_k\right), \tag{B.1}$$

where $\boldsymbol{\theta} \in \mathbb{R}^p$ is the vector of p parameters to be estimated, $\hat{y}_k \in \mathbb{R}$ is the modeled output corresponding to the measured output y_k, and $\mathbf{x}_k \in \mathbb{R}^m$ is a vector of m measured input variables at sample k.

An example of a model which can be expressed in the above parameter-affine form is the clutch load model $f_l\left(y\right)$, where the parameters are to be fitted to measurements of the nonlinear clutch characteristic in Section 3.1.

B.1.1 The standard least squares problem formulation

The least squares algorithm can be used to find the parameters that minimizes the sum of the squared output errors over the set $\{\mathbf{x}_k\}$, $\{y_k\}$, $k = 1, 2, \cdots, N$. The

output error at each sample is

$$e_k = y_k - \hat{y}_k = y_k - \boldsymbol{\theta}^T \cdot \boldsymbol{\phi}(\mathbf{x}_k). \tag{B.2}$$

The least squares cost criterion is given as

$$J = \sum_{k=1}^{N} e_k^2 = \sum_{k=1}^{N} \left(y_k - \boldsymbol{\theta}^T \cdot \boldsymbol{\phi}(\mathbf{x}_k)\right)^2, \tag{B.3}$$

which has its minimum at $\frac{\partial J}{\partial \theta} = 0$. This gives

$$\frac{\partial J}{\partial \theta} = -\sum_{k=1}^{N} \boldsymbol{\phi}(\mathbf{x}_k)\, 2\left(y_k - \boldsymbol{\theta}^T \cdot \boldsymbol{\phi}(\mathbf{x}_k)\right) = 0 \tag{B.4}$$

$$\Updownarrow \tag{B.5}$$

$$\sum_{k=1}^{N} \left(-\boldsymbol{\phi}(\mathbf{x}_k)\, y_k + \boldsymbol{\phi}(\mathbf{x}_k) \cdot \boldsymbol{\phi}(\mathbf{x}_k)^T \boldsymbol{\theta}\right) = 0, \tag{B.6}$$

which results in the *normal equation*

$$\sum_{k=1}^{N} \boldsymbol{\phi}(\mathbf{x}_k)\, y_k = \sum_{k=1}^{N} \boldsymbol{\phi}(\mathbf{x}_k) \cdot \boldsymbol{\phi}(\mathbf{x}_k)^T \boldsymbol{\theta}. \tag{B.7}$$

The normal equation can be expressed on the compact form

$$\mathbf{b} = \boldsymbol{\Phi} \cdot \boldsymbol{\theta}, \tag{B.8}$$

which is straightforwared to solve in *e.g.* Matlab. Note, however, that solving for $\boldsymbol{\theta}$ from the normal equation may result in a numerically bad conditioned problem. In the next section, an alternative problem formulation is given.

B.1.2 Alternative least squares problem formulation

The problem of fitting the parameters of models in the form (B.1), can alternatively be formulated in the following manner, usually resulting in a numerically better conditioned problem than the solution obtained via the normal equation.

For each data sample $\mathbf{x}_k, y_k$, the output y_k can be expressed by the modeled output $\hat{y}_k$ and the resulting modeling error e_k as

$$y_k = \hat{y}_k + e_k = \boldsymbol{\theta}^T \boldsymbol{\phi}(\mathbf{x}_k) + e_k. \tag{B.9}$$

Listing all the samples $k = 1, 2, \cdots, N$ gives

$$\begin{aligned} y_1 &= \boldsymbol{\theta}^T \cdot \boldsymbol{\phi}(\mathbf{x}_1) + e_1 \\ y_2 &= \boldsymbol{\theta}^T \cdot \boldsymbol{\phi}(\mathbf{x}_2) + e_2 \\ &\vdots \\ y_N &= \boldsymbol{\theta}^T \cdot \boldsymbol{\phi}(\mathbf{x}_N) + e_N, \end{aligned}$$

which can be formulated in vector form as

$$\begin{bmatrix} y_1 - e_1 \\ y_2 - e_2 \\ \vdots \\ y_N - e_N \end{bmatrix} = \begin{bmatrix} \boldsymbol{\phi}(x_1)^T \\ \boldsymbol{\phi}(x_2)^T \\ \vdots \\ \boldsymbol{\phi}(x_N)^T \end{bmatrix} \cdot \boldsymbol{\theta} \tag{B.10}$$

$$\Updownarrow$$

$$\mathbf{b}_e = \boldsymbol{\Phi} \cdot \boldsymbol{\theta}. \tag{B.11}$$

Since e_k is unknown, we instead attempt to find an approximate solution to the problem by letting $e_k = 0$, *i.e.*, the problem can be stated as

$$\mathbf{b} = \boldsymbol{\Phi} \cdot \boldsymbol{\theta}, \tag{B.12}$$

where $\mathbf{b} = [y_1, y_2, \cdots, y_N]^T$. Equation (B.12) has a solution only in the case when we have a perfect model, *i.e.*, when $\mathbf{b}$ lies in the range of $\boldsymbol{\Phi}$ ($\mathbf{b} \in \mathcal{R}(\boldsymbol{\Phi})$), such that a solution exists which gives $e_k = 0$, $k = 1, 2, \cdots, N$. In the case of parameter fitting to empirical data, we usually have an overdetermined problem, *i.e.*, there are more equations than parameters in $\boldsymbol{\theta}$. Thus, we are only able to solve the equation such that the output error e_k is minimized in some sense.

In Matlab, the unconstrained linear equation

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b}, \tag{B.13}$$

can be solved using the left division operator

```
>> x = A\b,
```

which for an overdetermined problem provides the least squares solution of (B.13), *i.e.*, the squared error $(\mathbf{A} \cdot \mathbf{x} - \mathbf{b})^2$ is minimized. The left division operator can also be used to find a solution when the problem is underdetermined, or when an exact solution exists.

In many cases we have *a priori* known bounds on the parameters, *i.e.*, we want to solve the problem (B.12) with lower and upper bounds on the parameters, given by

$$\boldsymbol{\theta}_{\text{lb}} \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_{\text{ub}}. \tag{B.14}$$

In this case, the constrained least squares problem can be solved in Matlab using the `lsqlin` function in the Matlab Optimization Toolbox. In addition to handle bounds on the parameters, the algorithm also handles equality and inequality constraints in the form

$$\mathbf{C}_{eq} \cdot \mathbf{x} = \mathbf{d}_{eq} \tag{B.15}$$

and

$$\mathbf{C}_1 \cdot \mathbf{x} \geq \mathbf{d}_1, \quad \mathbf{C}_2 \cdot \mathbf{x} \leq \mathbf{d}_2, \tag{B.16}$$

respectively.

B.2 Nonlinear least squares parameter fitting

The fitting of parameters that appear in a non-affine form can be performed by the use of a *nonlinear optimization algorithm* in order to search for a set of parameters which, in some sense, minimizes the error between the empirical data and the modeled outputs. The approach is to formulate a cost criterion which is a scalar measure of the *goodness of fit* of the model, and search for the model parameters which minimizes it. A cost criterion is usually formulated as a weighted sum of the model error, $\mathbf{e}_k = \hat{\mathbf{y}}_k - \mathbf{y}_k \in \mathbb{R}^m$, between the measurements $\mathbf{y}_k$, $k = 1, 2, \cdots, N$ and the modeled outputs

$$\hat{\mathbf{y}}_k = \mathbf{f}(\mathbf{x}_k) \tag{B.17}$$

at each sample $k = 1, 2, \cdots, N$.

Unlike the linear least squares problem, there might exist several local minima to the nonlinear minimization problem, consequently, the algorithm is not guaranteed to find the absolute minimum. As a result, nonlinear parameter fitting is usually a trial-and-error approach, where one has to try various initial parameter values in order to find a good fit.

Remark 46 *When fitting both affine and non-affine parameters of a model, a recursive, two-stage method, briefly described in [59], is often preferable. The method, called* separable least squares*, is straightforward: for each iteration step, the least squares solution of fitting the affine parameters are obtained using the estimated non-affine parameters from the previous step, and subsequently, their values are inserted into the nonlinear problem, and the non-affine parameters are found using a nonlinear optimization algorithm.*

B.2.1 Formulation of the nonlinear optimization problem

The nonlinear optimization algorithm is used to minimize a criterion in the form

$$J(\boldsymbol{\theta}) = \frac{1}{2}\sum_{k=1}^{N} \mathbf{e}_k(\boldsymbol{\theta})^T \mathbf{Q}\mathbf{e}_k(\boldsymbol{\theta}), \tag{B.18}$$

where $J(\boldsymbol{\theta})$ is the function to be minimized, N is the number of samples in the data set, m is the number of measured outputs, and p is the number of parameters. Furthermore, $\mathbf{e}_k = [e_{1,k}, e_{2,k}, \cdots, e_{m,k}]^T$ are the errors between the measured and the modeled outputs at sample k, the vector $\boldsymbol{\theta} = [\theta_1, \theta_2, \cdots, \theta_p]^T$ contains the parameters to be identified, and $\mathbf{Q} = \mathbf{q}\mathbf{q}^T$ is a diagonal weighting matrix where the vector $\mathbf{q} = [q_1, q_2, \cdots, q_m]^T$ contains the weights for each of the measured outputs. The minimization problem can be stated as

$$\hat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}\in\Theta} J(\boldsymbol{\theta}), \tag{B.19}$$

where the estimated parameter vector $\hat{\boldsymbol{\theta}}$ is restricted to be in the set

$$\Theta = \{\boldsymbol{\theta} \in \mathbb{R}^p : \boldsymbol{\theta}_{\text{lb}} \leq \boldsymbol{\theta} \leq \boldsymbol{\theta}_{\text{ub}}\}, \tag{B.20}$$

where $\boldsymbol{\theta}_{\text{lb}} = [\theta_{1,\text{lb}}, \theta_{2,\text{lb}}, \cdots, \theta_{p,\text{lb}}]^T$ and $\boldsymbol{\theta}_{\text{ub}} = [\theta_{1,\text{ub}}, \theta_{2,\text{ub}}, \cdots, \theta_{p,\text{ub}}]^T$ are the lower and upper bounds on the parameters, respectively.

The function `lsqnonlin` in the Matlab Optimization Toolbox may be used for parameter identification for the nonlinear model of the electro-pneumatic actuator. By default, the function uses the Levenberg-Marquardt algorithm (a modified version of the Gauss-Newton algorithm) in its search for the optimal parameter values [65]. For use with the function, the optimization criterion in (B.18) has to be rewritten on the form

$$J(\boldsymbol{\theta}) = \frac{1}{2}\sum_{k=1}^{N} \mathbf{e}_k(\boldsymbol{\theta})^T \mathbf{q}\mathbf{q}^T \mathbf{e}_k(\boldsymbol{\theta}) = \frac{1}{2}\mathbf{F}(\boldsymbol{\theta})^T \mathbf{F}(\boldsymbol{\theta}), \tag{B.21}$$

where a vector of weighted outputs are defined as
$\mathbf{F}(\boldsymbol{\theta}) \triangleq \left[\mathbf{q}^T\mathbf{e}_1(\boldsymbol{\theta}), \mathbf{q}^T\mathbf{e}_2(\boldsymbol{\theta}), \cdots, \mathbf{q}^T\mathbf{e}_N(\boldsymbol{\theta})\right]^T$.

B.2.2 Fitting of parameter of dynamic models

For the fitting of parameters of a dynamic model, the generation of the output error vector, given by $\mathbf{F}(\boldsymbol{\theta})$, is optionally obtained either by a *one-step-ahead prediction* approach, or a *ballistic simulation* approach[1]. The first method refers to simulating the system response between each sample, while resetting the initial output to the measured output at every sample k, *i.e.*, the error is given as $\mathbf{e}_k = \mathbf{y}(t_k) - \hat{\mathbf{y}}(t_k|t_{k-1})$. This approach is typically preferred when the model is used to design an observer, where the estimated output is corrected at each sample. The ballistic approach is simply a full simulation of the response of the model, where the error at each sample is computed as $\mathbf{e}_k = \mathbf{y}(t_k) - \hat{\mathbf{y}}(t_k|t_0)$. This approach may be preferred, *e.g.*, when the long-term prediction capabilities of the model are important. However, for nonlinear systems such as the pneumatic actuator, a problem with a ballistic approach may be that unstable states drifts off from its actual trajectory due to small errors. To avoid this, we may employ a *partitioned ballistic approach*, which means that we partition the measurements into smaller time intervals where the simulation are re-initialized (corrected according to the measurements) at the beginning of each interval. For example, a sequence of measurements of length 2.0 s may be partitioned into 20 subintervals of length 100 ms. In most cases, we use the ballistic simulation approach for identification of parameters of the model of the electro-pneumatic clutch actuator.

[1] The denotation *ballistic* means that the simulation is carried out without any corrections from the measurements during the simulation.

Appendix C

Derivation of Air Dynamics

In this section we derive the dynamic equations for the air dynamics of chamber A in detail. The derivation for chamber B is similar. Subscripts A and B are dropped for notational brevity. Figure 4.1 shows a schematic diagram of the flow control valve and pneumatic actuator. The following basic assumptions are made:

A1) At the attainable pressures, air behaves like an *ideal gas* which obeys the ideal gas *equation of state* (See (C.1)) with negligible error.

A2) The specific heats c_p and c_v of air are assumed to be constant, *i.e.*, not functions of temperature (or pressure[1]). For the attainable temperature range for this application, the deviations are insignificant, see *e.g.* [16, Section 3.7, pp. 182-134].

A3) Energy change in the fluid due to elevation is negligible.

A4) The thermodynamic properties are uniformly distributed (homogenous) within the control volume, *i.e.*, "perfectly mixed". This is reasonable due to the small dimensions of the system, and lets us simplify the analysis to a *one-dimensional* problem, rather than a more complex distributed problem formulation.

A5) The flow through pipes, valves and restrictions in the system is assumed to be *isentropic*[2]. That is, we assume *frictionless* flow, and disregard the effect of heat transfer on the flow (*adiabatic* flow).

[1] For an ideal gas (A1), the internal energy u and enthalpy h vary only with temperature. Furthermore, the specific heats c_v and c_p will in general vary with temperature: $c_v = c_v(T)$ and $c_p = c_p(T)$. However, this temperature dependence is insignificant for this application.

[2] This is a common approximation in compressible fluid analysis when the system dimensions are small, [101].

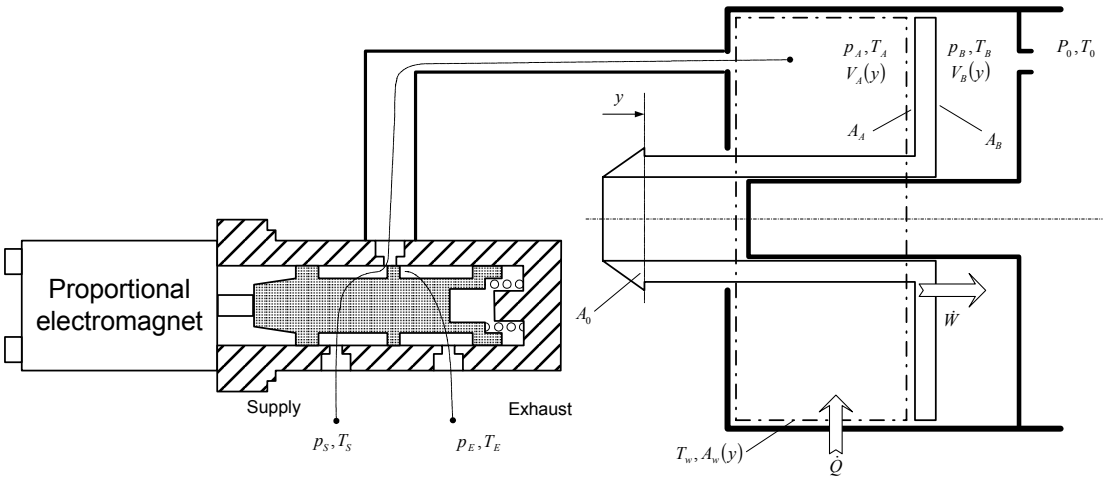

Figure C.1: Schematic diagram illustrating the pneumatic subsystem and control valve.

C.1 Thermodynamics

C.1.1 Properties

We first recapitulate some basic thermodynamic properties. The *equation of state* (or *perfect-gas law*) is given as

$$p = \rho R T, \tag{C.1}$$

where p, ρ and T are the pressure, the density, and the temperature of the fluid, respectively. The gas constant R is defined by the specific heats of the fluid as

$$R \triangleq c_p - c_v \tag{C.2}$$

which for air is $R = 288\,\mathrm{J}/(\,\mathrm{kg\,K})$. Furthermore, the ratio of specific heats is defined as

$$\kappa \triangleq \frac{c_p}{c_v} \tag{C.3}$$

The specific heat c_v is defined as the change of specific internal energy u with temperature at constant volume. Likewise, c_p is defined as the change of the specific enthalpy h with temperature at constant pressure:

$$c_v \triangleq \left(\frac{\partial u}{\partial T}\right)_{V=const.} \quad \text{and} \quad c_p \triangleq \left(\frac{\partial h}{\partial T}\right)_{p=const.} \tag{C.4}$$

For ideal gases, $u = u(T)$ and $h = h(T)$, such that the partial derivatives in (C.4) can be replaced by ordinary derivatives, giving

$$du = c_v dT \quad \text{and} \quad dh = c_p dT. \tag{C.5}$$

Furthermore, it is convenient to define zero points as $u(T = 0) \triangleq 0$ and $h(T = 0) \triangleq 0$. Integrating (C.5) we get the following useful expressions for the internal energy and enthalpy

$$u = c_v T \quad \text{and} \quad h = c_p T. \tag{C.6}$$

C.1.2 Conservation of mass

The conservation of mass is given by the *continuity equation*

$$\frac{d}{dt}(\rho V) = w_{in} - w_{out}, \tag{C.7}$$

where V is the total volume of the considered control volume, ρ is the density of the fluid, and w_{in} and w_{out} are the inlet and outlet mass flows, respectively.

C.1.3 Conservation of energy

The conservation of energy is governed by the *energy equation*

$$\frac{dE}{dt} = \varepsilon_{in} - \varepsilon_{out} + \dot{Q} - \dot{W}, \tag{C.8}$$

where E is the total energy in the control volume, ε_{in} and ε_{out} are the energies associated with the inlet and outlet flows, respectively, $\dot{Q}$ is the rate of heat transfer to the control volume, and $\dot{W}$ is the piston work done by the system.

C.1.4 Isentropic flow

In the case of isentropic flow, the energy equation reduces to the simple form

$$h_1 + \frac{1}{2}v_1^2 = h_2 + \frac{1}{2}v_2^2, \tag{C.9}$$

where the subscripts refer to flow from a point 1 to a point 2. It is important to note that in a reservoir, the kinetic energy is approximately zero ($v \approx 0$). This means that in the case of isentropic flow through short pipes from a reservoir (h_0) into a second reservoir (h_2), part of the enthalpy is first converted to kinetic energy in the pipe line (h_1), which is completely recovered in the second reservoir, i.e. $h_0 = h_1 + \frac{1}{2}v_1^2 = h_2$.

C.2 Pneumatic chamber

The net energy change due to inlet and outlet flows can be expressed as

$$\varepsilon_{in} - \varepsilon_{out} = w_{in} \cdot \left(h_{in} + \frac{1}{2}v_{in}^2\right) - w_{out} \cdot \left(h_{out} + \frac{1}{2}v_{out}^2\right), \tag{C.10}$$

where h_{in} and h_{out} are the enthalpies associated with w_{in} and w_{out}, respectively, and v_{in} and v_{out} are the mean inlet and outlet flow velocities (change in elevation has been neglected — assumption A3).

The heat transfer from the cylinder wall to the control volume is mainly governed by convective heat transfer, which in most cases can be adequately described by the empirical model

$$\dot{Q} = H_w A_w \cdot (T_w - T). \tag{C.11}$$

Here, H_w is the convective heat coefficient, A_w is the cylinder wall area (control volume boundary), T_w is the cylinder wall temperature, and T is the temperature within the control volume. For example, for still air (natural convection), the convective heat coefficient is in the range $H_w \in \langle 3, 23 \rangle$ W/ (m^2 K), while for moving air $H_w \in \langle 11, 55 \rangle$ W/ (m^2 K), [29, p. 219].

The piston work done by the system can be expressed as

$$\dot{W} = p\frac{dV}{dt}, \tag{C.12}$$

where dV/dt is the rate of volume change, and p is the chamber pressure acting against the piston.

Substituting (C.10), (C.11), and (C.12) into (C.8), the energy equation is expressed as

$$\frac{dE}{dt} = w_{in} \cdot \left(h_{in} + \frac{1}{2} v_{in}^2 \right) - w_{out} \cdot \left(h_{out} + \frac{1}{2} v_{out}^2 \right) + H_w A_w \cdot (T_w - T) - p\frac{dV}{dt} \tag{C.13}$$

A sketch illustrating the control volume and flow paths of the pneumatic system, is given in Figure C.1.

The pneumatic chambers can be considered to be reservoirs, where the kinetic energy within the control volume is negligible. Hence, we add the following assumption to our list:

A6) The kinetic energy ($\frac{1}{2}v^2$) within the pneumatic chambers (A and B) are negligible.

In chamber A, the air flows either from the supply reservoir to the pneumatic chamber, or from the pneumatic chamber to exhaust (atmosphere). The same applies to chamber B. Consequently, the kinetic energy term can be neglected in the energy equation. Furthermore, since the flow is isentropic (assumption A5), there is no energy loss in the flow, which means that the energy equation (C.13) simplifies to

$$\frac{dE}{dt} = w_{in} h_{in} - w_{out} h_{out} + H_w A_w \cdot (T_w - T) - p\frac{dV}{dt}. \tag{C.14}$$

C.2.1 Pressure dynamics

The dynamics of the pressure in the pneumatic chamber is derived from the energy equation (C.14). Since the change of potential and kinetic energy within the chamber is negligible (assumptions A3 and A6), the change of total energy within the control volume can be given as

$$\begin{aligned} \frac{dE}{dt} &= \frac{d}{dt}(\rho V u) \\ &= c_v \frac{d}{dt}(\rho V T). \end{aligned} \tag{C.15}$$

where we have substituted with the expression for u in (C.6). Furthermore, we substitute with the equation of state (C.1), giving

$$\begin{aligned} \frac{dE}{dt} &= \frac{c_v}{R}\frac{d}{dt}(Vp) \\ &= \frac{c_v}{R}\left(\frac{dV}{dt}p + V\frac{dp}{dt}\right). \end{aligned} \tag{C.16}$$

By substituting this relation with the left-hand side of the energy equation (C.14), we obtain

$$\frac{c_v}{R}\left(\frac{dV}{dt}p + V\frac{dp}{dt}\right) = w_{in}h_{in} - w_{out}h_{out} + H_w A_w \cdot (T_w - T) - p\frac{dV}{dt} \tag{C.17}$$

$$\Updownarrow$$

$$\dot{V}p + V\frac{dp}{dt} = \frac{R}{c_v}w_{in}h_{in} - \frac{R}{c_v}w_{out}h_{out} + \frac{R}{c_v}H_w A_w \cdot (T_w - T) - \frac{R}{c_v}p\dot{V} \tag{C.18}$$

$$\Updownarrow$$

$$\begin{aligned} V\frac{dp}{dt} &= -\dot{V}p + \frac{R}{c_v}w_{in}h_{in} - \frac{R}{c_v}w_{out}h_{out} + \frac{R}{c_v}H_w A_w \cdot (T_w - T) - \frac{R}{c_v}p\dot{V} \\ &= -\left(1 + \frac{R}{c_v}\right)p\dot{V} + \frac{R}{c_v}w_{in}h_{in} - \frac{R}{c_v}w_{out}h_{out} \\ &\quad + \frac{R}{c_v}H_w A_w \cdot (T_w - T). \end{aligned} \tag{C.19}$$

We replace the enthalpies in terms of the corresponding temperatures by substituting with the expression in (C.6), giving

$$V\frac{dp}{dt} = -\left(1 + \frac{R}{c_v}\right)p\dot{V} + \frac{R}{c_v}c_p T_{in}w_{in} - \frac{R}{c_v}c_p T_{out}w_{out} + \frac{R}{c_v}H_w A_w \cdot (T_w - T). \tag{C.20}$$

Furthermore, we substitute with the relation

$$\frac{R}{c_v} = \kappa - 1, \tag{C.21}$$

which is obtained by combining (C.3) and (C.2), and replace $T_{out} = T$, which results from the homogeneity assumption A4. Finally, the full dynamic equation of the pressure can be expressed as

$$\frac{dp}{dt} = -\frac{\kappa\dot{V}}{V}p + \frac{\kappa R T_{in}}{V}w_{in} - \frac{\kappa R T}{V}w_{out} + \frac{(\kappa - 1)\cdot H_w A_w}{V}(T_w - T), \tag{C.22}$$

where the volume V of the pneumatic chamber, and the area of effective heat transfer A_w, are both functions of the actuator position y.

C.2.2 Temperature dynamics

The derivation of the temperature dynamics proceeds in similar manner as the pressure dynamics. From (C.15), we can write

$$\frac{dE}{dt} = c_v \frac{d(\rho V)}{dt} T + c_v \rho V \frac{dT}{dt} \tag{C.23}$$

$$= c_v T w_{in} - c_v T w_{out} + c_v \frac{pV}{RT} \frac{dT}{dt} \tag{C.24}$$

where in the second step we have substituted with the continuity equation (C.7), and the equation of state (C.1).

As for the pressure dynamics, this relation is substituted for the left-hand side of the energy equation (C.14), giving

$$c_v T w_{in} - c_v T w_{out} + c_v \frac{pV}{RT} \frac{dT}{dt} = w_{in} h_{in} - w_{out} h_{out} + H_w A_w \cdot (T_w - T) - p \frac{dV}{dt} \tag{C.25}$$

$$\Updownarrow$$

$$c_v \frac{pV}{RT} \frac{dT}{dt} = -p\dot{V} + (c_p T_{in} - c_v T) w_{in} - (c_p T_{out} - c_v T) w_{out} + H_w A_w \cdot (T_w - T) . \tag{C.26}$$

In the first step above, we substituted with the expression for the specific enthalpy in (C.6), and the second step was just a rearranging of the equation. Isolating the term dT/dt, and substituting $T_{out} = T$ due to homogeneity, gives

$$\frac{dT}{dt} = -\frac{R\dot{V}}{c_v V} T + \left(\frac{c_p}{c_v} T_{in} - T \right) \frac{RT}{pV} w_{in} - \left(\frac{c_p}{c_v} - 1 \right) \frac{RT^2}{pV} w_{out} + \frac{RT H_w A_w}{c_v pV} (T_w - T) . \tag{C.27}$$

Then, by substituting with the ratio of specific heats (C.3) and the relation (C.21), we obtain the temperature dynamics in its final form

$$\frac{dT}{dt} = -\frac{(\kappa - 1)\dot{V}}{V} T + \frac{(\kappa T_{in} - T) RT}{pV} w_{in} - \frac{(\kappa - 1) RT^2}{pV} w_{out} + \frac{(\kappa - 1) T H_w A_w}{pV} (T_w - T) . \tag{C.28}$$

Bibliography

[1] T. Acarman and C. Hatipoglu. A robust nonlinear controller design for a pneumatic actuator. In *Proc. American Control Conference*, pages 4490–4495, Arlington, VA, June 25-27 2001.

[2] T. Acarman, U. Ozguner, C. Hatipoglu, and A.-M. Igusky. Pneumatic brake system modeling for systems analysis. SAE Technical paper series 2000-01-3414, SAE, Portland, Oregon, Decemeber 4-6 2000. Truck and bus meeting and exposition.

[3] B. Aloliwi and H. K. Khalil. Robust adaptive output feedback control of nonlinear systems without persistence of excitation. *Automatica*, 33(11):2025–2032, 1997.

[4] B. W. Andersen. *The Analysis and Design of Pneumatic Systems*. Robert E. Krieger Publishing Company, Inc., Florida, 1967.

[5] B. Armstrong-Hélouvry. Stick slip and control in low-speed motion. *IEEE Transactions on automatic control*, 38(10):1483–1496, October 1993.

[6] B. Armstrong-Hélouvry, P. Dupont, and C. C. de Wit. A survey of models, analysis tools and compensation methods for the control of machines with friction. *Automatica*, 30(7):1083–1138, 1994.

[7] N. Barabanov and R. Ortega. Necessary and sufficient conditions for passivity of the LuGre friction model. *IEEE Transactions on automatic control*, 45(4):830–832, 2000.

[8] M. Belgharbi, S. Sesmat, S. Scavarda, and D. Thomasset. Analytical model of the flow stage of a pneumatic servo-distributor for simulation and nonlinear control. In *6th Scandinavian Int. Conf. On Fluid Power*, pages 847–861, Tampere, Finland, May 26-28 1999.

[9] J. F. Blackburn, G. Reethof, and J. L. Shearer. *Fluid Power Control*. M.I.T. Press, Cambridge, 1960.

[10] J. E. Bobrow and B. W. McDonell. Modeling, identification, and control of a pneumatically actuated, force controllable robot. *IEEE Trans. on robotics and automation*, 14(5):732–742, October 1998.

[11] S. Boudart, E. Richard, and S. Scavarda. A modified linear reduced observer for a pneumatic servodrive. In *4th Bath Int. Fluid Power Workshop*, pages 357–370, Bath, UK, 1991.

[12] M. Bouri, D. Thomasset, E. Richard, and S. Scavarda. Nonlinear sliding mode control of an electropneumatic servodrive. In *7th Bath Int. Fluid Power Workshop, Innovations in Fluid Power*, pages 201–219, Bath, UK, Sept 21 to 23 1994.

[13] X. Brun, M. Belgharbi, S. Sesmat, D. Thomasset, and S. Scavarda. Control of an electropneumatic actuator: Comparison between some linear and nonlinear control laws. *Proceedings Instn Mech Engrs*, 213:387–407, 1999.

[14] C. R. Burrows and C. R. Webb. Use of root loci in design of pneumatic servo-motors. *Control*, pages 423–427, August 1966.

[15] C. Canudas de Wit, H. Olsson, K. J. Åström, and P. Lischinsky. A new model for control of systems with friction. *IEEE Transactions on automatic control*, 40(3):419–425, March 1995.

[16] Y. A. Cengel and M. A. Boles. *Thermodynamics: An Engineering Approach.* McGraw-Hill, Inc., Princeton Road, S-1, Hightown, NJ 08520, 2nd edition, 1994.

[17] J. T. Chen. *Linear System Theory and Design.* Oxford University Press, Inc., 3rd edition, 1999.

[18] S. Chen, H. Zhou, J. Jiang, and Y. Lu. On using nonlinear filter to new state feedback control system for electro-pneumatic servo system. In *Proc. 2nd. Int. Conf. On Fluid Power Transmissions and Control*, pages 149–151, 1989.

[19] W. Cheney and D. Kincaid. *Numerical Mathematics and Computing.* Brooks/Cole Publishing Company, Pacific Grove, California, 3rd edition, 1994.

[20] F. Det, S. Scavarda, and E. Richard. Simulation of an electropneumatic servovalve to study the feasability of a position control loop. In *2nd. Int. Conf. On Fluid Power Transmissions and Control*, pages 505–510, 1989.

[21] S. Drakunov, G. D. Hanchin, W. C. Su, and Ü. Özgüner. Nonlinear control of a rodless pneumatic servoactuator, or sliding mode versus coulomb friction. *Automatica*, 33(7):1401–1408, 1997.

[22] P. Dupont, V. Hayward, B. Armstrong, and F. Altpeter. Single state elastoplastic friction model. *IEEE Transactions on automatic control*, 47(5):787–792, May 2002.

[23] O. Egeland and J. T. Gravdahl. *Modeling and Simulation for Automatic Control.* Tapir Trykkeri, Trondheim, Norway, 2002.

[24] K. O. Ezal. *Disturbance Attenuating Control of Nonlinear Systems with Local Optimality*. PhD thesis, University of California, Santa Barbara, 1998.

[25] R. A. Freeman and P. V. Kokotovic. Design of 'softer' robust nonlinear control laws. *Automatica*, 19:1425–1437, 1993.

[26] R. A. Freeman and P. V. Kokotovic. *Robust Nonlinear Control Design. State-Space and Lyapunov Techniques*. Birkhäuser, Boston, 1996.

[27] R. A. Freeman, M. Krstic, and P. V. Kokotovic. Robustness of adaptive nonlinear control to bounded uncertainties. *Automatica*, 34(10):1227–1230, 1998.

[28] R. A. Freeman and L. Praly. Integrator backstepping for bounded controls and control rates. *IEEE Transactions on automatic control*, 43(2):258–262, February 1998.

[29] C. J. Geankoplis. *Transport Processes and Unit Operations*. Prentice-Hall International, Inc., 3rd edition, 1993.

[30] Z. Hong and L. Yongxiang. The further study on electro-pneumatic servo control system. In *Proceedings of the International Conference on Fluid Power Transmission and Control*, pages 415–420, 1989.

[31] C. Hsieh and Y.-C. Pan. Dynamic behavior and modelling of the pre-sliding static friction. *Wear*, 242:1–17, 2000.

[32] F. Ikhouane and M. Krstié. Adaptive backstepping with parameter projection: Robustness and asymptotic performance. *Automatica*, 34(4):429–435, 1998.

[33] P. A. Ioannou and J. Sun. *Robust Adaptive Control*. PTR Prentice-Hall, Upper Saddle River, NJ 07458, 1996.

[34] A. Isidori. *Nonlinear Control Systems*. Springer Verlag, London, 3rd edition, 1995.

[35] ISO—Int. organization for standardization. *ISO 6358 – Pneumatic Fluid Power – Components Using Compressible Fluids – Determination of Flow-Rate Characteristics*, 1. edition, October 1989.

[36] A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, New York and London, 1970.

[37] H. S. Jebar. *Design of Pneumatic Actuator Systems*. PhD thesis, University of Nottingham, 1977.

[38] T. A. Johansen. Computation of lyapunov functions for smooth nonlinear systems using convex optimization. *Automatica*, 36:1617–1626, 2000.

[39] A. Johansson. *Nonlinear Observers with Application in the Steel Industry*. PhD thesis, Luleå tekniska universitet, 2001.

[40] G.-O. Kaasa, P. J. Chapple, and B. Lie. An extended Kalman filter applied to a pneumatic servo system — velocity and acceleration estimation in a clutch actuation application. In C. R. Burrows and K. A. Edge, editors, *Proceedings, Power Transmission and Motion Control Conference PMTC2001*, London and Bury St. Edmunds, UK, September 12–14 2001. Professional Engineering Publishing Ltd.

[41] G.-O. Kaasa, P. J. Chapple, and B. Lie. Modeling of an electro-pneumatic cylinder actuator for nonlinear and adaptive control, with application to clutch actuation in heavy-duty trucks. In *3rd FPNI – PhD Symposium on Fluid Power*, pages 255–273, Terrassa, Spain, June 30–July 2, 2004 2004. Fluid Power Net International.

[42] G.-O. Kaasa and M. Takahashi. Adaptive tracking control of a clutch actuator. In *Int. Symp. On Advanced Control of Industrial Processes*, Kumamoto, Japan, June 2002.

[43] T. Kailath. *Linear Systems*. Prentice-Hall, Inc., Englewood Cliffs, N.J. 07632, 1980.

[44] H. Keller and R. Isermann. Model-based nonlinear adaptive control of a pneumatic actuator. *Control engineering practice*, 1(3):505–511, 1993.

[45] H. K. Khalil. Robust servomechanism output feedback controllers for a class of feedback linearizable systems. *Automatica*, 30(10):1587–1599, 1994.

[46] H. K. Khalil. Adaptive output feedback control of nonlinear systems represented by input-output models. *IEEE Transactions on automatic control*, 41(2):177–188, February 1996.

[47] H. K. Khalil. *Nonlinear Systems*. Prentice Hall, Upper Saddle River, New Jersey, 2nd edition, 1996.

[48] H. K. Khalil. High-gain observers in nonlinear feedback control. In H. Nijmeijer and T. I. Fossen, editors, *New Directions in Nonlinear Observer Design*, Lecture Notes in Control and Information Sciences, pages 249–268, London, Great Britain, 1999. Springer.

[49] T. Kimura, S. Hara, T. Fujita, and T. Kagawa. Feedback linearization for pneumatic actuator systems with static friction. *Control Eng. Practice*, 5(10):1385–1394, 1997.

[50] K. A. Knutsen. Nonlinear observer design for electropneumatic clutch actuators. Master's thesis, Norwegian University of Science and Technology, 2005.

[51] P. Kokotovic and M. Arcak. Constructive nonlinear control: A historical perspective. *Automatica*, 37:637–662, 2001.

[52] M. Krstic, I. Kanellakopoulos, and P. Kokotovic. *Nonlinear and Adaptive Control Design*. John Wiley & Sons, Inc., New York, 1995.

[53] M. Krstic and Z.-H. Li. Inverse optimal design of input-to-state stabilizing nonlinear controllers. *IEEE Transactions on automatic control*, 43:336–351, 1998.

[54] V. Lampaert, J. Swevers, and F. Al-Bender. Modification of the leuven integrated friction model structure. *IEEE Transactions on automatic control*, 47(4):683–687, April 2002.

[55] H. Lee and M. Tomizuka. Robust adaptive control using a universal approximator for SISO nonlinear systems. *IEEE trans. on fuzzy systems*, 8(1):95–106, February 2000.

[56] H. K. Lee, G. S. Choi, and G. H. Choi. A study on tracking position control of pneumatic actuators. *Mechatronics*, 12:813–831, 2002.

[57] W. S. Levine, editor. *The Control Handbook*. CRG Press and IEEE Press, 1995.

[58] S. Liu and J. E. Bobrow. An analysis of a pneumatic servo system and its application to a computer-controlled robot. *Journal of dynamic systems, measurements, and control*, 110:228–235, September 1988.

[59] L. Ljung. *System Identification*. Prentice Hall PTR, New Jersey, 2nd edition, 1999.

[60] D. G. Luenberger. Observing the state of a linear system. *IEEE Transactions on Military Electronics*, 8, April 1964.

[61] M. Maggiore. *Output Feedback Control: A State-Variable Approach*. PhD thesis, Ohio state university, 2000.

[62] N. A. Mahmoud and H. K. Khalil. Asymptotic regulation of minimum phase nonlinear systems using output feedback. *IEEE Transactions on automatic control*, 41(10):1402–1412, October 1996.

[63] N. A. Mahmoud and H. K. Khalil. Robust control for a nonlinear servomechanism problem. *Int. Journal of Control*, 66(6):779–802, 1997.

[64] J.-C. Maré, O. Geider, and S. Colin. An improved dynamic model of pneumatic actuators. *Int. journal of fluid power*, 1(2):39–47, 2000.

[65] The Math Works Inc. *Optimization Toolbox for Use with MATLAB*, 2000.

[66] J. P. T. Mo. Analysis of compressed air flow through a spool valve. *Proc instn mech engrs*, 203:121–131, 1988.

[67] O. Nelles. *Nonlinear System Identification*. Springer-Verlag, Berlin, Germany, 2001.

[68] H. Nijmeijer and T. I. Fossen, editors. *New Directions in Nonlinear Observer Design*. Lecture Notes in Control and Information Series 244. Springer, Great Britain, 1999.

[69] T. Noritsugu and M. Takaiwa. Robust positioning control of pneumatic servo systems with pressure control loop. In *IEEE Int. Conference on Robotics and Automation*, pages 2613–2618, 1995.

[70] B. M. Nouri, F. Al-Bender, J. Swevers, P. Vanherck, and H. V. Brussel. Modelling a pneumatic servo positioning system with friction. In *Proceedings of the Americal Control Conference*, pages 1067–1071, Chicago, Illinois, June 2000.

[71] S. Oh and H. K. Khalil. Nonlinear output-feedback tracking using high-gain observer and variable structure control. *Automatica*, 33(10):1845–1856, 1997.

[72] Z. Pan, K. Ezal, A. J. Krener, and P. V. Kokotovic. Backstepping design with local optimality matching. *IEEE Transactions on automatic control*, 46(7):1014–1027, July 2001.

[73] S. R. Pandian, Y. Hayakawa, Y. Kanazawa, Y. Kamoyama, and S. Kawamura. Practical design of a sliding mode controller for pneumatic actuators. *Journal of Dynamic Systems, Measurement, and Control*, 119:666–674, December 1997.

[74] S. R. Pandian, K. Leda, Y. Kamoyame, and S. Kawamura. Modelling and control of a pneumatic rotary actuator. In C. R. Burrows and R. Clifford, editors, *Bath Workshop on Power Transmissions and Motion Control*, pages 363–377, Bath, UK, 1998.

[75] S. R. Pandian, F. Takemura, Y. Hayakawa, and S. Kawamura. Practical design of adaptive model-based sliding mode control of pneumatic actuators. *Transactions of the Japan Hydraulics & Pneumatics Society*, 31(4):17–24, July 2000.

[76] M. J. D. Powell. Radial basis functions for multivariable interpolation: A review. In *IMA Conference on Algorithms for the Approximation of Functions and Data*, pages 143–167, Schrivenham, UK, 1985.

[77] R. Rajamani. *Observers for Nonlinear Systems, with Application to Active Automotive Suspensions*. PhD thesis, University of California, Berkeley, 1993.

[78] R. Rajamani and J. K. Hedrick. Adaptive observers for active automotive suspensions: Theory and experiment. *IEEE Transactions on control systems technology*, 3(1):86–93, March 1995.

[79] E. Richard. Non-linear control of a pneumatic servodrive. In C. R. Burrows and K. A. Edge, editors, *Bath Int. Fluid Power Workshop*, pages 59–74, 1989.

[80] E. Richard and S. Scavarda. Comparison between linear and nonlinear control of an electropneumatic servodrive. *Jrnl. of Dynamic systems, measurement, and control*, 118:245–252, November 1996.

[81] W. J. Rugh. *Linear System Theory.* Prentice-Hall, Inc., Upper Saddle River, New Jersey, 2nd edition, 1996.

[82] W. J. Rugh and J. S. Shamma. Research on gain scheduling. *Automatica*, (36):1401–1425, 2000.

[83] S. Scavarda. Linearized models for an electropneumatic cylinder servovalve system. In *3rd Int. Conf. On Advanced Robotics—Towards Third Generation Robotics*, pages 149–160, 1987.

[84] L. E. Schroeder and R. Singh. Experimental study of friction in a pneumatic actuator at constant velocity. In *Winter Annual Meeting*, Anaheim, CA, November 8-13 1992. ASME.

[85] S. Sesmat and S. Scavarda. Static characteristics of a three way electropneumatic servovalve. In *12. Aachener Fluidtechnisches Kolloquium*, pages 643–652, March 1996.

[86] J. L. Shearer. Study of pneumatic processes in the continuous control of motion with compressed air—i and II. *Trans. ASME*, 78:233–249, February 1956.

[87] M.-C. Shih and Y.-F. Huang. Pneumatic servo-cylinder position control using a self-tuning controller. *JSME international journal*, 35(2):247–254, 1992.

[88] J.-J. E. Slotine, J. K. Hedrick, and E. A. Misawa. On sliding observers for nonlinear systems. *Journal of dynamic systems, measurements and control*, 109:245–252, 1987.

[89] J.-J. E. Slotine and W. Li. *Applied Nonlinear Control.* Prentice-Hall, Inc., 1991.

[90] B. W. Surgenor and N. D. Vaughan. Continuous sliding model control of a pneumatic actuator. *Journal of Dynamic Systems, Measurement, and Control*, 119:578–581, 1997.

[91] J. Swevers, F. Al-Bender, C. G. Ganseman, and T. Prajogo. An integrated friction model structure with improved presliding behaviour for accurate friction compensation. *IEEE Transactions on automatic control*, 45(4):675–686, April 2000.

[92] F. Takemura, S. R. Pandian, Y. Hayakawa, and S. Kawamura. Observer design for control of pneumatic cylinder actuators. In *Bath Workshop on Power Transmission & Motion Control*, 1999.

[93] H. Tanaka and H. Wada. Fuzzy control of clutch engagement for automated manual transmissions. *Vehicle System Dynamics*, 24:365–376, 1995.

[94] G. Tao and P. V. Kokotovic. *Adaptive Control of Systems with Actuator and Sensor Nonlinearities*. John Wiley & Sons, Inc., U.S.A, 1996.

[95] A. R. Teel. Adaptive tracking with robust stability. In *Proceedings of the 32nd Conference on Decision and Control*, pages 570–575, San Antonio, Texas, December 1993.

[96] D. Thomasset, E. Richard, S. Scavarda, X. F. Lin, S. Sesmat, and A. Bouhal. Control of an electropneumatic servodrive: A state-affine or sliding approach. In *IFAC 12th Triennial World Congress*, pages 967–970, Sydney, Australia, 1993.

[97] T. Virvalo. *Modelling and Design of a Pneumatic Position Servo System Realized with Commercial Components*. PhD thesis, Tampere University of Technology, 1995.

[98] J. Wang, J. Pu, and P. Moore. A practical control strategy for servo-pneumatic actuator systems. *Control engineering practice*, 7:1483–1488, 1999.

[99] J. Wang, J. Pu, P. R. Moore, and Z. Zhang. Modelling study and servo-control of air motor systems. *International Journal of Control*, 71:459–476, 1998.

[100] J. Wang, D. J. D. Wang, P. R. Moore, and J. Pu. Modelling study, analysis and robust servocontrol of pneumatic cylinder actuator systems. *IEE proc. - control theory and applications*, 148(1):35–42, 2001.

[101] F. M. White. *Fluid Mechanics*. McGRAW-HILL, INC., 3rd edition, 1994.

[102] F. Xiang. *Block-Oriented Nonlinear Control of Pneumatic Actuator Systems*. PhD thesis, Royal institute of technology, KTH, Stockholm, Sweden, 2001.

[103] F. Xiang and J. Wikander. Modelling and control of the volvo pneumatic actuated truck clutch system. technical report, DAMEK Research group, Royal institute of technology, Sweden, 1997.

[104] F. Xiang, J. Wikander, and B. Erisson. Nonlinear control of pneumatic servo - a feedback linearization approach. In *1. Internationales Fluidtechnisches Kollogquium*, pages 207–221, Aachen, 1998.

[105] B. Yao. *Adaptive Robust Control of Nonlinear Systems with Application to Control of Mechanical Systems*. PhD thesis, Univ. California, Berkeley, CA, 1996.

[106] B. Yao and M. Tomizuka. Smooth robust adaptive sliding mode control of robot manipulators with guaranteed transient performance. *ASME Journal of dynamic systems, measurements, and control*, 118(4):764–775, 1996.

[107] B. Yao and L. Xu. Observer-based adaptive robust control of a class of nonlinear systems with dynamic uncertainties. *Int. J. of robust and nonlinear control*, 11:335–356, 2001.