Adaptation in Dynamical Systems

Ivan Tyukin

CAMBRIDGE

CAMBRIDGE www.cambridge.org/9780521198196

This page intentionally left blank

ADAPTATION IN DYNAMICAL SYSTEMS

In the context of this book, adaptation is taken to mean a feature of a system aimed at achieving the best possible performance when mathematical models of the environment and the system itself are not fully available. This has applications ranging from theories of visual perception and the processing of information to the more technical problems of friction compensation and adaptive classification of signals in fixed-weight recurrent neural networks.

Largely devoted to the problems of adaptive regulation, tracking and identification, this book presents a unifying system-theoretic view on the problem of adaptation in dynamical systems. Special attention is given to systems with nonlinearly parametrized models of uncertainty. Concepts, methods, and algorithms given in the text can be successfully employed in wider areas of science and technology. The detailed examples and background information make this book suitable for a wide range of researchers and graduates in cybernetics, mathematical modeling, and neuroscience.

IVAN TYUKIN is an RCUK Academic Fellow in the Department of Mathematics, University of Leicester. His research and scientific interests cover many areas, including the analysis, modeling, and synthesis of systems with fragile, nonlinear, chaotic, and meta-stable dynamics.

ADAPTATION IN DYNAMICAL SYSTEMS

IVAN TYUKIN

University of Leicester and

Saint-Petersburg State Electrotechnical University

CAMBRIDGE
UNIVERSITY PRESS

CAMBRIDGE UNIVERSITY PRESS
Cambridge, New York, Melbourne, Madrid, Cape Town,
Singapore, São Paulo, Delhi, Tokyo, Mexico City

Cambridge University Press
The Edinburgh Building, Cambridge CB2 8RU, UK

Published in the United States of America by Cambridge University Press, New York

www.cambridge.org
Information on this title: www.cambridge.org/9780521198196

First published 2011

Printed in the United Kingdom at the University Press, Cambridge

A catalog record for this publication is available from the British Library

Library of Congress Cataloging in Publication data
Tyukin, Ivan.
Adaptation in Dynamical Systems / Ivan Tyukin.
p. cm
Includes bibliographical references and index.
ISBN 978-0-521-19819-6 (hardback)
1. Dynamics. 2. Control theory–Mathematical models.
3. Neurosciences–Mathematics. I. Title.
QA845.T94 2011
515′.39–dc22 2010051427

ISBN 978 0 521 19819 6 Hardback

Contents

Preface

Adaptation is amongst the most familiar and wide spread phenomena in nature. Since the early days of the nineteenth century it has puzzled researchers in broad areas of science. Since it had often been observed in responsive behaviors of biological systems, adaptation was initially understood as a regulatory mechanism that helps an animal to survive in a changing environment. Later the notion of adaptation was adopted in wider fields of science and engineering.

As a theoretical discipline it began to emerge as a branch of control theory during the first half of the twentieth century. Its beginning was marked by publications discussing basic principles of adaptation and its merits for engineering. Imprecise technology and mechanisms were, perhaps, amongst the strongest practical motivations for such a theory at that time. Various notions of adaptation were adopted by engineers and theoreticians in order to grasp, understand, and implement relevant features of this phenomenon in practice. The first applications of the new theory were simple schemes for extremal control of mechanical systems; these systems could be described by just a few linear ordinary differential equations. Since then adaptive controllers have evolved to encompass substantially more complex devices. The controlling devices themselves can now be viewed as nonlinear dynamical systems with specific input–output properties. Methods for the design and analysis of such systems are currently recognized by many in terms of the theory of adaptive control and systems identification.

Because the initial motivation to develop a theory of adaptation was driven mainly by the demands of mechanical engineering and the need for robust design of otherwise imprecise machines, the domain of application of the theories of adaptation was naturally restricted to the realm of artificial devices and engineering. The focus of the developing theory was restricted, in particular, to the problems of control of a relatively narrow class of well-studied and modeled mechanical systems, many of which were stable in the Lyapunov sense, for which the values of some parameters and variables are unknown and cannot be measured explicitly.

Yet, the potential role of the theory of adaptation was much wider and broader. It has become evident recently that there exists a demand for a systematic theory of adaptation outside of the domain of engineering.

Understanding basic mechanisms and principles of adaptation and regulation is recognized as relevant in physics, chemistry, biology, and brain sciences (Sontag 2004; Fradkov 2005). Because of the huge complexity of the phenomena studied in these domains, using the standard language of each particular science for systematic studies of the phenomenon of adaptation might not be adequate. Therefore in these areas system-theoretic views, irrespective of the particular subjects of study, have exceptional potential.

Apart from in the natural sciences, the needs for further development of the theory of adaptation are evident in handling complex artificial systems. This is especially true when changes in the working environment cannot be predicted a priori or there is a substantial degree of uncertainty about the system's internal state. Although there is a large literature on the theory of adaptive systems, both in the theoretical and in the applied domain, there are several issues preventing explicit application of classical recipes of adaptive control in these fields. These issues with classical schemes are

- the necessity to have a precise mathematical model of a controlled system,
- the requirement that models of uncertainties are linear or convex with respect to unknown or uncertain variables,
- the assumption that the target dynamics is stable in the Lyapunov sense,
- the assumption that a corresponding Lyapunov function for the target motions is available (Sastry and Bodson 1989; Narendra and Annaswamy 1989; Krstić *et al.* 1995; Ljung 1999; Eykhoff 1975; Bastin *et al.* 1992; Fradkov *et al.* 1999).

Every one of these requirements alone limits the role of the existing theory of adaptive systems in solving relevant problems in science. Altogether they constitute the "standard" approach which applies to several canonical cases, which are limited even within the realm of engineering.

The purpose of this work is to contribute towards extending the existing theory of adaptation and adaptive control beyond the scope of its usual applications in engineering to new and non-conventional areas, such as neuroscience and mathematical modeling of biological systems. It is hoped that this extension will create additional opportunities for control theorists to apply their expertise in novel and still developing fields of science; it will also help to expand the synthetic and analytical functions of systems and control theory into the natural sciences.

The focus of this book on the analysis of possible adaptation mechanisms in systems with nonlinear parametrization and unstable target dynamics was influenced by the author's work in the Laboratory for Perceptual Dynamics, RIKEN Brain Science Institute, Japan from November 2001 to March 2007. Neural systems of living organisms, and ultimately the human brain, were the source of inspiration. It became clear very quickly that the standard tools and methods in the arsenal of conventional adaptive control theory do not offer an acceptable explanation for the versatility and robustness of neural systems working in an uncertain environment. The aim therefore was to enhance the theory by making it suitable for the analysis and synthesis of adaptive schemes for nonlinear dynamical systems:

- with potentially Lyapunov-unstable and non-equilibrium target dynamics;
- when explicit definition of the target sets is not possible;
- using minimal, *qualitative*, macro-information about the system, and also allowing substantial uncertainty about the specific mathematical model of the system;
- allowing uncertainty models that are maximally adequate to describe the physical laws of processes and phenomena in the system.

The necessary ingredients of this extended theory of adaptation follow naturally from the logic of its development: from basic principles of the system's organization in the presence of uncertainties to specific laws of regulation. These ingredients include

(1) *methods for analysis* of basic input–output properties of the nonlinear systems; they should allow incomplete knowledge of equations describing the system dynamics; and they also should apply both to stable and to unstable systems;
(2) *principles and methods* of adaptation to disturbances that are unknown a priori and unavailable for measurement; the principles should rely exclusively on the fundamental physical properties of the systems considered; and the adaptation mechanisms should be able to realize these principles using adequate physical models of uncertainties and requiring a minimal amount of measurement information.

The following topics received particular attention: analysis of the completeness, realizability, and state boundedness of interconnections of uncertain dynamical systems; conditions ensuring convergence of the system's state to the target sets and their neighborhoods; designing laws of adaptive regulation and parameter estimation of nonlinearly parametrized models; characterizing the quality of the transient dynamics in systems with uncertainties; and parametric, signal/functional perturbations. In order to provide the reader with the necessary background and also to support our own argumentation a brief review of major classical concepts of adaptation is included.

The content of the book is based largely on the work I had the privilege to carry out together with my colleagues and co-authors.[1] The structure of the book can be summarized as follows. The text is organized into three large parts. The first part (Chapters 1–3) contains mainly introductory and preliminary results. Proofs of lemmas and theorems presented in this introductory part are kept within the main text.

In Chapter 1 we provide an informal discussion of the notion of adaptation followed by an overview of the range of specific problems considered in the text.

Chapter 2 contains background and preliminary results such as basic notions of stability, a very brief introduction to the method of Lyapunov functions, and a particularly important result on the exponential stability of the origin for a class of linear systems of ordinary equations with skew-symmetric matrices.

In Chapter 3 we review and analyze conventional approaches to the problem of adaptive control of nonlinear systems. We formulate the main theoretical and practical issues arising in these standard approaches (Fomin *et al.* 1981; Fradkov 1990; Narendra and Annaswamy 1989; Krstić *et al.* 1995) and their mathematical statements of the problem. These issues include the ambiguity of standard mathematical notions of an adaptive system, performance measures, limitations on defining the system's target sets,[2] restricted classes of the uncertainty models, and requirements for precise knowledge of the mathematical model of a system.

The second part, Chapters 4 and 5, presents the main theoretical results developed in the monograph. In order to preserve the integrity of the text, proofs of statements formulated in this part are given in appendices at the ends of these chapters.

In Chapter 4 we consider nonlinear systems defined in terms of their "input-to-output" and "input-to-state" characterizations given by mappings, or operators in functional, L_p, spaces. We introduce mathematical tools for the analysis of interconnections of dynamical systems with input–output (input–state) operators that are locally bounded in state and provide a formal statement of the problem for functional synthesis of an adaptive system. We demonstrate how this problem can be solved. The solution to the problem of functional synthesis of an adaptive system allows us to formulate various principles of its organization at the macroscopic level: the separation principle, the bottle-neck principle, and the emergence of weakly attracting sets in the interconnections of systems with contracting and wandering dynamics. The latter result is based on Tyukin *et al.* (2008a).

[1] This includes earlier texts such as Tyukin and Terekhov (2008).

[2] One of the most severe restrictions is the requirement for the target dynamics to be globally stable in the sense of Lyapunov. In addition, there is a necessity to specify target sets of the adaptive system a priori. The latter condition either requires prior identification of the system, which contradicts the very essence of adaptive behavior, or leads to enforcing motions that are not necessarily inherent and, hence, optimal to a physical system itself.

In Chapter 5 we utilize the principles derived in the previous chapter in order to provide an adequate statement of the problem of adaptive control and regulation of nonlinear dynamical systems. Its distinctive features are that the uncertainty models are allowed to be nonlinearly parametrized, mathematical models of the system need not be known precisely, the target dynamics is not restricted exclusively to globally Lyapunov-stable motions, and the target sets could be defined implicitly – as invariant sets of an auxiliary dynamical system. Generally, the problem is stated as that of *regulating the influence of uncertainties on the target dynamics to some functional space*. This allows one to refrain from explicit use of the method of Lyapunov functions and, hence, avoid its limitations.

We also consider several specific problems that have substantial theoretical and practical interest:

- adaptive regulation to invariant sets;
- adaptation in interconnected systems;
- state and parameter inference for systems with nonlinear parametrization of uncertainty.

In order to solve these problems two synthesis strategies were developed: the method of the *virtual adaptation algorithm* presented in Tyukin *et al.* (2007b) and the strategy based on purposeful introduction of unstable attracting sets into the system's state space (Tyukin *et al.* 2008a).

In the third part of the book (Chapters 6–8) we illustrate how the theory can be used to solve a number of practical problems of control, processing of information, and identification in mechanics, experimental biophysics, and computer and cognitive science. In particular, we consider the problem of adaptive classification in neural networks with fixed weights, the problem of identifying the dynamics of neuronal cells, and the problem of invariant recognition of spatially distributed information. We discuss why existing techniques cannot be successfully applied to solve these problems, or their application yields practically inefficient outcomes. The content of this part is based on Tyukin *et al.* (2008b), Tyukin *et al.* (2009), and Fairhurst *et al.* (2010).

This book would never have seen the light of day without the continuous support, help, and encouragement I received from many people with whom I have had the honor of working. I would like to express my deep gratitude to Professor V. A. Terekhov, my teacher, friend, and co-author, for his help, fruitful and motivating discussions of the philosophical foundations of the problem of adaptation, and unlimited patience. I am grateful to my colleagues and co-authors Cees van Leeuwen, Danil Prokhorov, Henk Nijmeijer, Erik Steur, David Fairhurst, Alexey Semyanov, and Inseon Song who contributed to the development of the ideas in the monograph. I am grateful to Dr Steven Holt and his colleagues for proof-reading

and editing the monograph at the final stage of production. Finally, I am indebted to my dear wife Tanya, who contributed to the applied side of the project, assisted with the artwork, and also helped me enormously to summarize the results during the later stage of the production of the manuscript. My own personal role was limited to mere listening, interpretation, and writing. As is unfortunately the case in scientific endeavors, errors are inevitable companions. Even though I tried to avoid these unwelcome companions, my own journey is unlikely to be an exception, for which I fully accept sole responsibility. I would therefore be extremely grateful to readers, should they wish to help by contacting me when an error is found.

Notational conventions

Throughout the text the following notational conventions apply.

- Symbol $\mathbb{R}$ defines the field of real numbers and $\mathbb{R}_{\geq c} = \{x \in \mathbb{R} | x \geq c\}$; $\mathbb{N}$ defines the set of natural numbers; and $\mathbb{Z}$ denotes the set of whole numbers or integers.
- Symbol $\mathbb{R}^n$ stands for an n-dimensional linear space over the field of reals.
- $\mathcal{C}^k$ denotes the space of functions that are at least k times differentiable.
- Symbol $\mathcal{K}$ denotes the class of all strictly increasing functions $\kappa : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ such that $\kappa(0) = 0$; symbol $\mathcal{K}_\infty$ denotes the class of all functions $\kappa \in \mathcal{K}$ such that $\lim_{s\to\infty} \kappa(s) = \infty$.
- Let Ω be a set, then by $\mathcal{S}\{\Omega\}$ we denote the set of all subsets of Ω.
- $\|\mathbf{x}\|$ denotes the Euclidian norm of $\mathbf{x} \in \mathbb{R}^n$.
- The notation $|\cdot|$ stands for the absolute value of a scalar.
- The notation $\mathrm{sign}(\cdot)$ denotes the signum function.
- By $L_p^n[t_0, T]$, where $t_0 \geq 0$, $T \geq t_0$, $p \geq 1$, we denote the space of all functions $\mathbf{f} : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ such that

$$\|\mathbf{f}\|_{p,[t_0,T]} = \left(\int_{t_0}^{T} \|\mathbf{f}(\tau)\|^p \, d\tau \right)^{1/p} < \infty.$$

- The notation $\|\mathbf{f}\|_{p,[t_0,T]}$ denotes the $L_p^n[t_0, T]$-norm of $\mathbf{f}(t)$.
- By $L_\infty^n[t_0, T]$, $t_0 \geq 0$, $T \geq t_0$, we denote the space of all functions $\mathbf{f} : \mathbb{R}_{\geq 0} \to \mathbb{R}^n$ such that

$$\|\mathbf{f}\|_{\infty,[t_0,T]} = \operatorname{ess\,sup}\{\|\mathbf{f}(t)\|, t \in [t_0, T]\} < \infty,$$

  and $\|\mathbf{f}\|_{\infty,[t_0,T]}$ stands for the $L_\infty^n[t_0, T]$-norm of $\mathbf{f}(t)$.
- Let $\mathcal{A}$ be a set in $\mathbb{R}^n$, $\mathbf{x} \in \mathbb{R}^n$, and let $\|\cdot\|$ be the usual Euclidean norm in $\mathbb{R}^n$. By the symbol $\|\cdot\|_{\mathcal{A}}$ we denote the following induced norm:

$$\|\mathbf{x}\|_{\mathcal{A}} = \inf_{\mathbf{q}\in\mathcal{A}} \{\|\mathbf{x} - \mathbf{q}\|\}.$$

- Let $\Delta \in \mathbb{R}_{\geq 0}$, then the notation $\|\mathbf{x}\|_{\mathcal{A}_\Delta}$ stands for the following equality:

$$\|\mathbf{x}\|_{\mathcal{A}_\Delta} = \begin{cases} \|\mathbf{x}\|_{\mathcal{A}} - \Delta, & \|\mathbf{x}\|_{\mathcal{A}} > \Delta, \\ 0, & \|\mathbf{x}\|_{\mathcal{A}} \leq \Delta. \end{cases}$$

- The symbol $\|\cdot\|_{\mathcal{A}\infty,[t_0,t]}$ is defined as follows:

$$\|\mathbf{x}(\tau)\|_{\mathcal{A}\infty,[t_0,t]} = \sup_{\tau \in [t_0,t]} \|\mathbf{x}(\tau)\|_{\mathcal{A}}.$$

- Let $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ be given. The function $\mathbf{f}(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^m$ is said to be locally bounded if for any $\|\mathbf{x}\| < \delta$, $\delta \in \mathbb{R}_{>0}$ there exists a constant $D(\delta) > 0$ such that $\|\mathbf{f}(\mathbf{x})\| \leq D(\delta)$.
- Let Γ be an $n \times n$ square matrix, then $\Gamma > 0$ denotes a positive definite (symmetric) matrix. (Γ^{-1} is the inverse of Γ). By $\Gamma \geq 0$ we denote a positive semi-definite matrix.
- We reserve $\|\mathbf{x}\|_\Gamma^2$ to denote the quadratic form $\mathbf{x}^T\Gamma\mathbf{x}$, where $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{x}^T$ is the transpose of $\mathbf{x}$.
- Symbols $\lambda_{\min}(\Gamma)$ and $\lambda_{\max}(\Gamma)$ stand for the minimal and maximal eigenvalues of Γ, respectively.
- By the symbol I we denote the identity matrix.
- The solution of a system of differential equations $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t, \boldsymbol{\theta}, \mathbf{u}(t))$, $\mathbf{u} : \mathbb{R}_{\geq 0} \to \mathbb{R}^m$, $\boldsymbol{\theta} \in \mathbb{R}^d$ passing through point $\mathbf{x}_0$ at $t = t_0$ will be denoted for $t \geq t_0$ as $\mathbf{x}(t, \mathbf{x}_0, t_0, \boldsymbol{\theta}, \mathbf{u})$, or simply as $\mathbf{x}(t)$ if it is clear from the context what the values of $\mathbf{x}_0$ and $\boldsymbol{\theta}$ are and how the function $\mathbf{u}(t)$ is defined.
- Let $\mathbf{u} : \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}_{\geq 0} \to \mathbb{R}^m$ be a function of state $\mathbf{x}$, parameters $\hat{\boldsymbol{\theta}}$, and time t. Let in addition both $\mathbf{x}$ and $\hat{\boldsymbol{\theta}}$ be functions of t. Then, when the arguments of $\mathbf{u}$ are clearly defined by the context, we will simply write $\mathbf{u}(t)$ instead of $\mathbf{u}(\mathbf{x}(t), \hat{\boldsymbol{\theta}}(t), t)$.
- When dealing with vector fields and partial derivatives we will use the following extended notion of the Lie derivative of a function. Let it be the case that $\mathbf{x} \in \mathbb{R}^n$ and $\mathbf{x}$ can be partitioned as follows: $\mathbf{x} = \mathbf{x}_1 \oplus \mathbf{x}_2$, where $\mathbf{x}_1 \in \mathbb{R}^q$, $\mathbf{x}_1 = (x_{11}, \ldots, x_{1q})^T$, $\mathbf{x}_2 \in \mathbb{R}^p$, $\mathbf{x}_2 = (x_{21}, \ldots, x_{2p})^T$, $q + p = n$, and $\oplus$ denotes concatenation of two vectors. We define $\mathbf{f} : \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^n$ such that $\mathbf{f}(\mathbf{x}, \boldsymbol{\theta}, t) = \mathbf{f}_1(\mathbf{x}, \boldsymbol{\theta}, t) \oplus \mathbf{f}_2(\mathbf{x}, \boldsymbol{\theta}, t)$, where $\mathbf{f}_1 : \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^q$, $\mathbf{f}_1(\cdot) = (f_{11}(\cdot), \ldots, f_{1q}(\cdot))^T$, $\mathbf{f}_2 : \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}^p$, and $\mathbf{f}_2(\cdot) = (f_{21}(\cdot), \ldots, f_{2p}(\cdot))^T$. Then $L_{\mathbf{f}_i(\mathbf{x},\boldsymbol{\theta},t)}\psi(\mathbf{x}, t)$, $i \in \{1, 2\}$, denotes the Lie derivative of the function $\psi(\mathbf{x}, t)$ with respect to the vector field $\mathbf{f}_i(\mathbf{x}, \boldsymbol{\theta}, t)$:

$$L_{\mathbf{f}_i(\mathbf{x},\boldsymbol{\theta},t)}\psi(\mathbf{x}, t) = \sum_{j=1}^{\dim \mathbf{x}_i} \frac{\partial \psi(\mathbf{x}, t)}{\partial x_{ij}} f_{ij}(\mathbf{x}, \boldsymbol{\theta}, t).$$

- Let $\mathbf{f}, \mathbf{g} : \mathbb{R}^n \to \mathbb{R}^n$ be differentiable vector fields. Then the symbol $[\mathbf{f}, \mathbf{g}]$ stands for the Lie bracket:
$$[\mathbf{f}, \mathbf{g}] = \frac{\partial \mathbf{f}}{\partial \mathbf{x}} \mathbf{g} - \frac{\partial \mathbf{g}}{\partial \mathbf{x}} \mathbf{f}.$$
The adjoint representation of the Lie bracket is defined as
$$\mathrm{ad}_f^0 \mathbf{g} = \mathbf{g}, \qquad \mathrm{ad}_{\mathbf{f}}^k \mathbf{g} = [\mathbf{f}, \mathrm{ad}_{\mathbf{f}}^{k-1} \mathbf{g}].$$

Part I

Introduction and preliminaries

1

Introduction

Consider a living organism or an artificial mechanism, which we shall refer to for the moment as a system, aiming to perform optimally in an uncertain environment. Despite the fact that the environment may be uncertain, we will suppose that we know the structure of the physical laws of the environment determining plausible motions of the system. Suppose that we even know what the system's action might be and assume that criteria of optimality according to which the system must determine its actions are available. Would we be able to decide a priori which particular action a system must execute or how it should adjust itself in order to maintain its behavior at the optimum?

Depending on the language describing the system's behavior, environment, and uncertainties a number of theoretical frameworks can be employed to find an answer to this non-trivial question. If the available information about the system is limited to a statistical description of the events and their likelihoods are known, then a good methodological candidate is the theory of statistical decision making. On the other hand, if the more sophisticated and involved apparatus of stochastic calculus is used to formalize the behavior of a system in an uncertain environment then a reasonable way to approach the analysis of such an object is to employ the theory of stochastic control and regulation. Despite these differences in how the behavior of a system may be described in various settings, there is a fundamental similarity in the corresponding theoretical frameworks. This similarity, if expressed informally, is that every framework should contain a description of the system's *actions*, *mechanisms for maintaining* and *adjusting* its behavior, and *criteria of optimality* or *goals*. These are in essence components of what we usually understand when calling a system adaptive.

In biology, according to the *Encyclopedia Britannica*, adaptation is described as a "process by which an animal or plant species becomes fitted to its environment; it is the result of natural selection's acting upon heritable variation. Even the simpler organisms must be adapted in a great variety of ways: in their structure, physiology, and genetics, in their locomotion or dispersal, in their means of defense and attack, in their reproduction and development, and in other respects." Actions, regulation and adjustments, criteria of optimality (fitness) are all present in this definition.

In systems theory there is less consensus on what the term "an adaptive system" describes. According to Evleigh (1967) a system is called adaptive if it "is a system which is provided with a means of continuously monitoring its own performance in relation to a given figure of merit or optimal condition and a means of modifying its own parameters by a closed-loop action so as to approach this optimum." Other definitions of an adaptive system have been provided by e.g. L. Zadeh, R. Bellman and R. Kalaba, J. G. Truxall, and V. A. Yakubovich, which we will consider in detail in Chapter 3. Yet they all share the very same ingredients such as actions, adjustments, and criteria of optimality. In this book we will also use the same general understanding of what an adaptive system means, though we will allow some technical deviations from these classical definitions.

Because the phenomenon of adaptation is generally understood as a special regulatory process in which a system maintains its performance at the optimum by adjusting itself and its actions, a natural language to analyze the phenomenon of adaptation is the language of systems and control theories. There are many inspiring and excellent monographs covering the general topic of adaptation. A non-exhaustive list of influential texts includes Tsypkin (1968), Tsypkin (1970), Narendra and Annaswamy (1989), Sastry and Bodson (1989), Krstić *et al.* (1995), and Fradkov *et al.* (1999). Hence it is reasonable to ask whether anything new can be added to this wealth of intellectual resources by one more text. As is often the case in science, novelty is a frequent consequence of a new formulation of a known problem, or it emerges as a result of answering new questions about familiar objects.

The purpose of this monograph is to contribute to the theory of adaptive systems by presenting a list of challenging questions and providing a unified theory that would allow one to find answers to these questions in a rigorous and systematic way. There are numerous examples illustrating the benefits of mathematical analysis of the phenomenon of adaptation: they range from solving the problems of crisis predictions (Gorban *et al.* 2010) to explaining plausible mechanisms of cell functioning in biology (Moreau and Sontag 2003), understanding the evolution of species (Gorban 2007), and motor learning (Smith *et al.* 2006). It is the author's hope that the methods developed here will also be useful for addressing open questions in science.

Below we present several examples of these questions emerging across the disciplines ranging from brain modeling to the issues of precise perturbation compensation in engineering and the problems of signal classification and pattern analysis in artificial intelligence. These examples are split into two major groups related to the problems of observation and regulation. For each of these groups we provide informal statements of the corresponding adaptation problems. These statements are not to be considered final and we will reshape them later on in the text. The function of these statements is to emphasize different facets of the problem of adaptation. There was no specific reason for choosing particular subject areas from which the examples are taken except probably the author's personal interests and bias.

1.1 Observation problems

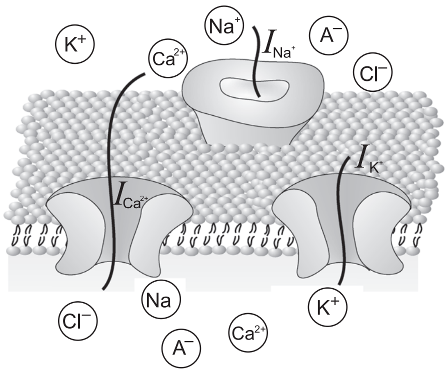
The problem of state and parameter reconstruction of dynamical systems from the values of just few variables is a common task in the domain of mathematical modeling. Despite the fact that this problem received substantial attention in the past (see e.g. Bastin and Dochain (1990) and Ljung (1999)), there are gray spots in the literature for which finding a computationally plausible and theoretically rigorous solution remains a non-trivial task. The usual sources of difficulties are the presence of nonlinear parametrization, and the fact that we are not allowed to influence the system's behavior by varying its inputs in a reasonably broad class of functions.

There are numerous observation problems of this kind in physics. We start by presenting two examples from the domains of biophysics and neuroscience.

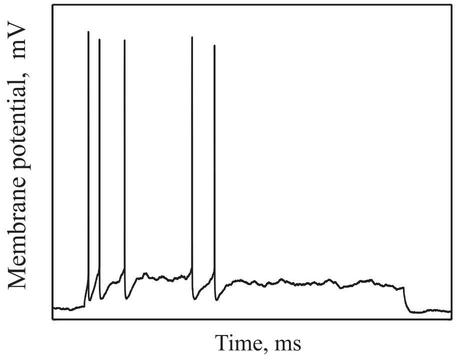
1.1.1 Example: quantitative modeling in biophysics and neuroscience

Let us consider the problem of simultaneous state and parameter reconstruction of models describing the dynamics of neural cells. Most of the available models of individual biological neurons are systems of ordinary differential equations describing the cell's response to stimulation; their parameters characterize variables such as time constants, conductances, and response thresholds, which are important for relating the model responses to the behavior of biological cells. Even the simplest models in this class, such as the Morris–Lecar model (Morris and Lecar 1981), are a great source of inspiration from the modeler's perspective (see Figure 1.1). This model is defined by the following system of equations:

$$
\begin{aligned}
\dot{V} &= \frac{1}{C}(-\bar{g}_{\mathrm{Ca}} m_{\infty}(V)(V - E_{\mathrm{Ca}}) - \bar{g}_{\mathrm{K}} w(V - E_{\mathrm{K}}) - \bar{g}_{\mathrm{L}}(V - E_{\mathrm{L}})) + I, \\
\dot{w} &= -\frac{1}{\tau(V)} w + \frac{w_{\infty}(V)}{\tau(V)},
\end{aligned}
\tag{1.1}
$$



Physical variables



Measured variable

Figure 1.1 Incompleteness of information in quantitative modeling of a cell's behavior. The diagram on the left shows a basic phenomenological description of how currents propagate through a patch of the cell's membrane. There is a number of voltage-dependent channels, such as for Ca, Na, and K depicted in the figure. These channels pump ions through the membrane, and each of these channels has its own dynamics. The problem is that recording currents through every single channel in the membrane simultaneously is not always possible. Thus they must be estimated from available measurements, such as the membrane potentials depicted in the right diagram.

where

$$m_\infty(V) = 0.5\left(1 + \tanh\left(\frac{V - V_1}{V_2}\right)\right),$$
$$w_\infty(V) = 0.5\left(1 + \tanh\left(\frac{V - V_3}{V_4}\right)\right),$$
$$\tau(V) = T_0 \frac{1}{\cosh\left((V - V_3)/2V_4\right)}.$$

The variable V in (1.1) corresponds to the measured membrane potential, and I models an external stimulation current. The parameters $\bar{g}_{\mathrm{Ca}}$, $\bar{g}_{\mathrm{K}}$, and $\bar{g}_{\mathrm{L}}$ stand for the maximal conductances of the calcium, potassium, and leakage currents, respectively; C is the membrane capacitance; V_1, V_2, V_3, and V_4 are the parameters of the gating variables; T_0 is the parameter regulating the time scale of ionic currents; E_{Ca} and E_{K} are the Nernst potentials of the calcium and potassium currents; and E_{L} is the rest potential.

The total number of parameters in system (1.1) is 12, excluding the stimulation current I. Some of these parameters can be considered typical. For example the values of the Nernst potentials for calcium and potassium channels, E_{Ca} and E_{K}, are known and usually are set as $E_{\mathrm{Ca}} = 100\,\mathrm{mV}$ and $E_{\mathrm{K}} = -70\,\mathrm{mV}$ (Koch 2002).

The value of the rest potential, E_{L}, can be measured explicitly. The values of the parameters, $\bar{g}_{\mathrm{Ca}}$, $\bar{g}_{\mathrm{K}}$, $\bar{g}_{\mathrm{L}}$, and T_0, however, may vary substantially from one cell to another, and in general they are dependent on the conditions of the experiment. For example, the values of $\bar{g}_{\mathrm{Ca}}$, $\bar{g}_{\mathrm{K}}$, and $\bar{g}_{\mathrm{L}}$ depend on the density of ion channels in a patch of the membrane; and the value of T_0 is dependent on temperature. Hence, to be able to model the dynamics of individual cells, we have to recover these values from data.

Another example of the same nature is a model predicting the force generated by rat skeletal muscles during brief isometric contractions (Wexler *et al.* 1997). The model consists of three coupled nonlinear differential equations,

$$\begin{aligned}\dot{F} &= aT\left(1 - \frac{F}{F_{\mathrm{m}}}\right) - \frac{F}{\tau_1 + \tau_2 T/T_0},\\ \dot{T} &= k_1 T_0 C^2 - (k_1 C^2 + k_2)T,\\ \dot{C} &= 2(k_1 C^2 + k_2)T - 2k_1 T_0 C^2 + kC_0 - (k + k_0)C,\end{aligned} \tag{1.2}$$

where F is the force generated by the muscles, T is the concentration of Ca^{2+}–troponin complex, and C is the concentration of Ca^{2+} in the sarcoplasmic reticulum. The parameters τ_1, C_0, and k are fixed, while the parameters k_0, k_1, k_2, τ_2, F_{m}, a, and T_0 are free. The values of T and C are not available for direct observation, and the values of F over time can be measured. The question is whether it is possible to reconstruct the free parameters of the model together with the values of the concentrations T and C from the measurements of F. As in the previous example, we are dealing with an uncertain system in which the unknown parameters enter the right-hand side of the corresponding differential equations nonlinearly.

1.1.2 Example: adaptive classification in neural networks

The problem of estimating parameters of ordinary differential equations is not limited to the domain of modeling. It has an important relative in the field of artificial intelligence, namely the problem of adaptive classification of signals. An example of this problem is provided below.

Consider a set of signals defined as

$$\begin{aligned}\mathcal{F} &= \{f_i(\xi(t), \theta_i)\},\ i \in \{1, \ldots, N_f\},\\ f_i &: \mathbb{R} \times \mathbb{R} \to \mathbb{R},\ f_i(\cdot, \cdot) \in \mathcal{C}^0,\\ \xi &: \mathbb{R}_{\geq 0} \to \mathbb{R},\ \xi(\cdot) \in \mathcal{C}^1 \cap L_\infty[0, \infty],\end{aligned} \tag{1.3}$$

where $\theta_i \in \Omega_\theta \subset \mathbb{R}$ are parameters of which the values are unknown a priori, $\Omega_\theta = [\theta_{\min}, \theta_{\max}]$ is a bounded interval, and $\xi(t)$ is a known and bounded function.

Signals $f_i(\xi(t), \theta_i)$ constitute the set of variables chosen to represent the state of an object.

Let $s \in \mathcal{F}$ be an element of class $\mathcal{F}$. The values of $s(t, \theta)$ are fed into the following system of differential equations:

$$
\begin{aligned}
\dot{x}_j &= \sum_{m=1}^{N} c_{j,m}\sigma(\mathbf{w}_{j,m}^{\mathrm{T}}\mathbf{x} + w_{s,j,m}s(t) + w_{\xi,j,m}\xi + b_{j,m}),\\
j &\in \{1, \dots, N_x\},\\
\mathbf{x} &= \mathrm{col}(x_1, \dots, x_{N_x}),\ \mathbf{x}(t_0) = \mathbf{x}_0.
\end{aligned}
\tag{1.4}
$$

System (1.4) is often referred to as the recurrent neural network with standard multilayer perceptron structure. Here $c_{j,m}$, $\mathbf{w}_{j,m}$, $w_{s,j,m}$, $w_{\xi,j.m}$, and $b_{j,m}$ are parameters of which the values are fixed, and the function $\sigma : \mathbb{R} \to \mathbb{R}$ is sigmoidal:

$$
\sigma(p) = \frac{1}{1 + e^{-p}}.
$$

The problem of classification can now be stated as follows: is there a network of type (1.4) that is able to recover uncertain parameters i and θ_i from the input $s(t)$ (see Figure 1.2)? Informally, this means that there exist two sets of functions of the network state $\mathbf{x}$ and input $s(t)$:

$$
\begin{aligned}
&\{h_{f,j}(\mathbf{x}(t), s(t))\},\ \{h_{\theta,j}(\mathbf{x}(t), s(t))\},\\
&h_{f,j} : \mathbb{R}^{N_x} \times \mathbb{R} \to \mathbb{R},\ h_{\theta,j} : \mathbb{R}^{N_x} \times \mathbb{R} \to \mathbb{R},\ j \in \{1, \dots, N_f\},
\end{aligned}
$$

such that the values of i and θ_i can be inferred from $\{h_{f,j}(\mathbf{x}(t), s(t))\}$ and $\{h_{\theta,j}(\mathbf{x}(t), s(t))\}$, respectively, within a given finite interval of time.

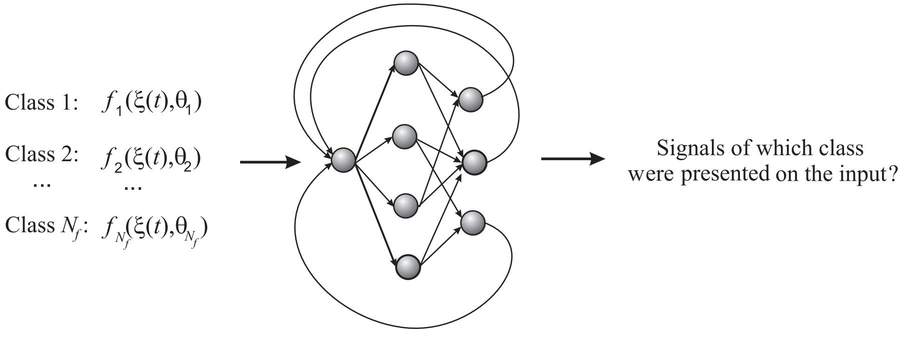


Figure 1.2 Adaptive classification of temporal signals in recurrent neural networks with fixed weights.

Networks (1.4) form an important class of computational structures of which the practical utility and capabilities are widely acknowledged in the literature (Haykin 1999). This class has been shown to be successful in dealing with a wide range of classification problems, including that of classifying signals from (1.3), provided that the values of θ_i in (1.3) are known. Empirical studies suggest that recurrent neural networks of this type are able to solve the classification problem (Feldkamp and Puskorius 1997; Prokhorov *et al.* 2002a) even if θ_i are unknown. The question, however, is how to show that this is indeed the case.

The problem of adaptive classification may look different from the previous examples in the domain of modeling. Indeed, here we have an existence question, whereas in the examples before we asked for a specific estimation algorithm. Despite these differences, there is substantial similarity in these problems. To be able to see this similarity, we would like to state the observation problem in a more general context below.

1.1.3 Preliminary statement of the problem

Let us generalize model (1.1) to the following class of dynamical systems:

$$\begin{aligned} &\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, \boldsymbol{\theta}) + \mathbf{g}(\mathbf{x}, \boldsymbol{\theta})u(t), \quad \mathbf{x}(t_0) \in \Omega_x \subset \mathbb{R}^n, \\ &y = h(\mathbf{x}), \quad \mathbf{x} \in \mathbb{R}^n, \quad \boldsymbol{\theta} \in \Omega_\theta,\ \Omega_\theta \subset \mathbb{R}^d, \quad y \in \mathbb{R}, \end{aligned} \tag{1.5}$$

where $\mathbf{f}, \mathbf{g} : \mathbb{R}^n \times \mathbb{R}^d \to \mathbb{R}^n$, $h : \mathbb{R}^n \to \mathbb{R}$, and $u : \mathbb{R} \to \mathbb{R}$ are continuous and differentiable functions. The variable x stands for the state vector, $u \in \mathcal{U} \subset \mathcal{C}^1[t_0, \infty)$ is the known input, $\boldsymbol{\theta}$ is the vector of unknown parameters, and y is the output of (1.5). Given that the right-hand side of (1.5) is differentiable, for any $\mathbf{x}' \in \Omega_x$, $u \in \mathcal{C}^1[t_0, \infty)$ there exists a time interval $\mathcal{T} = [t_0, t_1]$, $t_1 > t_0$ such that a solution $\mathbf{x}(t, \mathbf{x}')$ of (1.5) passing through $\mathbf{x}'$ at t_0 exists for all $t \in \mathcal{T}$. Hence, $y(t) = h(\mathbf{x}(t))$ is defined for all $t \in \mathcal{T}$. For the sake of convenience we will assume that the interval $\mathcal{T}$ of the solutions is large enough or even coincides with $[t_0, \infty)$ when necessary.

Taking these notations into account, we can now state the observation problem as follows: suppose that we are able to measure the values of $y(t)$ precisely; can the values of $\mathbf{x}'$ and the parameter vector $\boldsymbol{\theta}$ be recovered from the observations of $y(t)$, and, if so, how? In particular, we are interested in finding a computational algorithm

$$\dot{\boldsymbol{\xi}} = \mathbf{p}(\boldsymbol{\xi}, t, u(t), y(t)), \qquad \boldsymbol{\xi}_0 = \boldsymbol{\xi}(t_0) \in \Omega_\xi, \tag{1.6}$$

such that for some known functions $\mathbf{h}_x(\boldsymbol{\xi})$ and $\mathbf{h}_\theta(\boldsymbol{\xi})$ and given number $\varepsilon > 0$ the following property holds:

$$\begin{aligned} \limsup_{t\to\infty} \|\mathbf{h}_x(\boldsymbol{\xi}(t,\boldsymbol{\xi}_0)) - \mathbf{x}(t)\| &\leq \varepsilon, \\ \limsup_{t\to\infty} \|\mathbf{h}_\theta(\boldsymbol{\xi}(t,\boldsymbol{\xi}_0)) - \boldsymbol{\theta}\| &\leq \varepsilon \ \forall\, \boldsymbol{\xi}_0 \in \Omega_\xi. \end{aligned} \tag{1.7}$$

In order to see how this statement is related to the adaptive classification problem in neural networks it is sufficient to notice that (1) the right-hand side of (1.4) can approximate an arbitrary continuous function in a bounded domain (hence it can model the right-hand side of (1.6)), and (2) the function s in (1.4) may be modeled as an output of system (1.5).

System (1.5) can be viewed as an external object or environment, and computational algorithm (1.6) and the functions $\mathbf{h}_x(\boldsymbol{\xi})$ and $\mathbf{h}_\theta(\boldsymbol{\xi})$ constitute the adapting system. The system responds to changes in the environment so that its performance (defined here by (1.7)) reaches an acceptable level and is maintained at this level indefinitely. If (1.5) were linearly parametrized, i.e. the functions $\mathbf{f}(\mathbf{x},\boldsymbol{\theta})$ and $\mathbf{g}(\mathbf{x},\boldsymbol{\theta})$ were linear in $\boldsymbol{\theta}$, then in order to answer this question we could employ the well-developed machinery of standard adaptive observers design (Marino and Tomei 1995b). Yet, as model (1.1) illustrates, the assumption of linear parametrization does not always hold. Hence alternative methods are needed.

This question (as well as other related issues of parameter estimation of nonlinear ordinary equations) is discussed in detail in Chapter 5. In addition to presenting sufficient conditions stipulating the mere existence of solutions to the observation problem, we provide specific computational algorithms (1.6) satisfying the required asymptotic properties (1.7). Special attention is paid to the analysis of the convergence rates of these algorithms. One may expect that the rates of convergence are likely to depend on the classes of nonlinearities in the models. This is indeed the case, as we illustrate in Chapter 5.

1.2 Regulation problems

Suppose now that we are not interested in reconstructing the values of the state and parameters of system (1.5). We do, however, require that the system's state is regulated to a given set in the system's state space for all $\boldsymbol{\theta} \in \Omega_\theta$. Consider for example the following system:

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= -x_1 - x_2 + g(x_1, x_2, \boldsymbol{\theta}) + u, \end{aligned} \tag{1.8}$$

where x_1 and x_2 are the state variables, $\boldsymbol{\theta} \in \Omega_\theta$, $\Omega_\theta \subset \mathbb{R}^d$ is the vector of unknown parameters, $g : \mathbb{R} \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}$ is a continuous function, and $u : \mathbb{R} \to \mathbb{R}$ is an input. Equations (1.8) describe a large class of mechanical and chemical systems. If we accept a simplified interpretation in which x_1 is the position of an object in space and x_2 is its velocity then $g(x_1, x_2, \boldsymbol{\theta})$ could stand for the friction terms (Canudas de Wit and Tsiotras 1999). If (1.8) is a model of a bio-reactor then x_1 and x_2 are the substrate concentrations and $g(x_1, x_2, \boldsymbol{\theta})$ could stand for the standard Michaelis–Menten nonlinearity (Bastin and Dochain 1990). In all these cases the function $g(x_1, x_2, \boldsymbol{\theta})$ is nonlinear in $\boldsymbol{\theta}$. The question is whether there is a function $u(x_1, x_2, \hat{\boldsymbol{\theta}})$ such that the solutions of (1.8) converge to the origin for all $\boldsymbol{\theta} \in \Omega_\theta$.

1.2.1 Example: non-dominating adaptive regulation

If no additional constraints are imposed then the above problem can be easily solved within the framework of dominating functions (Lin and Qian 2002b; Putov 1993). In this framework the original nonlinearly parametrized uncertainty $g(x_1, x_2, \boldsymbol{\theta})$ is replaced by a dominating linearly parametrized one $|g(x_1, x_2, \boldsymbol{\theta})| \leq \bar{g}(x_1, x_2)^{\mathrm{T}} \boldsymbol{\eta}$ and the problem is then solved using the standard method of Lyapunov functions (see Lin and Qian (2002b) for details). Although practical, this approach is not necessarily optimal for systems with limited resources. If the system is a living organism then using resources excessively may be an important limiting factor. The same argument applies for artificial yet autonomous systems. For these classes of systems a reasonable assumption is that the system is penalized for excessive use of domination terms in control.

One of the simplest examples of such non-dominating control schemes is the compensatory control $u = -g(x_1, x_2, \boldsymbol{\theta})$. If the value of $\boldsymbol{\theta}$ were known then this feedback would be able to steer the system to the origin. The problem, however, is that the values of $\boldsymbol{\theta}$ are unknown and the function $g(x_1, x_2, \boldsymbol{\theta})$ is nonlinearly parametrized. A possible strategy would be to make an initial guess at $\boldsymbol{\theta}$ and then adjust its value over time. The question, however, is how should one do this? This is a typical example of the non-dominating adaptation problem, of which a more formal statement is provided at the end of this section.

1.2.2 Example: adaptive tuning to bifurcations

In the previous case the set to which the system solutions are to converge was a priori known. There are systems for which information of this kind is not explicitly available. Their goal is not to reach a given state in the system's state space but rather to maintain adaptively a certain functional property of the system. An interesting example is the problem of adaptive self-tuning of a hearing nerve cell

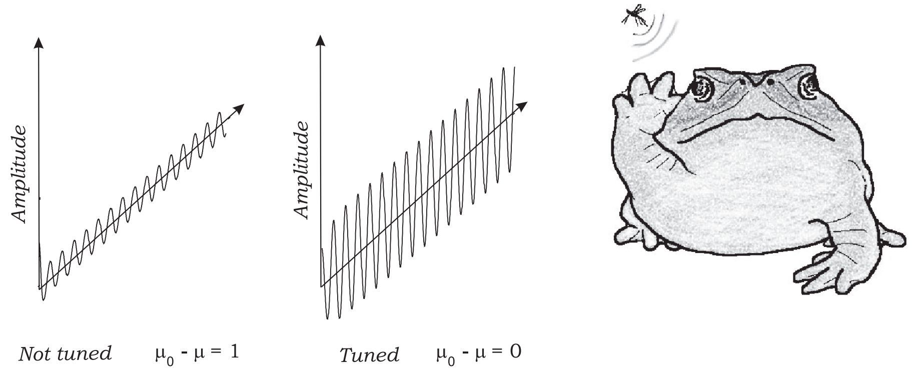


Figure 1.3 A diagram illustrating sensitivity control in the hearing cells via tuning to Andronov–Hopf bifurcation. The left panel shows response of the model, (1.9), $\lambda = 1, \omega = 1$, to a stimulus $u(t) = \sin(t)$ at $\mu_0 - \mu = 1$. The middle panel shows the response of the "tuned" model with $\mu_0 = \mu$ to the same stimulus. We see that the amplitude of oscillations in the tuned cell is many times larger than that in the untuned one.

(Moreau and Sontag 2003). The dynamics of the cell can be described by a nonlinear oscillator

$$\ddot{x} + (\mu_0 - \mu)\dot{x} + \lambda\dot{x}^3 + \omega^2 x = u(t),\ \lambda \in \mathbb{R}_{>0}, \tag{1.9}$$

where μ is the parameter to be adjusted and $u(t)$ is the input (stimulus). When the value of μ is set close to μ_0 the system's dynamics approaches supercritical Andronov–Hopf bifurcation. This leads to the possibility of substantial amplifications of signals in the specific frequency range (Camalet *et al.* 2000) (see Figure 1.3). The question, however, is what are the mechanisms ensuring that the system is always operating in close proximity to the bifurcation?

It has been shown in Moreau and Sontag (2003) that a simple adaptation procedure,

$$\dot{\mu} = -a \log \sqrt{x^2 + \dot{x}^2/\omega^2} - b,\ a, b \in \mathbb{R}_{>0},\ a < b^2,$$

provides the required property. The value of this and similar results is difficult to overestimate, for they provide plausible adaptation models that can be searched for and validated in experiments. Moreover, the example motivates us to generalize this question even further and ask whether there exists a general recipe for deriving such feedbacks. This is the class of problems also known as adaptive tuning to bifurcations.

1.2.3 Example: adaptive regulation to invariant sets

The question of adaptive tuning to bifurcations is closely related to another interesting problem, that of adaptive regulation to invariant sets. The need to pose the problem of regulation as that of steering to a given invariant set emerges under those conditions, when the target set is not completely known. For example, we may know that the target set is necessarily an equilibrium or a periodic orbit, yet the precise location and shape of these sets might not be known. In Chapter 5 we present a set of results that allow us to solve both the problem of adaptive tuning to bifurcation and the problem of regulation to invariant sets in a unified manner using the method of a virtual algorithm of adaptation.

1.2.4 Preliminary statement of the problem

Let us now summarize the regulation problems considered above. Suppose that the system's motions are governed by the following set of equations:

$$\dot{\mathbf{x}} = \mathbf{f}_0(\mathbf{x}) + \mathbf{f}(\mathbf{x}, \boldsymbol{\theta}) + \mathbf{g}(\mathbf{x})u, \tag{1.10}$$

where $\mathbf{f}_0$, $\mathbf{f}$ and $\mathbf{g}$ are continuous functions and $\boldsymbol{\theta}$ is the vector of unknown parameters. The standard adaptive regulation question is that of whether there is a feedback

$$\begin{aligned} u &= u(\mathbf{x}, \hat{\boldsymbol{\theta}}), \\ \dot{\hat{\boldsymbol{\theta}}} &= A(\mathbf{x}) \end{aligned} \tag{1.11}$$

such that the system's state is stirred asymptotically to the origin. As our examples motivate, in addition to this standard requirement, we may wish to require that a functional of $u(\mathbf{x}, \boldsymbol{\theta})$ is optimized over time. A reasonable requirement could be that

$$\lim_{t\to\infty} \int_0^t (u(\mathbf{x}(\tau, \mathbf{x}_0), \boldsymbol{\theta}) - u(\mathbf{x}(\tau, \mathbf{x}_0), \hat{\boldsymbol{\theta}}(\tau)))^2 \, d\tau \to \min.$$

In the next chapters we shall see when, how, and in what sense such requirements may be satisfied.

Let us now suppose that system (1.10) undergoes a certain bifurcation at $\boldsymbol{\theta} = 0$ and its operating conditions require that this regime is maintained adaptively. Yet the values of $\boldsymbol{\theta}$ may change abruptly. In this case the adaptive regulation problem can be stated as that of looking for the functions (1.11) such that

$$\lim_{t\to\infty} \mathbf{f}(\mathbf{x}(t, \mathbf{x}_0), \boldsymbol{\theta}) + \mathbf{g}(\mathbf{x}(t, \mathbf{x}_0))u(\mathbf{x}(t, \mathbf{x}_0), \hat{\boldsymbol{\theta}}(t)) = 0.$$

The question, however, is how do we find such algorithms?

1.3 Summary

The examples of problems presented above are a sample of the sorts of challenges we wish to attack in this book. Although they originate in rather different fields, they are connected together by the need for theoretical assessment of how adaptation may be organized and analyzed in these systems. The systems considered in the examples contain nonlinearly parametrized uncertainties, and their target behavior need not necessarily be stable. Even the definition of the target sets is allowed to bear a degree of uncertainty. The main theoretical focus of this book is to provide a systematic extension of the existing theories of adaptation so that all these challenging problems, irrespective of their field of origin, can be addressed in a rigorous and unified manner.

The main strategy in our quest to create such an extension can be described as that of looking beyond the usual presumptions in the domain of analysis and synthesis of adaptive systems. In particular, we will concentrate on breaking through the following constraints (presumptions) which are often implicitly or explicitly imposed in the standard statements of the problem of adaptation:

(1) a practically successful adaptive system must be stable in the sense of Lyapunov;
(2) the analysis and synthesis methods should operate exclusively and at all times with those variables of which the values are available for direct observations;
(3) a successful system should be able to maintain its optimal performance over infinitely long intervals of time.

In order to be able to avoid these constraints when their presence in the problem is not at all necessary, we shall present a systematic view on the problem of adaptation starting from the very basic principles of a system's organization and passing on to the laws implementing these principles in particular settings. As a result of this hierarchical approach, a likely object of our analysis would be a system that is adapting, albeit not necessarily being globally stable in the sense of Lyapunov.

A possible way to develop a feeling for why some constraints are important whereas others can be removed from the problem is to look at the problem retrospectively (Lakatos 1976). In the next two chapters we review the most influential concepts of adaptation in the literature of systems and control theories and justify the research program that guided the development of our own contribution.

2

Preliminaries

Determining asymptotic properties of dynamical systems, including the formulation of a qualitative picture of the system's trajectories over large intervals of time, is one of the central questions of modern theory for adaptive systems. This is not surprising, for the very reason for adaptation is the lack of available measurement information. If such information is not available a priori, and carrying out numerical or physical experiments is not a feasible option, assessment of the qualitative properties of the system's behavior is often the only way to characterize the system. What are these qualitative properties? Informally, from these properties we should be able to tell, for example, how a system might respond to external perturbations, or how the system's variables behave over long intervals of time. Formally, we may wish to know whether the system is stable in some sense, whether its trajectories are bounded, and to what sets these trajectories will be confined with time.

In this chapter we shall provide a brief summary and necessary background about these qualitative properties of dynamical systems. We do not wish, however, to present an exhaustive review of all concepts. There are many excellent texts devoted to detailed analysis of every single issue mentioned above. Here we will rather review these concepts with a level of detail and generality just sufficient for developing a qualitative understanding of the problem of adaptation and the basics of methods of adaptive regulation. Let us start with the simplest and at the same time the most difficult notion for analysis: the notion of an *attracting set*.

2.1 Attracting sets and attractors

In order to introduce the notion of an attracting set it is often useful to think of a system as a family of parametrized maps $\mathbf{x} : \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$. In the modeling language this will restrict our attention to the following models: $\mathbf{x}(t, \mathbf{x}_0)$, or *flows*, where t stands for the time instance and $\mathbf{x}_0$ is the value of the system's state at

$t = 0$. Usually an additional semi-group property is imposed on $\mathbf{x}(t, \mathbf{x}_0)$:

$$\mathbf{x}(t', \mathbf{x}(t'', \mathbf{x}_0)) = \mathbf{x}(t' + t'', \mathbf{x}_0).$$

Although this assumption is not entirely necessary for producing the definition, we shall keep this possibility in mind, for it provides us with a link to physical reality. We would also like to notice that models $\mathbf{x}(t, \mathbf{x}_0)$ do not yet have any explicit reference to any inputs or other factors acting on a real system externally. These factors are all hidden in the model. The main reason for this is that we want to keep the notation compact. However, should such a necessity arise, one can easily modify the definitions below so that all external variables are made explicitly visible.

Before we proceed with a formal definition explaining what we will understand under the term *attracting set* we will need to introduce one additional notion. This is the notion of an *invariant set* with respect to a given flow $\mathbf{x}(t, \mathbf{x}_0)$.

Definition 2.1.1 A set $\mathcal{A} \subset \mathbb{R}^n$ is called invariant with respect to the flow $\mathbf{x}(t, \mathbf{x}_0)$ iff for all $\mathbf{x}_0 \in \mathcal{A}$, $t \in \mathbb{R}$ the following property holds:

$$\mathbf{x}(t, \mathbf{x}_0) \in \mathcal{A}.$$

It is sometimes useful to distinguish between forward-invariant and backward-invariant sets, of which the definitions are provided below.

Definition 2.1.2 A set $\mathcal{A} \subset \mathbb{R}^n$ is called forward-invariant with respect to the flow $\mathbf{x}(t, \mathbf{x}_0)$ iff for all $\mathbf{x}_0 \in \mathcal{A}$, $t \in \mathbb{R}_{\geq 0}$ we have that $\mathbf{x}(t, \mathbf{x}_0) \in \mathcal{A}$. The set is backward-invariant iff $\mathbf{x}(t, \mathbf{x}_0) \in \mathcal{A}$ for all $\mathbf{x}_0 \in \mathcal{A}$, $t \in \mathbb{R}_{\leq 0}$.

Simple examples of invariant sets are equilibria, limit cycles, or just orbits of autonomous systems in the state space (see Figure 2.1). Whether a given set is invariant or not is an important item of information for the analysis of uncertain systems. Indeed, if we know that $\mathcal{A}$ is invariant then all trajectories passing through at least one point of $\mathcal{A}$ will necessarily remain there for all t. Despite its clear benefits for analysis, the notion of invariance of a set with respect to $\mathbf{x}(t, \mathbf{x}_0)$ is not a very instrumental property from the viewpoint of regulation. Suppose that we know that $\mathcal{A} \subset \mathbb{R}^n$ is forward-invariant with respect to $\mathbf{x}(t, \mathbf{x}_0)$, and let us suppose that the system's state passes through a point that does not belong to $\mathcal{A}$. Let us finally assume that for some reason we wish to know whether the system's state will reach $\mathcal{A}$ or its arbitrarily small vicinity in finite time. For example, if $\mathbf{x}(t, \mathbf{x}_0)$ models trajectories of an organism in space and $\mathcal{A}$ is the set of locations of food then a question of vital importance for this organism is whether it should employ its resources to initiate movements towards the set $\mathcal{A}$ or whether it should wait for some time until external forces such as the flow of water or wind eventually bring it to $\mathcal{A}$ in a reasonable amount of time. Answering this question might not be a feasible

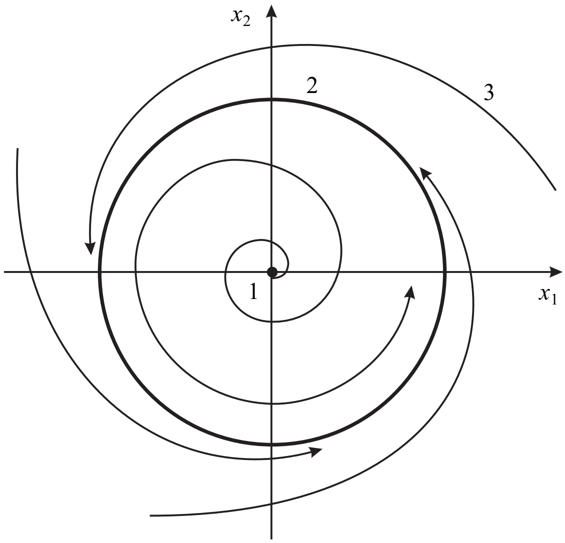


Figure 2.1 Examples of invariant sets of autonomous dynamical systems: 1, equilibrium; 2, a limit cycle; and 3, an orbit.

exercise in the absence of additional information about the system. We may still view this property as preferable and desirable. An invariant set possessing such a property is often referred to as *attracting*. Formally the notion of an attracting set is provided in the next definition

Definition 2.1.3 A closed[1] invariant set $\mathcal{A} \subset \mathbb{R}^n$ is called attracting iff

(1) there is a neighborhood $U(\mathcal{A})$ of $\mathcal{A}$ such that

$$\mathbf{x}(t, \mathbf{x}_0) \in U(\mathcal{A}) \ \forall \ \mathbf{x}_0 \in U(\mathcal{A}), t \in \mathbb{R}_{\geq 0}; \tag{2.1}$$

(2) the following limiting property holds

$$\lim_{t \to \infty} \|\mathbf{x}(t, \mathbf{x}_0)\|_{\mathcal{A}} = 0 \ \forall \ \mathbf{x}_0 \in U(\mathcal{A}). \tag{2.2}$$

According to Definition 2.1.3 a closed invariant set $\mathcal{A}$ is attracting if there is a forward-invariant neighborhood $U(\mathcal{A})$ such that all trajectories starting in $U(\mathcal{A})$ converge to $\mathcal{A}$ asymptotically. At first glance the definition is rather general and clear. Although this is indeed the case, there are situations in which a generalization of this notion may be required. Let us consider an example.

[1] Let us remind the reader that a set $\mathcal{A}$ is closed iff it contains all of its limit points. For example, if $\mathcal{A}$ is closed and $a_i \in \mathcal{A}$ $i = 1, \ldots, \infty$ is a sequence then $\lim_{i \to \infty} a_i$ (if exists) should also belong to $\mathcal{A}$. If $\mathcal{A}$ is an interval then it is closed iff $\mathcal{A}$ contains its boundaries. A point in $\mathbb{R}^n$ is obviously a closed set. In addition to these simple examples there are more exotic instances of closed sets such as the Cantor set (also known as "Cantor dust").

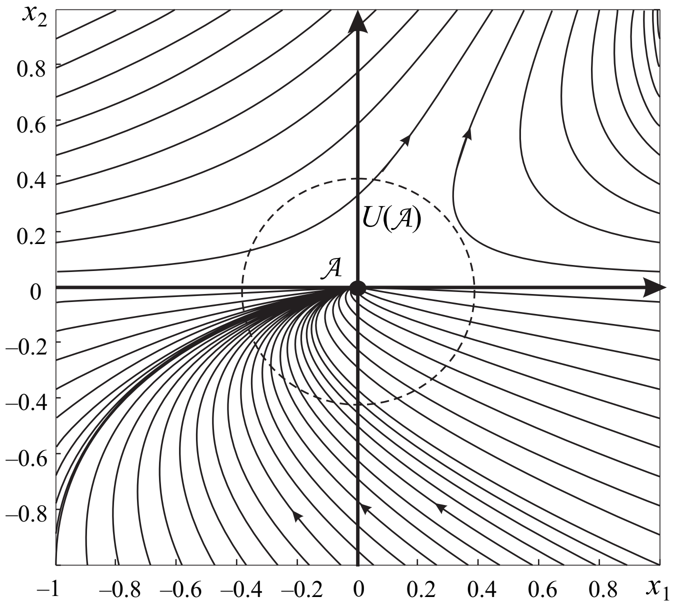


Figure 2.2 The phase portrait of system (2.3).

Example 2.1.1 Suppose that the system dynamics is governed, up to a coordinate transformation, by the following set of ordinary differential equations:

$$\begin{aligned} \dot{x}_1 &= -x_1 + x_2, \\ \dot{x}_2 &= |x_2|. \end{aligned} \tag{2.3}$$

The solution of the second equation in system (2.3) is a non-decreasing function of t for all initial conditions. Furthermore, for all $x_2(0) \leq 0$ we have $\lim_{t\to\infty} x_2(t, x_2(0)) = 0$; and $\lim_{t\to\infty} x_2(t, x_2(0)) = \infty$ for all $x_2(0) > 0$. From this simple analysis we can conclude that solutions of the system will necessarily approach the origin asymptotically for all $x_2(0) \leq 0$, and will move away from the equilibrium for arbitrarily large distances if $x_2(0) > 0$. This is illustrated in Figure 2.2 depicting the phase portrait of system (2.3). This figure demonstrates that for any neighborhood $U(\mathcal{A})$ of the origin $\mathcal{A}$ there are points $\mathbf{x}' \in U(\mathcal{A})$ such that solutions $\mathbf{x}(t, \mathbf{x}')$ escape the neighborhood $U(\mathcal{A})$ and never come back. Hence, according to Definition 2.1.3, $\mathcal{A}$ cannot be called an attracting set. On the other hand, there are points $\mathbf{x}'' \in U(\mathcal{A})$ such that $\lim_{t\to\infty} \mathbf{x}(t, \mathbf{x}'') = 0$. If $U(\mathcal{A})$ is an open circle, then the number of such points is as large as the number of points corresponding to the solutions escaping $U(\mathcal{A})$. Thus the set $\mathcal{A}$ bears an overall signature of attractivity.

Contradiction of the type we discussed in this example was noticed and analyzed by many authors, e.g. in Gorban and Cheresiz (1981).[2] This led to the emergence of the new notion of a *weakly attracting set*, which was formally defined by J. Milnor in his seminal work (Milnor 1985):

Definition 2.1.4 A set $\mathcal{A}$ is a weakly attracting, or Milnor attracting, set iff

(1) it is closed, invariant, and
(2) for some set $\mathcal{V}$ (not necessarily a neighborhood of $\mathcal{A}$) with *strictly positive measure* and for all $\mathbf{x}_0 \in \mathcal{V}$ the following limiting relation holds:

$$\lim_{t\to\infty} \mathbf{x}(t, \mathbf{x}_0) = \mathcal{A} \; \forall \, \mathbf{x}_0 \in \mathcal{V}(\mathcal{A}). \tag{2.4}$$

The key difference of the notion of a weakly attracting set from that provided in Definition 2.1.3 is that the domain of attraction $\mathcal{V}$ is not necessarily a neighborhood of $\mathcal{A}$. Despite the fact that this difference may look small and insignificant at first glance, it becomes very instrumental for successful statement and solution of particular problems of adaptation. In Chapter 5 we will present a large class of problems for which solutions might not even exist if the standard definition of attracting sets were exclusively used in the formal statement of the problem. Although we are not yet ready to provide these examples now, we still would like to point out the existence of these two rather different views on what an attracting set may mean.[3]

So far we have defined the notions of invariance and attractivity of a set with respect to a flow. In the context of adaptation, invariance and attractivity are often desirable asymptotic characterizations of the preferred domain to which the state of an adapting system must be able to move. The question, however, is whether these properties characterize the preferred state with minimal ambiguity. To some degree, thanks to the requirement of invariance in the definitions, this issue is already taken into account. In order to illustrate this point let us suppose that the invariance property in Definitions 2.1.3 and 2.1.4 is replaced with forward-invariance.

Consider system (2.3) from Example 2.1.1. If we replace the invariance requirement with forward-invariance in Definition 2.1.4 then the equilibrium of this system will still be weakly attracting. One can easily see that in this case the equilibrium will not be the only attracting set in the state space. In fact, if we were to replace invariance with mere forward-invariance, the bottom half of every disk centered at the point $(0, 0)$ would be a weakly attracting set too. Indeed, all sets defined in this

[2] See also Gorban (2004) for a more recent and extended review.
[3] We would like to note that Definitions 2.1.3 and 2.1.4 do not exhaust all of the possibilities for defining attracting sets of dynamical systems. There are many other alternatives, such as in Bhatia and Szego (1970). Our choice of particular notions is motivated mostly by the scope of the problems we will consider in this book.

way are forward-invariant according to Definition 2.1.2, and for every such set there exists a set $\mathcal{V}(\mathcal{O})$ (e.g. pick $\mathcal{V}(\mathcal{O}) = \{(x_1, x_2)|x_2 \leq 0\}$) satisfying condition (2.4). Thus the number of weakly attracting sets in system (2.3) would be infinite and not even countable. Hence an object specified in terms of mere forward-invariance and attraction can in principle bear a substantial degree of ambiguity.

In order to disambiguate the asymptotic behavior of dynamical systems even further, the attractivity property of a set is often considered, together with its minimality. Informally the minimality property can be viewed as a requirement that an attracting set $\mathcal{A}$ should not contain any other attracting sets strictly smaller than $\mathcal{A}$. Formally this can be stated as the requirement that for every $\mathbf{x}_0 \in \mathcal{A}$ the trajectory $\mathbf{x}(t, \mathbf{x}_0)$ is dense in $\mathcal{A}$. Attracting sets having this latter property are often referred to as *attractors*. Similarly to attracting sets, there are standard and weak attractors, and we shall be able to see the advantage of both notions in the next chapters.

So far we have provided formal definitions for invariance, attracting sets, and attractors. It is natural now to ask how we can tell whether a set is invariant, attracting, or is an attractor for a given dynamical system. In other words, in addition to the definitions we need to have instrumental criteria for establishing at least the existence of the sets with the aforementioned properties. The role of these criteria in the domain of analysis and synthesis of adaptive systems is that these criteria will provide specific *target constraints* an adapting system should implement in order to be able to fulfill its goals.

In the literature on adaptive control there are many criteria of this kind. Here we consider only those criteria that are necessary in order to understand state-of-the-art statements of the problem of adaptation which we discuss in Chapter 3. These are Barbalat's lemma, stability, persistency of excitation of a vector-function, and one special class of dynamical systems of which the asymptotic behavior can be easily analyzed analytically. Let us start with the simplest of them – Barbalat's lemma.

2.2 Barbalat's lemma

An inherent feature of many adaptive systems is that they operate in conditions under which information about the environment and their own dynamics is lacking. A simple example is that of an organism that may be able to measure its relative position in space with a certain tolerance but is not able to measure its velocity. Yet, it needs to detect conditions under which the velocity is converging to zero asymptotically. More generally, let $h : \mathbb{R} \to \mathbb{R}$ be a function of which the value is physically relevant, but we do not know this function precisely. Suppose that we know some integral characterization of the function, such as the upper and lower bounds of its integral over a family of intervals. What can we say about the asymptotic properties of the function? Is there a limit of $h(t)$ at $t \to \infty$, and, if so, what is its value? The answer to this question is partially provided by Barbalat's lemma.

In order to state the lemma let us recall the property of uniform continuity of a function of a real variable.

Definition 2.2.1 A function $h : \mathbb{R} \to \mathbb{R}$ is called uniformly continuous iff for every $\varepsilon > 0, \varepsilon \in \mathbb{R}$ there exists $\delta > 0, \delta \in \mathbb{R}$ such that for all $t, \tau \in \mathbb{R}$ the following inequality holds:

$$|t - \tau| < \delta \Rightarrow |h(t) - h(\tau)| < \varepsilon. \tag{2.5}$$

The lemma now can be formulated as follows.

Lemma 2.1 *Let $h : \mathbb{R} \to \mathbb{R}$ be a uniformly continuous function and suppose that the following limit exists:*

$$\lim_{t\to\infty} \int_{t_0}^{t} h(\tau)d\tau = a,\ t_0 \in \mathbb{R},\ a \in \mathbb{R}. \tag{2.6}$$

Then

$$\lim_{t\to\infty} h(t) = 0. \tag{2.7}$$

Proof of Lemma 2.1. Suppose that (2.7) does not hold. This implies that there exists a diverging sequence of $t_n, n = 1, \ldots, \infty$ such that

$$|h(t_n)| > \varepsilon,\ \varepsilon \in \mathbb{R},\ \varepsilon > 0.$$

Because the function $h(t)$ is uniformly continuous we have that

$$\forall\, \varepsilon_1 > 0, \varepsilon_1 \in \mathbb{R}\ \exists\, \delta_1 > 0,\ \delta_1 \in \mathbb{R} :\ |t - t_n| < \delta_1 \Rightarrow |h(t) - h(t_n)| < \varepsilon_1.$$

Let $\varepsilon_1 = \varepsilon/2$, then

$$|h(t)| + |h(t_n) - h(t)| \geq |h(t_n)| \Rightarrow |h(t)| \geq |h(t_n)| - |h(t) - h(t_n)| \geq \varepsilon/2 \tag{2.8}$$

$\forall\, t \in [t_n, t_n + \delta_1]$. Consider now

$$\left| \int_{t_0}^{t_n+\delta_1} h(\tau)d\tau - \int_{t_0}^{t_n} h(\tau)d\tau \right| = \left| \int_{t_n}^{t_n+\delta_1} h(\tau)d\tau \right|. \tag{2.9}$$

Given that (2.6) holds, we can conclude that there exists a number n' such that

$$\left| \int_{t_n}^{t_n+\delta_1} h(\tau)d\tau \right| \leq \varepsilon_2,\ \varepsilon_2 > 0,\ \varepsilon_2 \in \mathbb{R}\ \forall\, n \geq n', \tag{2.10}$$

where ε_2 is an arbitrarily small number. On the other hand, on applying the mean-value theorem to the right-hand side of (2.9) and using (2.8) we obtain that the estimate

$$\left|\int_{t_n}^{t_n+\delta_1} h(\tau)d\tau\right| = \left|\delta_1 h(t')\right|,\ t' \in [t_n, t_n+\delta] \Rightarrow$$
$$\left|\int_{t_n}^{t_n+\delta_1} h(\tau)d\tau\right| \geq \delta_1\varepsilon/2 \tag{2.11}$$

must hold for all $n = 1, \ldots, \infty$. Thus, taking (2.11) and (2.10) into account, we can conclude that

$$\varepsilon_2 > \left|\int_{t_n}^{t_n+\delta_1} h(\tau)d\tau\right| \geq \delta_1\varepsilon/2,\ \forall\, n \geq n'.$$

Given that the value of ε_2 can be chosen arbitrarily small and that $\delta_1\varepsilon/2 > 0$, we obtain a contradiction. Hence the assumption that $h(t)$ does not converge to zero is not true. □

An instrumental function of Lemma 2.1 in the domain of synthesis and analysis of adaptive systems is that it constitutes a simple convergence criterion. If we know that the state vector $\mathbf{x}$ of a system satisfies the integral inequality

$$\int_{t_0}^{t} \|\mathbf{x}(\tau,\mathbf{x}_0)\|^2\, d\tau < B,\ B \in \mathbb{R}_{\geq 0},\ \forall\, t \geq t_0,$$

and the derivative of $\mathbf{x}(t,\mathbf{x}_0)$ with respect to t is bounded, we can conclude that $\mathbf{x}(t,\mathbf{x}_0) \to 0$ at $t \to \infty$. In other words, the system's state will have to approach the origin asymptotically. Although simple, this argument is a common component of convergence proofs in the domain of adaptive regulation.

Despite their simplicity and practical utility, the analysis arguments based exclusively on Lemma 2.1 have obvious limitations. This is because the lemma does not characterize the transient properties of the converging functions. For example, we may be interested in knowing how fast a function approaches its limit values, or how large the excursions of the state vector in the system's state space may become before it will settle in close proximity to the origin. The answers to these important questions cannot be derived explicitly from Lemma 2.1. The lemma does not guarantee that the convergence is going to be fast or slow, or that the state does not deviate much from the origin over time. In order to be able to produce these more delicate predictions, additional characterizations of the system's flow rather than simply uniform continuity are needed. One such characterization is the notion of *stability*.

2.3 Basic notions of stability

Let us ask ourselves what we usually mean by referring to some system or process as being stable. Intuitively and in vague everyday language we link stability with the property of a system that a given variable or perhaps a set of variables will not change much in a certain sense in response to perturbations of some kind. In order to state the very same definition formally, one needs to clarify what these variables and perturbations are and what this phrase "will not change much" means. Fortunately, all necessary clarifications usually follow explicitly from the nature of the problem and our own understanding of the goals of the analysis. However, depending on the problem, these specific clarifications vary from one case to another. This gives rise to a rich family of stability definitions. Here we will consider only those few which from the author's point of view are immediately relevant for the analysis of classical mathematical statements of the problem of adaptive regulation provided. These are the notions of *Lyapunov stability*, *Poincaré stability*, and *Poisson stability*. Other basic stability notions, such as *input-to-state* and *input-to-output stability*, which will be instrumental for the further development of the problem of adaptation, are introduced in Chapter 4.

Definition 2.3.1 Let $\mathbf{x}(t,\mathbf{x}_0):\mathbb{R}\times\mathbb{R}^n\rightarrow\mathbb{R}^n$ be a solution of a dynamical system defined for all $t\geq t_0$, $t_0,t\in\mathbb{R}$ and passing through $\mathbf{x}_0\in\mathbb{R}^n$ at $t=t_0$. Solution $\mathbf{x}(t,\mathbf{x}_0)$ is globally stable in the sense of Lyapunov iff for every $\varepsilon>0$, $\varepsilon\in\mathbb{R}$ there exists $\delta>0$, $\delta\in\mathbb{R}$, such that the following holds:[4]

$$\|\mathbf{x}_0-\mathbf{x}_0'\|\leq\delta\Rightarrow\|\mathbf{x}(t,\mathbf{x}_0)-\mathbf{x}(t,\mathbf{x}_0')\|\leq\varepsilon\ \forall\ t\geq t_0. \tag{2.12}$$

Alternatively,

$$\|\mathbf{x}_0-\mathbf{x}_0'\|\leq\delta\Rightarrow\left\|\mathbf{x}(t,\mathbf{x}_0)-\mathbf{x}(t,\mathbf{x}_0')\right\|_{\infty,[t_0,\infty]}\leq\varepsilon. \tag{2.13}$$

If property (2.12) holds only in a neighborhood of $\mathbf{x}(t,\mathbf{x}_0)$ then the stability is local. The property of Lyapunov stability of a solution has a very simple interpretation. Let us view the flow $\mathbf{x}(t,\mathbf{x}_0)$ as a mapping from the space $\mathbb{R}^n$ of initial conditions $\mathbf{x}_0$ into the space of trajectories $\mathbf{x}(t,\mathbf{x}_0)$, and let the space of trajectories be endowed with the standard uniform norm $\|\cdot\|_{\infty,[t_0,\infty]}$. Then stability of a solution in the sense of Lyapunov is analogous to the usual notion of continuity of the mapping $\mathbf{x}:\mathbb{R}^n\rightarrow L^n_\infty[t_0,\infty]$. This is precisely what expression (2.13) in Definition 2.3.1 states. In other words, small variations of initial conditions lead to small variations of $\mathbf{x}(t,\mathbf{x}_0)$ over all $t\geq t_0$. If $\mathbf{x}(t,\mathbf{x}_0)$ is stable in the sense of Lyapunov then we can make sure that the value of an observed trajectory $\mathbf{x}(t,\mathbf{x}_0')$

[4] Here and in other definitions of stability, when this applies, we assume that $\mathbf{x}(t,\mathbf{x}_0')$ is also defined for all $t\geq t_0$.