

Mathematical Systems Theory

Mathematical Systems Theory

fourth edition

G.J. Olsder
J.W. van der Woude
J.G. Moks
D. Jeltsema

Faculty of
Electrical Engineering, Mathematics and Computer Science
Delft University of Technology

VSSD

© VSSD
Fourth edition 2011

Published by VSSD
Leeghwaterstraat 42, 2628 CA Delft, The Netherlands
tel. +31 15 27 82124, telefax +31 15 27 87585, e-mail: hlf@vssd.nl
internet: <http://www.vssd.nl/hlf>
URL about this book: <http://www.vssd.nl/hlf/a003.htm>

All rights reserved. No part of this publication may be reproduced, stored in a retrieval system, or transmitted, in any form or by any means, electronic, mechanical, photocopying, recording, or otherwise, without the prior written permission of the publisher.

| | |
|-----------------------------|------------------------|
| Printed version 4th edition | ebook version |
| ISBN 978-90-6562-280-8 | ISBN 978-90-6562-296-9 |

NUR 919

Key words: mathematical systems theory

Preface

Fourth edition

The major difference between this fourth edition and its predecessor is the presentation of the material in Chapter 3. The method of linearization of a system described in this chapter has now been restricted to solution-input pairs that are constant in time (equilibrium pairs) and the presentation of the analytical method of solving a linear system is restricted to the linear systems that are time-invariant. We firmly believe that both these restrictions are a great improvement from a didactic point of view. Another important change in this chapter concerns the introductory text, which we believe to be an improvement in the sense that the connection with the previous chapter on modeling is now described more explicitly. A change in Chapter 4 which is worth mentioning is given by the passage where we deal with the duality between the concepts of controllability and observability. We have given a qualitative interpretation of this phenomenon of duality which we believe to be a useful addition to the merely symbolic method using the transposition of matrices.

Delft, November 2011

J.W. van der Woude, J.G. Maks and D. Jeltsema

Third edition

Compared to the second edition, the presentation of material in this third edition has been changed significantly. For a start, based on feedback by students, certain topics, like linearization, Routh's criterion, interval stability, observer and compensator design, have been discussed in some more detail than in the second edition. Further, in each chapter theorems, lemmas, examples, and so on, are numbered consecutively now, and exercises have been moved towards the end of chapters. Also additional exercises have been included. Finally, errors and typos, found in the second edition, have been corrected. A.A. Stoorvogel and J.G. Maks are greatly acknowledged for their remarks on the second edition. We also thank VSSD for its willingness to publish these notes as a book. We hope that this third edition will be as successful as the previous ones.

Delft, November 2004

G.J. Olsder and J.W. van der Woude

Second edition

The main changes of this second edition over the first one are (i) the addition of a chapter with MATLAB[®]¹ exercises and possible solutions, and (ii) the chapter on ‘Polynomial representations’ in the first edition has been left out. A summary of that chapter now appears as a section in chapter 8. The material within the chapter on ‘Input/output representations’ has been shifted somewhat such that the parts dealing with frequency methods form one section now. Moreover, some exercises have been added and some mistakes have been corrected. I hope that this revised edition will find its way as its predecessor did.

Delft, December 1997

G.J. Olsder

First edition

These course notes are intended for use at undergraduate level. They are a substantial revision of the course notes used during the academic years 1983-’84 till 1993-’94. The most notable changes are an omission of some abstract system formulations and the addition of new chapters on modelling principles and on polynomial representation of systems. Also changes and additions in the already existing chapters have been made. The main purpose of the revision has been to make the student familiar with some recently developed concepts (such as ‘disturbance rejection’) and to give a more complete overview of the field.

A dilemma for any author of course notes, of which the total contents is limited by the number of teaching hours and the level of the students (and of the author!), is what to include and what not. One extreme choice is to treat a few subjects in depth and not to talk about the other subjects at all. The other extreme is to touch upon all subjects only very briefly. The choice made here is to teach the so-called state space approach in reasonable depth (with theorems and proofs) and to deal with the other approaches more briefly (in general no proofs) and to provide links of these other approaches with the state space approach.

The most essential prerequisites are a working knowledge of matrix manipulations and an elementary knowledge of differential equations. The mathematics student will probably experience these notes as a blend of techniques studied in other (first and second year) courses and as a solid introduction to a new field, viz. that of mathematical system theory, which opens vistas to various fields of application. The text is also of interest to the engineering student, who will, with his background in applications, probably experience these notes as more fundamental. Exercises are interspersed throughout the text; the student should not skip them. Unlike many mathematics texts, these notes contain more exercises (61) than definitions (31) and more examples (56) than theorems (36).

For the preparation of these notes various sources have been consulted. For the first edition such a source was, apart from some of the books mentioned in the bibliography, ‘Inleiding wiskundige systeemtheorie’ by A.J. van der Schaft, Twente University of

¹MATLAB is a registered trademark of The MathWorks, Inc.

Technology. For the preparation of these revised notes, also use was made of ‘Course d’Automatique, Commande Linéaire des Systèmes Dynamiques’ by B. d’Andréa-Novel and M. Cohen de Lara, Ecole Nationale Supérieure des Mines de Paris. The contents of Chapter 2 have been prepared by J.W. van der Woude, which is gratefully acknowledged. The author is also grateful to many of his colleagues with whom he had discussions about the contents and who sometimes proposed changes. The figures have been prepared by Mrs T. Tijanova, who also helped with some aspects of the \LaTeX document preparation system by means of which these notes have been prepared.

Parallel to this course there are computer lab sessions, based on MATLAB, by means of which the student himself can play with various examples such as to get a better feeling for concepts and for designing systems himself. This lab has been prepared by P. Twaalfhoven and J.G. Braker.

Delft, April 1994

G.J. Olsder

Contents

| | | |
|----------|---|-----------|
| 1 | Introduction | 1 |
| 1.1 | What is mathematical systems theory? | 1 |
| 1.2 | A brief history | 4 |
| 1.3 | Brief description of contents | 6 |
| 1.4 | Exercises | 7 |
| 2 | Some Modelling Principles | 8 |
| 2.1 | Conservation laws | 8 |
| 2.2 | Phenomenological principles | 8 |
| 2.3 | Physical principles and laws | 8 |
| 2.3.1 | Thermodynamics | 9 |
| 2.3.2 | Mechanics | 9 |
| 2.3.3 | Electromagnetism | 10 |
| 2.4 | Examples | 12 |
| 2.4.1 | Inverted pendulum | 12 |
| 2.4.2 | Model of a satellite | 13 |
| 2.4.3 | Heated bar | 15 |
| 2.4.4 | Electrical circuit | 15 |
| 2.4.5 | Population dynamics | 17 |
| 2.4.6 | Age dependent population dynamics | 18 |
| 2.4.7 | Bioreactor | 19 |
| 2.4.8 | Transport of pollution | 20 |
| 2.4.9 | National economy | 21 |
| 2.5 | Exercises | 22 |
| 3 | Linear Differential Systems | 25 |
| 3.1 | Input-State-Output Descriptions | 25 |
| 3.2 | Linearization | 26 |
| 3.3 | Solution of a system of linear differential equations | 30 |
| 3.4 | Impulse response and step response | 38 |
| 3.5 | Exercises | 43 |

| | | |
|----------|--|------------|
| 4 | System Properties | 49 |
| 4.1 | Stability | 49 |
| 4.1.1 | Stability in terms of eigenvalues | 49 |
| 4.1.2 | Routh's criterion | 52 |
| 4.1.3 | Lyapunov stability | 54 |
| 4.1.4 | Interval stability | 55 |
| 4.1.5 | Input-output stability | 56 |
| 4.2 | Controllability | 57 |
| 4.3 | Observability | 69 |
| 4.4 | Realization theory and Hankel matrices | 76 |
| 4.5 | Exercises | 77 |
| 5 | State and Output Feedback | 84 |
| 5.1 | Feedback and stabilizability | 84 |
| 5.2 | Observers and state reconstruction | 93 |
| 5.3 | Separation principle and compensators | 97 |
| 5.4 | Disturbance rejection | 102 |
| 5.5 | Exercises | 103 |
| 6 | Input/Output Representations | 109 |
| 6.1 | Laplace transforms and their use for linear time-invariant systems | 109 |
| 6.2 | Connection of systems | 112 |
| 6.3 | Rational functions | 113 |
| 6.4 | Transfer functions and transfer matrices | 116 |
| 6.5 | More on realizations | 121 |
| 6.5.1 | Flow diagrams | 121 |
| 6.5.2 | Alternative realizations | 123 |
| 6.5.3 | Example | 125 |
| 6.6 | Transfer functions and minimal realizations | 127 |
| 6.6.1 | Realizations of single-input single-output systems | 127 |
| 6.6.2 | Realizations of multiple-input multiple-output systems | 130 |
| 6.7 | Frequency methods | 133 |
| 6.7.1 | Oscillations | 133 |
| 6.7.2 | Nyquist and Bode diagrams | 134 |
| 6.8 | Exercises | 138 |
| 7 | Linear Difference Systems | 144 |
| 7.1 | Exercises | 158 |
| 8 | Extensions and Some Related Topics | 163 |
| 8.1 | Abstract system descriptions | 163 |
| 8.1.1 | Behavioral modelling | 167 |
| 8.2 | Polynomial representations | 167 |
| 8.3 | Examples of other kinds of systems | 171 |
| 8.3.1 | Nonlinear systems | 171 |
| 8.3.2 | Descriptor systems | 172 |

| | | |
|----------|---|------------|
| 8.3.3 | Stochastic systems | 172 |
| 8.3.4 | Automata | 173 |
| 8.3.5 | Distributed parameter systems | 174 |
| 8.3.6 | Discrete event systems | 175 |
| 8.4 | Optimal control theory | 177 |
| 8.5 | Parameter estimation | 180 |
| 8.6 | Filter theory | 181 |
| 8.7 | Model reduction | 182 |
| 8.8 | Adaptive and robust control | 183 |
| 8.9 | Exercises | 184 |
| 9 | MATLAB Exercises | 186 |
| 9.1 | Problems | 186 |
| 9.2 | Solutions | 190 |
| | Bibliography | 202 |
| | Index | 203 |

Chapter 1

Introduction

1.1 What is mathematical systems theory?

A system is part of reality which we think to be a separated unit within this reality. The reality outside the system is called the surroundings. The interaction between system and surroundings is realized via quantities, quite often functions of time, which are called input and output. The system is influenced via the input(-functions) and the system itself has an influence on the surroundings by means of the output(-functions).

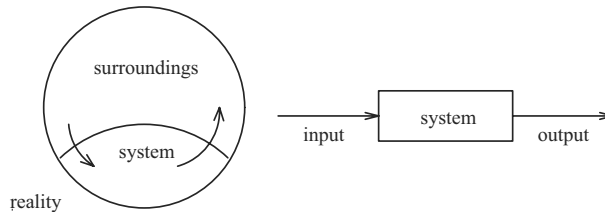


Figure 1.1 A system in interaction with its environment.

Three examples:

- How to fly an aeroplane: the position of the control wheel (the input) has an influence on the course (the output).
- In economics: the interest rate (the input) has an influence on investment behavior (the output).
- Rainfall (the input) has an influence on the level of the water in a river (the output).

In many fields of study, a phenomenon is not studied directly but indirectly through a model of the phenomenon. A model is a representation, often in mathematical terms, of what are felt to be the important features of the object or system under study. By the manipulation of the representation, it is hoped that new knowledge about the modelled phenomenon can be obtained without the danger, cost, or inconvenience of manipulating the real phenomenon itself. In mathematical system theory we only work with models and when talking about a system we mean a modelled version of the system as part of reality.

Most modelling uses mathematics. The important features of many physical phenomena can be described numerically and the relations between these features can be described by equations or inequalities. Particularly in natural sciences and engineering, properties such as mass, acceleration and forces can be described in mathematical terms.

To successfully utilize the modelling approach, however, knowledge is required of both the modelled phenomena and properties of the modelling technique. The development of high-speed computers has greatly increased the use and usefulness of modelling. By representing a system as a mathematical model, converting it into instructions for a computer and running the computer, it is possible to model larger and more complex systems than ever before.

Mathematical system(s) theory is concerned with the study and control of input/output phenomena. There is no difference between the terminologies ‘system theory’ and ‘systems theory’; both are used in the (scientific) literature and will be used interchangeably. The emphasis in system(s) theory is on the dynamic behavior of these phenomena, i.e., how do characteristic features (such as input and output) change in time and what are the relationships between them, also as functions of time. One tries to design control systems such that a desired behavior is achieved. In this sense mathematical system(s) theory (and control theory) distinguishes itself from many other branches of mathematics by the fact that is prescriptive rather than descriptive.

Mathematical system theory forms the mathematical base for technical areas such as automatic control and networks. It is also the starting point for other mathematical subjects such as optimal control theory and filter theory. In optimal control theory one tries to find an input function which yields an output function that satisfies a certain requirement as well as possible. In filter theory the input to the system, then being a so-called filter, consists of observations with measurement errors, while the system itself tries to realize an output which equals the ‘ideal’ observations, i.e., as much as possible without measurement errors. Mathematical system theory also plays a role in economics (specially in macro-economic control theory and time series analysis), theoretical computer science (via automaton theory and Petri-nets) and management science (models of firms and other organizations). Lastly, mathematical system theory forms the hard, mathematical, core of more philosophically oriented areas such as general systems theory and cybernetics.

Example 1.1 [Autopilot of a boat] An autopilot is a device which receives as input the heading $\alpha(t)$ of a boat at time t (measured by an instrument such as a magnetic compass or a gyrocompass) and the (fixed) desired heading α_c (reference point) by the navigator. Using this information, the device automatically yields, as a function of time t , the positioning command $u(t)$ of the rudder in order to achieve the smallest possible heading error $e(t) = \alpha_c - \alpha(t)$. Given the dynamics of the boat and the external perturbations (wind, swell, etc.) the theory of automatic control helps to determine a control input command $u = f(e)$ that meets the imposed technical specifications (stability, accuracy, response time, etc.). For example, this control might be bang-bang:

$$u(t) = \begin{cases} +u_{\max} & \text{if } e(t) > 0, \\ -u_{\max} & \text{if } e(t) < 0. \end{cases}$$

(The arrows in the left-hand picture in Figure 1.2 point in the positive direction of the quantities concerned.) Alternatively, the control might be proportional:

$$u(t) = Ke(t),$$

where K is a constant. It has tacitly been assumed here that for all e -values of interest, $-u_{\max} \leq Ke(t) \leq u_{\max}$. If this is not the case, some kind of saturation must be intro-

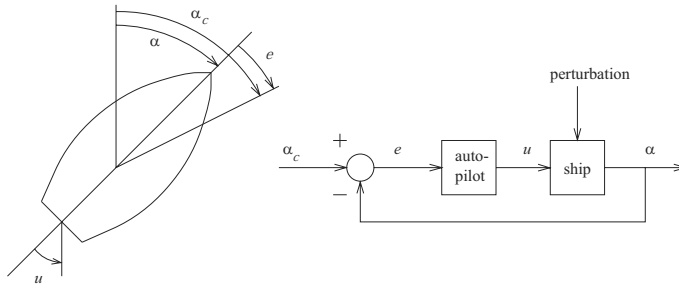


Figure 1.2 Autopilot of a boat.

duced. The control law might also consist of a proportional part, an integrating part and a differentiating part:

$$u(t) = Ke(t) + K' \int^t e(s) ds + K'' \frac{d}{dt} e(t), \quad (1.1)$$

where K , K' and K'' are constants. This control law is sometimes referred to as a PID controller, where P stands for the proportional part, I for the integrating part and D for the differentiating part. The lower bound of the integral in (1.1) has not been given explicitly; various choices are possible. In all these examples of a control law, a signal (the error in this case) is fed back to the input. One speaks of control by feedback.

Automatic control theory helps in the choice of the best control law. If the ship itself is considered as a system, then the input to the ship is the rudder setting u (and possibly perturbations) and the output is the course α . The autopilot is another system. Its input is the error signal e and output is the rudder setting u . Thus, we see that the output of one system can be the input of another system. The combination of ship, autopilot and the connection from α to α_c , all depicted in the right-hand side of Figure 1.2, can also be considered as a system. The inputs of the combined system are the desired course α_c and possible perturbations, and the output is the real course α . \square

Example 1.2 [Optimal control problem] The motion of a ship is described by the differential equation

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

where the so-called state vector $x = (x_1, x_2)^\top \in \mathbb{R}^2$ equals the ship's position with respect to a fixed reference frame, the vector $u = (u_1, u_2)^\top \in \mathbb{R}^2$ denotes the control and t is the time. The superscript \top refers to 'transposed'. If not explicitly stated differently, vectors are always supposed to be column vectors. Although not specifically indicated, both x and u are supposed to be functions of time. The notation \dot{x} refers to the time derivative of the (two) state components. In this example one control variable to be chosen is the ship's heading u_1 , whereas the other one, u_2 , is the ship's velocity. The problem now is to choose u_1 and u_2 in such a way that the ship uses as little fuel as possible such that, if it leaves Rotterdam at a certain time, it arrives in New York not more than 10 days later. The functions u_1 and u_2 may depend on available information such as time, weather forecast,

ocean streams, and so on. Formally, $u = (u_1, u_2)^\top$ must be chosen such that

$$\int_{t_0}^{t_f} g(x, u, t) dt$$

is minimized, where the (integral) criterion describes the amount of fuel used. The function g is the fuel consumption per time unit, t_0 is the departure time and t_f is the arrival time. \square

Example 1.3 [Filtering] NAVSAT is the acronym for NAVigation by means of SATellites. It refers to a worldwide navigation system studied by the European Space Agency (ESA). During the 1980s the NAVSAT system was in the development phase with feasibility studies being performed by several European aerospace research institutes. At the National Aerospace Laboratory (NLR), Amsterdam, the Netherlands, for instance, a simulation tool was developed by which various alternative NAVSAT concepts and scenarios could be evaluated.

Recently, the United States and the European Union have reached an agreement on sharing their satellite navigation services, i.e., the current U.S. Global Positioning System and Europe's Galileo system, which is scheduled to be in operation by 2008. NAVSAT can be seen as a forerunner of Galileo.

The central idea of satellite based navigation system is the following. A user (such as an airplane, a ship or a car) receives messages from satellites, from which he can estimate his own position. Each satellite broadcasts its own coordinates (in some known reference frame) and the time instant at which this message is broadcast. The user measures the time instant at which he receives this message on his own clock. Thus, he knows the time difference between sending and receiving the message, which yields the distance between the position of the satellite and the user. If the user can calculate these distances with respect to at least three different satellites, he can in principle calculate his own position. Complicating factors in these calculations are: (i) different satellites send messages at different time instants, while the user moves in the meantime, (ii) several different sources of error present in the data, e.g. unknown ionospheric and tropospheric delays, the clocks of the satellites and of the user not running exactly synchronously, the satellite position being broadcast with only limited accuracy.

The problem to be solved by the user is how to calculate his position as accurately as possible, when he gets the information from the satellites and if he knows the stochastic characteristics of the errors or uncertainties mentioned above. As the satellites broadcast the information periodically, the user can update also periodically the estimate of his position, which is a function of time. \square

1.2 A brief history

Feedback - the key concept of system theory - is found in many places such as in nature and in living organisms. An example is the control of the body temperature. Also, social and economic processes are controlled by feedback mechanisms. In most technical equipment use is made of control mechanisms.

In ancient times feedback was already applied in for instance the Babylonian water wheels and for the control of water levels in Roman aqueducts. According to historian

Otto Mayr, the first explicit use of a feedback mechanism has been designed by Cornelis Drebbel [1572–1633], both an engineer and an alchemist. He designed the ‘Athanor’, an oven in which he optimistically hoped to change lead into gold. Control of the temperature in this oven was rather complex and the method invented by Drebbel could be viewed as a feedback design.

Drebbel’s invention was then used for commercial purposes by his son in law, Augustus Kuffler [1595–1677], being a contemporary of Christian Huygens [1629–1695]. It was Christian Huygens who designed a fly-wheel for the control of the rotational speed of windmills. This idea was refined by R. Hooke [1635–1703] and J. Watt [1736–1819], the latter being the inventor of the steam engine. In the middle of the 19th century more than 75,000 James Watt’s fly-ball governors (see Exercise 1.4.2) were in use. Soon it was realized that these contraptions gave problems if control was too rigid. Nowadays one realizes that the undesired behavior was a form of instability due to a high gain in the feedback loop. This problem of bad behavior was investigated J.C. Maxwell [1831–1879] – the Maxwell of the electromagnetism – who was the first to perform a mathematical analysis of stability problems. His paper ‘On Governors’ can be viewed as the first mathematical article devoted to control theory.

The next important development started in the period before the Second World War, in the Bell Labs in the USA. The invention of the electronic amplification by means of feedback started the design and use of feedback controllers in communication devices. In the theoretical area, frequency-domain techniques were developed for the analysis of stability and sensitivity. H. Nyquist [1889–1976] and H.W. Bode [1905–1982] are the most important representatives of this direction.

Norbert Wiener [1894–1964] worked on the fire-control of anti-aircraft defence during the Second World War. He also advocated control theory as some kind of artificial intelligence as an independent discipline which he called ‘Cybernetics’ (this word was already used by A.M. Ampere [1775–1836]).

Mathematical system theory and automatic control, as known nowadays, found their feet in the 1950s. Classic control theory played a stimulating role. Initially mathematical system theory was more or less a collection of concepts and techniques from the theory of differential equations, linear algebra, matrix theory, probability theory, statistics, and, to a lesser extent, complex function theory. Later on (around 1960) system theory got its own face, i.e., ‘own’ results were obtained which were especially related to the ‘structure’ of the ‘box’ between input and output, see the right-hand side picture in Figure 1.1. Two developments contributed to that. Firstly, there were fundamental theoretical developments in the nineteen fifties. Names attached to these developments are R. Bellman (dynamic programming), L.S. Pontryagin (optimal control) and R.E. Kalman (state space models and recursive filtering). Secondly, there was the invention of the chip at the end of the nineteen sixties and the subsequent development of micro-electronics. This led to cheap and fast computers by means of which control algorithms with a high degree of complexity could really be used.

1.3 Brief description of contents

In the present chapter a very superficial overview is given of what system theory is and the relations with other (mainly: technically oriented) fields are discussed. One could say that in this chapter the ‘geographical map’ is unfolded and that in the subsequent chapters parts of the map are studied in (more) detail.

In Chapter 2 modelling techniques are discussed and as such the chapter, strictly speaking, does not belong to the area of system theory. Since, however, the starting point in system theory always is a model or a class of models, it is important to know about modelling techniques and the principles underlying such models. Such principles are for instance the conservation of mass and of energy. A classification of the variables involved into input (or: control) variables, output (or: measurement) variables, and variables which describe dependencies within the model itself, will become apparent.

In Chapters 3, 4 and 5 the theory around the important class of linear differential systems is dealt with. The reason for studying such systems in detail is twofold. Firstly, many systems in practice can (at least: approximately) be described by linear differential systems. Secondly, the theory for these systems has been well developed and has matured during the last forty years or so. Many concepts can be explained quite naturally for such systems.

The view on systems is characterized by the ‘state space approach’ and the main mathematical technique used is that of linear algebra. Besides linear algebra one also encounters matrix theory and the theory of differential equations. Chapter 3 deals specifically with linearization and linear differential systems. Chapter 4 deals with structural properties of linear systems. Specially, various forms of stability and relationships between the input, output and state of the system, such as controllability and observability, are dealt with. Chapter 5 considers feedback issues, both state feedback and output feedback, in order to obtain desired system properties. The description of the separation principle is also part of this chapter.

Chapter 6 also deals with linear systems, but now from the input/output point of view. One studies formulas which relate inputs to outputs directly. Main mathematical tools are the theory of the Laplace transform and complex function theory. The advantage of this kind of system view is that systems can easily be viewed as ‘blocks’ and that one can build larger systems by combining subsystems. A possible disadvantage is that this way of describing systems is essentially limited to linear time-invariant systems, whereas the state space approach of the previous chapters is also suited as a means of describing nonlinear and/or time-dependent systems.

In Chapters 3, 4, 5 and 6 ‘time’ was considered to flow continuously. In Chapter 7 one deals with ‘discrete-time’ models. Rather than differential equations one now has difference equations which describe the model from the state space point of view. The most crucial concepts of Chapters 4 and 5 are repeated here for such systems. The role of the Laplace transform is taken over by the so-called z -transform. The theories of continuous-time systems and of discrete-time systems are equivalent in many aspects, and therefore Chapter 7 has been kept rather brief. Some modelling pitfalls when approximating a continuous-time system by a discrete-time one are briefly indicated.

Chapter 8 shows some avenues towards related fields. There is an abstract point of

view on systems, characterizing them in terms of input space, output space, and maybe state space, and the mappings between these spaces. Also the more recently introduced ‘behavioral approach’ towards system theory is briefly mentioned. In this approach no distinction is made between inputs and outputs. It is followed by a brief introduction of polynomial matrices used to represent linear systems algebraically. Some remarks on nonlinear systems – a class many times larger than the class of linear systems – will be made together with some progress in this direction. Also other types of systems are mentioned such as descriptor systems, stochastic systems, finite state systems, distributed parameter systems and discrete event systems. Brief introductions to optimal control theory, filter theory, model reduction, and adaptive and robust control will be given. In those fields system theoretical notions introduced earlier are used heavily.

Lastly, Chapter 9 contains a collection of problems and their solutions that can be used for a course on system theory. The problems are solved using the software package MATLAB. For most of them also the MATLAB *Control Toolbox* must be used. The nature of this chapter is clearly different from that of the others.

Books mentioned in the text and some ‘classics’ in the field of systems theory are given in the bibliography. This book ends with an index.

1.4 Exercises

Exercise 1.4.1 *The water clock (‘clepsydra’) invented by Ktesibios, a Greek of the third century before Christ, is an old and very well known example of **feedback control** (i.e., the error is fed back in order to make corrections). Look this up and give a schematic drawing of the water clock with control.*

Exercise 1.4.2 *Another example of an old control mechanism is Watt’s centrifugal governor for the control of a steam engine. Consult the literature and find out how this governor works. See for instance [Faurre and Depeyrot, 1977].*

Exercise 1.4.3 *Determine how a float in the water reservoir of a toilet operates.*

Exercise 1.4.4 *Investigate the working of a thermostat in the central heating of a greenhouse. Specify the controls and the measurements.*

Exercise 1.4.5 *Describe how feedback plays a role when riding a bicycle. What are the inputs/controls and what are the outputs/measurements.*

Exercise 1.4.6 *Investigate the mechanism of your body to control its temperature. What is the control action?*

Chapter 2

Some Modelling Principles

In this chapter we present some tools that can be used in the modelling of dynamical phenomena. This chapter does not give an exhaustive treatment of such tools, but it is meant as an introduction to some of the underlying principles. One could argue that modelling principles do not belong to the domain of mathematical system theory. Indeed, in the latter theory one usually starts with a given model, perhaps built by an expert in the field of application.

2.1 Conservation laws

One of the most fundamental modelling principles is the notion of conservation. The laws derived from this notion follow from natural reasoning and can be applied everywhere.

For instance, when modelling physical phenomena, one often uses (even without realizing) conservation of matter, conservation of electrical charge, conservation of energy, and so on. But also in disciplines that are not so much physically oriented conservation principles are used. For instance, in describing the evolution of a population, it can be assumed that there is conservation of individuals, simply because no individuals can be created or lost without reason. Similarly in economy, there always has to be conservation of assets in one sense or the other.

Hence, conservation laws can be seen as laws based on reasoning and on counting.

2.2 Phenomenological principles

In addition to the conservation laws discussed above, often also so-called phenomenological laws are used. These laws are obtained in an empirical way and depend very much on the nature of the phenomenon that has to be modelled.

One example of such a law is Ohm's law $V = RI$ relating the voltage V over a resistor of value R with the current I that goes through the resistor. Ohm's law is of importance in modelling electrical networks. However, laws with a similar form occur in other disciplines like Fourier's law on heat conductivity and Fick's law on light diffusion. It is not by reasoning that laws like Ohm's law are derived; they are simply the result of experiments. There is no reason why the voltage, the current and the resistance should be related as they do in Ohm's law. Nevertheless, it turns out to be part of the physical reality and therefore it can be used in the modelling of dynamic phenomena. Many more phenomenological laws exist, some of which are discussed in the next section.

2.3 Physical principles and laws

In this section we briefly discuss some of the most important laws and principles that hold in (parts of) the physical reality.

2.3.1 Thermodynamics

When modelling a thermodynamical phenomenon we can make use of three very fundamental laws and principles.

1. *Conservation of energy.*
2. *The irreversibility of the behavior of a macroscopic system.*
3. *The absolute zero temperature cannot be reached.*

The second law is often also expressed by saying that the entropy of a system cannot decrease. The entropy is a measure for the disorder in a system.

We note that the first law is based on reasoning. If the law were not satisfied, then some form of energy would be missing, and the law could be made to hold by simply introducing the missing type of energy. The second and third law are based on experiments and describe phenomenological properties.

2.3.2 Mechanics

When modelling mechanical phenomena we often, without realizing this, use some very important laws and principles. One of these principles, the conservation of energy, has already been discussed. Other forms of the conservation principle are also often used. Furthermore, the following three laws (postulates) of Newton are very useful.

1. *If there is no force acting on a point mass, then this mass will stay at rest, or it will move with a constant speed along a straight line.*
2. *The force F on a point mass m and its position s are related by $F = m \frac{d^2 s}{dt^2}$.*
3. *action = – reaction.*

The first law was already known to Galileo, as the result of experiments that he had carried out. The second law was formulated by Newton, using the differential calculus he had developed.

Newton's laws, especially the first one, are inspired by experiments. Originally, the laws were developed for point masses and rectilinear movements. Gradually, versions of his laws were developed for continuous media, rotational motions, in fluids, in gasses, and so on. For instance, if a torque N with respect to some axis is applied to a body, and the moment of inertia of the body around the axis is J , then $N = J \frac{d^2 \varphi}{dt^2}$, where $\frac{d^2 \varphi}{dt^2}$ denotes the angular acceleration of the body around the used axis.

After Newton's laws were available, also other approaches to describe the general motion of mechanical structures were developed. One of these approaches, using the concepts of kinetic and potential energy, leads to equations of motion which are known as the Euler-Lagrange equations.

2.3.3 Electromagnetism

When modelling electromagnetic phenomena, versions of laws that are expressed by the four Maxwell equations can be used, complemented by the Lorentz equation.

In a medium with dielectric constant ϵ and magnetic susceptibility μ , the Maxwell equations relating an electric field E , a magnetic field B , a charge density ρ and a current density ι are the following

$$\operatorname{div} E = \frac{1}{\epsilon}\rho, \quad \operatorname{rot} E = -\frac{\partial B}{\partial t}, \quad \operatorname{div} B = 0, \quad \operatorname{rot} B = \mu \left(\iota + \epsilon \frac{\partial E}{\partial t} \right).$$

In these equations all variables depend on time t and, in general, position (x, y, z) . Furthermore, E , B and ι are vectorial quantities, whereas ρ is a scalar. The words ‘div’ and ‘rot’ stand for divergence and rotation, respectively. The first and third equation in the above Maxwell equations express in a sense the conservation of electrical charge and ‘magnetic charge’, respectively. In fact, $\operatorname{div} B = 0$ can be related to the fact that there do not exist magnetic monopoles (isolated charges).

The force F on a particle with charge q moving with velocity v in a medium as described above, with an electric field E and a magnetic field B , is given by the Lorentz equation

$$F = q(E + v \times B).$$

Here \times denotes the cross product. Both F and v are vectors, and q is a scalar. All three quantities will depend on time t and position (x, y, z) .

The above equations are very general in nature and are often too general for our purposes. Therefore, other (more simplified) laws have been obtained from these equations. Some of these laws for electrical networks are discussed below. These networks are built, amongst others, from basic elements like resistors, capacitors and coils. For these elements the following relations have been established.

1. If a current of strength I is led through a resistor with value R , then the voltage drop V over the resistor can be computed by Ohm’s law as illustrated in Figure 2.1.

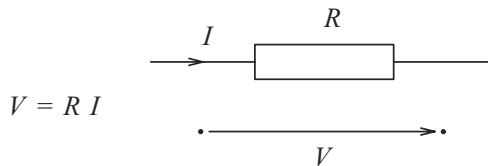


Figure 2.1 Ohm’s law.

2. If a current of strength I flows into a capacitor with capacity C , the voltage drop V over the capacitor is related to I and C in the way shown in Figure 2.2.
3. Finally, if a current of strength I goes through a coil with inductance L , the voltage drop V over the coil can be obtained as depicted in Figure 2.3.

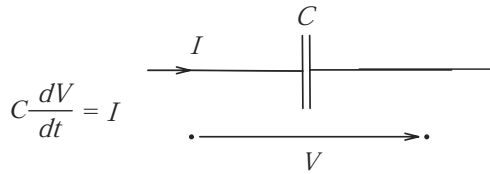


Figure 2.2 Law for capacitor.

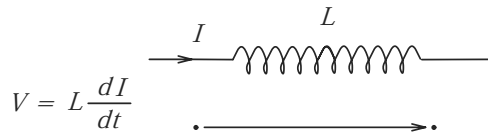


Figure 2.3 Law for inductor.

The variables V and I in Figures 2.1, 2.2 and 2.3 are functions of time. The values R , C and L are assumed to be time independent.

The above laws (rules) are phenomenological in nature. They are the results of experiments. In addition to these laws, two other laws (rules) play an important role in the area of electrical networks. These laws are called the ‘laws of Kirchhoff’, and can be formulated as follows.

4. In any node of the network the sum of all the currents is zero.
5. In any loop of the network the sum of all the voltage drops is zero.

In both laws the direction of currents and voltage drops have to be taken into account. Note that the Kirchhoff laws are of the conservation type. To explain these two laws we consider the abstract network in Figure 2.4, with a source over which the voltage drop is a constant equal to V . An arrow in the figure with an index i stands for an element through which a current I_i flows that induces a voltage drop V_i , both in the direction of the arrow.

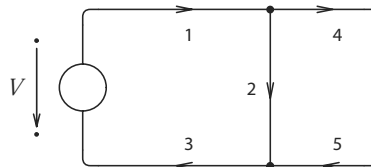


Figure 2.4 Electrical network.

Then in the four nodes (also the source is considered to be a node) the following holds

$$-I_1 + I_2 + I_4 = 0, \quad -I_2 - I_5 + I_3 = 0, \quad -I_4 + I_5 = 0, \quad I_1 - I_3 = 0.$$

For the three loops in the network it follows that

$$-V + V_1 + V_2 + V_3 = 0, \quad -V + V_1 + V_4 + V_5 + V_3 = 0, \quad -V_2 + V_4 + V_5 = 0.$$

2.4 Examples

In this section we give some examples of systems. The models underlying the examples can be derived using the physical principles and laws discussed in the previous.

2.4.1 Inverted pendulum

Consider the inverted pendulum in Figure 2.5. The pivot of the pendulum is mounted on a carriage which can move in the horizontal direction. The carriage is driven by a small motor that at time t exerts a force $u(t)$ on the carriage. This force is the input variable to the system. The mass of the carriage will be indicated by M , that of the pendulum by m .

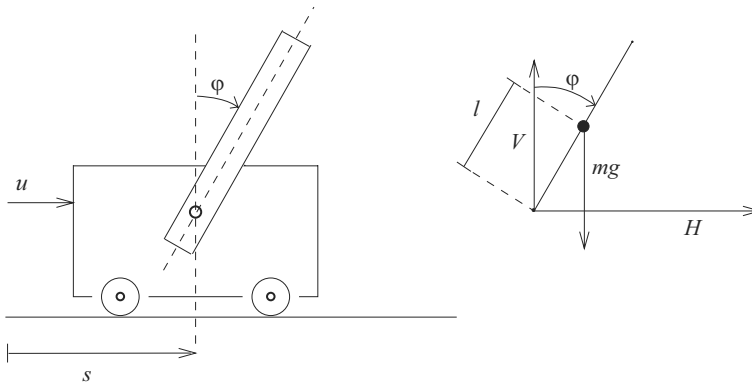


Figure 2.5 Inverted pendulum.

In the pendulum the distance between the pivot and the center of gravity is l . In Figure 2.5 the variable H denotes the horizontal reaction force and V is the vertical reaction force in the pivot. The angle that the pendulum makes with the vertical is indicated by φ . For the center of gravity of the pendulum we have the following equations, which are in the spirit of Newton's second law.

$$m \frac{d^2}{dt^2} (s + l \sin \varphi) = H, \quad m \frac{d^2}{dt^2} (l \cos \varphi) = V - mg, \quad (2.1)$$

$$J \frac{d^2 \varphi}{dt^2} = V l \sin \varphi - H l \cos \varphi. \quad (2.2)$$

The function s denotes the position of the carriage and J is the moment of inertia of the pendulum with respect to the center of gravity. Clearly, φ, s, H and V depend on time t , whereas m, l, g and J are constant. The pendulum has length $2l$ and if it has a uniform mass distribution of $\frac{m}{2l}$ per unit of length, then the moment of inertia around its center of gravity is given by

$$J = \frac{m}{2l} \int_{-l}^l \sigma^2 d\sigma = \frac{1}{3} ml^2.$$

The equation which describes the motion of the carriage is

$$M \frac{d^2 s}{dt^2} = u - H, \quad (2.3)$$

where u may depend on t , while M is constant. Elimination of H and V in the above equations leads to

$$\begin{aligned} \frac{4l}{3}\ddot{\varphi} - g \sin \varphi + \dot{s} \cos \varphi &= 0, \\ (M+m)\dot{s} + ml(\ddot{\varphi} \cos \varphi - \dot{\varphi}^2 \sin \varphi) &= u, \end{aligned} \quad (2.4)$$

where $\dot{}$ denotes the first derivative with respect to time, and $\ddot{}$ the second derivative. So, $\dot{s} = \frac{ds}{dt}$ and $\ddot{\varphi} = \frac{d^2\varphi}{dt^2}$.

The above two equations can also be written as a set of four first order differential equations in $\varphi, \dot{\varphi}, s$ and \dot{s} .

In order to distinguish the above type differential equations from partial differential equations, to be introduced shortly, one refers to the above type of differential equations also as ordinary differential equations.

The equations of motion of the inverted pendulum can also be obtained as the Euler-Lagrange equations using the following expressions for the total kinetic energy T and the potential energy V

$$T = \frac{1}{2}M\dot{s}^2 + \frac{1}{2}\frac{m}{2l} \int_0^{2l} ((\dot{s} + \sigma\dot{\varphi} \cos \varphi)^2 + (\sigma\dot{\varphi} \sin \varphi)^2) d\sigma,$$

$$V = \frac{m}{2l}g \int_0^{2l} \sigma \cos \varphi d\sigma = mgl \cos \varphi,$$

where T , in addition to the kinetic energy of the carriage, consists of the kinetic energy of all the infinitesimal parts $d\sigma$ of the pendulum at a distance σ from the pivot, $0 \leq \sigma \leq 2l$. A similar remark holds with respect to the potential energy.

With the Lagrangian L , defined as $L = T - V$, it follows after evaluation of the integrals that

$$L = \frac{1}{2}M\dot{s}^2 + \frac{1}{2}m\dot{s}^2 + ml\dot{s}\dot{\varphi} \cos \varphi + \frac{2}{3}ml^2\dot{\varphi}^2 - mgl \cos \varphi. \quad (2.5)$$

The Euler-Lagrange equations describing the motion of the inverted pendulum can now be obtained by working out the next equations

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\varphi}} \right) - \frac{\partial L}{\partial \varphi} = 0, \quad \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{s}} \right) - \frac{\partial L}{\partial s} = u.$$

In these equations the variable L is considered to depend on $\varphi, \dot{\varphi}, s$ and \dot{s} , whereas the latter variables depend on t . For instance, with T and V as above, this means that

$$\frac{\partial L}{\partial \dot{\varphi}} = ml\dot{s} \cos \varphi + \frac{4}{3}ml^2\dot{\varphi}, \quad \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\varphi}} \right) = ml\dot{s} \cos \varphi - ml\dot{s}\dot{\varphi} \sin \varphi + \frac{4}{3}ml^2\ddot{\varphi},$$

and similarly for the other (partial) derivatives.

2.4.2 Model of a satellite

Consider the motion of a satellite with mass m_s in a plane through the center of earth. See also the picture in Figure 2.6. As the satellite will orbit around the earth, it is natural

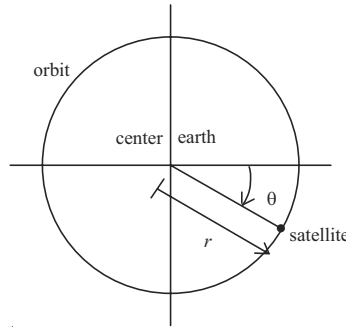


Figure 2.6 Satellite.

to give its position and velocity in terms of the polar coordinates r, θ , and their time derivatives $\dot{r}, \dot{\theta}$, with the earth's center located at the origin. Clearly, r, θ, \dot{r} and $\dot{\theta}$ depend on time t .

The velocity of the satellite has a radial component given by \dot{r} , and a tangential component equal to $r\dot{\theta}$. To apply Newton's laws, also the radial and tangential components of the acceleration of the satellite are required. The radial component of the acceleration is given by $\ddot{r} - r\dot{\theta}^2$, and the tangential component equals $2\dot{r}\dot{\theta} + r\ddot{\theta}$. The previous expressions for the radial and tangential components of the velocity and acceleration are elementary and can be found in any textbook on mechanics.

When in orbit the satellite is attracted by the earth by the gravitational force. This force has a radial direction and its magnitude equals $G\frac{m_e m_s}{r^2}$, where m_e denotes the mass of the earth and G stands for the gravitational constant. Assume that, in addition to gravity, the satellite is also subjected to a radially directed force F_r , and a tangentially directed force F_θ . The force F_r is assumed to be directed away from the earth. Both F_r and F_θ are thought to be caused by thrust jets mounted on the satellite.

Application of Newton's second law in the radial direction and the tangential direction results in

$$m_s (\ddot{r} - r\dot{\theta}^2) = -G\frac{m_e m_s}{r^2} + F_r, \quad m_s (2\dot{r}\dot{\theta} + r\ddot{\theta}) = F_\theta. \quad (2.6)$$

Remark 2.1 The above equations also can be obtained from the Euler-Lagrange equations. For that purpose, note that the kinetic energy T and the potential energy V of the satellite are given as follows

$$T = \frac{1}{2}m_s (\dot{r}^2 + (r\dot{\theta})^2), \quad V = -G\frac{m_e m_s}{r}.$$

Now define the Lagrangian as $L = T - V$, then the equations in (2.6) follow by working out the next equations

$$\frac{d}{dt} \left(\frac{\partial L}{\partial \dot{r}} \right) - \frac{\partial L}{\partial r} = F_r, \quad \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{\theta}} \right) - \frac{\partial L}{\partial \theta} = rF_\theta,$$

where rF_θ must be interpreted as a torque due to the tangential force F_θ . □

2.4.3 Heated bar

Consider a metal bar of length L which is insulated from its environment, except at the left side where the bar is heated by a jet with heat transfer u at time t . For a picture, see Figure 2.7. The temperature of the bar at time t and position r , with $0 \leq r \leq L$, is denoted

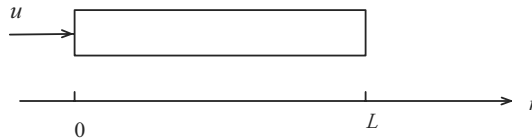


Figure 2.7 Heated bar.

by $T(t, r)$, i.e., r is the spatial variable. In order to be able to determine the thermal behavior of the bar one must know $T(t_0, r)$, $0 \leq r \leq L$, the initial temperature distribution at time $t = t_0$, and the heat transfer $u(t)$ for $t \geq t_0$. The state of the system is the function $T(t, \cdot) : [0, L] \rightarrow \mathbb{R}$. From physics it is known that T satisfies a partial differential equation

$$\frac{\partial T(t, r)}{\partial t} = c \frac{\partial^2 T(t, r)}{\partial r^2}, \quad (2.7)$$

where c is a characteristic constant of the bar. At the left side of the bar we have

$$-A \frac{\partial T(t, r)}{\partial r} \Big|_{r=0} = u(t), \quad (2.8)$$

where A is a measure for the area of the cross section of the bar. At the right side we have

$$\frac{\partial T(t, r)}{\partial r} \Big|_{r=L} = 0, \quad (2.9)$$

because of the insulation there. The evolution of the state is described by the partial differential equation (2.7) with boundary conditions (2.8) and (2.9). In this example the input enters the problem only via the boundary conditions. In other problems the input can also be distributed; see Exercise 2.5.10.

2.4.4 Electrical circuit

Consider the electrical network depicted in Figure 2.8, consisting of a resistor R , a capacitor C and a coil L . The network is connected to a source with constant voltage drop V and the voltage drop over the capacity is measured. The current is denoted by I . If V_R , V_C and V_L denote the voltage drops over the resistor, the capacitor and the coil, respectively, then it follows from the laws of electricity mentioned in the previous section, that

$$V_R = RI, \quad V_C = \frac{1}{C}Q, \quad V_L = L \frac{dI}{dt},$$

where the variable Q denotes the electrical charge on the capacitor, which satisfies $I = \frac{dQ}{dt}$. According to the Kirchhoff laws, $V = V_R + V_C + V_L$. Hence,

$$V = RI + \frac{1}{C}Q + L \frac{dI}{dt}, \quad I = \frac{dQ}{dt}.$$

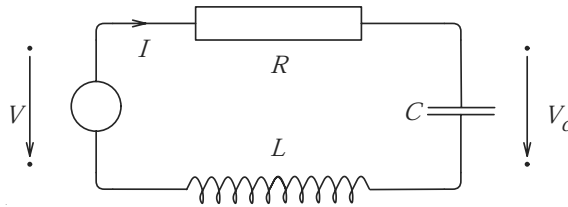


Figure 2.8 RLC network.

Now rearranging these equations, it follows that

$$\frac{d}{dt} \begin{pmatrix} Q \\ I \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -\frac{1}{LC} & -\frac{R}{L} \end{pmatrix} \begin{pmatrix} Q \\ I \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{1}{L} \end{pmatrix} V, \quad V_C = \begin{pmatrix} \frac{1}{C} & 0 \end{pmatrix} \begin{pmatrix} Q \\ I \end{pmatrix}.$$

Define $u = V$, $y = V_C$ and

$$x = \begin{pmatrix} Q \\ I \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 \\ -\frac{1}{LC} & -\frac{R}{L} \end{pmatrix}, \quad B = \begin{pmatrix} 0 \\ \frac{1}{L} \end{pmatrix}, \quad C = \begin{pmatrix} \frac{1}{C} & 0 \end{pmatrix},$$

where it must be emphasized that the newly defined C is a matrix (more specifically, here a row matrix with two elements). It should not be confused with the capacity C . This is an instance of the same symbol being used for different quantities.

With the above way of writing, the following description of the system is obtained

$$\dot{x} = Ax + Bu, \quad y = Cx.$$

Remark 2.2 Elimination of I from the equations above yields the following ordinary linear differential equation with constant coefficients

$$L \frac{d^2 Q}{dt^2} + R \frac{dQ}{dt} + \frac{1}{C} Q = V.$$

This type of equation not only occurs in the modelling of electrical networks. Also in other disciplines this type of equation may arise. For instance, when modelling a mechanical structure as depicted in Figure 2.9. The structure consists of a mass M connected

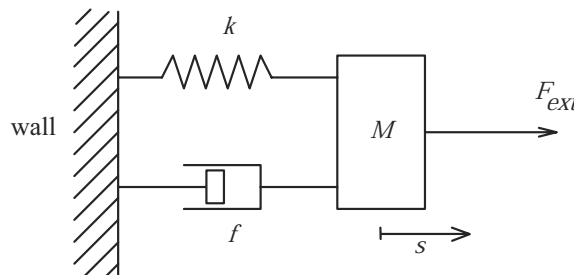


Figure 2.9 Mass-damper-spring system.

to a vertical wall by means of a spring with constant k and a damper with damping factor

f. On the mass an external force F_{ext} may be exerted. As the mass is moving horizontally only, gravity does not play a role. If s denotes the displacement of the mass from its equilibrium position, it follows from Newton's second law that $M\ddot{s} = -ks - f\dot{s} + F_{\text{ext}}$. Hence,

$$M\ddot{s} + f\dot{s} + ks = F_{\text{ext}}.$$

This equation is similar to the one derived for the electrical network above. Other examples of equations of this type can be found in modelling phenomena in disciplines like acoustics, chemistry and hydraulics. \square

2.4.5 Population dynamics

Consider a closed population of humans in a country, or animals or organisms in nature. Let $N(t)$ denote the number of individuals in the population at time t . Assume that $N(t)$ is so large that it can be thought of as being a continuously changing variable. If $B(t, t + \delta)$ and $D(t, t + \delta)$ denote the number of births and deaths, respectively, in the interval $(t, t + \delta]$, then conservation of individuals means that

$$N(t + \delta) - N(t) = B(t, t + \delta) - D(t, t + \delta).$$

Let

$$B(t, t + \delta) = b(t)\delta + o(\delta), \quad D(t, t + \delta) = d(t)\delta + o(\delta),$$

where $o(\delta)$ stands for a function that tends to zero faster than δ . The functions $b(t)$ and $d(t)$ are called the birth rate and death rate, respectively. Moreover, assume that $b(t)$ and $d(t)$ depend on $N(t)$ in a proportional way, independent of time. Hence,

$$b(t) = bN(t), \quad d(t) = dN(t),$$

for some constants b and d . This means that

$$N(t + \delta) - N(t) = (b - d)N(t)\delta + o(\delta).$$

Define $r = b - d$, divide by δ and take the limit for δ to zero. Then it follows that

$$\dot{N}(t) = rN(t).$$

This equation has as solution $N(t) = N(t_0)e^{r(t-t_0)}$. Hence, the number of individuals is increasing (decreasing) when $r > 0$ ($r < 0$).

In general, the growth rate of a population depends on more factors than the above mentioned birth and death rates alone. In particular, it often depends on how the internal interaction is. For instance, if a country is densely populated, then the death rate may increase due to the effects of competition for space and resources, or due to the high susceptibility for deceases. Assuming that the population cannot consist of more than $K > 0$ individuals, the above model might be modified as

$$\dot{N} = r \left(1 - \frac{N}{K} \right) N,$$

where in the equation the dependency of N on t is omitted. The equation is also known as the 'logistic equation'.

The model can further be modified in the following way. Assume that the species of the above population are the prey for a second population of predators consisting of $M(t)$ individuals at time t . It is then reasonable to assume that $r > 0$, and that the previous equation has to be changed into

$$\dot{N} = r \left(1 - \frac{N}{K} \right) N - \alpha NM,$$

with $\alpha > 0$. The modification means that the rate of decrease of prey due to the presence of predators is proportional to the number of predators, but also to the number of prey itself. As a model for the predators the following can be used

$$\dot{M} = -cM + \beta NM,$$

with $c > 0$ and $\beta > 0$. Together these two equations form a so-called ‘predator-prey model’. Note that $r > 0$ means that the population of the prey has a natural tendency to increase, whereas because of $c > 0$ the population of predators has a natural tendency to decrease.

Now assume that the number of prey is unbounded ($K = \infty$). Think of anchovy as prey and of salmon as predator. Assume that due to fishing at time t a fraction $u_1(t)$ of the anchovy is caught, and a fraction $u_2(t)$ of the salmon. The previously derived predator-prey model then has to be changed as follows

$$\begin{aligned} \dot{N} &= rN - \alpha NM - Nu_1 = (r - \alpha M - u_1)N, \\ \dot{M} &= \beta NM - cM - Mu_2 = (\beta N - c - u_2)M. \end{aligned}$$

This type of model is well-known, and is also called a **Volterra-Lotka model**. If the number of salmon is monitored in some way and is denoted $y(t)$, then the above model can be described as a system

$$\dot{x} = f(x, u), \quad y = h(x, u),$$

with $x = (x_1, x_2)^\top = (N, M)^\top$, $u = (u_1, u_2)^\top$ and $y = M$, and functions

$$f(x, u) = \begin{pmatrix} (r - \alpha x_2 - u_1)x_1 \\ (\beta x_1 - c - u_2)x_2 \end{pmatrix}, \quad h(x, u) = x_2.$$

2.4.6 Age dependent population dynamics

Consider again a population and let its size be denoted by N . To express N as a function of the birth rate b , let $P(t, r)$ be the probability that somebody, born at time $t - r$, is still alive at time t (at which he/she has an age of r). Then

$$N(t) = \int_{-\infty}^t P(t, t-s)b(s)ds,$$

where s represents the time of birth. Assume that the functions P and b are such that this integral is well defined. It is reasonable to assume that $P(t, r) = 0$ for $r > L$ for some L

(nobody will become older than L). Then

$$N(t) = \int_{t-L}^t P(t, t-s)b(s)ds.$$

If P is continuous in its arguments and if b is piecewise continuous (a description of piecewise continuity is given later), then the above integral exists.

Returning to the original integral and assuming that a function g exists such that $g(r) = P(t, r)$, it follows that

$$N(t) = \int_{-\infty}^t g(t-s)b(s)ds.$$

If this integral exists for all admissible birth rates b , then it will be shown later that it can be associated with a time-invariant, strictly causal input/output system. (The notions of time-invariance and (strict) causality will be made precise later (in Sections 3.2 and 3.4). Heuristically, time-invariance means that the absolute (calendar) time does not play any role and causality means that the future does not influence the current behavior.) For such a system the probability that somebody is still alive at age r is determined by r only and not by the time of birth.

2.4.7 Bioreactor

Consider a bioreactor as depicted in Figure 2.10. In the reactor there is biomass (or-

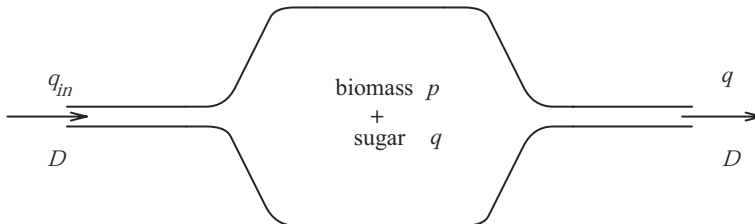


Figure 2.10 Bioreactor.

ganisms) that is nourished with sugar (nutrition). Further, extra nutrition is supplied and products are withdrawn. At time t denote

$p(t)$ for the concentration of biomass in the reactor (g/l),

$q(t)$ for the concentration of sugar in the reactor (g/l),

$q_{in}(t)$ for the concentration of sugar in the flow into the reactor (g/l),

$D(t)$ for the flow of 'sugar water' through the reactor (1/sec), i.e., the fraction of its contents that flows through the reactor per second.

The equations that govern the reaction inside the reactor are given as follows

$$\frac{d}{dt} \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} \text{natural growth} -Dp \\ -\text{natural consumption} -Dq + Dq_{in} \end{pmatrix}.$$

Note that Dp and Dq are products (in mathematical sense). They stand for the amounts the biomass and sugar, respectively, that are withdrawn from the reactor. The product Dq_{in} stands for the amount of sugar that is supplied to the reactor. To complete the mathematical description some empirical laws (or rules of thumb) on the relation between biomass and sugar concentration will be used. Here these laws state that the growth of biomass is proportional to its concentration and that its consumption of sugar is also proportional to its concentration. Furthermore, it is assumed that these proportionalities only depend on the sugar concentration. Hence, there are functions μ and ν , depending on the sugar concentration, that determine the rate of growth of biomass and the consumption rate of sugar, respectively, in the following way

$$\frac{d}{dt} \begin{pmatrix} p \\ q \end{pmatrix} = \begin{pmatrix} \mu(q)p - Dp \\ -\nu(q)p - Dq + Dq_{in} \end{pmatrix}.$$

2.4.8 Transport of pollution

Consider a 'one dimensional' river, contaminated by organic material that is dissolved in the water, see Figure 2.11. Once in the water, the material is degraded by the action of bacteria. Denote

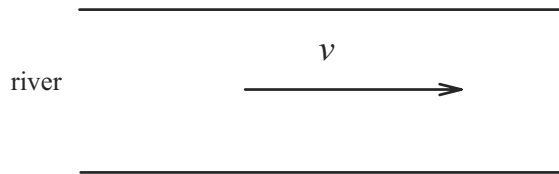


Figure 2.11 Pollution in river.

$\rho(r,t)$ for the density of pollutant in the river at place r and at time t (kg/m),

$v(r,t)$ for the speed of pollutant and water in the river at place r and at time t (m/sec),

$q(r,t)$ for the flux of pollutant in the river at place r and at time t (kg/sec),

$k(r,t)$ for the rate of change by which the density of the pollutant is increased in the river at place r and at time t (kg/(m sec)).

Conservation of mass can be expressed as

$$\frac{\partial \rho}{\partial t} + \frac{\partial q}{\partial r} = k,$$

which has been obtained by considering the infinitesimal equality

$$\rho(r,t+dt)dr = \rho(r,t)dr + q(r,t)dt - q(r+dr,t)dt + k(r,t)dt dr.$$

Now two extreme cases can be considered.

1. There is only advection. Then ρ , q and v are related by $q = \rho v$. This means that the flux of pollutant is only due to transportation phenomena.

2. There is only diffusion. Then ρ and q are related by $q = -\mu \frac{\partial \rho}{\partial r}$, where μ is some constant depending on the place r and the time t . Diffusion means that everything is smoothed.

When both diffusion and advection are taken into account then $q = \rho v - \mu \frac{\partial \rho}{\partial r}$. Assuming that μ is a constant, independent of r and t , and that v does not depend on r , but only on t , the conservation of mass equation can be written as

$$\frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial r} \left(\rho v - \mu \frac{\partial \rho}{\partial r} \right) + k = \mu \frac{\partial^2 \rho}{\partial r^2} - v \frac{\partial \rho}{\partial r} + k.$$

To model the action of bacteria that degrade the pollution, and to model the role of industry, assume that $k = -v\rho + \beta$ with v independent of r and t , and with β a measure for the pollution in the river caused by the industry. Then it follows that

$$\frac{\partial \rho}{\partial t} = \mu \frac{\partial^2 \rho}{\partial r^2} - v \frac{\partial \rho}{\partial r} - v\rho + \beta.$$

Remark 2.3 With μ , v and β constant the last equation can also formally be written as

$$\dot{x} = Ax + \beta,$$

where $x = \rho$ and $A = \mu \frac{\partial^2}{\partial r^2} - v \frac{\partial}{\partial r} - v$ is a linear mapping between appropriate function spaces. □

2.4.9 National economy

Consider the following simplified model of the national economy of a country. Let

$y(k)$ be the total national income in year k ,

$c(k)$ be the consumer expenditure in year k ,

$i(k)$ be the investments in year k ,

$u(k)$ be the government expenditure in year k .

For the model of the national economy the following assumptions are made.

1. $y(k) = c(k) + i(k) + u(k)$,
2. The consumer expenditure is a fixed fraction of the total income of the previous year:
 $c(k) = my(k-1)$ with $0 \leq m < 1$,
3. The investments in year k depend on the increase in consumer expenditure from year $k-1$ to year k : $i(k) = \mu(c(k) - c(k-1))$, where μ is some positive constant.

Note the first assumption is of the conservation type, whereas the other two assumptions may be based on observations.

With the above assumptions the evolution of the national economy can be described as follows:

$$\begin{aligned}i(k+1) - \mu c(k+1) &= -\mu c(k), \\c(k+1) &= my(k) = m(i(k) - \mu c(k)) + m(1 + \mu)c(k) + mu(k).\end{aligned}$$

If a state vector is defined as $x(k) = (x_1(k), x_2(k))^T$, with $x_1(k) = i(k) - \mu c(k)$ and $x_2(k) = c(k)$, then the state evolution equation is given by

$$\begin{pmatrix} x_1(k+1) \\ x_2(k+1) \end{pmatrix} = \begin{pmatrix} 0 & -\mu \\ m & m(1+\mu) \end{pmatrix} \begin{pmatrix} x_1(k) \\ x_2(k) \end{pmatrix} + \begin{pmatrix} 0 \\ m \end{pmatrix} u(k),$$

and the output equation by

$$y(k) = (1 \quad 1 + \mu) \begin{pmatrix} x_1(k) \\ x_2(k) \end{pmatrix} + u(k).$$

Thus, a linear time-invariant discrete-time system has been obtained as a model for the national economy.

2.5 Exercises

Exercise 2.5.1 Consider the inverted pendulum in Section 2.4.1. Assume that the angle φ of the pendulum with the vertical is measured. Let this measurement be denoted by the variable y . So, $y = \varphi$. Note that y as well as all the other variables $\varphi, \dot{\varphi}, s, \dot{s}$ and u are functions of time. Consider the vector $x = (\varphi, \dot{\varphi}, s, \dot{s})^T$, and find functions $f(x, u)$ and $h(x, u)$ such that the inverted pendulum can be described as

$$\dot{x} = f(x, u), \quad y = h(x, u).$$

Here $\dot{x} = \frac{d}{dt}x = (\dot{\varphi}, \ddot{\varphi}, \dot{s}, \ddot{s})^T$.

Exercise 2.5.2 Take the variable L as in (2.5) and derive the equations of motion of the inverted pendulum in Section 2.4.1 by working out the given Euler-Lagrange equations.

Exercise 2.5.3 In the above exercise, the pendulum is assumed to be able to rotate around its end point. Now assume that the pendulum can rotate around a given point somewhere on its longitudinal axis, not necessarily the end point. Derive the equations of motion of this modified inverted pendulum. Start with the (direct) approach of Section 2.4.1 and verify your results with the approach using the Euler-Lagrange equations.

Exercise 2.5.4 In the inverted pendulum example of Section 2.4.1 the input is a force exerted on the carriage. Now assume that the input is a torque exerted on the pendulum around its pivot. Determine how the equations change with respect to those in Section 2.4.1.

Exercise 2.5.5 In Section 2.4.1 the carriage moves horizontally. Now assume that the carriage moves only in the vertical direction and that only vertical forces can be exerted, while the gravity remains to act vertically. Investigate how the equations change with respect to those in Section 2.4.1.

Exercise 2.5.6 Consider the model of the satellite in Section 2.4.2. Assume that the distance r is measured and is denoted y . Further, introduce the vectors $x = (r, \theta, \dot{r}, \dot{\theta})^\top$ and $u = (\frac{F_r}{m_s}, \frac{F_\theta}{m_s})^\top$, and find functions $f(x, u)$ and $h(x, u)$ such that the model of the satellite can be described as

$$\dot{x} = f(x, u), \quad y = h(x, u).$$

Exercise 2.5.7 In Section 2.4.2, starting from the Lagrangian $L = T - V$, work out the given Euler-Lagrange equations to obtain the equations of the motion of the satellite.

Exercise 2.5.8 Consider the electrical network depicted in Figure 2.12. Take V_{in} as in-

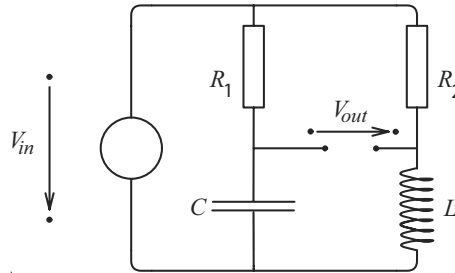


Figure 2.12 Bridge network.

put and V_{out} as output, and derive a state space model for the network using the laws introduced in the Section 2.3.3. Note that V_{out} can be seen as a voltage drop in the ‘loop’ containing just the two resistors. Clearly, there are more such loops containing V_{out} .

Exercise 2.5.9 Consider the electrical network in Figure 2.13. Take the source voltage

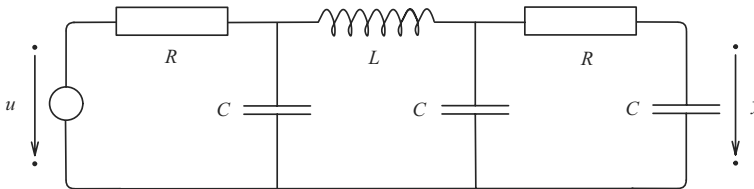


Figure 2.13 Electrical network.

as input u , the voltage over the most right capacitor as output y and derive a state space model for the network using the methods of the Section 2.3.3.

Exercise 2.5.10 In the context of Section 2.4.3, consider the partial differential equation

$$\frac{\partial T(t, r)}{\partial t} = c \frac{\partial^2 T(t, r)}{\partial r^2} + u(t, r),$$

and give an interpretation of $u(t, r)$, seen as a distributed input/control function.

Exercise 2.5.11 For each of the models in Section 2.4.5, find the stationary situations. These are situations in which the variables remain at a constant level and therefore have (time) derivatives that are identically equal to zero.

Exercise 2.5.12 Let p denote the population density, and let it depend on time t and age r . The number of people of ages between r and $r + dr$ at a certain time t is given by $p(t, r)dr$. Define the mortality rate $\mu(t, r)$ in the following way: $\mu(t, r)p(t, r)drdt$ is the number of people in the age class $[r, r + dr]$ who die in the time interval $[t, t + dt]$. Prove the infinitesimal equality

$$p(t + dt, r + dt)dr - p(t, r)dr = -\mu(t, r)p(t, r)drdt,$$

and show that p satisfies the following partial differential equation

$$\frac{\partial p}{\partial r} + \frac{\partial p}{\partial t} = -\mu p. \quad (2.10)$$

Let the initial age distribution be given as

$$p(0, r) = p_0(r), \quad 0 \leq r \leq 1,$$

and the birth rate function as the boundary condition

$$p(t, 0) = u(t), \quad t \geq 0.$$

Here it is assumed that the age r is scaled in such a way that nobody reaches an age $r > 1$. One can consider $u(t)$ as the input to the system and as output $y(t)$ for instance the number of people in the working age, say between the ages a and b , $0 < a < b < 1$. This means that

$$y(t) = \int_a^b p(t, r)dr.$$

Exercise 2.5.13 In Section 2.4.7, assume that the flow D of 'sugar water' into the reactor is fixed, but that the sugar concentration q_{in} in this flow can be controlled. Further, assume that the concentration of sugar of the outgoing flow is measured. Now describe the above process as a system with state, input and output.

Exercise 2.5.14 The same as the above question, but now the sugar concentration q_{in} in the incoming flow is fixed, and the amount of flow D can be controlled.

Exercise 2.5.15 In Section 2.4.9, suppose that the government decides to stop its expenditure from the year $k = 0$ on. Hence, $u(k) = 0$ for all $k \geq 0$. Furthermore, suppose that in the year $k = 0$ the consumers do not spend any money and that the investments are 1 (scaled). So, $c(0) = 0$, $i(0) = 1$. Investigate how the total national income changes for $k \geq 0$.

Exercise 2.5.16 For the same model of the economy as in the above question, find the stationary situations when $u(k) = 1$ for all k , i.e., find those situations that will not change anymore as the years pass by, when $u(k) = 1$ for all k .