BIRKHÄUSER

Systems & Control: Foundations & Applications

Franco Blanchini
Stefano Miani

Set-Theoretic Methods in Control

Birkhäuser
Boston • Basel • Berlin

Franco Blanchini
Dipartimento di Matematica
e Informatica
Università degli Studi di Udine
Via delle Scienze 208
33100 Udine
Italy

Stefano Miani
Dipartimento Ingegneria Elettrica,
Gestionale e Meccanica
Università degli Studi di Udine
Via delle Scienze 208
33100 Udine
Italy

Mathematics Subject Classification: 34D20, 49L20, 93D09, 93D15

Library of Congress Control Number: 2007932665

ISBN-13: 978-0-8176-3255-7 e-ISBN-13: 978-0-8176-4606-6

Printed on acid-free paper.

9 8 7 6 5 4 3 2 1

www.birkhauser.com (KeS/SB)

To Ulla Tahir

Franco Blanchini

To Christina, Giovanna, Pietro, and Lorenza

Stefano Miani

Preface

Many control problems can be naturally formulated, analyzed, and solved in a set-theoretic context. Sets appear naturally when three aspects, which are crucial in control systems design, are considered: constraints, uncertainties, and design specifications. Furthermore, sets are the most appropriate language to specify several system performances, for instance when we are interested in determining the domain of attraction, in measuring the effect of a persistent noise in a feedback loop or in bounding the error of an estimation algorithm.

From a conceptual point of view, the peculiarity of the material presented in this book lies in the fact that sets are not only terms of the formulation, but they play an active role in the solution of the problems as well. Generally speaking, in the control theory context, all the techniques which are theoretically based on some properties of subsets of the state-space could be referred to as set-theoretic methods. The most popular and clear link is that with Lyapunov theory and positive invariance. Lyapunov functions are positive-definite energy-type functions of the state variables, which have the property of being decreasing in time and are fundamental tools to guarantee stability. Besides, their sublevel sets are positively invariant and thus their shape is quite meaningful to characterize the system dynamics, a key point which will be enlightened in the present book. The invariance property will be shown to be fundamental in dealing with problems such as saturating control, noise suppression, model-predictive control, and many others.

The main purpose of this book is to describe the set-theoretic approach for the control and analysis of dynamic systems from both a theoretical and practical standpoint. The material presented in the book is only partially due to the authors' work. Most of it is derived from the existing literature starting from some seminal works of the early 1970s concerning a special kind of dynamic games. By its nature, the book has many intersections with other areas in control theory including constrained control, robust control, disturbance rejection, and robust estimation. None of these is fully covered, but for each of them we will present a particular view only. However, when necessary, the reader will be referred to specialized literature for a complementary reading.

The present work could be seen as a new book on Lyapunov methods, but this would not be an accurate classification. Although Lyapunov's name, as well as the string "set", will appear hundreds of times, our aim is that of providing a different view with respect to the existing excellent work which typically introduces the invariance concept starting from that of Lyapunov functions. Here we basically do the opposite: we show how to synthesize Lyapunov functions starting from sets which are specifically constructed to face relevant problems in control.

Although the considered approach is based on established mathematical and dynamic programming concepts, it is apparent that the approach is far from being considered obsolete. The reason is that these methods, proposed several decades ago, were subsequently abandoned because they were clearly unsuitable for the limited computer technology of the time.

In the authors' mind, it was important to revise those techniques in a renewed light, especially in view of modern computing possibilities. Besides, many connections with other theories which have been developed in recent years (often based on the same old ideas) have been pointed out.

Concerning the audience, the book is mostly oriented towards faculty and advanced graduate students. A good background on control-and-system theory is necessary for the reader to access the book. Although, for the sake of completeness, some of its parts are mathematically involved, the "hard-to-digest" initial mathematical digressions can be left to an intuitive level without compromising the reading and understanding of the sequel. To this aim, an introduction has been written to simplify as much as possible the comprehension of the book. In this chapter, the reasons for dealing with non-differentiable Lyapunov functions are discussed and preliminary examples are proposed to make the (potentially intimidating) notation of the following sections more reader-friendly. In the same spirit, many exercises have been put at the end of each chapter.

The outline of the book, depicted in the figure at the end of the present section, is as follows.

Basic mathematical notations and acronyms, an intuitive description of the main book content, and the link with Lyapunov theory and Nagumo's theorem are provided in Chapter 1.

In Chapter 2, Lyapunov's methods, including nonsmooth functions and converse stability results, are detailed together with their connections to invariant set theory. Some links with differential games and differential inclusion theories are also indicated.

Background material on convex sets and convex analysis, used in the rest of the book, is presented in Chapter 3.

Set invariance theory fundamentals are developed in Chapter 4 along with methods for the determination of appropriate invariant sets, essentially ellipsoids and polytopes, for dynamic systems analysis and design.

Dynamic programming ideas and techniques are presented in Chapter 5, and some algorithms for backward computation of Lyapunov functions are derived.

The ideas presented in Chapters 4 and 5 are at the basis of the following three chapters. Their application to dynamic system analysis is reported in Chapter 6, where it is shown how to compute reachable sets and how these tools prove extremely helpful in the stability and performance analysis of polytopic systems.

The control of parameter varying systems by means of robust or gain-scheduled controllers is looked at in Chapter 7, where it is shown how to derive such controllers starting from quadratic or polytopic functions.

Time constraints are dealt with in Chapter 8. Special emphasis is put on controllability and reachability issues and on the computation of a domain of attraction under bounded or rate constrained inputs. An extension of such techniques to tracking problems is presented.

Chapter 9 presents a set-theoretic solution to different optimal and sub-optimal control problems such as the minimum-time, bounded disturbance rejection, constrained receding horizon, and the recently developed notion of relatively optimal control.

Basic ideas in the set-theoretic estimation area are reported in Chapter 10, where it is more or less shown how it is possible to bound the error estimate via sets, though paying a high price in terms of computational complexity, especially when polytopes are to be considered.

Finally, some topics, which can be solved by set-theoretic methods, are presented in Chapter 11: adaptive control, estimation of the domain of attraction, switched, and planar systems.

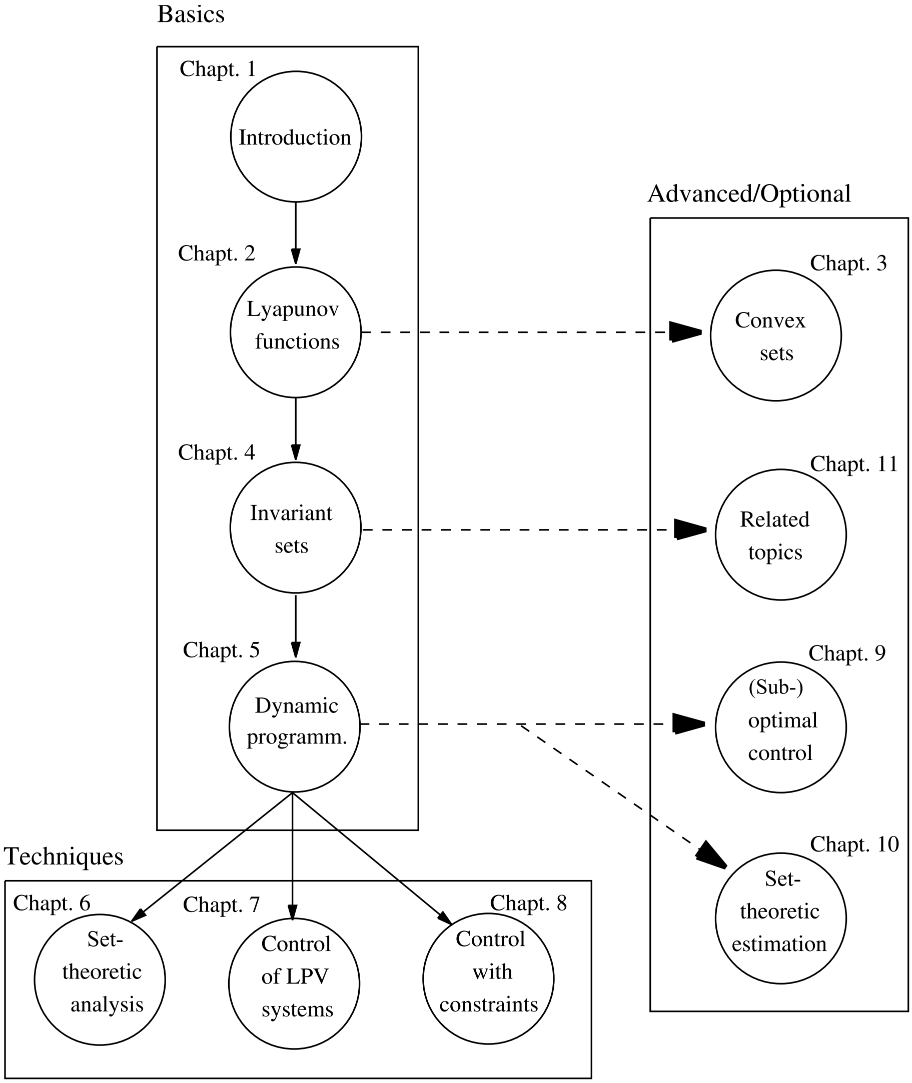
A concluding Chapter 12 illustrates some interesting properties of the Euler auxiliary system, the discrete-time dynamic system which is used throughout the book in many proofs, and the basic functioning of the numerical algorithm used for the backward computation of polytopes for linear parameter varying systems.

There are many people the authors should thank (including the members of their own families) and a full citation would be impossible. Special thanks are due to Dr. Sasa Raković of ETH, Zurich, and to Prof. Fabio Zanolin, University of Udine, for their invaluable help. We also thank Dr. Felice Andrea Pellegrino, from the University of Trieste, Dr. Angelo Alessandri, from C.N.R. Genova, and Mirko Fiacchini from the University of Sevilla, for their comments. We also thank Dr. Carlo Savorgnan, from the University of Udine, who wrote the appendix on the MAXIS–G code, and the anonymous reviewers. We finally thank the editorial staff at Birkhäuser for their excellent work and their careful proofreading, which eliminated many typos. Further errors detected by readers will be posted at

http://users.dimi.uniud.it/~franco.blanchini/SetTheoCon.htm

Udine
August 2007

Franco Blanchini
Stefano Miani



Description of the book's contents

Contents

1

Introduction

1.1 Notation

This book will cover several topics requiring many different mathematical tools. Therefore, adopting a completely coherent notation is impossible. Several letters will have different meanings in different sections of the book. Coherence is preserved inside single sections as long as it is possible. Typically, but not exclusively, Greek letters α, β, ... will denote scalars, Roman letter a, b, ..., vectors, Roman capital letters A, B, matrices, and script letters $\mathcal{A}$, $\mathcal{B}$, ..., sets. A_i will denote both the ith row or the ith column of matrix A. Besides the conventional mathematical conventions, the following notations will be used.

- $\mathbb{R}$ is the set of real numbers.
- $\mathbb{R}_+$ is the set of nonnegative real numbers.
- A^T denotes the transposed of matrix A.
- $eig(A)$ denotes the set of the eigenvalues of the matrix A.
- Given function $\Psi : \mathbb{R}^n \to \mathbb{R}$ and $\alpha \leq \beta$ we denote the sets

  $$\mathcal{N}[\Psi, \alpha, \beta] \doteq \{x : \ \alpha \leq \Psi(x) \leq \beta\}$$

  and

  $$\mathcal{N}[\Psi, \beta] \doteq \mathcal{N}[\Psi(x), -\infty, \beta] = \{x : \ \Psi(x) \leq \beta\}$$

- Given a smooth function $\Psi : \mathbb{R}^n \to \mathbb{R}$, its gradient $\nabla\Psi(x)$ is the column vector

  $$\nabla\Psi(x) = \begin{bmatrix} \frac{\partial\Psi}{\partial x_1}(x) & \frac{\partial\Psi}{\partial x_2}(x) & \dots & \frac{\partial\Psi}{\partial x_n}(x) \end{bmatrix}^T$$

- If $x, z \in \mathbb{R}^n$, we denote the directional upper derivative of $\Psi : \mathbb{R}^n \to \mathbb{R}$

  $$D^+\Psi(x, z) = \lim\sup_{h \to 0+} \frac{\Psi(x + hz) - \Psi(x)}{h}$$

(in the case of a smooth function $\Psi(x)$ it simply reduces to $\nabla\Psi(x)^T z$). We will also (ab)use (of) this notation when a function $z = f(x, w, u)$ has to be considered and we will write

$$D^+\Psi(x, w, u) = D^+\Psi(x, f(x, w, u))$$

for the simple notation, to mean the upper directional derivative with respect to $f(x, w, u)$.

- If A and B are matrices of the same dimensions (or vectors), then

$$A < (\leq, >, \geq) B$$

has to be intended componentwise $A_{ij} < (\leq, >, \geq) B_{ij}$ for all i and j.

- In the space of symmetric matrices,

$$Q \prec (\preceq, \succ, \succeq) P$$

denotes that $P - Q$ is positive definite (positive semi-definite, negative definite, negative semi-definite).

- We will denote by $\|\cdot\|$ a generic norm. We will use this notation in all cases in which specifying the norm is of no importance.
- More specifically, $\|x\|_p$, with integer $1 \leq p < \infty$, denotes the p-norm

$$\|x\|_p = \sqrt[p]{\sum_{i=1}^{n} |x_i|^p}$$

and

$$\|x\|_\infty = \max_i \; |x_i|$$

- If $P \succ 0$ is a symmetric square matrix then

$$\|x\|_P = \sqrt[2]{x^T P x}$$

- Given any vector norm $\|\cdot\|_*$, the corresponding induced matrix norm is

$$\|A\|_* \doteq \sup_{x \neq 0} \; \frac{\|Ax\|_*}{\|x\|_*}$$

- For $x \in \mathbb{R}^n$, the saturation and sign vector functions $\text{sat}(x)$ and $\text{sgn}(x)$ are defined, respectively, by the component-wise assignations

$$[\text{sgn}(x)]_i \doteq \begin{cases} 1 & \text{if } \; x_i > 0 \\ 0 & \text{if } \; x_i = 0 \\ -1 & \text{if } \; x_i < 0 \end{cases}$$

$$[\text{sat}(x)]_i \doteq \begin{cases} x_i & \text{if } \; |x_i| \leq 1 \\ \text{sgn}(x_i) & \text{if } \; |x_i| > 1 \end{cases}$$

The saturation function can be generalized to the weighted case $\text{sat}_a[x]$ or the unsymmetrical case where $\text{sat}_{a,b}[x]$ where a and b are vectors as follows

$$[\text{sat}_{a,b}(x)]_i \doteq \begin{cases} x_i & \text{if } a_i \leq x_i \leq b_i \\ a & \text{if } x_i < a_i \\ b & \text{if } x_i > b_i \end{cases}$$

and $\text{sat}_a[x] \doteq \text{sat}_{-a,a}[x]$.

- With a slight abuse of notation, we will often refer to a function $y(\cdot) : \mathbb{R}^q \mapsto \mathcal{Y} \subset \mathbb{R}^p$ by writing "the function $y(t) \in \mathcal{Y}$".
- A locally Lipschitz function $\Psi : \mathbb{R}^n \to \mathbb{R}$ is positive definite if $\Psi(0) = 0$ and $\Psi(x) > 0$ for all $x \neq 0$. It is positive semi-definite if the strict inequality is replaced by the weak one. A function $\Psi(x)$ is negative (semi-)definite if $-\Psi(x)$ is positive (semi-)definite.
  The above definitions admit local versions in a neighborhood $\mathcal{S}$ of the origin. In this case, the statement "for all $x \neq 0$" is replaced by "for all $x \in \mathcal{S}$".

1.1.1 Acronyms

In this book, very few acronyms will be used, with a few exceptions. We report some of the acronyms below.

EAS	Euler Auxiliary System
LMI(s)	Linear Matrix Inequality (Inequalities)
LPV	Linear Parameter Varying
RAS	Region of Asymptotic Stability
DOA	Domain of Attraction
GUAS	Globally Uniformly Asymptotically Stable
UUB	Uniformly Ultimately Bounded

1.2 Basic ideas and motivations

The goal of this book is to provide a broad overview of important problems in system analysis and control that can be successfully faced via set-theoretic methods.

1.2.1 The spirit of the book

We immediately warn the reader who is mainly interested in plug-and-play solutions to problems or in "user-friendly" recipes for engineering problems that she/he might be partially disappointed by this book. The material presented in most parts of the book is essentially conceptual. By no means does the book lack numerical examples and numerical procedures presented in detail, however it turns out that in some cases, the provided examples evidence the limit of the approach if thought in the "toolbox" manner. However, we hope that if the reader will be patient enough to read the following subsections, she/he will be convinced that the book can provide useful support.

The set-theoretic approach applies naturally in many contexts in which its language is essential even to state the problem. Therefore, the set-theoretic framework is not only a collection of methods, but it is mainly a natural way to formulate, study, and solve problems.

As a simple example consider the "problem" of actuator limitations whose practical meaning is out of questions. The main issue in this regard is indeed a different one: how to formulate the "problem" in a meaningful way. It is indeed known that, as long as a controlled system is close to the desired equilibrium point, actuator limitation is not an issue at all. Clearly, troubles arise when the state is "far" from the target. However, to properly formulate the problem in an engineering spirit, one must decide what "far" means and provide the problem specification. A possible way to proceed is the typical analysis problem in which a control is fixed and its performance is evaluated by determining the domain of attraction under the effect of saturation. If one is interested in a synthesis problem, then a possible approach is that of trying to find a controller which includes in its domain of attraction a certain initial condition or a set of initial conditions. A more ambitious problem is that of determining a controller which maximizes the domain of attraction. From the above simple problem formulation, it is apparent that the set of states which can be brought to the origin, the domain of attraction, is essential in the problem specification. The same considerations can be done if output or state constraints are considered, since quite a natural requirement is meeting the constraints for specified initial conditions. As will be seen, this is equivalent to the requirement that these initial states belong to a proper set in the state space which is a domain of attraction for the closed-loop system.

Constrained control can be actually solved in the disturbance rejection framework by seeking a stabilizing compensator which guarantees constraints satisfaction when a certain disturbance signal (or a class or disturbance signals) is applied with zero initial conditions. Although, in principle, no sets are involved at all in this problem, there are strong relations with the set-theoretic approach. For instance, if one considers the tracking problem of reaching a certain constant reference without constraints violation, the problem can be cast in the set-theoretic language after state translation, assuming the target state as the new origin and by checking if the initial state (formerly the origin) is in the domain of attraction.

In other contexts, such as the rejection of unknown-but-bounded disturbances under constraints, though the connection with the set-theoretic theory is not so obvious, still the theory plays a central role. Indeed, classical results on dynamic programming show how the problem of keeping the state inside prescribed constraint-admissible sets under the effect of persistent unknown-but-bounded disturbances can be formulated and solved exactly (up to computational complexity limits) in the set-theoretic framework.

Beside the fact that the set-theoretic language is the natural one to state several important problems, it also provides the natural tool for solving them as, for instance, in the case of uncertain systems with unknown-but-bounded time varying parameters, for which Lyapunov's theory plays a fundamental role. One key point of Lyapunov's work is that the designer has to choose a class in which a candidate Lyapunov function has to be found. Several classes of functions are available and, without a doubt, the more popular ones are those based on quadratic forms. Very powerful tools are available to handle these functions. However, it is known (as it will clearly be evidenced) that quadratic functions have strong theoretical limitations. Other classes of functions have no such limitations, for instance the polyhedral ones, and several methods to compute them are based on the set-theoretic approach, as we will see later.

In this book, several problems will be considered, without entering into deep details of any of them. Indeed we are describing tools that can be exploited in several different situations (although, for space reasons, some of these will be only sketched).

1.2.2 Solving a problem

It is quite useful to briefly dwell on the sentence "solving a problem", since this is often used with different meanings. As long as we are talking about a problem which is mathematically formulated, a distinction between its "general formulation" and "the instance of the problem", being the latter referred to a special case, namely to a problem with specific data, has to be made. When we say that a problem "is solved" (or can be "solved") we are referring to the general formulation. For instance, the analytic integration problem which consists on finding a primitive of a function is a generically unsolved problem although many special instances ($\int x dt = x^2/2 + C$) are solvable.

The meaning of solving a problem could be discussed for years. Physicists, doctors, mathematicians, and engineers have different feelings about this. Therefore, we decided to insert a "pseudo-definition" of problem solving in order to render clear our approach.

Our pseudo-definition of "solving a problem" is as follows. Given a general problem, mathematically formulated, we say that this is solved if there exists an algorithm that can be implemented on a computer such that, given any instance of the problem, in a finite number of steps (no matter how many), leads to one of the following conclusions:

- the instance can be solved (and, hopefully, a solution is provided);
- there is no solution with the given data.

The discussion here would be almost endless since nothing about the computability has been said, and indeed computability will not be the main issue of the book. Certainly, we will often consider the computational complexity of the proposed algorithms, but we will not assume that an "algorithm" must necessarily possess good "computational complexity", as, for instance, that of being solvable in a time which is a polynomial function of the data dimensions.

We remark that, although we absolutely do not underestimate the importance of the complexity issue, complexity will not be considered of primary importance in this book. Basically, we support this decision by two considerations:

- if we claimed that a problem can be solved if there exists a polynomial algorithm, then we would implicitly admit that the major part of problems is unsolvable;
- complexity analysis is quite useful in all disciplines in which large instances are the normal case (operation research and networks). We rather believe that this is not the case of control area.

Unfortunately, as will be shown later, finding tools which solve a problem in a complete way require algorithms that can be very demanding from a computational viewpoint and therefore complexity aspects cannot be completely disregarded. In particular, the issue of the trade-off between conservativeness and complexity, that will be discussed next, will be a recurring theme of the book.

1.2.3 Conservative or intractable?

Constructive control theory is based on mathematical propositions. Typical conditions have the form "condition **C** implies property **P**" or, in lucky cases, "condition **C** is equivalent to property **P**", where **P** is any property pertaining to a system and **C** is any checkable (at least by means of a computer) mathematical condition. Clearly, when the formulation is of the equivalence type (often referred to as characterization), the control theoretician is more satisfied. For instance, for linear discrete time constant systems, stability is equivalent to the state matrix having eigenvalues with modulus strictly less than 1. This condition is often called a characterization since the family of asymptotically stable systems is the same family of systems whose matrix A has only eigenvalues included in the open unit disk.

There is a much simpler condition, which can be stated in terms of norms and which states that a discrete-time linear system is asymptotically stable if $\|A\| < 1$, where $\|\cdot\|$ is any matrix induced norm, for instance $\|A\|_\infty \doteq \max_i \sum_j |A_{ij}|$. This gain-type condition is generically preferable since, from the computational point of view, it is easier to compute the norm of A rather than its eigenvalues. However, the gain condition is a sufficient condition only since if $\|A\| \geq 1$, nothing can be inferred about the system stability as in the next case

$$A(\mu,\nu) = \begin{bmatrix} 0 & \mu \\ \nu & 0 \end{bmatrix}$$

In the book we will say that a criterion based on a condition **C** is *conservative* to establish property **P** if **C** implies **P**, but it is not equivalent to. In lucky cases it is possible to establish a measure of conservativeness. We say that a criterion based on a condition **C** is *arbitrarily conservative* to establish property **P** if, besides being conservative, there are examples in which condition **C** is "arbitrarily violated", but still they have property **P**. This is the case of the previous example, since $\|A(\mu,\nu)\|_\infty = \max\{\mu,\nu\}$ so $\|A(\mu,\nu)\| < 1$ can be arbitrary violated (for instance, for $\nu = 0$ and arbitrarily large μ) and still the matrix could be asymptotically stable. If we can "measure the violation" then we can also measure the conservativeness.

The counterpart of conservativeness is intractability. Certainly the example provided is not so significant since computing the eigenvalues of a matrix is not a problem as long as the computers will work. However, we can easily become trapped in the complexity issue if we consider a more sophisticated problem, for instance that of establishing the stability of a system of the form $x(k+1) = A(w(k))x(k)$ where $A(w(k))$ takes its values in the discrete set $\{A_1, A_2\}$ (this is a switching system, a family that will be considered in the book). Since $A(w(k))$ is time-varying, the eigenvalues play a marginal role[1]. Conversely, the condition $\|A(w)\| < 1$ remains valid as a conservative sufficient condition for stability. If we are interested in a nonconservative (sufficient and necessary) condition, we can exploit the following result: $x(k+1) = A(w(k))x(k)$ is stable if and only if there exists a full column rank matrix F such that the norm $\|A\|_F \doteq \|FA\|_\infty \leq 1$ [Bar88a] [Bar88b] [Bar88c]. The matrix F can be numerically computed and it will be shown how to manage the computation via set-theoretic algorithms. However, it will also be apparent that the number of rows forming F, which depends on the problem, can be very large. Actually, it turns out that the problem of establishing stability of $x(k+1) = A(w(k))x(k)$ or, equivalently, computing the spectral radius of the pair $\{A_1, A_2\}$, is computationally intractable [TB97].

It is known that computer technology has improved so much[2] that hard problems can be faced in at least reasonable instances. However, there is a further issue. Assume that we are considering a design problem and we are interested in finding an optimal compensator. Assume that we can spend two days and two nights in computing a compensator of order 200 which is "optimal". It is expected that no one (or few people) will actually implement this compensator since in many cases she/he will be satisfied by a simple compensator, for instance a PID. The trade-off between conservativeness and intractability is one of the major concerns in control community. A brilliant idea for coming up with a solution of the dilemma is inspired by the book [BB91]: we can use the hard solutions to evaluate the approximate solutions. It is almost a paradox, but a frequent case, well described by the situation

[1] $|\lambda| < 1$ for $\lambda \in \sigma(A_1) \bigcup \sigma(A_2)$ is a necessary condition only.

[2] Otherwise this book would not have reason to exist.

below in which it happens that reasonably simple solutions are quite close to the optimal ones, motivates our approach.

Quite frequently, situations of this kind arise: an “optimal” compensator of order 200 is computed and it is then established that by means of a PID one can achieve a performance which is 5% worse than the optimal one so that, seemingly, the “optimal control evaluation” has been almost useless. However, we can find a solid argument (and a very good motivation to proceed): we should be happy to use a simple PID-based controller because, thanks to the fact that the optimal solution was found, we are now aware *of the limits of performance*, and that the PID is just 5% sub-optimal.

However, there are cases in which the simple solution is not so close to the “optimal” one and therefore it is reasonable and recommendable to seek a compromise. A typical case is the use of a linear compensator, which is adopted since they normally suffer from the fact that they “react proportionally” to the distance of the state from the target point, which is known to be a source of performance degradation. If a high gain feedback is adopted, things works smoothly when close to the origin but can deteriorate when this is not the case. Conversely, reducing the gain to limit the saturation leads to a weak action. A simple way to overcome the problem is to use small gains when far from the origin and large gains as the origin is approached. This can be achieved in several ways, for instance by switching among compensators, and in this case the switching law must be suitably coordinated, as we will see later on, to ensure the stability of the scheme.

1.2.4 How to avoid reading this book

The book is not structured as a manual or a collection of recipes to be accessed in case of specific problems, but rather the opposite: it is a collection of ideas and concepts. Organizing and ordering these ideas and concepts has certainly been a major effort as will be pointed out soon.

To avoid a waste of time for the reader who is not interested in the details, we have introduced Section 1.3, in which the essentials of the book are presented in form of a summary with examples. Therefore, Section 1.3 could be very useful in deciding whether to continue or to drop further reading[3]. In such a section we have sketched, in a very intuitive way, which is the context of the book, which are the main results and concepts and which is the spirit of the presentation.

We think that accessing Section 1.3 could be sufficient at least in understanding the basics of the message the authors are trying to send, even in the case of postponed (or abandoned) reading.

[3] With the hope that the final decision will be the former.

1.2.5 How to benefit from reading this book

If, eventually, the decision is to read, we would like to give the reader some advice:

- Do not be too intimidated by the mathematics you will find at the beginning. It has been introduced for the sake of completeness. For instance, if you do not like the Dini superior derivative just think in terms of regular derivative of a differentiable function.
- Do not be too concerned with proofs. We have, clearly, inserted them (or referred to available references), but we have not spent too much effort in elegance. We have rather concentrated on enlightening the main ideas.
- If you find the book interesting, please look at the exercises at the end of each chapter while reading. We have tried our best to stimulate ideas.
- Please note that a strong effort has been made to emphasize the main concepts. We could not avoid the details, but do not sacrifice the time to follow them if this compromises the essential.
- Always remember that we are humans and therefore error-prone. We are 100% sure that the book will include errors, questionable sentences, or opinions.

1.2.6 Past work referencing

This has been a crucial aspect, especially in view of the fact that the book includes material which has been known for more than 30 years. As the reader can see, the reference list is full of items, but we assume as an unavoidable fact, that some relevant references will be missing. This is certainly a problem that can have two types of consequences:

- misleading readers who will ignore some work;
- disappointing authors who will not see their work recognized.

The provided references are our good-faith best knowledge of the literature (up to errors or specific decisions of not including some work for which we will accept responsibility).

Clearly, any comment/remark/complaint concerning forgotten or improperly cited references will be very much appreciated.

1.3 Outline of the book

We generically refer to all the techniques which exploit properties of suitably chosen or constructed sets in the state space as set-theoretic methods. The set-theoretic approach appears naturally or can be successfully employed in many problems of a different nature. As a consequence, it was absolutely not obvious how to present the material and how to sequence the chapters (actually this was the major concern in structuring this work). Among the several aspects which are related to the set-theoretic approach, the dominant one is certainly Lyapunov's theory which is considered next. Other fundamental issues are constrained control problems and robust analysis and design.

1.3.1 The link with Lyapunov's theory

Lyapunov 's theory is inspired by the concept of energy and energy-dissipation (or preservation). The main idea of the theory is based on the fact that if an equilibrium point of a dynamical system is the local minimum of an energy function and the system is dissipative, then the equilibrium is (locally) stable. There is a subsequent property that comes into play, and more precisely, the fact that the sublevel sets of a Lyapunov function $\Psi(x)$ (i.e., the sets $\mathcal{N}[\Psi, \kappa] = \{x : \Psi(x) \leq \kappa\}$) are positively invariant for κ small enough. This means that if the initial state is inside this set at time t, then it will be in the set for all $t' \geq t$.[4] This fact turns out to be very useful in many applications which will be examined later on.

The concept of positive invariance is, in principle, not associated with a Lyapunov function. There are examples of invariant sets that do not derive from any Lyapunov function. Therefore the idea of set-invariance can originate a theory which is, in some sense, even more general than the Lyapunov one. For instance, the standard definition of a Lyapunov function requires positive definiteness. As a consequence, the sublevel sets $\{x : \Psi(x) \leq \kappa\}$, for $\kappa > 0$, are bounded sets which include the origin as an interior point. But this is not necessary in many problems in which suitable invariant sets do not need to have or should not have this property.

Consider the case of a positive system, precisely a system such that if the initial state has nonnegative components then the same property is preserved in the future. This property can be alternatively stated by claiming that the state-space positive orthant is positively invariant. It is clear that the claim has no stability implications, since a positive system can be stable or unstable. Still the positive invariance conditions are quite close (at least from a technical standpoint) to the known derivative conditions in Lyapunov theory.

As a preliminary example we borrow a simple example from nonlinear mechanics. See http://users.dimi.uniud.it/∼franco.blanchini/double.jpg for details.

Example 1.1. Consider the following nonlinear system

$$\ddot{\theta}(t) = \alpha \sin(\mu\theta(t)) - \beta \sin(\nu\theta(t))$$

A standard procedure to investigate the behavior of the system is that of multiplying both members by $\dot{\theta}$

$$\dot{\theta}\ddot{\theta} - \alpha \sin(\mu\theta(t))\dot{\theta} + \beta \sin(\nu\theta(t))\dot{\theta} = 0$$

[4] If $t = 0$, then it will belong to the set for positive values of t', hence the name "positive invariance".

and integrating the above so as to achieve

$$\Psi(\theta, \dot{\theta}) \doteq \frac{1}{2}\dot{\theta}^2 + \frac{\alpha}{\mu}\cos(\mu\theta(t)) - \frac{\beta}{\nu}\cos(\nu\theta(t)) = C$$

This means that $\Psi(\theta, \dot{\theta})$ is constant along any system trajectory and thus, a qualitative investigation of such trajectories can be obtained by simply plotting the level curves of the function Ψ in the θ–$\dot{\theta}$ space. For $\alpha = 2$, $\beta = 1$, $\mu = 2$, and $\nu = 1$, the level curves are depicted in Fig. 1.1. From the picture it

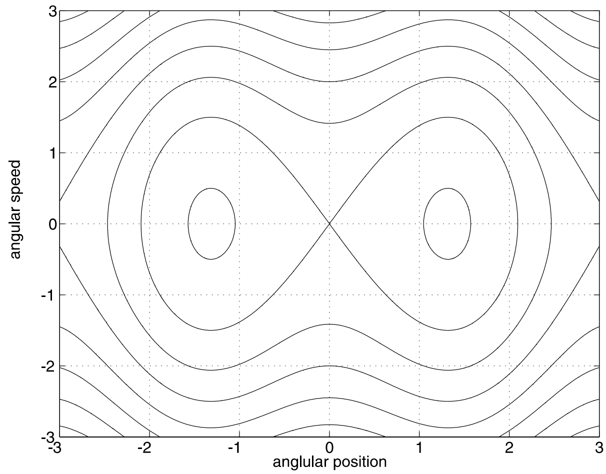


Fig. 1.1. The level surfaces of the function Ψ.

can be inferred that the equilibrium point $\theta = 0$ is unstable (a conclusion that can be derived via elementary analysis), and that there are another two equilibrium points which are stable (not asymptotically), $(\pm\bar{\theta}, 0)$, with $\bar{\theta} \approx 1.3$. However, if the system is initialized close to the origin there are two types of trajectories. For instance, if $\theta(0) = -\epsilon(\epsilon)$, $\epsilon > 0$ and $\dot{\theta}(0) = 0$, then the system trajectories are periodic and encircle the left (right) equilibrium point. Conversely, for any initial condition $\theta(0) = 0$ and $\dot{\theta}(0) = \epsilon$ $(-\epsilon)$, the trajectory encircles both equilibria.

The type of investigation in the example can be clearly extended to cases which are not so lucky and the property that the trajectories evolve along the set $\Psi = C$ in not true anymore. The invariance property of some suitably chosen set can provide useful information about the qualitative behavior. For instance, if a damping is introduced in the nonlinear system

$$\ddot{\theta}(t) = \alpha\sin(\mu\theta(t)) - \beta\sin(\nu\theta(t)) - \gamma\dot{\theta}(t),$$

with $\gamma > 0$, one gets

$$\frac{d}{dt}\Psi(\theta, \dot{\theta}) = -\gamma\dot{\theta}(t)^2$$

so that, due to the energy dissipation, the system will eventually "fall" in one of the stable equilibrium points (with the exception of a zero-measure set of initial conditions from which the state converges to the origin in an unrealistic behavior).

The next natural question concerning the link between set invariance and Lyapunov theory is the following: since the existence of a Lyapunov function implies the existence of positively invariant sets, is the opposite true? More precisely, given an invariant set, is it possible to derive a Lyapunov function from it? The answer is negative, in general. For instance, for positive systems the positive orthant is invariant, but it does not originate any Lyapunov function. However, for certain classes of systems (e.g., those with linear or affine dynamics) it is actually possible to derive a Lyapunov function from a compact invariant set which contains the origin in its interior, as in the following example.

Example 1.2. Consider the linear system $\dot{x} = Ax$ with

$$A = \begin{bmatrix} -1 & \alpha \\ -\beta & -1 \end{bmatrix}$$

where $0 \leq \alpha \leq 1$ and $0 \leq \beta \leq 1$ are uncertain, constant, parameters. To check whether such a system is stable for any of the above values of α and β in the given range, it is sufficient to consider the unit circle (actually any circle) and check whether it is positively invariant. An elementary way to achieve this is to use the Lyapunov function $\Psi(x) = x^T x/2$ and notice that the Lyapunov derivative

$$\dot{\Psi}(x) = x^T\dot{x} = x^T Ax = -x_1^2 - x_2^2 + (\alpha - \beta)x_1x_2 < 0$$

for $(x_1, x_2) \neq 0$. An interpretation of the inequality can be deduced from Fig. 1.2 (left). The time derivative of $\Psi(x(t))$ for x on the circle is equal to the scalar product between the gradient, namely the vector which is orthogonal to the circle surface and points outside, and the velocity vector $\dot{x}$ (represented by a thick arrow in the figure). Intuitively, the fact that such a scalar product is negative, namely that the derivative points inside, implies that any system trajectory originating on the circle surface goes inside the circle (the arrowed curve). This condition will be referred to as sub-tangentiality condition.

Up until now, standard quadratic functions have been considered together with standard derivatives. As an alternative, one might think about, or for some mysterious reasons be interested in, other shapes. If, for example, a unit square is investigated, it is possible to reason in a similar way but with a fundamental difference: *it has corners!* An important theorem, due to Nagumo in 1942 [Nag42], comes into play. Consider the right top vertex of the square which is $[1\ 1]^T$ (the other three vertices can be handled in a similar way). The corresponding derivative is

$$\dot{x} = \begin{bmatrix} (-1+\alpha) \\ (-1-\beta) \end{bmatrix}$$

This means that it "points towards the interior of the square as long as $0 \leq \alpha \leq 1$, $0 \leq \beta \leq 1$". Intuitively, this means that any trajectory passing through the vertex "goes inside". It is also very easy to see that, for any point of any edge, the trajectory points inside the square. Consider, for instance, any point $[x_1\ x_2]^T$ on the right edge, $x_1 = 1$ and $|x_2| \leq 1$. The time derivative of x_1 results in

$$\dot{x}_1 = -x_1 + \alpha x_2 \leq -1 + \alpha|x_2| \leq -1 + \alpha \leq 0.$$

This means that $x_1(t)$ is nonincreasing when the state x is on the right edge, so that no trajectory can cross it from the left to right. By combining the edge and the vertex conditions, one can expect that no trajectory originating in the square will leave it (it will be shown later that for linear uncertain systems one needs only to check "vertex conditions"). It is rather obvious that, by homogeneity, the same consideration can be applied to any scaled square. This fact allows one to consider the norm

$$\Psi(x) = \|x\|_\infty = \max_i |x_i|$$

which is such that $\Psi(x(t))$ is nonincreasing along the trajectories of the system and then results in a Lyapunov function for the system. Given that different shapes (say, different level sets) can be considered, the obvious next question is the following: how can we deal with this kind of functions since $\Psi(x)$ is nondifferentiable and the standard Lyapunov derivative cannot be applied? We will reply to this question in two ways. From a theoretical standpoint we will introduce a powerful tool, the Dini derivative, which is suitable for locally Lipschitz Lyapunov functions (therefore including all kinds of norms). From a practical standpoint, it will be shown that for the class of piecewise linear positive definite functions there exist *linear programming* conditions which are equivalent to the fact that $\Psi(x(t))$ is nonincreasing. These conditions are basically derived by the same type of analysis sketched before and performed on the unit ball of Ψ (this type of functions are called set-induced).

1.3.2 Uncertain systems

Enlightening the importance of Lyapunov's theory for the analysis and control of uncertain systems is definitely not an original contribution of this book. However, the issue of uncertainty is considered of primary importance in the book and will be deeply investigated. Uncertainty will be analyzed not only in the standard way, i.e., by Lyapunov's second method, but also by means of a set-theoretic approach which will allow for a broader view and the facing of

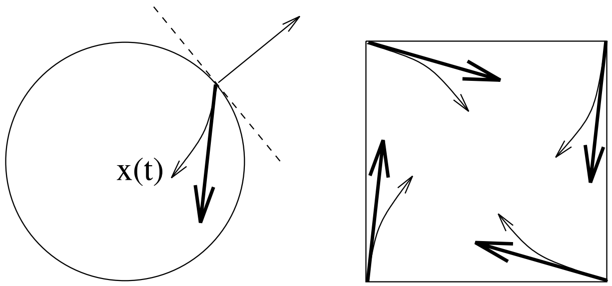


Fig. 1.2. The subtangentiality conditions for the circle and the square.

several problems which are not directly solvable by means of the standard Lyapunov theory. In particular, reachability and controllability problems under uncertainties and their applications will be considered. These will be faced by means of a dynamic programming approach. To provide a very simple example consider the next inventory problem.

Example 1.3. The following equation

$$x(k+1) = x(k) - d(k) + u(k)$$

represents a typical (and, probably, the simplest) inventory model. The variable u is the control representing the production rate while d is the demand rate. The state variable x is the amount of stored good. Consider the problem of finding a control u over an horizon $0, 1, \ldots, T-1$ such that, given $x(0) = x_0$, the following constraints will be satisfied: $0 \leq x(k)$, $x(T) = \bar{x}$, and $0 \leq u(k) \leq \bar{u}$. If $d(k)$ is assumed to be a known function, then the problem is a standard reachability problem in which the following constraints have to be taken into account

$$x(k) = x_0 - \sum_{i=0}^{k-1} d(i) + \sum_{i=0}^{k-1} u(i) \geq 0$$

along with the control constraints $0 \leq u(k) \leq \bar{u}$ and the final condition $x(T) = \bar{x}$. The situation is completely different if one assumes $d(k)$ uncertain and bounded, for instance $d^-(k) \leq d(k) \leq d^+(k)$. This a typical unknown-but-bounded uncertainty specification and the scenario changes completely. Three kind of policies basically can be considered. The first is the open-loop, in which the whole sequence is chosen as a function of the initial state $u(\cdot) = \Phi(x_0)$. The second is the state feedback strategy, precisely $u(k) = \Phi(x(k))$, while the third is the full information strategy, $u(k) = \Phi(x(k), d(k))$, in which the controller is granted the knowledge of $d(k)$, at the current time. These three strategies are strictly equivalent if d is known in advance, in the sense that if the problem is solvable by one of them then it is solvable by the other two, but under uncertainty the situation changes. It is immediate that only the third type of strategy can lead to the terminal goal $x(T) = \bar{x}$. In the hopes of producing something useful by means of the other two strategies we have to relax our request to a more reasonable target like $|x(T) - \bar{x}| \leq \beta$, where β is a tolerance factor.

The open loop problem can then be solved if and only if one can find an open-loop sequence such that $0 \leq u(k) \leq \bar{u}$ and

$$x(k) = x_0 - \sum_{i=0}^{k-1} d^+(i) + \sum_{i=0}^{k-1} u(i) \geq 0, \quad K = 1, 2, \ldots, T,$$

$$x(T) = x_0 - \sum_{i=0}^{T-1} d^+(i) + \sum_{i=0}^{T-1} u(i) \geq \bar{x} - \beta$$

$$x(T) = x_0 - \sum_{i=0}^{T-1} d^-(i) + \sum_{i=0}^{T-1} u(i) \leq \bar{x} + \beta$$

In this case the solution is simple since, in view of the fact that the problem is scalar, one can consider the "worst case" action of the disturbance (which is $d^+(i)$ for the upper bound and $d^-(i)$ for the lower bound). For multidimensional problems, the situation is more involved because there is no clear way to detect the "worst case". Then a possibility, in the linear system case, is to compute the "effect of the disturbance", namely the reachability set at time k. In this simple case we have that such a set is

$$\mathcal{D}_k = \left\{ \sum_{i=0}^{k-1} d(i), \quad \text{for all possible sequences} \quad d(i) \right\}$$

namely, the interval, $\left[\sum_{i=0}^{T-1} d^-(i), \sum_{i=0}^{T-1} d^+(i)\right]$ (as we will see the situation is more involved in the general case). Then, for instance, nonnegative constraint satisfaction reduces to the condition

$$x_0 - \delta + \sum_{i=0}^{k-1} u(i) \geq 0, \quad \text{for all} \quad \delta \in \mathcal{D}_k$$

We will see how this kind of trick is very useful in model predictive control (a technique which embeds an open-loop control computation in a feedback scheme) in the presence of uncertainties.

The feedback problem is more involved and the solution procedure works as follows. Consider the set of all nonnegative states at time $T-1$ which can be driven in one step to the interval $[x_T^-, x_T^+] \doteq [\bar{x} - \delta, \bar{x} + \delta]$. This is the set

$$\mathcal{X}_{T-1} = \{x \geq 0 : \exists u, \ 0 \leq u \leq \bar{u}, \text{ such that } x - d + u \in [x_T^-, x_T^+],$$
$$\forall d^-(T-1) \leq d \leq d^+(T-1)\}$$

It will be seen that such a set is convex. In the scalar case it is an interval $\mathcal{X}_{T-1} = [x_{T-1}^-, x_{T-1}^+]$ (the impatient reader can have fun in determining the extrema). Once $\mathcal{X}_{T-1}$ has been computed, the procedure repeats exactly backward by determining the set of all nonnegative states at time $T-2$ that can be driven in one step to the interval $\mathcal{X}_{T-1} = [x_{T-1}^-, x_{T-1}^+]$ and so on.

It is apparent that the control strategy requires two stages:

- **off-line stage**: the sequence of sets $\mathcal{X}_k$ is sequentially determined backward in time and stored;
- **on-line stage**: the sets of the sequence, $\mathcal{X}_k$, are used at time k to determine the control value $u(k) = \Phi(x(k))$ inside the following set

  $$\Omega_k(x) = \{u : \ 0 \leq u \leq \bar{u}, \ x(k) - d(k) + u(k) \in \mathcal{X}_{k+1}, \ \forall d^-(k) \leq d \leq d^+(k)\}$$

  which is referred to as control map. It is not difficult to realize that feasibility is assured by construction if and only if $x(0) \in \mathcal{X}_0$.

Though the operation of storing the sets $\mathcal{X}_k$ can be computationally demanding, this solution (being of the feedback nature) presents several advantages over the open-loop one. It is very simple to find examples in which the open-loop solution does not exist while the feedback one does. For instance, for $1 \leq d(k) \leq 3$, $\bar{u} = 4$ and the target interval $0 \leq x \leq 4$, the target can be met for arbitrary $T > 0$ and all $0 \leq x_0 \leq 4$ by using the feedback strategy, but no open loop strategy exists for $T > 4$.

The previous closed-loop solution is a typical dynamic programming algorithm [Ber00]. The basic idea of dynamic programming is that of determining the solution backward in time by starting from the target.

There is an interesting connection between dynamic programming and the construction of Lyapunov functions for uncertain systems. Let us consider the case of a simple linear time-varying uncertain system

$$x(k+1) = A(w(k))x(k)$$

where $A(w)$, for $w \in \mathcal{W}$, and $\mathcal{W}$ is a compact set. Being w time-varying, the natural way to check robust stability of the system is to seek for a Lyapunov function. If we resort to quadratic functions, then the problem is basically that of checking if for some positive definite P it is possible to assure

$$x^T A(w)^T P A(w) x < x^T P x, \quad \forall w \in \mathcal{W}$$

However, this is a sufficient condition only. Indeed there are examples of linear uncertain systems for which no quadratic Lyapunov function can be found, but are indeed stable. Then the question is shifted to the following one: is it possible to find a suitable nonquadratic Lyapunov function? The set-theoretic approach provides a constructive way to do this. Again, one possibility is to apply dynamic programming ideas. Given an arbitrary convex set containing