อสมการเมตริกซ์เชิงเส้นในการควบคุมแบบโรบัส

LINEAR MATRIX INEQUALITIES IN ROBUST CONTROL

โดย

นาย การุณ สัญญา

นาย คามิรอน พรหมปลัด

ปริญญานิพนธ์นี้เป็นส่วนหนึ่งของการศึกษาตามหลักสูตรปริญญาวิศวกรรมศาสตรบัณฑิต

สาขาวิศวกรรมระบบควบคุม

สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

ปีการศึกษา 2549

ปริญญานิพนธ์ปีการศึกษา 2549

ภาควิชาวิศวกรรมระบบควบคุม

คณะวิศวกรรมศาสตร์ สถาบันเทคโนโลยีพระจอมเกล้าเจ้าคุณทหารลาดกระบัง

เรื่อง อสมการเมตริกซ์เชิงเส้นในการควบคุมแบบโรบัส

Linear Matrix Inequalities in Robust Control

ผู้จัดทำ นายการุณ สัญญา 46010034

นายคามิรอน พรหมปลัด 46010085

P. Pramayanuntana อาจารย์ที่ปรึกษา

(ผศ. ดร. ปรเมษฐ์ ประนยานันทน์)

อสมการเมตริกซ์เชิงเส้นในการควบคุมแบบโรบัส

โดย

นายการุณ สัญญา

นายคามิรอน พรหมปลัด

อาจารย์ที่ปรึกษา

ผศ. ดร. ปรเมษฐ์ ประนยานันทน์

บทคัดย่อ

โปรเจคนี้เราได้ทำการศึกษาเกี่ยวกับทฤษฎีคณิตศาสตร์ทางด้านเมตริกซ์เชิงเส้นซึ่งสามารถนำไปใช้ในโรบัสท์คอนโทรล โดยได้ศึกษาทฤษฎีเบื้องต้นของเมตริกซ์, โครเนคเกอร์โพรดัคและเสถียรภาพลาปูนอฟในบทที่ 1, ศึกษาเกี่ยวกับอสมการเมตริกซ์เชิงเส้นในบทที่ 2, ศึกษาเกี่ยวกับระบบ Dissipative dynamical ในบทที่ 3,ได้ทำการศึกษาเรื่องระบบที่มีตัวแปรที่ไม่แน่นอน หรือ System with paramatric uncertainty ในบทที่ 4 และความรู้พื้นฐานด้านพีชคณิตเชิงเส้นในภาคผนวก ในโครงงานนี้เราได้ศึกษาคุณสมบัติต่างๆ ของอสมการเมตริกซ์เชิงเส้นในด้านที่สามารถนำไปประยุกต์ใช้ในงานควบคุมซึ่งจะเป็นกลุ่มโปรเจ็คอีกสองกลุ่มที่จะนำพื้นฐานทางด้านอสมการเมตริกซ์เชิงเส้นไปใช้ โดยจะศึกษาการประยุกต์ใช้ต่างๆไม่ว่าจะเป็นเสถียรภาพลาปูนอฟและอื่นๆ ซึ่งสามารถที่จะช่วยให้การวิเคราะห์เป็นไปอย่างสะดวกยิ่งขึ้นและรวดเร็วยิ่งขึ้น

LINEAR MATRIX INEQUALITIES IN ROBUST CONTROL

By

Mr. Karoon Sanya

Mr. Camiron Prompalad

Advisor

Asst. Prof. Dr. Poramate Pranayanuntana

Academic Year 2006

ABSTRACT

A linear matrix inequality (LMI)-based criterion for the global asymptotic stability and uniqueness of the equilibrium point is presented. The criterion turns out to be a generalization and improvement over some previous criteria. The objective of this thesis is to study about LMI apply in control theory. This thesis describes theorem, dissipative system and the application of a technique based on linear matrix inequalities (LMI) use in control system. The problem is reduced to finding a solution of some linear matrix inequalities (LMI). The conditions obtained guarantee that all the control system's solutions starting in some neighborhood of a zero equilibrium vanish as time increases. The neighborhood description is found from the LMI.

LINEAR MATRIX INEQUALITIES IN ROBUST CONTROL

PROJECTREPORT

for the Degree of

BACHELOR OF ENGINEERING

KAROON SANYA ID 46010034

CAMIRON PROMPALAD ID 46010085

March 2007

LINEAR MATRIX INEQUALITIES IN ROBUST CONTROL

PROJECT REPORT

Submitted in Partial Fulfillment
of the REQUIREMENTS for the

Degree of

Bachelor of Engineering

Control Engineering

at the

KING'S MONGKHUT INSTITUTE TECHNOLOGY OF LADKRABANG

by

Karoon Sanya ID 46010034

Camiron Prompalad ID 46010085

March 2007

Copy No. ____________

Approved by the Project Advisor:

P. Pranayanuntana

Dr. Poramate Pranayanuntana

Assistant Professor of Control Engineering

MAR 2, 2007

Date

This work is dedicated to my mother, Wunhareesa Sanya,

my family,

my father Praden Sanya,

my grandmother Kusaera Paduka ,

my older brother Hasungarman.

my sister Huda,

my elder brother Nazeem.

and to our beloved.

ACKNOWLEDGMENT

An enormous debt of gratitude is owed Assistant Professor Poramate Pranayanuntana for his invaluable help, both in the selection of this fantastic project topic and in its writing. Without him this project seems to be untouchable. We spend a lots of time with him at his place and sometimes he even asked us to go to his home to finish up the work for the weekend. Therefore, we have learned many things from him. He taught us basically everything we need to know, from how to use LaTex and editor program that makes our project report more beautiful, how to readd difficult book, how to prove difficult theorems, Linear Algebra, Differential Equation, Linear system theory, Nonlinear system theory and how to do conference paper etc. If it weren't because he were our senior advisor we couldn't learn the thing he taught us from anyone else.

Unforgettable, we would like to thank his family, P' A, Brooklyn and Brave for their very warm welcome to their place and for sharing with us. We wish to thank Dr.Poramate's graduate students, P' Paear and P' Keng, for his suggestion. We also thank our friends, Mai, Jack, Gigg, Nuan and Fuse for helping us.

Of course, We must also thank all teachers who enhance our knowledge and sharpen our skills.

Special thank to our closed friend, Nurzita and Duangporn for the documentation.

Last but not least we would like to thank to each other for the term work bring to success.

AN ABSTRACT

LINEAR MATRIX INEQUALITIES IN ROBUST CONTROL

by

Karoon Sanya ID 46010034

Camiron Prompalad ID 46010085

Advisor: Poramate Pranayanuntana

Submitted in Partial Fulfillment of the Requirements for the Degree of Bachelor of Engineering

March 2007

A linear matrix inequality (LMI)-based criterion for the global asymptotic stability and uniqueness of the equilibrium point is presented. The criterion turns out to be a generalization and improvement over some previous criteria.

The objective of this thesis is to study about LMI apply in control theory. This thesis describes theorem, dissipative system and the application of a technique based on linear matrix inequalities (LMI) use in control system. The problem is reduced to finding a solution of some linear matrix inequalities (LMI). The conditions obtained guarantee that all the control system's solutions starting in some neighborhood of a zero equilibrium vanish as time increases. The neighborhood description is found from the LMI.

Contents

Chapter 1

Introduction

1.1 History of LMI

Contrary to what many authors nowadays seem to suggest, the study of linear matrix inequalities in the context of dynamical systems and control goes back a long way in history and probably starts with the fundamental work of Aleksnadr Mikhailovich Lyapunov on the stability of motion. Lyapunov was a school friend of Markov (yes, the one of the Markov parameters) and later a student of Chebyshev. Around 1980, Lyapunov made a systematic study of the local expansion and contraction properties of motions of dynamical systems around and attractor. He worked out the idea that an invariant set of a differential equation is stable in the sense that it attracts all solution outside the invariant set.

Aleksandr Mikhailovic Lyapunov was born on May 25 , 1857 and published in 1982 his work 'The General Problem of the Stability of Motion' in which he analyzed the question of stability of equilibrium motions of mechanical systems. This work serve as his doctoral dissertation and was defended on September 1982 in Moscow University. Put into modern jargon, he studied stability of differential equations of the form

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

where $A : \mathbb{R}^n \to \mathbb{R}^n$ is some analytic function and x is a vector of position and velocities of material taking values i a finite dimension state space $\mathcal{X} = \mathbb{R}^n$. As Theorem it contains the statement that

> *if the differential equation of the disturbed motion is such that it is possible to find a definite function V of which the derivative V' is a function of fixed sign which is opposite to that of V, or reduces identically to zero, the undisturbed motion is stable.*

The intuitive idea behind this result is that the so called *Lyapunov function* V can be viewed as a generalized 'energy function' (in the context of mechanical system the kinetic and potential energies always served as typical Lyapunov functions). A system is then stable if it is 'dissipative' in the sense that the Lyapunov function decreases. Actually, this intuitive idea turns out to be extremely fruitful in understanding the role of linear matrix inequalities in may problems related to analysis and synthesis of systems. We will consider stability issues in much more further.

1.2 Review of useful matrix theory

1.2.1 Introduction

Matrix theory is a field of basic interest and has applications in scientific computing, differential equation, probability and statistics, control system theory, operational research, optimization, mathematical physics, economics and engineering disciplines. A lot of theorems in matrix theory appear in the form of inequalities. Given any complex valued function defined on matrices, there are inequalities for it.

Application in control system theory, we can reduce a very wide variety of problems arising in control system theory to a few standard convex or quasiconvex optimization problem involving linear matrix inequalities (LMI) which is the form

$$F(x) := F_0 + \Sigma_{i=1}^{m} x_i F_i > 0. \tag{1.2.1}$$

where $x \in \mathbb{R}^m$ is the variable and the symmetric matrices $F_i = F_i^T \in M_n(\mathbb{R}), i = 0, ..., m$ are given. The symbol $F(x) > 0$ means that $F(x)$ is positive definite which is come from Lowner partial order. These LMI are the constraints of the optimization.

1.2.2 Definitions, theorem and literature reviews

The purpose of this subsection is to catalog briefly, without proof, a number of useful concept and facts about the matrix which is the main portion of this thesis

A square matrix A is symmetric if $A^T = A$. A is Hermitian if $A^* = A$. A is orthogonal if $AA^T = I$. A is unitary if $A^*A = I$. A is real orthogonal if A is unitary and real. And A is normal if $A^*A = AA^*$.

Definition 1.2.1 *Let $A, B \in M_{m,n}(\mathbb{F})$. For index sets $\alpha \subseteq \{1, 2, ..., m\}$ and $\beta \subseteq \{1, 2, ..., n\}$. We denoted the submatrix that lies in the rows of A indexed by α and the columns of A indexed by β as $A[\alpha, \beta]$. If $m = n$ and $\alpha = \beta$, the submatrix $A[\alpha, \beta]$ is called a principal submatrix of A and is abbeviated by $A[\alpha]$*

Let $A \in M_n$ scalar $\lambda \in C$ is called eigenvalue of A if there exists a nonzero vector $x \in \mathbb{C}^n$ such that $Ax = \lambda x$ and the vector x is called the corresponding eigenvector of

A associated by with λ. The set of all eigenvalues of A is called the spectrum of A and is denoted by $\sigma(A)$. The spectral radius of A denoted by $\rho(A)$, is the nonnegative real number $\rho(A) := max\{|\lambda| : \lambda \in \sigma(A)\}$.

Definition 1.2.2 *Let $p(t) = a_0 + a_1t + \cdots + a_nt^n$ be a polynomial in $t(a_n \neq 0)$. A polynomial in $A \in M_n$ is defined as $p(A) := a_0I + a_1A + \cdots + a_nA^n$.*

Theorem 1.2.3 *Let $p(\cdot)$ be given polynomial. If λ is an eigenvalue of $A \in M_n$ and $x \in \mathbb{C}^n$ is the corresponding eigenvector of A. Then equation $p(\lambda)$ is an eigenvalue of $p(A)$ and x is the corresponding eigenvector of $p(A)$.*

Definition 1.2.4 *Thought of as a formal polynomial in t, the characteristic polynomial of $A \in M_n$ is defined by $p_A(t) := det(tI - A)$. The equation $det(tI - A) = 0$ is called the characteristic equation of A.*

Theorem 1.2.5 *Let $p_A(t)$ be the characteristic polynomial of $A \in M_n$. Then $p_A(A) = 0$.*

Definition 1.2.6 *The trace of square matrix A is the sum of all its main diagonal's entries, and is denoted by $tr(A)$.*

Proposition 1.2.7 *Let $A, B \in M_n$ and α, β be scalars. Then*

(a) $tr(\alpha A + \beta B) = \alpha tr(A) + \beta(B), i.e., tr(\cdot)$ *is linear.*

(b) $tr(AB) = tr(BA)$.

(c) $tr(A^*A) = \Sigma_{i=1}^n |a_{ij}|^2$.

1.2.3 Function of Matrix

Theorem 1.2.8 *Given a complex-valued function f of complex variable t. Let $A \in M_n, \lambda_1, ..., \lambda_n \in \Sigma(A)$ and $p(t)$ be the characteristic polynomial of A.*

$$p(t) = \prod_{i=1}^{m}(t-\lambda_i)^{n_i}, n = \sum_{i=1}^{m} n_i.$$

define $h(t) := \beta_0 + \beta_1 t + \cdots + \beta_{n-1}t^{n-1}$. $\beta_0, \beta_1, ..., \beta_{n-1}$ *are to be solved from the following set of equations:*

$$f^{(l)}(\lambda_i) = h^{(l)}(\lambda_i).$$

for all $l = 0, ..., n-1$, *for all* $i = 0, ..., m$. *Then* $f(A) = h(A)$ *and* $h(t)$ *is said to equal* $f(t)$ *on the spectrum of* A.

Theorem 1.2.9 *If* $A \in M_n$ *is a block diagonal matrix,*

$$A = diag[A_1, A_2, ..., A_k].$$

and the function f *is defined on the spectrum of* A, *then*

$$f(A) = diag[f(A_1, f(A_2), ..., f(A_k)].$$

Theorem 1.2.10 *If* $A, B, P \in M_n$ *where* $B = PAP^{-1}$ *and the function* f *is defined on the spectrum of* A, *then*

$$f(B) = Pf(A)P^{-1}$$

Definition 1.2.11 *The exponential of* $A \in M_n$ *is depend on the power series:*

$$e^A = \sum_{n=0}^{x} \frac{1}{n!}A^n.$$

1.2.4 Positive definite matrix

Definition 1.2.12 *A matrix $A \in M_n$ is said to be positive definite if for all nonzero $x =\in \mathbb{C}^n$*

$$x^T A x > 0$$

If the strict inequality required in is weaken to $x^T Ax \geq 0$, then A is said to be positive semidefinite.

Proposition 1.2.13 *Let $A, B \in M_n$ be positive definite matrix. Then*

(a) *A is Hermitian.*

(b) *A is positive semidefinite.*

(c) *$A^T, \bar{A}$ and A^* are positive definite.*

(d) *The positive linear combination of A and B are also positive definite.*

(e) *AB is positive definite, A^k is also positive definite for $k = 0, 1, 2, ...$*

(f) *Any principal submatrix of A is positive definite.*

(g) *All eigenvalues of A are positive real numbers.*

(h) *The determinant, the trace, all principal minors and the main's diagonal entries of A are positive real numbers.*

(i) *A is nonsingular, A^{-1} is also positive definite.*

(j) *A is full rank.*

Theorem 1.2.14 *Let $A \in M_n$. The following statements are equivalent*

(a) *A is positive definite.*

(b) $x^*Ax > 0$ *for all nonzero* $x \in \mathbb{C}^n$.

(c) $H(A) := \frac{1}{2}(A^* + A)$ *is positive definite.*

(d) *A is positive semidefinite and nonsingular.*

(e) *A is positive semidefinite and full rank.*

(f) *A is symmetric and* $x^T Ax > 0$ *for all nonzero* $x \in \mathbb{R}^n$ *(in this case, A is real)*

(g) *A is symmetric and have all eigenvalues positive (in this case, A is real)*

(h) *A is Hermitian and* $\exists$ *a nonsingular matrix* $C \in M_n$ *such that* C^*AC *is positive definite.*

(i) $\exists$ *a nonsingular matrix* $B \in M_n$ *such that* $A = B^*B$

(j) $\exists$ *a nonsingular upper triangular matrix* $L \in M_n$ *with positive diagonal entries such that* $A = L^*L$. *If A is real, L may be taken to be real.*

Theorem 1.2.15 *Let* $A \in M_n$ *be positive semidefinite and let* $k \geq 1$ *be given integer. Then there exists a unique positive semedefinite matrix X such that* $X^k = A$. *We also have*

(a) $AX = XA$ *and there is a polynomial* $p(t)$ *such that* $X = p(A)$.

(b) *rank* $X =$ *rank* A, *so X is positive definite if A is real.*

(c) *X is real if A is real.*

The most useful case of the preceding Theorem is for $k = 2$. The unique positive (semi)definite square root of the positive (semi)definite matrix A is usually denoted by $A^{\frac{1}{2}}$. Similarly, $A^{\frac{1}{k}}$ denote the unique positive (semi)definite $k-th$ root of A for each case $k = 1, 2, ..., n$.

1.3 Kronecker product

A notion that is useful in the study of matrix equations and others applications, and is of interest in it own right, is the Kronecker product, direct product or tensor product of matrices.

Definition 1.3.1 *The Kronecker product of* $A = [a_{ij}] \in M_{m,n}(\mathbb{F})$ *and* $B = [b_{ij} \in M_{p,q}(\mathbb{F})$ *is denoted by* $A \otimes B$ *is defined to be a block matrix.*

$$A \otimes B := \begin{bmatrix} a_1 1B & \cdots & a_1 B \\ \vdots & \ddots & \vdots \\ b_{m1}B & \cdots & a_{mn}B \end{bmatrix}$$

Definition 1.3.2 *Let* $A = [a_{ij} \in M_{m,n}(\mathbb{F}$ *The* $k-th$ *Kronecker power and* $A^{\otimes}$ *is defined inductively for all positive integer* k *by* $A^{\otimes} := A$ *and for* $k = 1, 2, \ldots$ $B = [b_{ij} \in M_{p,q}(\mathbb{F})$

$$A^{\otimes K} := A \otimes A \otimes A^{\otimes k-1}$$

Proposition 1.3.3 *In general,* $A \otimes B \neq B \otimes A$*, that is, the Kronecker product is not commutative. Some very basic properties of the Kronecker product include*

(a) $(\alpha A) \otimes B = \alpha(A \otimes B) = A \otimes (\alpha B)$, *for all* $\alpha \in (\mathbb{F}), A \in M_{m,n}(\mathbb{F})$ *and* $B \in M_{p,q}(\mathbb{F})$.

(b) $(A \otimes B)^T = A^T \otimes B^T$, *for all* $A \in M_{m,n}(\mathbb{F})$ *and* $B \in M_{p,q}(\mathbb{F})$.

(c) $(A \otimes B)^* = A^* \otimes B^*$, *for all* $A \in M_{m,n}(\mathbb{F})$ *and* $B \in M_{p,q}(\mathbb{F})$.

(d) $(A \otimes B) \otimes C = A \otimes (B \otimes C)$, *for all* $A \in M_{m,n}(\mathbb{F}), B \in M_{p,q}(\mathbb{F})$ *and* $C \in M_{r,s}(\mathbb{F})$.

(e) $(A + B) \otimes C = (A \otimes C) + (B \otimes C)$, *for all* $A, B \in M_{m,n}(\mathbb{F})$ *and* $C \in M_{p,q}(\mathbb{F})$.

(f) $A \otimes (B + C) = (A \otimes B) + (A \otimes C)$, *for all* $A \in M_{m,n}(\mathbb{F})$ *and* $B, C \in M_{p,q}(\mathbb{F})$.

Lemma 1.3.4 *Let* $A = [a_{ij} \in M_{m,n}(\mathbb{F}), B \in M_{p,q}(\mathbb{F}), C \in M_{m,k}(\mathbb{F})$ *and* $D \in M_{q,r}(\mathbb{F})$. *Then* $(A \otimes B)(C \otimes D) = AC \otimes BD$.

Corollary 1.3.5 *If* $A = [a_{ij} \in M_{m,n}(\mathbb{F})$ *and* $B \in M_{p,q}(\mathbb{F})$ *are nonsingular, then* $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$

Theorem 1.3.6 *Let* $A = [a_{ij} \in M_{m,n}(\mathbb{F})$ *and* $B \in M_{p,q}(\mathbb{F})$. *If* $\lambda \in \sigma(A)$ *and* $x \in \mathbb{C}^m$ *is the corresponding eigenvector of* A, *and if* $\mu \in \sigma(B)$ *and* $y \in \mathbb{C}^n$ *is the corresponding eigenvector of* B. *Then* $\lambda\mu \in \sigma(A \otimes B)$ *and* $x \otimes y \in \mathbb{C}^{mn}$ *is corresponding eigenvector of* $A \otimes B$. *Every eigenvalue of* $A \otimes B$ *arises as the product of the eigenvalues of* A *and* B. *If* $\sigma(A) = \lambda\{\lambda_1, ..., \lambda_m\}$ *and* $\sigma(B) = \{\mu_1, ..., \mu_n\}$, *then* $\sigma(A \otimes B) = \{\lambda_i \mu_j, i = 1, 2, ..., m, j = 1, 2, ..., n\}$. *In particular,* $\sigma(A \otimes B) = \sigma(B \otimes A)$.

Proposition 1.3.7 *If* $A \in M_m, B \in M_n$, *then*

(a) $\sigma(A \otimes B) = \sigma(B \otimes A)$.

(b) $det A \otimes B) = det(B \otimes A) = (det(A))^m (det(B))^n$.

(c) $tr(A \otimes B) = tr(B \otimes A) = tr(A)tr(B)$.

(d) $rank(A \otimes B) = rank(B \otimes A) = rank(A)rank(B)$.

(e) $A \otimes B$ *and* $B \otimes A$ *have the same singular values.*

(f) $A \otimes B$ *is normal if and only if* $B \otimes A$ *is normal.*

(g) $A \otimes B$ *is unitary if and only if* $B \otimes A$ *is unitary.*

(h) $A \otimes B$ *is always permutation equivalent to* $B \otimes A$, *moreover* $B \otimes A$ *is permutation similarity to* $B \otimes A$.

Definition 1.3.8 *With each matrix* $A = [a_{ij} \in M_{m,n}(\mathbb{F})$, *we associate the vector* $vec(A) \in \mathbb{F}^{mn}$ *defined by* $vec(A) := [a_{11}, ..., a_{1n}, a_{21}, ..., a_{2n}, a_{m1}, ..., a_{mn}]^T$

Proposition 1.3.9 *$vec(\cdot)$ is linear, i.e., $vec(\alpha A + \beta B) = \alpha vec(A) + \beta vec(B)$.*

Lemma 1.3.10 *Let $A = \in M_{m,n}(\mathbb{F}), B \in M_{p,q}(\mathbb{F})$ and $X \in M_{n,p}(\mathbb{F})$. Then $vec(AXB) = (B^T \otimes A)vec(X)$.*

1.4 Applications of the Kronecker product to matrix Equations

We start by studying the general linear matrix equation

$$A_1XB_1 + A_2XB_2 + \cdots + A_pXB_p = C. \tag{1.4.1}$$

where $A_j \in \mathbb{C}^{m\times m}, B_j \in \mathbb{C}^{n\times n}(j = 1, 2, \ldots, p)$. $X, C \in \mathbb{C}^{m\times n}$. and its particular cases. The method we use relies on the assertion of Kronecker Product proposition and reduce (1.4.1) to a matrix vector equation of the form $Gx = c$, where $G \in \mathbb{C}^{mn\times mn}$ and $c \in \mathbb{C}^{mn}$.

Theorem 1.4.1 *A matrix $X \in \mathbb{C}^{m\times n}$ is a solution of (1.4.1) if and only if the vector $x =$vec X is a solution of the equation*

$$Gx = c. \tag{1.4.2}$$

with $G = \sum_{j=1}^{p}(B_j^T \otimes A_j)$ and c=vec C.

Proof *By proposition of Kronecker Product, $vec(A_jXB_j) = (B_j^T \otimes A_j)$ vec X for each $(j = 1, 2, \ldots, p)$. Since the function vec A is linear. Then*

$$vec\ C = \sum_{j=1}^{p}(B_j^T \otimes A_j)\ vec\ X$$

and the result follows.

Having transformed the linear matrix equation to and equation of the form $Gx = c$, we may now apply the results developed further to determine criteria for the existence and uniqueness of a solution.

Corollary 1.4.2 *Equation* (1.4.1) *has a solution X if and only if rank $[G \quad c]$ =rank G.*

Corollary 1.4.3 *Equation* (1.4.1) *has a unique solution if and only if matrix G in* (1.4.2) *is nonsingular.*

We now consider an important particular case of (1.4.1), the matrix equation

$$AX + XB = C \tag{1.4.3}$$

Theorem 1.4.4 *Equation* (1.4.3) *has a unique solution if and only if the matrices A and $-B$ have no eigenvalues in common.*

Proof *In the case being considered, the matrix G in* (1.4.2) *is $(I_n \otimes A) + (B^T \otimes (I_m)$ and its eigenvalues are the number $\lambda_r + \mu_s, r = 1, 2, \ldots, m, s = 1, 2, \ldots, n$. It remains now to apply Corollary* (1.4.2) *and recall that a matrix nonsingular if and only if all its eigenvalues are nonzero.*

There is a special case in which we can obtain an explicit form for the solution matrix X of Equation (1.4.3).

Theorem 1.4.5 *If all eigenvalues of the matrices $A \in \mathbb{C}^{m\times m}$ have negative real parts (that is, A and B are stable), then the unique solution X of Equation* (1.4.3) *is given by*

$$X = -\int_0^\infty e^{At}Ce^{Bt}dt. \tag{1.4.4}$$

Proof *First observe that the conditions of the theorem imply that A and $-B$ have no eigenvalue in common, and hence there is a unique solution X of Equation* (1.4.3).

Consider a matrix-valued function $Z(t)$ defined as solution of the initial-value problem

$$dZdt = AZ + ZB, Z(0) = C. \tag{1.4.5}$$

The solution of Equation (1.4.5) *is*

$$Z(t) = e^{At}Ce^{Bt}$$

Now observe that, integrating the differential equation in Equation (1.4.5)*from $t = 0$ to $t = \infty$ under the hypothesis that matrix* (4) *exists, we obtain*

$$Z(x) - Z(0) = A\int_0^\infty Z(t)dt + (\int_0^\infty Z(t)dt)B.$$

Hence assuming, in addition, that

$$Z(\infty) = \lim_{t\to\infty} e^{At}Ce^{Bt} = 0. \tag{1.4.6}$$

it is found that the matrix given by Equation (1.4.4) *is the (unique) solution of Equation* (1.4.3). *Thus it suffices to check that, for stable A and B. The assumptions we have had to make are valid. To this end, write the spectral resolution for the function $e^{At}(0 < t < \infty)$.*

$$e^{At} = \sum_{k=1}^{s}\sum_{j=0}^{m_k 1} t^j e^{\lambda_k t} Z_{kj}. \tag{1.4.7}$$

where $\lambda_1, \lambda_2, \ldots, \lambda_s$. *Since* $x_k < 0(k = 1, 2, \ldots, s)$, *it follows that, for each* $k, exp(\lambda_k t) \to 0$ *as* $t \to \infty$. *Hence (see Equation* (1.4.7)*)* $\lim_{t\to\infty} e^{At} = 0$ *and similarly,* $\lim_{t\to\infty} e^{Bt} = 0$. *Thus, the relation as well as the existence of the integral in follow.*

1.5 Literature reviews

Theorem 1.5.1 *If* $A, B \in M_n$ *are positive definite. Then*

(a) $A \geq B$ *if and only if* $B^{-1} \geq A^{-1}$.

(b) *If* $A \geq B$, *then* $det(A) \geq det(B)$ *and* $tr(A) \geq tr(B)$.

(c) *If* $A > B$, *then* $det(A) \geq det(B)$ *and* $tr(A) \geq tr(B)$.

(d) *More generally, if* $A \geq B$, *then* $\lambda_k(A) \geq \lambda_k(B)$ *for all* $k = 1, 2, ..., n$ *if the respective eigenvalues of* A *and* B *are arranged in the same order.*

Theorem 1.5.2 *If* $A, B \in M_n$ *are positive definite. Then*

(a) $A^{-1} \circ B^{-1} \geq (A \circ B)^{-1}$.

(b) $A^{-1} \circ A^{-1} \geq (A \circ A)^{-1}$.

(c) $A^{-1} \circ A \geq I \geq (A^{-1} \circ A)^{-1}$.

Lemma 1.5.3 *Let* $A > 0$. *Then*

$$\begin{bmatrix} A & B \\ B^* & C \end{bmatrix} \geq 0$$

if and only if the ***Schur complement***

$$C - B^* A^{-1} B \geq 0$$

Lemma 1.5.4 *Let Φ be a unital positive linear map from M_m to M_n. Then*

$$\Phi(A^2) \geq \Phi(A)^2 \ , (A \geq 0). \tag{1.5.1}$$

$$\Phi(A^{-1}) \geq \Phi(A)^{-1}, (A > 0). \tag{1.5.2}$$

Theorem 1.5.5 *Let Φ be a unital positive linear map from M_m to M_n and f is an operator convex monotone function on $[0, \infty)$. Then for all $A \geq 0$,*

$$f(\Phi(A)) \geq \Phi(f(A)).$$

Theorem 1.5.6 *Let Φ be a unital positive linear map from M_m to M_n and f is an operator convex monotone function on $[0, \infty)$. Then for all $A \leq 0$,*

$$f(\Phi(A)) \leq \Phi(f(A)).$$

Lemma 1.5.7 *There exists a unital positive linear map from M_{n^2} to M_n such that $\Phi(A \otimes B) = A \circ B$ for all $A, B \in M_n$.*

Theorem 1.5.8 *(Löwner-Heinze). If $A \geq B \geq 0$ and $0 \leq r \leq 1$. Then $A^r \geq B^r$*

$$f(\Phi(A)) \leq \Phi(f(A)).$$

Corollary 1.5.9 *Let Φ be a unital positive linear map from M_m to M_n. Then*

$$\begin{aligned} \Phi(A^r) &\leq \Phi(A)^r \quad , \ A \geq 0, 0 < r \leq 1. \\ \Phi(A^r) &\geq \Phi(A)^r \quad , \ A > 0, -1 \leq r \leq 0 \ \text{or} \ \leq r \leq 2. \\ \Phi(\log A) &\leq \log(\Phi(A)) \ , \ A > 0. \end{aligned}$$

Corollary 1.5.10

$$\begin{aligned}
A^r \circ B^r &\leq (A \circ B)^r \quad , \ A, B \geq 0, 0 < r \leq 1. \\
A^r \circ B^r &\geq (A \circ B)^r \quad , \ A, B > 0, -1 \leq r \leq 0 \ \textit{or} \ \leq r \leq 2. \\
(\log A + \log B) \circ I &\leq \log(A \circ B) \ , \ A, B > 0.
\end{aligned}$$

Theorem 1.5.11 *If* $A \geq B \geq 0$ *then*

$$\begin{aligned}
(B^r A^p B^r)^{\frac{1}{q}} &\geq B^{\frac{(p+2r)}{q}}. \\
A^{\frac{(p+2r)}{q}} &\geq (A^r B^p A^r)^{\frac{1}{q}}
\end{aligned}$$

for $r \geq 0, p \geq 0, q \geq 1$ *with* $(1 + 2r)q \geq p + 2p$.

Theorem 1.5.12 *Let* $A, B, X, Y \in M_n$ *with* $A, B > 0$. *Then*

$$\begin{aligned}
(X^* A^{-1} X) \circ (Y^* B^{-1} Y) &\geq (X \circ Y)^* (A \circ B)^{-1} (X \circ Y). \\
(X^* A^{-1} X) + (Y^* B^{-1} Y) &\geq (X + Y)^* (A + B)^{-1} (X + Y).
\end{aligned}$$

Corollary 1.5.13 *Let* $A, B, X, Y \in M_n$ *with* $A, B > 0$. *Then*

$$\begin{aligned}
(X^* X) \circ (Y^* Y) &\geq (X \circ Y)^* (X \circ Y). \\
A^{-1} \circ B^{-1} &\geq (A \circ B)^{-1}.
\end{aligned}$$

1.6 Stability

1.6.1 Lyapunov stability

As mentioned in late content, Aleksandr Mikhailovich Lyapunov studied contraction and expansion phenomena of the motions of a mechanical system around and equilibrium.Translated in modern jargon, the study of Lyapunov stability concerns the asymptotic behavior of the state of an autonomous dynamical system. The main contribution of Lyapunov has been to define the concept of stability, asymptotic stability and instability of such systems and to give a method of verification of these concepts in terms of the existence of functions, called Lyapunov functions. Both his definition and his verification method characterize, in a local way, the stability properties of an autonomous dynamical system. Unfortunately, for the general class of nonlinear systems there are no systematic procedures for finding Lyapunov functions. However, we will see that for linear systems the problem of finding Lyapunov functions can be solved adequately as a feasibility test of a linear matrix inequality.

Let $\mathcal{X}$ be a set and $T \subseteq \mathbb{R}$ A flow is a mapping $\phi : T \times T \times \mathcal{X} \to \mathcal{X}$ which satisfies the

(a) consistency property: $\phi(t, t, x) = x$ and the

(b) semi-group property: $\phi(t_1, t_{-1}, x_0) = \phi(t_1, t_0, \phi(t_0, t_{-1}, x_0))$ for all $t_{-1} \leq t_0 \leq t_1$ and $x_0 \in \mathcal{X}$.

The set $\mathcal{X}$ is called the state space (or the phase space) and we will think of a flow as a state evolution map. A flow defines an unforced or autonomous dynamical system in the sense that the evolution of a flow is completely determined by an initial state and not by any kind of external input. A typical example of a flow is the solution of a differential equation of the form

$$\dot{x} = f(x(t), t) \tag{1.6.1}$$

with finite dimensional state space $\mathcal{X} = \mathbb{R}^n$ and where $f : \mathcal{X} \times T \to \mathcal{X}$ is a function. By a solution of Equation (1.6.1) over a (finite) time interval $T \subset \mathbb{R}$ we will mean function $x : T \to \mathcal{X}$ which is differentiable every where and which satisfies Equation (1.6.1) for all $t \in T$. By a solution over the unbounded interval $\mathbb{R}_+, \mathbb{R}_-$ or $\mathbb{R}$ we will mean a solution over all finite interval T contained in either of these sets. Let us denote by $\phi(t, t_0, x_0)$ the solution of Equation (1.6.1) at time t with initial condition $x(t_0) = x_0$. If for all pairs $(t_0, x_0) \in T \times \mathcal{X}$ one can guarantee existence and uniqueness of a solution $\phi(t, t_0, x_0)$ of (1.6.1) with $t \in T$, then the mapping ϕ satisfies the consistency and semi-group property, and therefore defines flow. In that case, ϕ is said to be the flow associated with the differential equation (1.6.1). This flow is called *time-invariant* if

$$\phi(t + \tau, t_0 + \tau, x_0) = \phi(t, t_0, x_0)$$

for any $\tau \in T$. It is said to be linear if $\mathcal{X}$ is a vector space and $\phi(t, t_0, \cdot)$ is linear mapping for all time instances t and t_0. For time-invariant flows we usually take (without loss of generality) $t_0 = 0$ as initial time. We remark that the condition on existence and uniqueness of solutions of Equation (1.6.1) holds whenever f satisfies a global lipschitz condition. That is, if $\mathcal{X}$ has the structure of a normed vector space with norm $\|\cdot\|$, and if for some $L > 0$, the inequality

$$\|f(x,t) - f(y,t)\| \leq L\|x - y\|$$

holds for $x, y \in \mathcal{X}, t \in T$.

An element x^* is a fixed point or an equilibrium point of Equation (1.6.1) if $f(x^*, t) = 0$ for all $t \in T$. It is easy to see that x^* is a fixed point if and only if $\phi(t, t_0, x^*) = x^*$ is a solution of Equation (1.6.1) for all t and t_0. In other words, solution of Equation (1.6.1) remain in x^* once they started there.

There exists a wealth of concepts to define the stability of a flow ϕ. The various notion of Lyapunov stability pertain to distinguished fixed point x^* of a flow ϕ and express to what extend and other trajectory $\phi(t, t_0, x_0)$, whose initial state x_0 lies in the neighborhood of x^* at time t_0, remains or gets close to $\phi(t, t_0, x^*)$ for all time instances $t \geq t_0$.

Definition 1.6.1 *(Lypunov stability) Let $\phi : T \times T \times \mathcal{X}$ be a flow and suppose that $T = \mathbb{R}$ and $\mathcal{X}$ is a normed vector space. The fixed point x^* is said to be*

(a) *stable (in the sense of Lyapunov) if given any $\varepsilon > 0$ and $t_0 \in T$, there exists $\delta = \delta(\varepsilon, t_0) > 0$ (not depending on t) such that*

$$\|x_0 - x^*\| \leq \delta \Rightarrow \|\phi(t, t_0, x_0) - x^*\| \leq \varepsilon \text{ for all } t \geq t_o. \quad (1.6.2)$$

(b) *attractive if for all $t_0 \in T$ there exists $\delta = \delta(t_0) > 0$ with the property that*

$$\|x_0 - x^*\| \leq \delta \Rightarrow \lim_{t\to\infty} \|\phi(t, t_0, x_0) - x^*\| = 0. \quad (1.6.3)$$

(c) *exponentially stable if for all $t_0 \in T$ there exists $\delta = \delta(t_0), \alpha = \alpha(t_0) > 0$ and $\beta = \beta(t_0) > 0$ such that*

$$\|x_0 - x^*\| \leq \delta \Rightarrow \|\phi(t, t_0, x_0) - x^*\| \leq \beta \|\phi(t, t_0, x_0) - x^*\| e^{-\alpha(t-t_0)} \ \forall t \geq t_o. \quad (1.6.4)$$

(d) *asymptotically stable (in the sense of Lyapunov) if it is both stable (in the sense of Lyapunov) and attractive.*

(e) *unstable if it not stable (in the sense of Lyapunov)*

(f) *uniformly stable (in the sense of Lyapunov) if given any $\varepsilon > 0$ there exists $\delta = \delta(\varepsilon) > 0$ (not depending on t_0) such that* (1.6.2) *holds for all $t_0 \in T$.*

(g) *uniformly attractive if there exists $\delta > 0$ (not depending on t_0) such that* (1.6.3) *holds for all $t_0 \in T$.*

(*h*) *uniformly exponentially stable (in the sense of Lyapunov) if there exists* $\delta > 0$ *(not depending on* t_0*) such that* (1.6.3) *holds for all* $t_0 \in T$.

(*i*) *uniformly asymptotically stable (in the sense of Lyapunov) if it both uniformly stable (in the sense of Lyapunov) and uniformly attractive.*

In words, a fixed point is stable if the graphs of all flow that initiate sufficiently close to x^* at time t_0, remain as close as desired to x^* for all time $t \geq t_0$. Stated otherwise, a fixed point is stable if the mapping $\phi(t, t_0, \cdot)$ is continuous at x^*, uniformly in $t \geq t_0$. The region of attraction associated with a fixed point x^* is defined to be the set of all initial states $x_0 \in \mathcal{X}$ for which $\phi(t, t_0, x_0) \to x^*$ as $t \to \infty$. If this region does not depend on t_0, it is said to be uniform, if it coincides with $\mathcal{X}$ then x^* is globally attractive. Similarly, we can define the region of stability, the region of asymptotic stability and the region of exponential stability associated with x^*. If these regions cover the entire state space $\mathcal{X}$, then the fixed point is called globally asymptotically stable, globally stable or globally exponentially stable.

We remark that there exist examples of stable fixed points that are not attractive, and examples of attractive fixed point that are not stable. The notion of exponential stability is the strongest in the sense that an exponentially stable fixed point is also asymptotically stable (i.e. a stable attractor). Similarly, uniform exponential stability implies uniform asymptotic stable.

For regulation task, the designed system is time invariant (TI)

$$\dot{x} = f(x) \tag{1.6.5}$$

in which case we are interest in its invariant sets.

Definition 1.6.2 *A set* M *is invariant if*

$$x(t_1) \in M \Rightarrow x(t) \in M, \forall t \geq t_1. \tag{1.6.6}$$

Proposition 1.6.3 *Let $\phi : T \times T \times \mathfrak{X}$ be a flow with $T = \mathbb{R}$ and suppose that x^* is a fixed point. If ϕ is linear then*

(a) *x^* is attractive if and only if x^* is globally attractive.*

(b) *x^* is asymptotically stable if and only if x^* is globally asymptotically stable.*

(c) *x^* is exponential stable if and only if x^* is globally exponentially stable.*

If ϕ is time-invariant then

(a) *x^* is stable if and only if x^* is uniformly stable.*

(b) *x^* is asymptotically stable if and only if x^* is uniformly asymptotically stable.*

(c) *x^* is exponentially stable if and only if x^* is uniformly exponentially stable.*

Definition 1.6.4 *(Definite functions) Let $\delta \subseteq \mathbb{R}^n$ have the point x^* in its interior and let $T \subseteq \mathbb{R}$. A function $V : \delta \times T \to \mathbb{R}$ is said to be*

(a) *positive definite (with respect to x^*) if there exist a continuous, strictly increasing function $a : \mathbb{R}_+ \to \mathbb{R}_+$ with $a(0) = 0$ such that $V(x,t) \geq a(\|x - x^*\|$ for all $(x,t) \in \delta \times T$.*

(b) *positive semi-definite if $V(x,t) \geq 0$ for all $(x,t) \in \delta \times T$.*

(c) *negative definite (around x^*)or negative semi-definite if $-V$ is positive definite or positive semi-definite, respectively.*

We introduced positive definite and positive semi-definite matrices in late Chapter. The use of terminology is consistent with Definition (1.6.4) in the following sense. If X is a real symmetric matrix then for all vector x we have that

$$\lambda_{min}(X)\|x\|^2 \leq x^T X x \leq \lambda_{max}(X)\|x\|^2.$$

Hence, the function $V(x) := x^T X x$ (not depending on time) is positive definite with respect to the origin if and only if the smallest eigenvalue $\lambda_{min}(X)$ is positive. Hence, X deserves to be called a positive definite matrix in that case. Similarly, V is negative definite with respect to the origin if and only if $\lambda_{max} < 0$. We denoted this by $X < 0$ and $X > 0$, respectively.

Consider the system (1.6.1) and suppose that x^* is and equilibrium point. Let δ be a set which has x^* in its interior and suppose that $V : \delta \times T \to \mathbb{R}$ has continuous partial derivatives (i.e., V is continuously differentiable). Consider, for $(x_0, t_0) \in \delta \times T$, the function $V^* : T \to \mathfrak{X}$ defined by the composition

$$V^*(t) := V(\phi(t, t_0, x_0), t).$$

Then this function is differentiable and its derivative reads

$$\dot{V}^*(t) = \partial_x V(\phi(t, t_0, x_0), t) f(\phi(t, t_0, x_0), t) + \partial_t V(\phi(t, t_0, x_0), t).$$

Now introduce the mapping $V' : \delta \times T \to \mathbb{R}$ by setting

$$V^*(t) = \partial_x V(x, t) f(x, t) + \partial_t V(x, t). \tag{1.6.7}$$

V' is called derivative of V along trajectories of Equation (1.6.1) and, by construction, we have that $\dot{V}^*(t) = V'(\phi(t, t_0, x_0), t)$ for all $t \in T$. It is very important to observe that V'

not only depends on V but also on the differential Equation (1.6.1). It is rather common to write $\dot{V}$ for V' and even more common to confuse $\dot{V}^*$ with V'. Formally, these objects are different as V^* is a function of time, whereas V' is a function of the state and time.

The main stability results for autonomous systems of the form Equation (1.6.1) are summarized in the following result.

Theorem 1.6.5 *(Lyapunov theorem) Consider the differential Equation* (1.6.1) *and let $x^* \in \mathcal{X}$ be and equilibrium point which belongs to the interior of a set δ.*

(*a*) *If there exists a positive definite, continuously differentiable function $V : \delta \times T \to \mathbb{R}$ with $V(x^*, t) = 0$ and V' negative semi-definite, then x^* is stable. If, in addition, V is decrescent, then x^* is uniformly stable.*

(*b*) *If there exists a positive definite decrescent and continuously differentiable function $V : \delta \times T \to \mathbb{R}$ with $V(x^*, t) = 0$ and V' negative definite, then x^* is uniformly asymptotically stable.*

Storage function, introduced in Chapter Dissipative system, and Lyapunov function are closely related as can been seen by comparing (1.6.7) with the differential dissipation inequality (3.2.4). Indeed, if $w(t) = w^*$ is taken to a constant input in the input-state-output system (3.2.1) then we obtain the autonomous time-invariant system

$$\begin{aligned} \dot{x} &= f(x, w^*) \\ z &= g(x, w^*). \end{aligned}$$

If x^* is an equilibrium point of this system and if this system is known to be dissipative with respect to a supply function that satisfies

$$s(w^*, g(x, w^*)) \leq 0.$$

for all x in neighborhood of x^*, then any (differentiable) storage function V satisfies the differential dissipation inequality (3.2.4) and is monotone non-increasing along solutions in a neighborhood of x^*. If the storage function is moreover non-negative in this neighborhood, then it follows from Theorem (1.6.5) that x^* is a stable equilibrium at x^*. In that case, the storage function V is nothing else than a Lyapunov function defined in a neighborhood of x^*. In particular, the dissipativity of the system implies the stability of the system provided the storage function is non-negative.

Unfortunately, of this general class of nonlinear differential equations there are no systematic procedures for actually finding such functions. We will see next that more explicit results can be obtained for the analysis of stability of solutions of linear differential equations.

Together with the autonomous system (1.6.1) let us consider the linear autonomous system

$$\dot{x} = Ax \tag{1.6.8}$$

where $A : \mathbb{R}^n \to \mathbb{R}^n$ is a linear map obtained as the linearization of $f : \mathcal{X} \to \mathcal{X}$ around an equilibrium point $x^* \in \mathcal{X}$ of (1.6.1). Precisely, for $x^* \in \mathcal{X}$ we write

$$f(x) = f(x^*) + \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x^*)[x - x^*] + \ldots$$

where we assume that f is at least once differentiable. The linearization of f around x^* is defined by the system (1.6.8) with A defined by the real $n \times n$ matrix

$$A := \partial_x f(x^*).$$

All elements is ker A are equilibrium points, but we will investigate the stability of (1.6.8) at the equilibrium $x^* = 0$. The positive definite quadratic function $V : \mathcal{X} \to \mathbb{R}$ defined by

$$V(x) = x^T X x$$

serves as a quadratic Lyapunov function. Indeed, V is continuous at $x^* = 0$, is assumes a strong local minimum at $x^* = 0$ (actually this is a strong global minimum of V), while the derivative of $V(x)$ in the direction of the vector field Ax is given by

$$V_x Ax = x^T [A^T X + XA] x$$

which should be negative to guarantee that teh origin is ans asymptotic stable equilibrium point of (1.6.8). We thus obtain the following result:

Proposition 1.6.6 *Let the linear system* (1.6.8) *be a linearization of* (1.6.1) *at the equilibrium* x^*. *The following statement are equivalent*

(a) *The origin is an asymptotic stable equilibrium for* (1.6.8).

(b) *The origin is a global asymptotic stable equilibrium for* (1.6.8).

(c) *All eigenvalues* $\lambda(A)$ *of* A *have strictly negative real part.*

(d) *The linear matrix inequalities*

$$A^T X + XA < 0, X > 0$$

are feasible.

Moreover, if one of these statements hold, then the equilibrium x^* *of the flow* (1.6.1) *is asymptotically stable.*

As the most important implication of Proposition (1.6.6), asymptotic stability of the equilibrium x^* of the nonlinear system (1.6.1) can be concluded from the asymptotic stability of its linearization at x^*.

Theorem 1.6.7 *(La Salle) Let Ω be a positive invariant set of* (1.6.5). *Let $V : \Omega \to \mathbb{R}_+, v \in C^1$ such that $\dot{v}(x) \le 0, \ \forall x \in \Omega$. Let $E\{x \in \Omega | \dot{v}(x) = 0\}$ and let M be the largest invariant set contained in E.*

Then, every bounded solution $x(t)$ starting in Ω converges to M as $t \to \infty$.

Corollary 1.6.8 *(Asymptotic Stability) Let $x = 0$ be the only equilibrium point of* (1.6.5). *Let $V : \mathbb{R}^n \to \mathbb{R}_+$ be a continuously differentiable ($V \in C^1$),be positive definite,R.U. such that $\dot{V} \le 0, \forall x \in \mathbb{R}^n$.*

Let $E\{x \in \mathbb{R}^n | \dot{v}(x) = 0\}$, and suppose that no solution other than $x(t) \equiv 0$ can say forever in E. Then the origin is globally asymptotically stable (GAS.).

Assumption 1.6.9 *Consider the system*

$$\dot{x} = f(x) + g(x)u, f(0) = 0. \tag{1.6.9}$$

where

$$x \in \mathbb{R}^n : \textit{state}$$
$$u \in \mathbb{R} : \textit{control input}$$

$$\exists \textit{ control law } u = \alpha(x), \alpha(0) = 0, \alpha(x) \in C^1 \tag{1.6.10}$$

$\exists$ C^∞, *be positive definite, R.U.,$V : \mathbb{R}^n \to \mathbb{R}$ such that*

$$\dot{V} = \frac{\partial V}{\partial X}(x)\dot{x} = \frac{\partial V}{\partial X}(x)[f(x) + g(x)u(x)] \leq -W(x) \leq 0, \forall x \in \mathbb{R}^n$$

where $W : \mathbb{R}^n \to \mathbb{R}$ *is positive semi-definite (P.S.D.).*

Under this assumption, the control (1.6.10), applied to the system (1.6.9), guarantees global boundedness of $x(t)$, and, via the LaSalle-Yoshizawa theorem, the regulation of $W(x(t))$:

$$\lim_{t\to\infty} W(x(t)) = 0. \tag{1.6.11}$$

A strong convergence result is obtained using LaSalle's theorem with $\Omega = \mathbb{R}^n$: $x(t)$ converges to the largest invariant set M contained in

$$E = \{x \in \mathbb{R}^n | W(x) = 0\}.$$

Clearly, if $W(x)$ is positive definite, the control (1.6.10) renders $x = 0$ the globally asymptotically stable (GAS.) equilibrium of (1.6.9).

Lemma 1.6.10 *(Integrator Backstepping)* (1.6.9) *augmented by an integrator:*

$$\dot{x} = f(x) + g(x)\xi \tag{1.6.12}$$

$$\dot{\xi} = u \tag{1.6.13}$$

and suppose that (1.6.12) *satisfies Assumption* (1.6.9) *with* $\xi \in \mathbb{R}$ *as its control.*

(*i*) *If* $W(x)$ *is positive definite, then*

$$V_a(x,\xi) = V(x) + \frac{1}{2}[\xi - \alpha(x)]^2 \tag{1.6.14}$$

is a clf for the full system (1.6.12), *that is, there exists* $u = \alpha_a(x,\xi)$ *which renders* $x = 0$, $\xi = 0$ *the GAS equilibrium of* (1.6.12). *One such control is*

$$u = -c(\xi - \alpha(x)) + \frac{\partial\alpha}{\partial x}[f(x) + g(x)\xi] - \frac{\partial V(x)}{\partial x}g(x), c > 0 \tag{1.6.15}$$

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

Chapter 2

Linear matrix inequality

2.1 Linear matrix inequalities

2.1.1 What are they?

A *linear matrix inequality* is an expression of the form

$$F(x) := F_0 + x_1 F_1 + \ldots + x_n F_n < 0 \qquad (2.1.1)$$

where

- $x = (x_1, \ldots, x_n)$ is a vector of n real numbers called the decision variables.
- $F_0, \ldots, F_n$ are real symmetric matrices, i.e., $F_j = F_j^T$,for $j = 0, \ldots, n$.
- the inequality < 0 in Equation (2.1.1) means 'negative definite'. That is, u^T $F(x)u < 0$ for all non-zero real vectors u. Because all eigenvalues of a real symmetric matrix are real, Equation (2.1.1) is equivalent to saying that all eigenvalues $\lambda(F(x))$ are negative. Equivalently, the maximal eigenvalue $\lambda_{max}(F(x)) < 0$.

It is convenient to introduce some notation for symmetric and Hermitian matrices. A matrix A is *Hermitian* if it is square and $A = A^* = \bar{A}^T$ where the bar denotes taking the complex conjugate of each entry in A. If A is real then this amounts to saying that $A = A^T$ and we call A *symmetric.* The sets of all $m \times m$ Hermitian and symmetric matrices will be denoted by $\mathbb{H}^m$ and $\mathbb{S}^m$, respectively, and we will omit the superscript m if it is not relevant for the context.

Definition 2.1.1 *(Linear Matrix Inequality) A* linear matrix inequality *(LMI) is an inequality*

$$F(x) < 0 \tag{2.1.2}$$

where F is an affine function mapping a finite dimensional vector space $\mathbb{X}$ to either $\mathbb{H}$ or $\mathbb{S}$.

Remark 2.1.2 *An affine mapping $F : \mathbb{X} \to \mathbb{S}$ necessarily takes the form $F(x) = F_0 + T(x)$ where $F_0 \in \mathbb{S}$(i.e., F_0 is real symmetric) and $T : \mathbb{X} \to \mathbb{S}$ is a linear transformation. Thus if $\mathbb{X}$ is finite dimensional, say of dimension n, and $e_1, ..., e_n$ constitutes a basis for $\mathbb{X}$, then every $x \epsilon \mathbb{X}$ can be represented as $x = \sum_{j=1}^{n} x_j e_j$ and we can write*

$$T(x) = T(\sum_{j=1}^{n} x_j e_j) \sum_{j=1}^{n} x_j F_j$$

where $F_j = T(e_j) \in \mathbb{S}$. Hence we obtain (2.1.1) as a special case.

Remark 2.1.3 *In most control applications, LMI's arise as functions of* matrix variables *rather than scalar valued decision variables. This means that we consider inequalities of the form* (2.1.2) *where $\mathbb{X} = \mathbb{R}^{m_1 \times m_2}$. A simple example with $m_1 = m_2 = m$ is the Lyapunov inequality $F(x) = A^T X + XA + Q < 0$. Here, A, $Q \in \mathbb{R}^{m \times m}$ are assumed to be given and X is the unknown matrix variable of dimension $m \times m$. Note that this defines an LMI only if $Q \in \mathbb{S}^m$. We can view this LMI as a special case of* (2.1.1) *by defining an arbitrary basis $e_1, ..., e_n$ of $\mathbb{X}$ and writing $X = \sum_{j=1}^{n} x_j e_j$. Then*

$$F(x) = F(\sum_{j=1}^{n} x_j e_j) = F_0 + \sum_{j=1}^{n} x_j F(e_j) = F_0 + \sum_{j=1}^{n} x_j F_j$$

which is of the form (2.1.1). The coefficients x_j in the expansion of X define the decision variables. Note that the number of decision variables n is at most m^2 and depend on the structure imposed on the matrix variable X. For example, of the matrix variable X is required to be symmetric, $\mathbb{X} = \mathbb{S}^m$ and $n = m(m+1)/2$.

Remark 2.1.4 A non-strict LMI *is a linear matrix inequality where* $<$ *in* (2.1.1) *and* (2.1.2) *is replaced by* $\leq$. *The matrix inequalities* $F(x) > 0$, *and* $F(x) < G(x)$ *with* F *and* G *affine functions are obtained as special cases of* (2.1.1) *as they can be rewritten as the linear matrix inequalities* $-F(x) < 0$ *and* $F(x) - G(x) < 0$.

2.1.2 Why are they interesting?

the linear matrix inequality (2.1.2) defines a *convex constraint* on x. That is, the set

$$\varrho := x \mid F(x) < 0$$

of solutions of the LMI $F(x) < 0$ is convex. Indeed, if $x_1, x_2 \in \varrho$ and $\alpha \in (0,1)$ then

$$F(ax_1 + (1-\alpha)x_2) = \alpha F(x_1) + (1-\alpha)F(x_2) < 0$$

where we used that F is affine and where the inequality follows from the fact that $\alpha > 0$ and $(1-\alpha) > 0$.

Although the convex constraint $F(x) < 0$ on x may seem rather special, it turns out that many convex sets can be represented in this way and that these sets have more attractive properties than general convex sets. In this subsection we discuss some seemingly trivial properties of linear matrix inequalities which turn out to be of eminent help to reduce multiple constraints on an unknown variable to an equivalent constraint involving a single linear matrix inequality.

Definition 2.1.5 *(System of LMI's)* A system of linear matrix inequalities *is a finite set of linear matrix inequalities*

$$F_1(x) < 0, ..., F_k(x) < 0. \tag{2.1.3}$$

From Proposition 1.4 we inter that the intersection of the feasible sets of each of the inequalities (2.1.3) is convex. In other words, the set of all x that satisfy (2.1.3) is convex. The question now arises whether or not this set can be represented as the feasibility set of another LMI. Ther answer is yes. Indeed, $F_1(x) < 0, ..., F_k < 0$ if and only if

$$\begin{bmatrix} F_1(x) & 0 & \cdots & 0 \\ 0 & F_2(x) & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & F_k(x) \end{bmatrix} < 0.$$

The last inequality indeed makes sense as $F(x)$ is symmetric(or Hermitian)for any x. Further, since the set of eigenvalues of $F(x)$ is simply the union of the eigenvalues of $F_1(x), ..., F_k(x)$,any x that satisfies $F(x) < 0$ also satisfies the system of LMI's(2.1.3) and vice versa. conclude that multiple LMI constraints can always be converted to a single LMI constraint.

A second important property amounts to incorporating *affine constraints* in linear matrix inequalities. By this, we mean that *combined constraints* (in the unknown x) of the form

$$\begin{cases} F(x)<0 \\ Ax=b \end{cases} or \begin{cases} F(x)<0 \\ x=Bu+c forsome u \end{cases}$$

where the affine function $F : \mathbb{R}^n \to \mathbb{S}$, matrices A and B and vectors b and c are given can be *lumped* in one linear matrix inequality $\hat{F}(\hat{x}) < 0$. More generally, the combined equations

$$\begin{cases} F(x)<0 \\ x \in \mathcal{M} \end{cases} \tag{2.1.4}$$

where $\mathcal{M}$ is an *affine set* in $\mathbb{R}^n$ can be rewritten in the form of one single linear matrix inequality $\hat{F}(\hat{x}) < 0$. To do this, recall that affine sets $\mathcal{M}$ can be written as

$$\mathcal{M} = x | x = x_0 + m, m \in \mathcal{M}_0$$

with $x_0 \in \mathbb{R}^n$ and $\mathcal{M}_0$ a linear subspace of $\mathbb{R}^n$. Suppose that $\hat{n} = dim(\mathcal{M}_0)$ and let $e_1, ..., e_{\hat{n}} \in \mathbb{R}$ be a basis of $\mathcal{M}_0$. Let $F(x) = F_0 + T(x)$ be decomposed as in Remark (2.1.2). Then (2.1.4) can be rewritten as

$$0 > F(x) = F_0 + T(x_0 + \textstyle\sum_{j=1}^{\hat{n}} x_j e_j) =$$

$$= F_0 + T(x_0) + \textstyle\sum_{j=1}^{\hat{n}} x_j T(e_j) = \hat{F}_0 + x_1 \hat{F}_1 + ... + x_{\hat{n}} \hat{F}_{\hat{n}} =: \hat{F}(\hat{x})$$

where $\hat{F}_0 = F_0 + T(x_0), \hat{F}_j = T(e_j)$ and $\hat{x} = col(x_1, ..., x_{\hat{n}})$ are the coefficients of $x - x_0$ in the basis of $\mathcal{M}_0$. This implies that $x \in \mathbb{R}^n$ satisfies (2.1.4) if and only if $\hat{F}(\hat{x}) < 0$. Note that the dimension $\hat{n}$ of $\hat{x}$ is at most equal to the dimension n of x.

A third property of LMI's is obtained from a simple exercise in algebra. It turns out to be possible to convert some *non-linear* inequalities to *linear inequalities*. Suppose that we partition a matrix $M \in \mathbb{R}^{n \times n}$ as

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$$

where M_{11} has dimension $r \times r$. Assume that M_{11} is non-singular. Then the matrix

$$S := M_{22} - M_{21} M_{11}^{-1} M_{12}$$

is called the *Schur complement* of M_{11} in M. If M is symmetric then we have that

$$M < 0 \Longleftrightarrow \begin{bmatrix} M_{11} & 0 \\ 0 & S \end{bmatrix} < 0 \Longleftrightarrow \begin{cases} M_{11} < 0 \\ S < 0 \end{cases}$$

For the interested reader, the result is obtained as follows. For any $F \in \mathbb{R}^{r \times (n-r)}$ there holds

$$\mathrm{M} = \begin{bmatrix} I & -F \\ 0 & I \end{bmatrix}^T \begin{bmatrix} I & F \\ 0 & I \end{bmatrix}^T M \begin{bmatrix} I & F \\ 0 & I \end{bmatrix} \begin{bmatrix} I & -F \\ 0 & I \end{bmatrix}$$

$$= \begin{bmatrix} I & -F \\ 0 & I \end{bmatrix}^T \begin{bmatrix} M_{11} & M_{11}F + M_{12} \\ F^T M_{11} + M_{21} & F^T M_{11} F + F^T M_{12} + M_{21} F + M_{22} \end{bmatrix} \begin{bmatrix} I & -F \\ 0 & I \end{bmatrix}.$$

Hence, $M < 0$ if and only if the matrix in the middle factor of the last expression is negative definite. Now take $F = -M_{11}^{-1} M_{12}$ to obtain the result. The following proposition is an immediate consequence of this observation.

Proposition 2.1.6 *(Schur complement) Let $F : \mathbb{X} \to \mathbb{S}$ be an affine function which is partitioned according to*

$$F(x) = \begin{bmatrix} F_{11}(x) & F_{12}(x) \\ F_{21}(x) & F_{22}(x) \end{bmatrix}$$

where $F_{11}(x)$ is square. Then the following statements are equivalent.

(a)

$$F(x) < 0.$$

(b)

$$\begin{cases} F_{11}(x) & < 0 \\ F_{22}(x) - F_{21}(x)[F_{11}(x)]^{-1} F_{12}(x) & < 0. \end{cases} \tag{2.1.5}$$

(c)

$$\begin{cases} F_{22}(x) & < 0 \\ F_{11}(x) - F_{12}(x)[F_{22}(x)]^{-1} F_{21}(x) & < 0. \end{cases} \tag{2.1.6}$$

Note that the second inequalities in (2.1.5) and (2.1.6) are *non-linear* constraints in x. Using this result, it follows that non-linear matrix inequalities of the form (2.1.5) or (2.1.6) can be converted to linear matrix inequalities. In particular, non-linear inequalities of the form (2.1.5) or (2.1.6) define convex constraints on the variable x in the sense that the solution set of these inequalities is convex.

2.1.3 What are they good for?

As we will see, many optimization problems in control, identification and signal processing can be formulated (or reformulated) using linear matrix inequalities. Clearly, it only makes sense to cast these problem in an LMI setting if these inequalities can be solved on an efficient and reliable way. Since the linear matrix inequality $F(x) < 0$ defines a *convex constraint* on the variable x, optimization problems involving the minimization (or maximization) of a performance function $f : \varrho \to \mathbb{R}$ with $\varrho : x|F(x) < 0$ belong to the class of *convex optimization problems.* Casting this in the setting of the previous section, it may be apparent that the full power of convex optimization theory can be employed if the performance function f is known to be convex.

Suppose that $F : \mathbb{X} \to \mathbb{S}$ is affine. There are two generic problems related to the study of linear matrix inequalities:

(*a*) **Feasibility:** The question whether or not there exist elements $x \in \mathbb{X}$ such that $F(x) < 0$ is called a *feasibility problem.* The LMI $F(x) < 0$ is called *feasible* if such x exists, otherwise it is said to be *infeasible.*

(*b*) **Optimization:** Let an *objective function* $f : \varrho \to \mathbb{R}$ be given and suppose that $\varrho = x|F(x) < 0$. The problem to determine

$$V_{opt} = \inf_{x \in \varrho} f(x)$$

is called an *optimization problem with an LMI constraint.* This problem involves the determination of V_{opt}, the calculation of an *almost optimal solution* x (i.e., for arbitrary $\varepsilon > 0$ the calculation of an $x \in \varrho$ such that $V_{opt} \leq f(x) \leq V_{opt} + \varepsilon$), or the calculation of a optimal solutions x_{opt}(elements $x_{opt} \in \varrho$ such that $V_{opt} = f(x_{opt})$).

Let us give some simple examples to motivate the study of these problems.

Example 1 : stability

consider the problem to determine asymptotic stability of the linear autonomous system

$$\dot{x} = Ax \tag{2.1.7}$$

where $A \in R^{n \times n}$. By this, we mean the problem to decide whether or not all functions $x : \mathbb{R} \to \mathbb{R}^n$ which satisfy (2.1.7) have the property that $\lim_{t\to\infty} x(t) = 0$. Lyapunov taught us that this system is asymptotically stable if and only if there exists $X = X^T$ such that $X > 0$ and $A^T X + XA < 0$. Indeed, in that case the function $V(x) := X^T X x$ qualifies as a Lyapunov function in that it is positive for all non-zero x and strictly decaying along solutions x of (2.1.7). Thus, asymptotic stability of the system (2.1.7) is equivalent to feasibility of the LMI

$$\begin{bmatrix} -X & 0 \\ 0 & A^T X + XA \end{bmatrix} < 0.$$

Example 2 : eigenvalue minimization

Let $F : \mathbb{X} \to \mathbb{S}$ be an affine function and let $\lambda_{max}(.)$ denote the maximal eigenvalue of a real symmetric matrix. Consider the problem to minimize $f(x) := \lambda_{max}(F(x))$ over x. clearly,

$$f(x) < \gamma \Longleftrightarrow \lambda max(F^T(x)F(x)) < \gamma^2 \Longleftrightarrow 1/\gamma F^T(x)F(x) - \gamma I < 0$$

$$\iff \begin{bmatrix} \gamma I & F(x) \\ F^T(x) & \gamma I \end{bmatrix} > 0$$

where the latter inequality is obtained by taking a Schur complement. If we define

$$y := \begin{bmatrix} \gamma \\ x \end{bmatrix}, G(\gamma) := -\begin{bmatrix} \gamma I & F(x) \\ F^T(x) & \gamma I \end{bmatrix}, g(y) := \gamma$$

then G is an *affine function* of y and the problem to minimize f over x is equivalent to the problem to minimize g over all y such that $G(\gamma) < 0$. Hence, this is an optimization problem with an LMI constraint and a *linear objective function* g.

Example 3 : a Leontief economy

A manufacturer is able to produce n different products from m different resources. Assume that the selling price of product j is p_j and that it takes the manufacturer a_{ij} units of resource i to produce one unit of product j. Let x_j denote the amount of product j that is to be produced and let a_i denote the amount of available units of resource i, $i = 1, ..., m$. A smart manager advised the manufacturer to maximize his profit

$$p(x_1, ..., x_n) := p_1x_1 + p_2x_2 + ... + p_nx_n.$$

but the manager can do this only subject to the production constraints

$$a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n \leq a_1$$

$$a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n \leq a_2$$

$$\vdots$$

$$a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n \leq a_m$$

and $x_j \geq 0, j = 1, ..., n$. Note that the manufacturer faces an optimization problem subject to a system of non-strict linear matrix inequalities.

Chapter 3

Dissipative dynamical systems

3.1 Introduction

The notion of dissipative is a most important concept in systems theory both for theoretical considerations as well as from a practical point of view. Especially in physical sciences, dissipativity is closely related to the notion of energy. Roughly speaking, a dissipative system is characterized by the property that at any time the amount of energy which the system can conceivably supply to its environment can not exceed the amount of energy that has been supplied to it. Stated otherwise, when time evolves, a dissipative system absorbs a fraction of its supplied energy and transforms it for example into heat, and increase of entrophy, mass, electro-magnetic radiation, or other kinds of energy 'losses'. In many applications, the question whether a system is dissipative or not can be answered from physical considerations on the way the system interacts with its environment. For example, by observing that the system is an interconnection of dissipative components or by considering systems in which a loss of energy is inherent to the behavior of the system due to friction, optical dispersion, evaporation losses, etc.

In this chapter we will formalize the notion of a dissipative system for a very general class of systems. It will be shown that the linear matrix inequalities occur in a very natural way in the study of linear dissipative systems. Perhaps the most appealing setting for studying LMI's in system and control theory is within the framework of dissipative dynamical systems. It will be shown that solutions of LMI's have a natural interpretation as *storage functions* associated with a dissipative system. This interpretation will play a key role in understanding the importance of LMI's in questions related to stability, performance, robustness, and a large variety of controller design problems.

3.2 Dissipative dynamic systems

Definitions and examples

Consider a continuous time, time invariant dynamical system Σ described by the equations

$$\dot{x} = f(x,w) \tag{3.2.1}$$

$$z = g(x,w) \tag{3.2.2}$$

Here, x is the state which takes its values in *a state space* X, w is the input taking its values in an *input space* W and z denotes the output of the system which assumes its values in the *output space* Z. Throughout this section, the precise representation of the system will not be relevant. What we need, though, is that for any initial condition $x_0 \in X$ and for any input w belonging to an input class W, there exists uniquely defined signals $x : \mathbb{R}_+ \to X$ and $x : \mathbb{R}_+ \to Z$ which satisfy (3.2.1) subject to $x(0) = x_0$. Here, $\mathbb{R}_+ = [0, \infty]$ is the time set. In addition, the output z is assumed to depend on w in a *causal* way; that is, if $w_1 \in W$ and $w_2 \in W$ are two input signals that are identical on $[0, T]$ then the outputs z_1 and z_2 of (3.2.1) corresponding to inputs w_1 and w_2 and the same (but arbitrary) initial condition $x(0) = x_0$ are also identical on $[0, T]$. The system

(3.2.1) therefore generates outputs from inputs and initial conditions while future values of the inputs do not have an effect on the past outputs. Let

$$s : W \times Z \to \mathbb{R}$$

be a mapping and assume that for all $t_0, t_1 \in \mathbb{R}$ and for all input-output pairs (w, z) satisfying (3.2.1) the composite function $s(w(t), z(t))$ is locally absolutely integrable, i.e., $\int_{t_0}^{t_1} |s(w(t), z(t))|dt < \infty$. The mapping s will be referred to as the supply function.

Definition 3.2.1 *(Dissipativity) The system Σ with supply function s is said to be* dissipative *if there exists a function $V : X \to \mathbb{R}$ such that*

$$V(x(t_0)) + \int_{t_0}^{t_1} s(w(t), z(t))dt \geq V(x(t_1)) \tag{3.2.3}$$

for all $t_0 \leq t_1$ and all signals (w, x, z) which satisfy (3.2.1). *The pair (Σ, s) is said to be* conservative *if equality holds in* (3.2.3) *for all $t_0 \leq t_1$ and all (w, x, z) satisfying* (3.2.1).

Interpretation 3.2.2 *The supply function (or supply rate) s should be interpreted as the supply delivered to the system. This means that $s(w(\cdot), z(\cdot))$ represents the rate at which supply flows into the system if the system generates the input-output pair $(w(\cdot), z(\cdot))$. In other words, in the time interval $[0, T]$ work has been done on the system whenever $\int_0^T s(w(t), z(t))dt$ is positive, while work is done by the system if this integral is negative. The function V is called a storage function and generalizes the notion of an energy function for a dissipative system. With this interpretation, inequality* (3.2.4) *formalizes the idea that a dissipative system is characterized by the property that the change of internal storage $V(x(t_1)) - V(x(t_0))$ in any time interval $[t_o, t_1]$ will never exceed the amount of supply that flows into the system. This means that part of what is supplied to the system is stored, while the remaining part is dissipated. Inequality* (3.2.4) *is known as the dissipation inequality.*

If the composite function $\bar{V}(t) := V(x(t))$ is differentiable, then

$$\frac{d\bar{V}}{dt}(t) = V_x(x(t))f(x(t)), \;\; w(t) \leq s(w(t), z(t))$$

for all t and all solution (w, x, z) of (3.2.1). Here, V_x denotes the gradient of V. This observation makes dissipativity of a dynamical system a local property in the sense that (Σ, s) is dissipative if and only if

$$V_x(x)f(x, w) \leq s(w, g(x, w)) \tag{3.2.4}$$

holds for all points $x \in \mathbb{R}^n$ and $w \in \mathbb{R}^m$. We will refer to (3.2.5) as the *differential dissipation inequality.*

The classical motivation for the study of dissipativity comes from circuit theory. In the analysis of electrical networks the product of voltages and currents at the external branches of a network, i.e. the power, is an obvious supply function. Similarly, the product of forces and velocities of masses is a candidate supply function in mechanical systems. For those familiar with the theory of bond-graphs we remark that every bond-graph can be viewed as a representation of a dissipative dynamical system where inputs and outputs are taken to be effort and flow variables and the supply function is the product of these two variables. A bond-graph is therefore a special case of a dissipative system.

Example 2.1 Consider a thermodynamic system at uniform temperature T on which mechanical work is being done at rate W and which is being heated at rate Q. Let (T, Q, W) be the external variables of such a system and assume that -either by physical of chemical principles or through experimentation- the mathematical model of the thermodynamic system has been decided upon and is given by the time invariant system (3.2.1). The first and second law of thermodynamics may then be formulated in the sense of Definition (3.2.1) by saying that the system Σ is *conservative* with respect to the supply function $s_1 := (W + Q)$ and dissipative with respect to the supply function $s_2 := -Q/T$. Indeed, the first law of thermodynamics states that for all system trajectories (T, Q, W) and all time instants $t_0 \leq t_1$,

$$E(x(t_0)) + \int_{t_0}^{t_1} Q(t) + W(t)dt = E(x(t_1))$$

(conservation of thermodynamic energy). The second law of thermodynamics states that all system trajectories satisfy

$$S(x(t_0)) + \int_{t_0}^{t_1} -\frac{Q(t)}{T(t)}dt \geq S(x(t_1))$$

for all $t_0 \leq t_1$. Here, E is called the *internal energy* and S the *entropy*. The first law promises that the change of internal energy is equal to the heat absorbed by the system and the mechanical work which is done on the system. The second law states that the entropy decreases at a higher rate than the quotient of absorbed heat and temperature. It follows that thermodynamic systems are dissipative with respect to two supply functions! Nernst's third law of thermodynamics -the entropy of any object of zero temperature is zero- is only a matter of scaling of the entropy function S and does not further constrain the trajectories of the system.

Example 2.2 Other examples of supply function $s : W \times Z \to \mathbb{R}$ are quadratic forms

$$\begin{aligned} s(w,z) &= w^T x, \\ s(w,z) &= \|z\|^2 + \|w\|^2, \\ s(w,z) &= \|z\|^2 - \|w\|^2, \\ s(w,z) &= \|w\|^2 \end{aligned}$$

which arise in network theory, bond-graph theory, scattering theory, H_∞ theory, game theory and LQ-optimal control and H_2-optimal control theory. We will come across these examples in more detail below.

There are a few refinements to Definition (3.2.1) which are worth mentioning. Definition (3.2.1) can be generalized to *time-varying* systems by letting the supply rate s explicitly depend on time. Another refinement consists of the idea that a system may be dissipative with respect to more than one supply function. See example 2.1 for an example. Also, a notion of robust dissipativity may be developed in which the system description (3.2.1) is not assumed to be perfectly known, but uncertain to some well defined extend. An uncertain system is called robustly dissipative if (3.2.3) holds for all $t_0 \leq t_1$ and all trajectories (w, x, z) that can possibly be generated by the uncertain system. The notion of strict dissipativity is a refinement of Definition (3.2.1) which will prove useful in the sequal. It is defined as follows

Definition 3.2.3 *(Strict dissipativity) The system Σ with supply rate s is to be strictly dissipative if there exists a storage function $V : X \to \mathbb{R}$ and $\varepsilon > 0$ such that*

$$V(x(t_0)) + \int_{t_0}^{t_1} (s(w(t), z(t)) - \varepsilon^2 \|w(t)\|^2) dt \geq V(x(t_1)) \tag{3.2.5}$$

for all $t_0 \leq t_1$ and all trajectories (w, x, z) which satisfy (3.2.1).

A strictly dissipative system satisfies (3.2.3) with strict inequality, which justifies its name. As a final comment we mention the notion of *cyclo dissipativity* which has been introduced. For $T > 0$, the function $w : \mathbb{R} \to W$ is said to be $T-periodic$ if for all $t \in \mathbb{R}$ we have that $w(t) = w(t+T)$. A system Σ with supply function s is called *cyclo dissipative* if for all $T > 0$ there holds

$$\int_0^T s(w(t), z(t)) dt \geq 0 \tag{3.2.6}$$

for all T-periodic trajectories $(w(\cdot), z(\cdot))$ which satisfy (3.2.1) *Cyclo dissipativity* is therefore a system property defined in terms of T-periodic trajectories only. The importance of this notion lies in the fact that it avoids reference to the internal state space structure of the system and requires a condition on single in the external (input-output) behavior of the system only. It is easily seen that a dissipative system is cyclo dissipative whenever the state x is observable from (w, z), that is, whenever x is uniquely defined by any (w, z) which satisfies (3.2.1). Conversely, under some mild minimality and connectability condition on the state X, a cyclo dissipative system is also dissipative.

3.3 A classification of storage functions

Suppose that (Σ, s) is dissipative and let $x^* \in X$ be a fixed reference point in the state space of Σ. Instead of considering the set of all possible storage function associated with (Σ, s), we will restrict attention to set of normalized storage function defined by

$$\mathcal{V}(x^*) := \{V : X \to \mathbb{R} | V(x^*) = 0 \ \textit{and} \ (3.2.3) \ \textit{holds}\}.$$

Hence, x^* is reference point of neutral storage. Clearly, if V is a storage function satisfying (3.2.3), the we can obtain a normalized storage function by defining $\tilde{V}(x) := V(x) - V(x^*)$ by noting that $\tilde{V} \in \mathcal{V}(x^*)$.

Two mappings $V_{av} : X \to \mathbb{R} \cup [+\infty]$ and $V_{req} : X \to \mathbb{R} \cup [-\infty]$ will play a crucial role in the sequel. They are defined by

$$V_{av}(x_0) := sup\{-\int_0^{t_1} s(t)dt | t_1 \geq 0 \ ; (w, x, z) \ satisfy \ (3.2.1) \tag{3.3.1}$$

$$\text{with } x(0) = x_0 \ \textit{and} \ x(t_1) = x^*\}$$

$$V_{req}(x_0) := inf \int_{t-1}^{0} s(t)|t_{-1} \leq 0 \;;(w,x,z) \;\; satisfy \;\; (3.2.1) \tag{3.3.2}$$

$$\text{with } x(0) = x_0 \;\; and \;\; x(t_{-1}) = x^*\}$$

Here, $V_{av}(x)$ denotes the maximal amount of internal storage that may be recovered from the system over all state trajectories starting in x and eventually ending in X^*. Similarly, $V_{req}(x)$ reflects the minimal supply which needs to be delivered to the system in order to x via any trajectory originating in x^*. We refer to V_{av} and V_{req} as the available storage and the required supply, (measured with respect to x^*). In (3.3.1) it is silently assumed the for $x_0 \in X$ there exists an input $w \in W$ which steers the state from x^* at some time constant $t_{-1} < 0$ to x_0 at time $t = 0$ and back to x^* at time $t_1 > 0$. We call x_0 connectable with x^* if this property holds. If such a circle can be run in finite time for any $x_0 \in X$, then we say that every state is connectable with x^*. the following characterization is the main result of this section

Proposition 3.3.1 *Let the system Σ be represented by (3.2.1) and let s be a supply function. Suppose that every state is connectable with x^* for some $x^* \in X$. Then the following statements are equivalent*

(a) (Σ, s) *is dissipative.*

(b) $-\infty < V_{av}(x) < \infty$ *for all* $x \in X$

(c) $-\infty < V_{req}(x) < \infty$ *for all* $x \in X$

moreover, if one of these equivalent statements hold, then

(a) $V_{av}, V_{req} \in \mathcal{V}(x^*)$.

(b) $\{V \in \mathcal{V}(x^*)\} \Rightarrow \{$*for all* $x \in X$ *there holds* $V_{av}(x) \leq V(x) \leq V_{req}(x)\}$

(c) $\mathcal{V}(x^*)$ *is a convex set. In particular,* $V_\alpha := \alpha V_{req} + (1-\alpha)V_{req} \in \mathcal{V}(x^*)$ *for all* $\alpha \in (0,1)$.

Interpretation 3.3.2 *Proposition* (3.3.1) *confirms the intuitive idea that a dissipative system can neither supply nor store and infinite amount of energy during any experiment that starts or ends in a state of neutral storage. Proposition* (3.3.1) *shows that a system is dissipative if and only if the available storage and the required supply are real (finite) valued functions. Moreover, both the available storage and required supply are possible storage function of a dissipative system, these functions are normalized and define extremal storage functions in* $\mathcal{V}(x^*)$ *in the sense that* V_{av} *is the smallest and* V_{req} *is the largest element in* $\mathcal{V}(x^*)$*. In particular, for any state of a dissipative system, the available storage can not exceed its required supply. In addition, convex combinations of the available storage and the required supply are candidate storage functions.*

3.4 Dissipation functions

If the system Σ is dissipative with respect to the supply function s the, by (3.2.4),

$$d(x,w) := s(w, g(x,w)) - V_x(x)f(x,w)$$

is a non-negative quantity that can be interpreted as the amount of supply that is dissipated in the system when it finds itself in stat x while the input w is supplied.

Definition 3.4.1 *(Dissipation functions) A function* $d : X \times W \to \mathbb{R}$ *is called a dissipation function for* (Σ, s) *if there exists a storage function* $V : X \to \mathbb{R}$ *such that*

$$V(x(t_0)) + \int_{t_0}^{t_1} [s(w(t), z(t)) - d(x(t), w(t))]dt = V(x(t_1))$$

holds for all $t_0 \le t_1$ *and all trajectories* (w, x, z) *which satisfy* (3.2.1)

Obviously, the pair (Σ, s) is dissipative if and only if there exists a dissipation function d for (Σ, s) which is non-negative. Note that strictly dissipative if and only if there exists $\varepsilon > 0$ such that $d(x, w) := \varepsilon^2 \|w\|^2$ is a dissipation function for (Σ, s).

3.5 Linear dissipative system with quadratic supply rates

In the previous section we analyzed the notion of dissipativity at a fairly high level of generality. In this section we will apply the above theory to linear input-output system Σ described by

$$\begin{bmatrix} \dot{x} \\ z \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix} \tag{3.5.1}$$

with state space $X = \mathbb{R}^n$, input space $W = \mathbb{R}^m$ and output space $Z = \mathbb{R}^p$. Let $x^* = 0$ be the point of neutral storage and consider supply functions that are general quadratic functions $s : W \times Z \to \mathbb{R}$ defined by

$$s(w, z) = \begin{bmatrix} w \\ z \end{bmatrix}^T \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} w \\ z \end{bmatrix} = w^T Q w + w^T S z + z^T S^T w + z^T R z. \tag{3.5.2}$$

Here, the matrix

$$P := \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix}$$

is a real symmetric matrix (that is, $P \in \mathbb{S}^{m+p}$) which is partitioned conform with w and z. No a priori definiteness assumptions are made on P.

Substituting the output equation $z = Cx + Dw$ in (3.5.2) can equivalently be viewed as a quadratic function in the variables x and w. Indeed,

$$s(w,z) = s(w, Cx + Dw) = \begin{bmatrix} x \\ w \end{bmatrix}^T \begin{bmatrix} 0 & I \\ C & D \end{bmatrix}^T \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} 0 & I \\ C & D \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix}$$

This quadratic form in x and w will be used frequently in the sequel.

3.5.1 Main results

The following theorem is the main result of this chapter. It provides necessary and sufficient conditions for the pair (Σ, s) to be dissipative. In addition, it provides a complete parametrization of all normalized storage functions, together with a useful frequency domain characterization of dissipativity.

Theorem 3.5.1 *Suppose that the system Σ described by Equation* (3.5.1) *is controllable and let the supply function s be defined by Equation* (3.5.2). *Then the following statement are equivalent*

(*a*) (Σ, s) *is dissipative.*

(*b*) (Σ, s) *admits a quadratic storage function $V(x) := x^T K x$ with $K = K^T \in \mathbb{R}^{n \times n}$.*

(*c*) *There exists $K = K^T$ such that*

$$F(K) = \begin{bmatrix} A^T K + KA & KB \\ B^T K & 0 \end{bmatrix} - \begin{bmatrix} 0 & I \\ C & D \end{bmatrix}^T \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} 0 & I \\ C & D \end{bmatrix} \leq 0 \tag{3.5.3}$$

(*d*) *There exists $K_- = K_-^T$ such that $V_{av}(x) = x^T K_- x$.*

(*e*) *There exists $K_+ = K_+^T$ such that $V_{av}(x) = x^T K_+ x$.*

(*f*) *For all* $w \in \mathbb{R}$ *with* $det(iwI - A) \neq 0$, *the transfer function* $T(s) := C(Is - A)^{-1}B + D$ *satisfies*

$$\begin{bmatrix} I \\ T(iw) \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} I \\ T(iw) \end{bmatrix} \geq 0 \qquad (3.5.4)$$

Moreover, if one of the above equivalent statements holds, then $V(x) := x^T K x$ *is a quadratic storage function in* $\mathcal{V}(0)$ *if and only if* $F(K) \leq 0$.

We recognize in (3.5.3) a non-strict linear matrix inequality. The matrix $F(K)$ is usually called the dissipation matrix. Observe that in the above theorem the set of quadratic storage functions in $\mathcal{V}(0)$ is completely characterized by the linear matrix inequality $F(K) \leq 0$. In other words, the set of normalized quadratic storage functions associated with (Σ, s) coincides with the feasibility set of the system of LMI $F(K) \leq 0$. In particular, the available storage and the required supply are quadratic storage functions and hence K_- and K_+ also satisfy $F(K_-) \leq 0$ and $F(K_+) \leq 0$. Using Proposition (3.3.1), it moreover follows that any solution $K = K^T$ of $F(K) \leq 0$ has the property that

$$K_- \leq K \leq K_+.$$

In other words, among the set of symmetric solutions K of the LMI $F(K) \leq 0$ there exists a smallest and a largest element. The inequality (3.5.6) is called the frequency domain inequality. The equivalence between statements 1 and the frequency domain characterization in statement 6 has a long history in system theory. The result goes bach to Popov (1962), V.A. Yakubovich (1962) and R. Kalman (1963) and is an application of the 'Kalman-Yakubovich-Popov Lemma', which we present and discuss subsection below.

For conservative system with quadratic supply functions a similar characterization can be given. The precise formulation is evident from Theorem (3.5.1). Strictly dissipative systems are characterized as follows.

Theorem 3.5.2 *Suppose that the system Σ is described by* (3.2.1) *and let s be a supply function defined by* (3.2.3). *Then following statements are equivalent.*

(*a*) (Σ, s) *is dissipative.*

(*b*) *There exists $K = K^T$ such that*

$$F(K) = \begin{bmatrix} A^TK + KA & KB \\ B^TK & 0 \end{bmatrix} - \begin{bmatrix} 0 & I \\ C & D \end{bmatrix}^T \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} 0 & I \\ C & D \end{bmatrix} < 0 \tag{3.5.5}$$

(*c*) *For all $w \in \mathbb{R}$ with $det(iwI - A) \neq 0$, the transfer function $T(s) := C(Is - A)^{-1}B + D$ satisfies*

$$\begin{bmatrix} I \\ T(iw) \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} I \\ T(iw) \end{bmatrix} > 0 \tag{3.5.6}$$

Moreover, if one of the above equivalent statements holds, then $V(x) := x^TKx$ is a quadratic storage function satisfying (3.2.5) *for some $\epsilon > 0$ if and only if $F(K) < 0$.*

Remark 3.5.3 *Contrary to Theorem* (3.5.1), *the system Σ is not assumed to be controllable in Theorem* (3.5.1).

3.5.2 The Kalman-Yakubovich-Popov lemma

As mentioned in the proofs of Theorem (3.5.1) and Theorem (3.5.2), the Kalman-Yakubovich-Popov lemma is at the basis of the relation between frequency dependent matrix inequalities and algebraic feasibility property of a linear matrix inequality. We will use this important result at various instance. The Lemma originates from a stability criterion of nonlinear feedback system given by Popov in 1962. Yakubovich and Kalman introduced the lemma by showing that the frequency condition of Popov is equivalent to the existence of Lyapunov function.

We present a very general statement of the lemma which is free of any assumption on the system parameters. The proof which we present here is an elementary proof based on the dualization result.

Lemma 3.5.4 *(Kalman-Yakubovich-Popov) Let* $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}$ *and* $M \in \mathbb{S}^{(n+m)\times(n+m)}$. *Then the following statements are equivalent.*

(a) *There exists* $K \in \mathbb{S}^{n\times n}$ *such that*

$$\begin{bmatrix} A^TK + KA & KB \\ B^TK & 0 \end{bmatrix} + M < 0 \tag{3.5.7}$$

(b) *For all* $x \in \mathbb{C}^n, w \in \mathbb{C}^m$ *for which* $(iwI - A)x = Bw, w \in \mathbb{R}$ *and* $col(x, w) \neq 0$

$$\begin{bmatrix} x \\ w \end{bmatrix}^* M \begin{bmatrix} x \\ w \end{bmatrix} < 0 \tag{3.5.8}$$

If (A, B) *is controllable, the corresponding equivalence also holds for non-strict inequalities.*

If

$$M = -\begin{bmatrix} 0 & I \\ C & D \end{bmatrix}^T \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} 0 & I \\ C & D \end{bmatrix}$$

then statement 2 *is equivalent to the condition that for all* $w \in \mathbb{R}$, *with* $det(iwI - A) \neq 0$,

$$\begin{bmatrix} I \\ C(iwI - A)^{-1}B + D \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^T & R \end{bmatrix} \begin{bmatrix} I \\ C(iwI - A)^{-1}B + D \end{bmatrix} < 0.$$

lemma (3.5.4) therefore reduces to the equivalence between the linear matrix inequality and the frequency domain inequality in Theorem (3.5.1) and Theorem (3.5.2). The Kalman-Yakubovich-Popov lemma therefore completes the proofs of these theorems.

3.5.3 The positive real lemma

Consider the system (3.5.1) together with the quadratic supply function $s(w, z) = z^T w + w^T z$. Then the following is an immediate consequence of Theorem (3.5.1).

Corollary 3.5.5 *Suppose that the system Σ described by* (3.5.1) *is controllable and has transfer function T. Let $s(w, z) = z^T w + w^T z$ be a supply function. Then equivalent statements are*

(*a*) (Σ, s) *is dissipative.*

(*b*) *the LMI*

$$\begin{bmatrix} -A^T K - KA & -KB + C^T \\ -B^T K + C & D + D^T \end{bmatrix} \geq 0.$$

is feasible

(*c*) *For all $w \in \mathbb{R}$ with det$(iwI - A) \neq 0$ one has $T(iw) + T(iw) \geq 0$.*

Moreover, $V(x) = x^T K x$ defines a quadratic storage function if and only if K satisfies the above LMI.

Remark 3.5.6 *Corollary* (3.5.5) *is know as the positive real lemma and it has played a crucial role in questions related to the stability of control systems and synthesis of passive electrical networks. Transfer functions which satisfy the third statement are generally called positive real.*

3.5.4 The bounded real lemma

Consider the quadratic supply function

$$s(w,z) = \gamma^2 w^T w - z^T z \tag{3.5.9}$$

where $\gamma \geq 0$. We obtain the following result as an immediate consequence of Theorem (3.5.1).

Corollary 3.5.7 *Suppose that the system Σ described by (3.5.1) is controllable and has transfer function T. Let $s(w,z) = \gamma^2 w^T w - z^T z$ be a supply function where $\gamma \geq 0$. Then equivalent statements are*

(*a*) *(Σ, s) is dissipative.*

(*b*) *the LMI*

$$\begin{bmatrix} A^T K + KA + C^T C & KB + C^T D \\ B^T K + D^T C & D^T D - \gamma^2 I \end{bmatrix} \geq 0.$$

is feasible

(*c*) *For all $w \in \mathbb{R}$ with $det(iwI - A) \neq 0$ one has $T(iw)^* T(iw) \leq \gamma^2 I$.*

Moreover, $V(x) = x^T K x$ defines a quadratic storage function if and only if K satisfies the above LMI.

Let us analyze the importance of this result. If Σ is dissipative with respect to the supply function (3.5.9) then (3.2.4) implies that for any quadratic storage function $V(x) = x^T K x$,

$$V_x(Ax + Bw) \leq \gamma^2 w^T w - z^T z \tag{3.5.10}$$

for all vector w, x and $z = Cx + Dw$. Suppose that $x(0) = 0, A$ has all its eigenvalues in the open left-half complex plane and the input w is taken from the set $\mathcal{L}_2$ of square integrable function, i.e., w is such that

$$\|w\|_2^2 := \int_0^\infty w^T(t)w(t)dt < \infty$$

Then both the state x and output z of (3.5.1) are square integrable function and $\lim_{t\to\infty} x(t) = 0$. Now integrate (3.5.10) from $t = 0$ till ∞ to obtain that for all $w \in \mathcal{L}_2$

$$\gamma^2\|w\|_2^2 - \|z\|_2^2 \geq 0.$$

Equivalently, for all $w \in \mathcal{L}_2$ we have that $\|z\|_2 \leq \gamma$ or, stated otherwise,

$$\sup_{w\in\mathcal{L}_2} \frac{\|z\|_2}{\|w\|_2} \leq \gamma \tag{3.5.11}$$

The left-hand side of (3.5.11) is the so called $\mathcal{L}_2$-induced norm or $\mathcal{L}_2$-gain of the system (3.5.1) (or of its transfer function T). The $\mathcal{L}_2$-gain is therefore characterized by the smallest $\gamma \geq 0$ for which (3.5.11) holds.

Chapter 4

Systems with parametric uncertainty

First principle models of physical systems are often represented by state space description in which the various component of the state variable represent a well defined physical quantity. Variations, perturbations or uncertainties in specific physical parameters lead to uncertainty in the model. Often, this uncertainty is reflected by variations in well distinguished parameters or coefficients in the model,while in addition the nature and/or range of the uncertain parameters may be known, or partially known. Since very small parameter variations may have a major impact on the dynamics of a system it is of evident importance to analyze parametric uncertainties of dynamical systems. This will be the subject of this chapter.

4.1 Parametric uncertainties

Suppose that $\delta = (\delta_1, \ldots, \delta_p)$ is the vector which expresses the ensemble of all uncertain quantities in a given dynamical systems. There are at least two distinct cases which are of independent interest:

(a) **time-invariant parametric uncertainties**: the vector δ is a fixed but unknown element of a uncertainty set $\Delta \subseteq \mathbb{R}^p$.

(*b*) **time-varying parametric uncertainties**: the vector δ is an unknown time varying function $\delta : \mathbb{R} \rightarrow \mathbb{R}^k$ whose values $\delta(t)$ belong to an uncertainty set $\Delta \subseteq \mathbb{R}^p$, and possibly satisfy additional constraints on rates of variation, continuity, spectral content, etc.

The first case typically appears in models in which the physical parameters are fixed but only approximately known up to some level of accuracy. The second case typically captures models in which the uncertainty in parameters, coefficients, or other physical quantities is time-dependent. One may object that in many practical situations both time-varying and time-invariant uncertainties occur so that the distinction between the two cases may seem somewhat artificial. This is true, but since time-invariant uncertainties can equivalently be view as time-varying uncertainties with a zero rate constraint, combined time-varying and time-invariant uncertainties are certainly not excluded. A rather general class of *uncertain* continuous time, dynamical systems is described by the state apace equations.

$$\dot{x} = f(x, w, \delta) \tag{4.1.1}$$

$$z = g(x, w, \delta) \tag{4.1.2}$$

where δ may or may not be time-varying.This constitutes a generalization of the model described in (3.2.1). If the uncertainties δ are fixed but unknown elements of an uncertainty set $\delta \subseteq \mathbb{R}^p$ then one way to think of equations of this sort is to view them as a set of time-invariant systems, parametrized by $\delta \in \Delta$. However, if δ is time-varying, then (4.1.1) is to be interpreted as $\dot{x} = f(x(t), w(t), \delta(t))$ and (4.1.1) is better viewed as a time-varying dynamical system. If the components of $\delta(t)$ coincide, for example, with state components then (4.1.1) defines a non-linear system, even when the mapping f and g are linear. If $\delta(t)$ is scalar valued and assumes values in a finite set $\Delta = 1, \ldots, k$ then (4.1.1) defines a hybrid system of K modes whose kth mode is defined by the dynamics.

$$\begin{aligned} \dot{x} &= f_k(x,w) := f(x,y,k) \\ z &= g_k(x,w) := g(x,w,k) \end{aligned}$$

and where the time-varying behavior of $\delta(t)$ defines the switching events between the various modes. In any case, the systems (4.1.1) is of considerable theoretical and practical interest as it covers quite some relevant classes of dynamical systems.

4.1.1 Affine parameter dependent systems

If f and g in (4.1.1) are linear in x and w then the uncertain model assumes a representation

$$\begin{bmatrix} \dot{x} \\ z \end{bmatrix} = \begin{bmatrix} A(\delta) & B(\delta) \\ c(\delta) & D(\delta) \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix} \tag{4.1.3}$$

in which δ may or may not be time-varying. Of particular interest will be those systems in which the systems matrices *affinely depend on* δ. This means that

$$\begin{aligned} A(\delta) &= A_0 + \delta_1 A_1 + \ldots + \delta_p A_p \\ B(\delta) &= B_0 + \delta_1 B_1 + \ldots + \delta_p B_p \\ C(\delta) &= C_0 + \delta_1 C_1 + \ldots + \delta_p C_p \\ D(\delta) &= D_0 + \delta_1 D_1 + \ldots + \delta_p D_p \end{aligned}$$

or, written in a more compact form,

$$S(\delta) = S_0 + \delta_1 S_1 + \ldots + \delta_p S_p$$

where

$$S(\delta) = \begin{bmatrix} A(\delta) & B(\delta) \\ C(\delta) & D(\delta) \end{bmatrix}$$

is the system matrix associated with (4.1.3). We call these models *affine parameter dependent models.*

4.1.2 Polytopic parameter dependent systems

In Introduction Chapter we introduced the notion of a convex combination of a finite set of points. This notion gets relevance in the context of dynamical systems if 'points' become systems. Consider a time-varying dynamical systems

$$\begin{aligned}\dot{x}(t) &= A(t)x(t) + B(t)w(t)\\ z(t) &= C(t)x(t) + D(t)w(t)\end{aligned}$$

with input w,output Z and state x. Suppose that its *systemmatrix*

$$S(t) := \begin{bmatrix} A(t) & B(t) \\ C(t) & D(t) \end{bmatrix}$$

is a time varying object which for any time instant $t \in \mathbb{R}$ can be written as a *convex combination* of the p system matrices $S_1, ..., S_p$. This means that there exist functions $\alpha_j : \mathbb{R} \to [0,1]$ such that for any time instant $t \in \mathbb{R}$ we have that

$$S(t) := \sum_{j=1}^{p} \alpha_j(t) S_j$$

where $\sum_{j=1}^{p} \alpha_j(t) = 1$ and

$$S_j := \begin{bmatrix} A_j & B_j \\ C_j & D_j \end{bmatrix}, j = 1, ..., p$$

are constant system matrices of equal dimension. In particular, this implies that the system matrices $S(t)$, $t \in \mathbb{R}$ belong to the convex hull of $S_1, ..., S_p$, i.e.,

$$S(t) \in conv(S_1, ..., S_p), t \in \mathbb{R}.$$

Such model are called *polytopic linear differential inclusions* and arise in a wide variety of modeling problems.

4.2 Time-invariant parametric uncertainty

An important issue in the design of control systems involves the question as to what extend the stability and performance of the controlled system is robust against perturbations and uncertainties in the parameters of the system. In this section we consider the linear time-invariant system defined by

$$\dot{x} = A(\delta)x \tag{4.2.1}$$

where the state matrix $A(.)$ is a *continuous* function of a real valued time-invariant parameter vector $\delta = col(\delta_1, ..., \delta_p)$ which we assume to be contained in an *uncertainty set* $\Delta \subseteq \mathbb{R}^p$. Let $X = \mathbb{R}^n$ be the state apace of this system. We will analyze the *robust stability* of the equilibrium point $x^* = 0$ of this system. Precisely, we address the question when the equilibrium point $x^* = 0$ of (4.2.1) is asymptotically stable in the sense of Definition (1.6.1) for all $\delta \in \Delta$.

Example 4.2.1 *As an example, let*

$$A(\delta) = \begin{bmatrix} -1 & 2\delta_1 & 2 \\ \delta_2 & -2 & 1 \\ 3 & -1 & \frac{\delta_3 - 10}{\delta_1 + 1} \end{bmatrix}$$

where δ_1, δ_2 and δ_3 are bounded as $\delta_1 \in [-0.5, 1], \delta_2 \in [-2, 1]$ and $\delta_3 \in [-0.5, 2]$. Then the uncertainty set Δ is polytopic and defined by

$$\Delta = col(\delta_1, \delta_2, \delta_3)|\delta_1 \in [-0.5, 1], \delta_2 \in [-2, 1], \delta_3 \in [-0.5, 2]$$

and Δ = conv (Δ_0) with

$$\Delta_0 = col(\delta_1, \delta_2, \delta_3)|\delta_1 \in [-0.5, 1], \delta_2 \epsilon [-2, 1], \delta_3 \in [-0.5, 2]$$

the set of vertices (or generators) of Δ.

In the case of time-invariant parametric uncertainties, the system $\dot{x} = A(\delta)x$ is asymptotically stable if and only if A(δ) is Hurwitz for all $\delta \in \Delta$. That is, if and only if the eigenvalues of A(δ) lie in the open left-half complex plane for all admissible perturbations $\delta \in \Delta$. Hence, using Proposition (1.6.6), the verification of robust stability amounts to checking whether

$$\rho(A(\delta)); = \max Re(\lambda) < 0 \ for \ all \ \delta \in \Delta.$$

If Δ is a continuum in $\mathbb{R}^p$ this means verifying an inequality at an infinite number of points. Even if Δ is a polytope, it will not suffice to check the inequality on the vertices of the uncertainty set only. Since $\rho(A(\delta))$ is not a convex or concave function of δ (except in a few trivial cases) it will be difficult to find its global maximum over Δ.

4.2.1 Quadratic stability with time-invariant uncertainties

We will apply Theorem (1.6.5) to infer the asymptotic stability of the equilibrium point $x^* = 0$ of (4.2.1). For this, we consider quadratic Lyapunov functions $V(x) = x^T K x$ where K is a real symmetric $n \times n$ matrix.

Definition 4.2.2 *(Quadratic Stability) The systems* (4.2.1) *is said to be* quadratically stable for perturbations $\delta \in \Delta$ *if there exists a matrix* $K = K^T > 0$ *such that*

$$A(\delta)^T K + KA(\delta) < 0 \;\; for \;\; all \;\; \delta \in \Delta. \tag{4.2.2}$$

The importance of this definition becomes apparent when considering the time-derivative V' of the composite function $V(x(t)) = x(t)^T Kx(t)$ along solutions of (4.2.1) as defined in (1.6.7). If $K > 0$ satisfies (4.2.2) then there exists an $\varepsilon > 0$ such that

$$A(\delta)^T K + KA(\delta) + \varepsilon K \leq 0 \;\; for \;\; all \;\; \delta \in \Delta.$$

Using the notation in (1.6.7), this means that

$$V'(x) + \varepsilon V(x) = \partial_x V(x) A(\delta)x + \varepsilon V(x) = x^T[A(\delta)^T K + KA(\delta) + \varepsilon K]x \leq 0.$$

for all $x \in \mathcal{X}$ and all $\delta \in \Delta$. But then the composite function $V^*(t) = V(x(t))$ with $x(\cdot)$ a solution of (4.2.1) satisfies

$$\dot{V}^*(t) + \varepsilon V^*(t) = V'(x(t)) + \varepsilon V(x(t)) = x(t)^T[A(\delta)^T K + KA(\delta) + \varepsilon k]x(t) \leq 0$$

for all $t \in \mathbb{R}$ and all $\delta \in \Delta$. If we integrate this expression over an interval $[t_0, t_1]$ it follows that V^* has exponential decay according to $V^*(t_1) \leq V^*(t_0)e^{-\varepsilon(t_1 - t_0)}$ for all $t_1 \geq t_0$ and all $\delta \in \Delta$. Now use that

$$\lambda_{min}(K)\|x\|^2 \leq x^T Kx \leq \lambda_{max}(K)\|x\|^2$$

to see that

$$\|x(t_1)\|^2 \leq \frac{1}{\lambda_{min}(K)} V(x(t_1)) \leq \frac{1}{\lambda_{min}(K)} V(x(t_0))e^{-\varepsilon(t_1 - t_0)} \leq \frac{\lambda_{max}(K)}{\lambda_{min}(K)} \|x(t_0)\|^2 e^{-\varepsilon(t_1 - t_0)}$$

That is,

$$\|x(t)\| \leq \| x(t_0) \| \sqrt{\frac{\lambda_{max}(K)}{\lambda_{min}(K)}} e^{-\frac{\epsilon}{2}(t-t_0)} \quad for \ all \ \delta \in \Delta$$

which shows that the origin is globally exponentially stable (and hence globally asymptotically stable) for all $\delta \in \Delta$.

4.2.2 Verifying quadratic stability for affine systems

It is important to note that verification of quadratic stability of the system places an infinite number of constraints on the symmetric matrix K in (4.2.2). It is the purpose of this section to make additional assumptions on the way the uncertainty enters the system, so as to convert the problem to find a positive definite K that satisfies (4.2.2) into a numerically tractable one.

Suppose that the state matrix $A(\delta)$ is an $affine$ function of the parameter vector δ. That is, suppose that there exist real matrices $A_0, ..., A_p$, all of dimension $n \times n$, such that

$$A(\delta) = A_0 + \delta_1 A_1 + ... + \delta_p A_p \tag{4.2.3}$$

with $\delta \in \Delta$. Then (4.2.1) is referred to as an *affine parameter dependent model.* In addition suppose that the uncertainty set Δ is convex and coincides with the convex hull of a set $\Delta_0 \subset \mathbb{R}^p$. With this structure on A and Δ we state the following result.

Proposition 4.2.3 *If $A(\cdot)$ is an affine function and $\Delta = conv(\Delta_0)$ with $\Delta_0 \subset \mathbb{R}^p$, then the system* (4.1.1) *is quadratically stable if and only if there exists $K = K^T > 0$ such that*

$$A(\delta)^T K + K A(\delta) < 0, \qquad for \ all \ \delta \in \Delta_0. \tag{4.2.4}$$

Example 4.2.4 *Following Up on Example* (4.2.1) *, the matrix $A(\delta)$ is not affine in δ, but by setting $\delta_4 = \frac{\delta_3 - 10}{\delta_1 + 1} + 12$ we obtain that*

$$A(\delta) = \begin{bmatrix} -1 & 2\delta_1 & 2 \\ \delta_2 & -2 & 1 \\ 3 & -1 & \delta_4 - 12 \end{bmatrix}, \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_4 \end{bmatrix} \in \Delta = [-0.5, 1] \times [-2, 1] \times [-9, 8].$$

Since $\Delta = conv(\Delta_0)$ with $\Delta_0 = \delta^1, ..., \delta^N$ consisting of the $N = 2^3 = 8$ corner points of the uncertainty set, the verification of the quadratic stability of (4.2.1) is a feasibility test of the (finite) system of **LMI's**

$$K > 0, \;\; A(\delta^j)^T K + KA(\delta^j) < 0, \;\; j = 1, ..., 8.$$

The test will not pass. By Proposition (4.2.3), the system is not quadratically stable for the given uncertainty set.

4.2.3 Parameter-dependent Lyapunov functions

The main disadvantage in searching for one quadratic Lyapunov function for a class of uncertain models is the conservatism of the test to prove stability of a class of models. Indeed, the last example shows that (4.2.1) may not be quadratically stable, but no conclusions can been drawn from this observation concerning the in-stability of the uncertain system. To reduce conservatism of the quadratic stability test we will consider quadratic Lyapunov functions for the system (4.2.1) Which are *parameter dependent*, i.e., Lyapunov functions $V : X \times \Delta \to \mathbb{R}$ of the form

$$V(x, \delta) := x^T K(\delta) x$$

Where $K(\delta)$ is a matrix valued function that is allowed to dependent on the uncertain parameter δ. A sufficient condition for robust asymptotic stability can be stated as follows.

Proposition 4.2.5 *Let the uncertainty set Δ be compact and suppose that $K(\delta)$ is continuously differentiable on Δ and satisfies*

$$K(\delta) > 0 \quad for\ all\ \delta \in \Delta. \tag{4.2.5}$$

$$A(\delta)^T K(\delta) + K(\delta)A(\delta) < 0 \quad for\ all\ \delta \in \Delta. \tag{4.2.6}$$

The search for matrix valued *functions* that satisfy the conditions of Proposition (4.2.5) is much more involved and virtually intractable from a computational point of view. We view therefore be interested in converting Proposition (4.2.5) to a numerically efficient scheme to compute parameter varying Lyapunov functions. For this, first consider Lyapunov functions that are *affine* in the parameter δ, i.e.,

$$K(\delta) = K_0 + \delta_1 K_1 + ... + \delta_p K_p$$

where $K_0, ..., K_p$ are real symmetric matrices of dimension $n \times n$ and $\delta = col(\delta_1, ..., \delta_p)$ is the time-invariant uncertainty vector. Clearly, with $K_1 = ... = K_p = 0$ we are back to the case of parameter *independent* quadratic Lyapunov functions as discussed in the previous subsection. The System (4.2.1) is called *affine quadratically stable* if there exists matrices $K_0, ..., K_p$ such that $K(\delta)$ satisfies the conditions (4.2.5) of Proposition (4.2.5)

As in the previous subsection,suppose that $A(\cdot)$ is affine and represented by (4.2.3) and suppose that Δ is convex with $\Delta = conv(\Delta_0)$ and Δ_0 a set of vertices of Δ. Then the expression

$$L(\delta) := A(\delta)^T K(\delta) + K(\delta)A(\delta)$$

in (4.2.5) is, in general, not affine in δ. As a consequence, for fixed $x \in \mathbb{R}^n$, the function $f_x : \Delta \to \mathbb{R}$ defined by

$$f_x(\delta) := x^T L(\delta) x \tag{4.2.7}$$

will not be convex so that the uniform bound $f_x(\cdot) < 0$ on Δ can not be verified by considering the vertices Δ_0 of Δ. Expanding $L(\delta)$ yields

$$\begin{aligned}
L(\delta) &= [A_0 + \sum_{j=1}^{p} \delta_j A_j]^T [K_0 + \sum_{j=1}^{p} \delta_j K_j] + [K_0 + \sum_{j=1}^{p} \delta_j K_j][A_0 + \sum_{j=1}^{p} \delta_j A_j] \\
&= A_0^T K_0 + K_0 A_0 + \sum_{j=1}^{p} \delta_j [A_j^T K_0 + K_0 A_j + A_0^T K_j + K_j A_0] \\
&+ \sum_{j=1}^{p} \sum_{i=1}^{j-1} \delta_i \delta_j [A_i^T K_j + K_j A_i + A_j^T K_i + K_i A_j] + \sum_{j=1}^{p} \delta_j^2 [A_j^T k_j + K_j A_j] \\
&= \sum_{j=0}^{p} \sum_{j=0}^{p} \delta_i \delta_j [A_i^T K_j + K_j A_i]
\end{aligned}$$

where, for the latter compact notation, we introduced $\delta_0 = 1$. Consequently, (4.2.7) takes the form

$$f_x(\delta) = c_0 + \sum_{j=1}^{p} \delta_j c_j + \sum_{j=1}^{p} \sum_{i=1}^{j-1} \delta_i \delta_j c_{ij} + \sum_{j=1}^{p} \delta_j^2 d_j$$

where c_0, c_j, c_{ij} and d_j are constants. Motivated to convert (4.2.5) into a numerically verifiable test, we will be interested in finding conditions for the implication

$$\{f_x(\delta) < 0 \ \ for \ \ all \ \ \delta \in \Delta_0\} \implies \{f_x(\delta) < 0 \ \ for \ \ all \ \ \delta \in \Delta\} \tag{4.2.8}$$

to hold. Consider the following sets

$$\Delta = \{\delta \in \mathbb{R}^p | \delta_k \in [\underline{\delta}_k, \bar{\delta}_k]\}, \ \ \Delta_0 = \{\delta \in \mathbb{R}^p | \delta_k \in \{\underline{\delta}_k, \bar{\delta}_k\}\} \tag{4.2.9}$$

where $\underline{\delta}_k \leq \bar{\delta}_k$. Then $\Delta = conv(\Delta_0)$ and it is easily seen that a sufficient condition for the implication (4.2.8) for the uncertainty sets (4.2.8) is that $f_x(\delta_1, ..., \delta_j, ...\delta_p)$ is *partially convex*, that is convex in each of its arguments δ_j, $j = 1, ..., p$. Since f_x is a twice differentiable function, this is the case when

$$d_j = \frac{\partial^2}{\partial \delta_j^2} f_x(\delta) = x^T [A_j^T K_j + K_j A_j] x \geq 0$$

for $j = 1, ..., p$. Since x is arbitrary,we therefore obtain that

$$A_j^T K_j + K_j A_j \geq 0, \;\; j = 1, ..., p$$

is a sufficient condition for (4.2.8) to hold on the uncertainty sets (4.2.9). This leads to the following main result.

Theorem 4.2.6 *If* $A(\cdot)$ *is an affine function described by* (4.2.3) *and* $\Delta = cov(\Delta_0)$ *assume the form* (4.2.9), *then the systems* (4.2.1) *is affine quadratically stable if there exist real matrices* $K_0, ..., K_p$ *such that* $K(\delta) = K_0 + \sum_{j=1}^{p} \delta_j K_j$ *satisfies*

$$A(\delta)^T K(\delta) + K(\delta) A(\delta) < 0 \quad for\ all\ \delta \in \Delta_0 \tag{4.2.10}$$

$$K(\delta) > 0 \quad for\ all\ \delta \in \Delta_0 \tag{4.2.11}$$

$$A_j^T K_j + K_j A_j \geq 0 \quad for\ j = 1, ..., p. \tag{4.2.12}$$

In that case, the parameter varying function satisfies the condition (4.2.5) and $V(x, \delta) := x^T K(\delta) x$ is a quadratic parameter-dependent Lyapunov function of the system.

Theorem (4.2.6) reduces the problem to verify affine quadratic stability of the system (4.2.1) to a feasibility problem of a (finite) set of linear matrix inequalities. The letter problem is a numerically tractable one.

4.3 Time-varying parametric uncertainty in autonomous systems

Robust stability against time-varying perturbations is generally a more demanding requirement then robust stability against time-invariant parameter uncertainties. In this section we consider the question of robust stability for the system

$$\dot{x}(t) = A(\delta(t))x(t) \tag{4.3.1}$$

where the values of the time-varying parameter vector $\delta(t)$ belong to the uncertainty set $\Delta \subset \mathbb{R}^p$ for all time $t \in \mathbb{R}$. In this section we assess the *robust stability* of the fixed point $x^* = 0$. It is important to remark that, unlike the case with time-invariant uncertainties, robust stability of the origin of the time-varying system (4.3.1) is *not equivalent* to the condition that the (time-varying) eigenvalues $\lambda(A(\delta(t)))$ belong to the stability region $\mathbb{C}^-$ for all admissible perturbations $\delta(t) \in \Delta$.

Proposition 4.3.1 *The uncertain system system* (4.3.1) *with time-varying uncertainties* $\delta(\cdot) \in \Delta$ *is asymptotically stable if there exists a matrix* $K = K^T > 0$ *such that* (4.2.2) *holds.*

The inequality (4.2.2) is a *time-invariant* condition that needs to be verified for all points in the uncertainty set Δ so as to conclude asymptotic stability of the *time-variant* uncertainties that occur in (4.3.1). Since Proposition (4.3.1) is obtained as a special case of Theorem (4.3.2) below, we defer its proof.

An important observation related to Proposition (4.3.1) is that the existence of a real symmetric matrix K with $K > 0$ and $A(\delta)^T K + KA(\delta) < 0$ for all $\delta \in \Delta$ not only yields quadratic stability of the system (4.2.1) with $\delta \in \Delta$ but also the asymptotic stability of (4.3.1) with $\delta(t) \in \Delta$. Hence, the existence of such K implies that *arbitrary fast variations* in the time-varying parameter vector $\delta(\cdot)$ may occur so as to guarantee asymptotic stability of (4.3.1). If additional a *prior* information on the time-varying parameters is known, the sufficient condition of Proposition (4.3.1) may become conservation and we therefore may like to resort to different techniques to incorporate information about the parameter trajectories $\delta(\cdot)$.

One way to do this is to assume that the trajectories $\delta(\cdot)$ are continuously differentiable while

$$\delta(t) \in \Delta, \quad \dot{\delta}(t) \in \Lambda \ \ for\ all\ time\ t \in \mathbb{R}. \tag{4.3.2}$$

Here, Δ and Λ are subsets of $\mathbb{R}^p$ and we will assume that these sets are compact. We will therefore assume that not only the *values* but also the *rates* of the parameter trajectories are constrained.

4.3.1 Quadratic stability with time-varying uncertainty

A central result for achieving robust stability against this class of uncertainties is given in the following theorem.

Theorem 4.3.2 *Suppose that the function $K : \Delta \to \mathbb{S}^n$ is continuously differentiable on Δ and satisfies*

$$\begin{aligned} K(\delta) &> 0 \quad for\ all\ \delta \in \Delta & (4.3.3)\\ \sum_{k=1}^{p} \delta_k K(\delta) v_k + A(\delta)^T K(\delta) + K(\delta) A(\delta) &< 0 \quad for\ all\ \delta \in \Delta\ and\ v \in \Lambda. & (4.3.4)\end{aligned}$$

Then the origin of the system (4.3.1) *is asymptotically stable against all time-varying uncertainties $\delta : \mathbb{R} \to R^p$ that satisfy* (4.3.2). *Moreover, in that case the function $V(x, \delta) := x^T K(\delta)$ qualifies as quadratic parameter depending Lyapunov function for the system.*

Theorem (4.3.2) involves a search for matrix functions satisfying the inequalities (4.3.3) to guarantee robust asymptotic stability. Note that the result is a sufficient algebraic test only that provides a quadratic parameter dependent Lyapunov function, when it exists. The result is not easy to apply or verify by a computer program as it involves(in general) an infinite number of conditions on the inequalities (4.3.3). We will therefore focus on a number of special cases that convert Theorem (4.3.2) in a feasible algorithm.

For this, first consider the case where the parameters are time-invariant. This is equivalent to saying that $\Lambda = \{0\}$. The conditions (4.3.3) then coincide with (4.2.5) and we therefore obtain Proposition (4.2.5) as a special case.

If we assume arbitrary fast time-variations in $\delta(t)$ then we consider rate constraints of the form $\Lambda = [-r, r]^p$ with $r \to \infty$. For (4.3.3) to hold for any v_k with $\mid v_k \mid > r$ and $r \to \infty$ it is immediate that $\partial_k K(\delta)$ needs to vanish for $k = 1, ..., p$ and all $\delta \in \Delta$. Consequently, in this case K can not depend on δ and the Theorem (4.3.2) reduces to Proposition (4.3.1). In particular, this argument proves Proposition (4.3.1) as a special case.

4.3.2 Verifying quadratic stability for affine systems

In this section we will assume that $A(\cdot)$ in (4.3.1) is an affine function represented by (4.2.3). The uncertainty sets Δ and Λ are assumed to be convex sets defined by 'boxes'

$$\Delta = \{\delta \in \mathbb{R}^p | \delta_k \in [\underline{\delta}_k, \bar{\delta}_k]\}, \quad \Lambda = \{\lambda \in \mathbb{R}^p | \lambda_k \in [\underline{\lambda}_k, \bar{\lambda}_k]\}$$

This means that the uncertainty regions are the convex hulls of the sets

$$\Delta_0 = \{\delta \in \mathbb{R}^p | \delta_k \in [\underline{\delta}_k, \bar{\delta}_k]\}, \quad \Lambda_0 = \{\lambda \in \mathbb{R}^p | \lambda_k \in [\underline{\lambda}_k, \bar{\lambda}_k]\}$$

In addition, we will restrict the search of a parameter dependent $K(\delta)$ to the class of *affine function* $K(\delta)$ represented by

$$K(\delta) = K_0 + \delta_1 K_1 + \ldots + \delta_p K_p$$

where $K_j \in \mathbb{S}^n, j = 1, \ldots, p$ is symmetric.For this class of parameter dependent functions we have that $\partial_k K(\delta) = K_k$ so that (4.3.3) reads

$$\sum_{k=1}^{p} \partial_k K(\delta) v_k \;\; + A(\delta)^T K(\delta) + K(\delta) A(\delta) = \sum_{k=1}^{p} K_k v_k + \sum_{v=0}^{p} \sum_{\mu=1}^{p} \delta_v \delta_\mu (A_v K_\mu + K_\mu A_v).$$

Here, we set $\delta_0 = 1$ to simplify notation. The Latter expression is

- affine in $K_1, \ldots, K_p$
- affine in $\nu_1, \ldots, \nu_p$
- quadratic in $\delta_1, \ldots, \delta_p$ due to the mixture of constant, linear and quadratic terms.

Again, similar to (4.2.7), introduce for arbitrary $x \in \mathcal{X}$ the function $f_x : \Delta \times \Lambda \to \mathbb{R}$ defined by

$$f_x(\delta, v) := x^T [\sum_{k=1}^{p} K_k \nu_k + \sum_{v=0}^{p} \sum_{\mu=0}^{p} \delta_v \delta_\mu (A_v K_\mu + K_\mu A_v)] x.$$

A sufficient condition for the implication

$$\{f_x(\delta, v) < 0 \;\; for \;\; all \;\; (\delta, v) \in \Delta_0 \times \Lambda_0\} \Rightarrow \{f_x(\delta, v) < 0 \;\; for \;\; all \;\; (\delta, v) \in \Delta \times \Lambda\}$$

is that f_x is *partially convex* in each of its arguments $\delta_j, j = 1, \ldots, p$. As before, this condition translates in the requirement that

$$A_j^T K_j + K_j A_j \geq 0, \;\; j = 1, \ldots, p.$$

which brings us the following main result.

Theorem 4.3.3 *If there exist* $K_0, ..., K_p$ *with* $A_j^T K_j + K_j A_j \geq 0, j = 1, ..., p,$ *and*

$$\sum\ {}_{k=0}^{p} K_k \delta_k > 0$$
$$\sum\ {}_{k=1}^{p} K_k v_k + \sum k = 0^p \sum t = 0^p \delta_k \delta_\ell (A_k K_\ell + K_\ell A_k) < 0$$

for all $\delta \in \Delta_0$ *and* $v \in \Lambda_0$, *then the origin of the system* (4.3.1) *is robustly asymptotically stable against all time-varying uncertainties that satisfy* (4.3.2). *Moreover, in that case,* $V(x, \delta) := x^T K(\delta) x$ *defines a quadratic Lyapunov function for the system where* $K(\delta) = K_0 + \sum_{k=1}^{p} \delta_k K_k$.

Theorem (4.3.3) provides an LMI feasibility test to verify robust asymptotic stability against uncertainty described by (4.3.2).

It is interesting to compare the numerical complexity of the conditions of Theorem (4.2.6) with the conditions mentioned in Theorem (4.3.3). If the uncertainty vector δ is p-dimensional then the vertex set Δ_0 has dimension 2^p so that the verification of conditions (4.3.2) requires a feasibility test of

$$2^p + 2^p + p$$

linear matrix inequalities.In this case,also the vertex set Λ_0 has dimension 2^k which implies that the condition of Theorem (4.3.3) require a feasibility test of

$$2^{2p} + 2^p + p \quad = \quad 4^p + 2^p + p$$

linear matrix inequalities.

Appendix A

Linear algebra, vector space and norm

Norm of signal and systems is an essential definition for robust control. One way to measure the performance of the system is the size of certain signals of interest (Doyle, Francis and Tannenbaum 1992). To introduce norm we first review some background; see also (Zhou, Doyle and Glover 1996).

A.1 Vector space

Definition 1 *Consider a closed set of vectors $\mathcal{V}$ and the field $\mathcal{F}$ (it is usually $\mathbb{R}$ or $\mathbb{C}$) with the de.ned vector addition " + " and scalar multiplication " $\cdot$ ". $\mathcal{V}$ is a vector space (also called a linear space), if the following conditions hold for all elements $x, y, z \in \mathcal{V}$ and any scalars $a, \beta \in F$:*

1 *Commutativity:* $x + y = y + x$.

2 *Associativity of vector addition:* $(x + y) + z = x + (y + z)$.

3 Additive identity: $\forall x, 0 + x = x + 0 = x$.

4 *Existence of additive inverse:* $\forall x$, *there exist* $a - x$ *such that* $x + (-x) = 0$.

5 *Associativity of scalar multiplication:* $\alpha(\beta x) = (\alpha\beta x)$.

6 *Distributivity of scalar sums:* $(\alpha + \beta)x = \alpha x + \beta x$.

7 *Distributivity of vector sums:* $\alpha(x + y) = \alpha x + \alpha y$.

8 *Scalar multiplication identity:* $1x = x$.

Example 1 Some examples of vector spaces are: the set of vectors in $\mathbb{R}^n$, the set of matrices in $R^{m\times n}$, the set of all polynomials, all stable rational and proper transfer functions, all unstable rational and proper transfer functions.

Remark 1 A subset of a vector space is also a vector space if it is closed and it contains zero.

Remark 2 The set of all functions $x(t)$ defined on the interval I (for instance $(-\infty, \infty)$ or $[0, \infty)$) with the property $(f_I |x(\dot{x})t|^p dt)^{\frac{1}{p}} < \infty$ is a vector space and denoted as $\mathcal{L}_p(I)$

Definition 2 The vectors $v1, v2, ..., v_m \in \mathcal{V}$ are linearly dependent if there exists the constants $a1, a2, \cdots, am$ that are not all zero such that $a_1v_1 + a_2v_2 + ... + a_mv_m = \mathbf{0}$. If $a_1v_1 + a_2v_2 + ... + a_mv_m = 0$ implies that $a_1 = a_2 = \cdots = a_m = 0$ then the vectors are said to be linearly independent.

Remark 3 A set of n vectors $v_1, v_2, \cdots, v_n \in \mathbb{R}^n$ is linearly independent if and only if the rank of the matrix $M = \begin{bmatrix} v_1 & v_2 & \cdots & v_n \end{bmatrix}$ is n or equivalently $det(M) \neq 0$.

A vector space can be identified by its basis which are linearly independent vectors span the whole space. The trivial basis in $\mathbb{R}^n$ is the set of unit vectors,

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix}, \cdots, e_n = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix}$$

Definition 3 The dimension of a vector space is the number of vectors in the basis. It is in fact the smallest number of independent vectors that can be found in a vector space.
Example 2 The dimension of the vector space $\mathbb{R}^n$ is n. The dimension of the vector space $\mathbb{R}^{m\times n}$ is mn. The dimension of the vector space of all stable, rational and proper transfer functions is infinity.

A.2 Singular Value Decomposition (SVD)

A given nonsquare matrix $\mathbf{A} \in \mathbb{C}^{m\times n}$ does not have an eigenvalue decomposition. However, $\mathbf{A}$ can be written using a so-called Singular Value Decomposition (or SVD) in a similar form $A = \mathbf{USV}^H$. The notation $(\cdot)^H$ denotes hermitian of a matrix which is conjugate of the transpose. This is a generalization of transpose to matrices with complex elements.

Here, U$\in \mathbb{C}^{m\times m}$,$\mathbf{V} \in \mathbb{C}^{n\times n}$ have orthogonal columns so that $\mathbf{U^H U} = \mathbf{I}_m, \mathbf{V}^H\mathbf{V} = \mathbf{I}_n$. Such Matrices are called Unitary matrices (or orthogonal if the matrices are real). The matrix $S \in \mathbb{C}^{m\times n}$ has the following structure,

$$\mathtt{for}\, m > n, S = \begin{bmatrix} \Sigma \\ 0 \end{bmatrix}$$

$$\mathtt{for}\, m < n, S = \begin{bmatrix} \Sigma & 0 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_{min(m,n)} \end{bmatrix}$$

The values of $\sigma_i s$ are usually in decreasing order,$\sigma_1 \geq \sigma_2 \geq \cdots$ and are called *Singular Values* of the matrix **A**. The largest and smallest singular values are denoted as $\overline{\sigma}$ and $\underline{\sigma}$, respectively.

Remark 4 If **A** is hermitian ($\mathbf{A} = \mathbf{A}^H$) with positive eigenvalues, then the SVD is the same as eigenvalue decomposition. In this case $\mathbf{U} = \mathbf{V}$ and the matrix posseses a set of orthogonal eigenvectors.

Theorem 1 *Consider that* $\mathbf{U} = \begin{bmatrix} u_1 & \cdots & u_m \end{bmatrix}, \mathbf{V} = \begin{bmatrix} v_1 & \cdots & v_m \end{bmatrix}$ *and rank of the matrix* **A** *is p. Then the following results can be proved:*

1 $\sigma_i = 0$ *for* $i = p+1, \cdots, min(m,n)$.

2 *The singular values of the matrix* **A** *are square root of the nonzero eigenvalues of* $\mathbf{A}\mathbf{A}^H$ *and* $\mathbf{A}^H\mathbf{A}$ *with the corresponding eigenvectors* u_i *and* v_i*, respectively.*

3 *The vectors* $v_{p+1}, \cdots, v_n$ *form a basis for the null space of* **A**.

4 *The vectors* $u_1, \cdots, u_p$ *form a basis for the image of* **A**.

A.3 Semi-definite matrices

Definition 4 *Consider the symmetric matrix $A \in \mathbb{C}^{n\times n}$. This matrix is called positive semidefinite (negative-semidefinite) denoted as $A \succeq 0(D\,(A\,[\preceq])\,0)$ if,*

$$\forall X \in \mathbb{C}^n : X^T AX \geq 0(\leq 0).$$

Definition 5 *Consider the symmetric matrix $A \in \mathbb{C}^{n\times n}$. This matrix is called positive definite (negative-definite) denoted as $A \succ 0(A \prec 0)$ if,*

$$\forall X \in \mathbb{C} : X^T AX > 0(< 0).$$

It is clear that a definite-matrix is nonsingular. Positive definiteness is the generalization of positive numbers to the matrix theory. It is also clear that $\mathbf{AA}^T \succeq 0$ and $-\mathbf{AA}^T \preceq 0$ for every matrix $\mathbf{A} \in \mathbb{C}^{m\times n}$. If A is full row rank then $\mathbf{AA}^T \succ 0$ and $\mathbf{AA}^T \prec 0$.

Theorem 2 *Consider the matrix $\mathbf{A} \succeq 0$. There exists a matrix $\mathbf{B}$ such that $\mathbf{BB}^T = \mathbf{A}$. The matrix $\mathbf{B}$ is called the square root of $\mathbf{A}$ denoted as $\mathbf{A}^{\frac{1}{2}}$.*

The square root of a positive-definite matrix can be calculated using SVD of the matrix $\mathbf{A}$ as, $\mathbf{A}^{\frac{1}{2}} = \mathbf{U}\Sigma^{\frac{1}{2}}\mathbf{U}^T$.

A positive definite matrix $\mathbf{A} \succ 0$ has the following properties:

- There exists a matrix $\mathbf{A}^{\frac{1}{2}}$ such that $\mathbf{A}^{\frac{1}{2}}\mathbf{A}^{\frac{1}{2}} = \mathbf{A}$.
- All eigenvalues are positive.
- All diagonal elements are positive.

- $\det(\mathbf{A}) > 0$.

- $\mathbf{A}^{-1} \succ 0$.

Definition 6 *Consider the symmetric matrix* $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}$. $\mathbf{A}, \mathbf{B} \in \mathbb{C}^{n\times n}, \mathbf{A} \prec \mathbf{B}$ *if and only if* $\mathbf{A} - \mathbf{B} \prec 0$. *The definition is generalized to other variations in similar way.*

A.4 Normed space

Norm of a vector is a concept to identify the length or size of a vector. It is an important concept as it is a tool to measure the distance of elements in a vector space.

Definition 7 *The real-valued function* $\|.\|$ *defined on the vector space* $\mathcal{V}$ *over* $\mathbb{C}$ *is said to be a norm on* $\mathcal{V}$, *if it satisfies the following properties for any* $x, y \in \mathcal{V}, \alpha \in \mathbb{R}$

i. Positivity: $\|x\| \geq 0$

ii. Positive-definiteness: $\|x\| = 0$, *if and only if* $x = 0$

iii. Homogeneity: $\|\alpha x\| = |\alpha|\|x\|$.

iv. Triangle inequality: $\|x + y\| \leq \|x\|\|y\|$.

Remark 5 The most common norm used in vector space $\mathbf{C}^n$ is the so called p-norm defined as,

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p \right)^p, for 1 \leq p \leq \infty.$$

The Euclidean norm or two-norm is defined for $p = 2$ as $\| \cdot \|_2$ which is the length of the vector and is usually denoted as $\| \cdot \|$.

Remark 6 The most common norm used for matrices in the vector space C^{mn} is the so called induced norm defined based on a vector norm as,

$$\mathbf{A} = \sup_{x \neq 0} \frac{\|\mathbf{A}x\|}{x}.$$

Some special cases for p-matrix norm are:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$$\|\mathbf{A}\|_1 = \max_{1 \leq j \leq n} \sum_{i=l}^{m} |a_{ij}|, \texttt{column sum}$$

$$\|\mathbf{A}\|_2 = \sqrt{\lambda_{max}(\mathbf{A}^H \mathbf{A})} = \max_i \sigma^i = \overline{\sigma}$$

$$\|\mathbf{A}\|_\infty = \max_{1 \leq i \leq m} \sum_{j=1}^{n} |a_{ij}|, \texttt{row sum}$$