Lecture Notes DISC Course on

Linear Matrix Inequalities in Control

Carsten Scherer and Siep Weiland

Preface

Objectives

In recent years linear matrix inequalities (LMI's) have emerged as a powerful tool to approach control problems that appear hard if not impossible to solve in an analytic fashion. Although the history of LMI's goes back to the fourties with a major emphasis of their role in control in the sixties (Kalman, Yakubovich, Popov, Willems), only recently powerful numerical interior point techniques have been developed to solve LMI's in a practically efficient manner (Nesterov, Nemirovskii 1994). Several Matlab software packages are available that allow a simple coding of general LMI problems and of those that arise in typical control problems (LMI Control Toolbox, LMI-tool).

Boosted by the availability of fast LMI solvers, research in robust control has experienced a paradigm shift – instead of arriving at an analytical solution, the intention is to reformulate a given problem to verifying whether an LMI is solvable or to optimizing functionals over LMI constraints.

The main emphasis of the course is

- to reveal the basic principles of formulating desired properties of a control system in the form of LMI's,

- to demonstrate the techniques to reduce the corresponding controller synthesis problem to an LMI problem,

- to get familiar with the use of software packages for performance analysis and controller synthesis using LMI tools.

The power of this approach is illustrated by several fundamental robustness and performance problems in analysis and design of linear and certain nonlinear control systems.

Topics

The lecture notes for this course have been written for a course of eight weeks on the subject. Within the DISC graduate program, two class hours are taught once per week during a period of eight weeks. The topics covered for this course are the following.

1. Examples. Facts from convex analysis. Interior point methods in convex programming and their efficiency. Linear Matrix Inequalities: History. The three basic problems and their solution with LMI-Lab.

2. Lyapunov functions for invariance, stability, performance, robust performance. Quadratic stability and performance. Considered criteria: Dissipativity, Integral quadratic constraints, H_2-norm, H_∞-norm, upper bound of peak-to-peak norm. LMI stability regions.

3. Frequency domain techniques for the robustness analysis of a control system. Integral Quadratic Constraints. Multipliers. Relations to classical tests and to μ-theory.

4. A general technique to proceed from LMI analysis to LMI synthesis. State-feedback and output-feedback synthesis algorithms for robust stability, nominal performance and robust performance using general scalings.

5. Mixed control problems. Multi-model control problems. Lyapunov shaping technique.

6. Extension to linear parametrically varying systems. Gain-scheduling. Examples of occurrence. Solution of design problem with scaling and gridding techniques.

7. Extension to certain nonlinear analysis and design problems.

Material

The notes for this course will be made available on-line on the internet site

```
http://www.er.ele.tue.nl/SWeiland/lmi99.htm
```

Apart from the notes, the main reference material for the course will be

[1] S. Boyd, L. El Ghaoui, E. Feron and V. Balakrishnan
    Linear Matrix Inequalities in System and Control Theory, SIAM studies in Applied Mathematics, Philadelphia, 1994.

[2] Gahinet et al., *LMI-lab* Matlab Toolbox for Control Analysis and Design.

Chapter 1

Convex optimization and linear matrix inequalities

1.1 Introduction

Optimization questions and decision making processes are abundant in daily life and invariably involve the selection of the best decision from a number of options or a set of candidate decisions. Many examples of this theme can be found in technical sciences such as electrical, mechanical and chemical engineering, in architecture and in economics, but also in the social sciences, in biological and ecological processes and organizational questions. For example, important economical benefits can be realized by making proper decisions in production processes: the waste of resources has to be minimized, flexibility of production methods has to be maximized, the supply of products has to be adapted to demand, etc. Due to increasing requirements on the safety and flexibility of production processes, environmental measures and economic trade agreements there is a constant need for a further optimization and improvement of production processes.

Casting an optimization problem in mathematics involves the specification of the candidate decisions and, most importantly, the formalization of the concept of *best* or *optimal decision*. If the (finite or infinite) set of candidate decisions is denoted by $\mathcal{S}$, then one approach to quantify the performance of a decision $x \in \mathcal{S}$ is to express its value in terms of a single real quantity $f(x)$ where f is some real valued function $f : \mathcal{S} \to \mathbb{R}$. The value of decision $x \in \mathcal{S}$ is then given by $f(x)$. Depending on the interpretation of f, we may wish to minimize or maximize f over all possible candidates in $\mathcal{S}$. An optimal decision is then simply an element of $\mathcal{S}$ that minimizes or maximizes f over all possible alternatives.

The optimization problem to *minimize* the criterion f over $\mathcal{S}$ involves various specific ques-

tions:

1. How to determine the optimal value (or optimal performance)

$$V_{\text{opt}} := \inf_{x \in \mathcal{S}} f(x) = \inf\{f(x) \mid x \in \mathcal{S}\}$$

2. How to determine an almost optimal solution, i.e., for arbitrary $\epsilon > 0$, how to determine $x_\epsilon \in \mathcal{S}$ such that

$$f(x_\epsilon) \leqslant V_{\text{opt}} + \epsilon.$$

3. Does there exist an optimal solution $x_{\text{opt}} \in \mathcal{S}$ such that $f(x_{\text{opt}}) = V_{\text{opt}}$?

4. If such an optimal solution x_{opt} exists, how can it be computed?

5. Finally, is the optimal solution x_{opt} unique?

1.2   Facts from convex analysis

In view of the optimization problems just formulated, we are interested in finding conditions for optimal solutions to exist. It is therefore natural to resort to a branch of analysis which provides such conditions: convex analysis. The results and definitions in this subsection are mainly basic, but they have important applications as we will see later.

We start with summarizing some definitions and elementary properties from linear algebra and functional analysis. We assume the reader to be familiar with the basic concepts of vector spaces and normed linear spaces.

1.2.1   Convexity

Definition 1.1 (Continuity) A function f mapping a normed space $\mathcal{S}$ into a normed space $\mathcal{T}$ is *continuous at* $x_0 \in \mathcal{S}$ if for every $\epsilon > 0$ there exist $\delta > 0$ such that $\|x - x_0\| < \delta$ implies that $\|f(x) - f(x_0)\| < \epsilon$. The function f is called *continuous* if it is continuous at each $x_0 \in \mathcal{S}$.

Obviously, continuity depends on the definition of the norm in the normed spaces $\mathcal{S}$ and $\mathcal{T}$. We remark that a function $f : \mathcal{S} \to \mathcal{T}$ is continuous at $x_0 \in \mathcal{S}$ if and only if for every sequence $\{x_n\}_{n=1}^\infty$, $x_n \in \mathcal{S}$, which converges to x_0 as $n \to \infty$, there holds that $f(x_n) \to f(x_0)$.

Definition 1.2 (Compactness) A set $\mathcal{S}$ in a normed linear space $\mathcal{X}$ is called *compact* if for every sequence $\{x_n\}_{n=1}^\infty$ in $\mathcal{S}$ there exists a subsequence $\{x_{n_m}\}_{m=1}^\infty$ which converges to an element $x_0 \in \mathcal{S}$.

If the normed linear space $\mathcal{X}$ is finite dimensional then compactness has an equivalent characterization as follows.

Proposition 1.3 *If $\mathcal{X}$ is finite dimensional then $\mathcal{S} \subset \mathcal{X}$ is compact if and only if $\mathcal{S}$ is closed and bounded[1].*

The well-known Weierstrass theorem provides a useful tool to determine whether an optimization problem admits a solution. It provides an answer to the third question raised in the previous subsection for special sets $\mathcal{S}$ and special performance functions f.

Proposition 1.4 (Weierstrass) *If $f : \mathcal{S} \to \mathbb{R}$ is a continuous function defined on a compact subset $\mathcal{S}$ of a normed linear space $\mathcal{X}$, then there exists $x_{\min}, x_{\max} \in \mathcal{S}$ such that*

$$\inf_{x \in \mathcal{S}} f(x) = f(x_{\min}) \leqslant f(x) \leqslant f(x_{\max}) = \sup_{x \in \mathcal{S}} f(x)$$

for all $x \in \mathcal{S}$.

Proof. Define $V_{\min} := \inf_{x \in \mathcal{S}} f(x)$. Then there exists a sequence $\{x_n\}_{n=1}^{\infty}$ in $\mathcal{S}$ such that $f(x_n) \to V_{\min}$ as $n \to \infty$. As $\mathcal{S}$ is compact, there must exist a subsequence $\{x_{n_m}\}_{m=1}^{\infty}$ of $\{x_n\}$ which converges to an element $x_{\min} \in \mathcal{S}$. Then $f(x_{n_m}) \to V_{\min}$ and the continuity of f implies that $f(x_{\min}) \leqslant \lim_{n_m \to \infty} f(x_{n_m}) = V_{\min}$. By definition of $V_{\min}$, it then follows that $f(x_{\min}) = V_{\min}$. The proof of the existence of a maximizing element is similar. ∎

Note that proposition 1.4 does not give a constructive method to find the extremizing solutions $x_{\min}$ and $x_{\max}$. It only guarantees the existence of these elements for continuous functions defined on compact sets. For many optimization problems these conditions (continuity and compactness) turn out to be overly restrictive. We will therefore resort to convex sets.

Definition 1.5 (Convex sets) A set $\mathcal{S}$ in a linear vector space is said to be *convex* if

$$\{x_1, x_2 \in \mathcal{S}\} \implies \{x := \alpha x_1 + (1 - \alpha) x_2 \in \mathcal{S} \text{ for all } \alpha \in (0, 1)\}.$$

In geometric terms, this states that for any two points of a convex set also the line segment connecting these two points belongs to the set. In general, the *empty set* is considered to be convex. The point $\alpha x_1 + (1 - \alpha) x_2$ with $\alpha \in (0, 1)$ is called a *convex combination* of the two points x_1 and x_2. More generally, convex combinations are defined for any finite set of points as follows.

Definition 1.6 (Convex combinations) Let $\mathcal{S}$ be a subset of a normed vector space and let $x_1, \ldots, x_n \in \mathcal{S}$. If $\alpha_1, \ldots, \alpha_n$, is a set of non-negative real numbers with $\sum_{i=1}^{n} \alpha_i = 1$ then

$$x := \sum_{i=1}^{n} \alpha_i x_i$$

---

[1] A set is $\mathcal{S}$ is bounded if there exists a number B such that for all $x \in \mathcal{S}$, $\|x\| \leqslant B$.

is called a *convex combination* of $x_1, \ldots, x_n$.

If $x_1, \ldots x_n \in \mathcal{S}$, then it is easy to see that the set of all convex combinations of $x_1, \ldots, x_n$ is itself convex, i.e.,

$$\mathcal{C} := \{x \mid x \text{ is a convex combination of } x_1, \ldots, x_n\}$$

is convex. We next define the notion of interior points and closure points of sets.

Definition 1.7 (Interior points) Let $\mathcal{S}$ be a subset of a normed space $\mathcal{X}$. The point $x \in \mathcal{S}$ is called an *interior point* of $\mathcal{S}$ if there exists an $\epsilon > 0$ such that all points $y \in \mathcal{X}$ with $\|x - y\| < \epsilon$ also belong the $\mathcal{S}$. The *interior* of $\mathcal{S}$ is the collection of all interior points of $\mathcal{S}$.

Definition 1.8 (Closure points) Let $\mathcal{S}$ be a subset of a normed space $\mathcal{X}$. The point $x \in \mathcal{X}$ is called a *closure point* of $\mathcal{S}$ if, for all $\epsilon > 0$, there exists a point $y \in \mathcal{S}$ with $\|x - y\| < \epsilon$. The *closure* of $\mathcal{S}$ is the collection of all closure points of $\mathcal{S}$. $\mathcal{S}$ is said to be *closed* if it is equal to its closure.

We summarize some elementary properties pertaining to convex sets in the following proposition.

Proposition 1.9 *Let $\mathcal{S}$ and $\mathcal{T}$ be convex sets in a normed vector space $\mathcal{X}$. Then*

1. *the set $\alpha \mathcal{S} := \{x \mid x = \alpha s, s \in \mathcal{S}\}$ is convex for any scalar α.*

2. *the sum $\mathcal{S} + \mathcal{T} := \{x \mid x = s + t, s \in \mathcal{S}, t \in \mathcal{T}\}$ is convex.*

3. *the closure and the interior of $\mathcal{S}$ (and $\mathcal{T}$) are convex.*

4. *the intersection $\mathcal{S} \cap \mathcal{T} := \{x \mid x \in \mathcal{S} \text{ and } x \in \mathcal{T}\}$ is convex.*

The last property actually holds for the intersection of an arbitrary collection of convex sets, i.e, if $\mathcal{S}_\alpha$, $\alpha \in A$ is a family of convex sets then $\cap_{\alpha \in A} \mathcal{S}_\alpha$ is convex. This property turns out to be useful when we wish to consider the smallest convex set that contains a given set.

Definition 1.10 (Convex hull) The *convex hull* $\mathrm{co}(\mathcal{S})$ of a set $\mathcal{S}$ is the intersection of all convex sets containing $\mathcal{S}$.

Convex hulls have the following property.

Proposition 1.11 (Convex hulls) *For any subset $\mathcal{S}$ of a linear vector space $\mathcal{X}$, the convex hull $\mathrm{co}(\mathcal{S})$ is convex and consists precisely of all convex combinations of the elements of $\mathcal{S}$.*

Definition 1.12 (Convex functions) A function $f : \mathcal{S} \to \mathbb{R}$ is called *convex* if

1. $\mathcal{S}$ is convex and

2. for all $x_1, x_2 \in \mathcal{S}$ and $\alpha \in (0,1)$ there holds that

$$f(\alpha x_1 + (1 - \alpha)x_2) \leqslant \alpha f(x_1) + (1 - \alpha)f(x_2). \qquad (1.2.1)$$

f is called *strictly convex* if the inequality in (1.2.1) is strict for $x_1 \neq x_2$.

Note that in our definition the domain of a convex function is by definition a convex set. Simple examples of convex functions are $f(x) = x^2$ on $\mathbb{R}$, $f(x) = \sin x$ on $[\pi, 2\pi]$ and $f(x) = |x|$ on $\mathbb{R}$.

Instead of minimizing the function $f : \mathcal{S} \to \mathbb{R}$ we can set our aims a little lower and be satisfied with considering all possible $x \in \mathcal{S}$ that give a guaranteed upper bound of f. For this, we introduce, for any number $\alpha \in \mathbb{R}$, the *sublevel sets* associated with f as follows

$$\mathcal{S}_\alpha := \{x \in \mathcal{S} \mid f(x) \leqslant \alpha\}.$$

Obviously, $\mathcal{S}_\alpha = \varnothing$ if $\alpha < \inf_{x \in \mathcal{S}} f(x)$ and $\mathcal{S}_\alpha$ coincides with the set of global minimizers of f if $\alpha = \inf_{x \in \mathcal{S}} f(x)$. Note also that $\mathcal{S}_\alpha \subseteq \mathcal{S}_\beta$ whenever $\alpha \leqslant \beta$. As you could have guessed, convex functions and convex sublevel sets are related to each other:

Proposition 1.13 *If $f : \mathcal{S} \to \mathbb{R}$ is convex then the sublevel set $\mathcal{S}_\alpha$ is convex for all $\alpha \in \mathbb{R}$.*

Proof. Suppose f is convex, let $\alpha \in \mathbb{R}$ and consider $\mathcal{S}_\alpha$. If $\mathcal{S}_\alpha$ is empty then the statement is trivial. Suppose therefore that $\mathcal{S}_\alpha \neq \varnothing$ and let $x_1, x_2 \in \mathcal{S}_\alpha$, $\lambda \in [0, 1]$. Then, as $\mathcal{S}$ is convex, $\lambda x_1 + (1 - \lambda)x_2 \in \mathcal{S}$ and by definition of $\mathcal{S}_\alpha$ we have that $f(x_1) \leqslant \alpha$, $f(x_2) \leqslant \alpha$. Convexity of f now implies that

$$f(\lambda x_1 + (1 - \lambda)x_2) \leqslant \lambda f(x_1) + (1 - \lambda)f(x_2) \leqslant \lambda \alpha + (1 - \lambda)\alpha = \alpha$$

i.e., $\lambda x_1 + (1 - \lambda)x_2 \in \mathcal{S}_\alpha$. ∎

We emphasize that it is *not* true that convexity of the sublevel sets $\mathcal{S}_\alpha$, $\alpha \in \mathbb{R}$ implies convexity of f. However, the class of functions for which all sublevel sets are convex are that important that they deserve their own name. The following concept is probably the most important generalization of convex functions.

Definition 1.14 (Quasi-convex functions) A function $f : \mathcal{S} \to \mathbb{R}$ is called *quasi-convex* if the sublevel set $\mathcal{S}_\alpha$ is convex for all $\alpha \in \mathbb{R}$.

It is easy to see that f is quasi-convex if and only if

$$f(\alpha x_1 + (1 - \alpha)x_2) \quad \leqslant \quad \max[f(x_1), f(x_2)]$$

for all $\alpha \in [0, 1]$ and for all $x_1, x_2 \in \mathcal{S}$. In particular, every convex function is also quasi-convex.

1.2.2   Convex optimization

We hope that the following propositions will convince the most sceptical reader why the convexity of sets and functions is such a desirable property. Anyone who gained experience with numerical optimization methods got familiar with the pitfalls of local minima and local maxima. One reason for studying convex functions is related to the *absence of local minima*.

Definition 1.15 (Local and global minima) Let $\mathcal{S}$ be a subset of a normed space $\mathcal{X}$. The function $f : \mathcal{S} \to \mathbb{R}$ is said to have a *local minimum* at $x_0 \in \mathcal{S}$ if there exists $\epsilon > 0$ such that

$$f(x_0) \leqslant f(x) \tag{1.2.2}$$

for all $x \in \mathcal{S}$ with $\|x - x_0\| < \epsilon$. It is a *global minimum* of f if (1.2.2) holds for all $x \in \mathcal{S}$.

In other words, f has a local minimum at $x_0 \in \mathcal{S}$ if there exists a neighborhood $\mathcal{N}$ of x_0 such that $f(x_0) \leqslant f(x)$ for all points $x \in \mathcal{S} \cap \mathcal{N}$. Note that according to this definition every global minimum is a local minimum as well. The notions of *local maximum* and *global maximum* of a function f are similarly defined. Here is a simple result which provides one of our main interests in convex functions.

Proposition 1.16 *Suppose that* $f : \mathcal{S} \to \mathbb{R}$ *is convex. If* f *has a local minimum at* $x_0 \in \mathcal{S}$ *then* $f(x_0)$ *is also the global minimum of* f. *If* f *is stricly convex, then* x_0 *is moreover unique.*

Proof.   Let f be convex and suppose that f has a local minimum at $x_0 \in \mathcal{S}$. Then for all $x \in \mathcal{S}$ and $\alpha \in (0, 1)$ sufficiently small,

$$f(x_0) \leqslant f((1 - \alpha)x_0 + \alpha x) = f(x_0 + \alpha(x - x_0)) \leqslant (1 - \alpha)f(x_0) + \alpha f(x). \tag{1.2.3}$$

This implies that

$$0 \leqslant \alpha(f(x) - f(x_0)) \tag{1.2.4}$$

or $f(x_0) \leqslant f(x)$. So $f(x_0)$ is a global minimum. If f is strictly convex, then the second inequality in (1.2.3) is strict so that (1.2.4) becomes strict for all $x \in \mathcal{S}$. Hence, x_0 is unique. ∎

Interpretation 1.17 It is very important to emphasize that proposition 1.16 does not make any statement about *existence* of optimal solutions $x_0 \in \mathcal{S}$ which minimize f. It merely says that all local minima of f are also global minima. It therefore suffices to compute *local minima* of a convex function f to actually determine its global minimum.

Remark 1.18 Proposition 1.16 does not hold for quasi-convex functions.

Let $\mathcal{S}_0$ be a set and suppose that the domain $\mathcal{S}$ of a convex function $f : \mathcal{S} \to \mathbb{R}$ is generated by $\mathcal{S}_0$ as follows

$$\mathcal{S} = \{x \mid x \text{ is a convex combination of } x_1, \dots, x_n \in \mathcal{S}_0, \ n > 0\} = \mathrm{co}(\mathcal{S}_0).$$

As we have seen before $\mathcal{S}$ is then convex and we have the following simple property.

Proposition 1.19 $f(x) \leqslant \gamma$ for all $x \in \mathcal{S}$ if and only if $f(x) \leqslant \gamma$ for all $x \in \mathcal{S}_0$

Proof. Every $x \in \mathcal{S}$ can be written as a convex combination $x = \sum_{i=1}^{n} \alpha_i x_i$ where $n > 0$, $\alpha_i \geqslant 0$, $x_i \in \mathcal{S}_0$, $i = 1, \dots, n$ and $\sum_{i=1}^{n} \alpha_i = 1$. Then, using convexity of f and non-negativity of the α_i's we obtain

$$f(x) = f(\sum_{i=1}^{n} \alpha_i x_i) \leqslant \sum_{i=1}^{n} \alpha_i f(x_i) \leqslant \sum_{i=1}^{n} \alpha_i \gamma = \gamma.$$

which yields the result. ∎

Interpretation 1.20 Proposition 1.19 states that the uniform bound $f(x) \leqslant \gamma$ on $\mathcal{S}$ can equivalenty be verified on the set $\mathcal{S}_0$. This is of great practical relevance especially when $\mathcal{S}_0$ contains only a *finite number of elements*. It then requires a *finite number of tests* to conclude whether or not $f(x) \leqslant \gamma$ for all $x \in \mathcal{S}$. In addition, the supremal value $\gamma_0 := \sup_{x \in \mathcal{S}} f(x)$ can then be determined from $\gamma_0 = \max_{x \in \mathcal{S}_0} f(x)$.

Our next characterization of convex functions is motivated by the geometric idea that through any point on the graph of a convex function we can draw a line such that the entire graph lies above or on the line. For functions $f : \mathcal{S} \to \mathbb{R}$ with $\mathcal{S} \subseteq \mathbb{R}$, this idea is pretty intuitive and is formalized as follows.

Proposition 1.21 *Suppose that $\mathcal{S} \subset \mathbb{R}$ is open. Then $f : \mathcal{S} \to \mathbb{R}$ is convex if and only if for all $x_0 \in \mathcal{S}$ there exists $g \in \mathbb{R}$ such that*

$$f(x) \geqslant f(x_0) + g(x - x_0) \tag{1.2.5}$$

for all $x \in \mathcal{S}$.

Proof. Let f be convex and $x_0 \in \mathcal{S}$. Choose $g \in [f'_-(x_0), f'_+(x_0)]$ where

$$f'_-(x_0) := \lim_{x \uparrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

$$f'_+(x_0) := \lim_{x \downarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}.$$

These limits actually exist as for any triple x_{-1}, x_0, x of points in $\mathcal{S}$ with $x_{-1} < x_0 < x$ we have that

$$\frac{f(x_0) - f(x_{-1})}{x_0 - x_{-1}} \leqslant \frac{f(x) - f(x_{-1})}{x - x_{-1}} \leqslant \frac{f(x) - f(x_0)}{x - x_0}.$$

Hence, $(f(x) - f(x_0))/(x - x_0)$ is a decreasing function of x which is bounded from below by $(f(x_0) - f(x_{-1}))/(x_0 - x_{-1})$. The limit $f'_+(x_0)$ therefore exists. a similar argument proves existence of the limit $f'_-(x_0)$. Now the existence of the limits has been proved, it follows that $\frac{f(x) - f(x_0)}{x - x_0}$ is $\geqslant g$ or $\leqslant g$ depending on whether $x > x_0$ or $x < x_0$. In either case we obtain (1.2.5). Conversely, suppose that for all $x_0 \in \mathcal{S}$ there exists $g \in \mathbb{R}$ such that (1.2.5) holds for all $x \in \mathcal{S}$. Let $x_1, x_2 \in \mathcal{S}$, $\alpha \in [0, 1]$ and put $x_0 = \alpha x_1 + (1 - \alpha)x_2$. By asssumption there exists $g \in \mathbb{R}$ such that $f(x_0) \leqslant f(x_i) + g(x_i - x_0)$, $i = 1, 2$. But then also

$$\begin{aligned}
f(x_0) &= \alpha f(x_0) + (1 - \alpha)f(x_0) \\
&\leqslant \alpha f(x_1) + (1 - \alpha)f(x_2) + g[\alpha x_1 + (1 - \alpha)x_2 - x_0] \\
&= \alpha f(x_1) + (1 - \alpha)f(x_2)
\end{aligned}$$

which shows that f is convex. ∎

Remark 1.22 The right-hand side of (1.2.5) is sometimes called a *support functional* for f at $x_0 \in \mathcal{S}$.

As can be deduced from the above proof, if f happens to be differentiable at x_0 then g is uniquely given by the derivative $f'(x_0)$. We turn now to the more general situation where $\mathcal{S} \subseteq \mathbb{R}^n$. The natural extension of the right hand side of (1.2.5) involves the introduction of an *affine function* through the point $(x_0, f(x_0))$ on the graph of f.

Definition 1.23 (Affine functions) A function $f : \mathcal{S} \to \mathcal{T}$ is *affine* if $f(x) = T(x) + f_0$ where $f_0 \in \mathcal{T}$ and $T : \mathcal{S} \to \mathcal{T}$ is a linear map, i.e.,

$$T(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 T(x_1) + \alpha_2 T(x_2)$$

for all $x_1, x_2 \in \mathcal{S}$ and $\alpha_1, \alpha_2 \in \mathbb{R}$.

Note that a function $f : \mathbb{R}^n \to \mathbb{R}$ is affine if and only if there exist $x_0 \in \mathbb{R}^n$ such that the mapping $x \to f(x) - f(x_0)$ is linear. This means that all affine functions $f : \mathbb{R}^n \to \mathbb{R}$ can be represented as $f(x) = f(x_0) + g^\top(x - x_0)$ where g is some vector in $\mathbb{R}^n$.

Proposition 1.24 *Let $\mathcal{S} \subseteq \mathbb{R}^n$. If $f : \mathcal{S} \to \mathbb{R}$ is convex then for all $x_0 \in \mathcal{S}$ there exists a subgradient $g \in \mathbb{R}^n$ such that*

$$f(x) \geqslant f(x_0) + g^\top(x - x_0) \tag{1.2.6}$$

for all $x \in \mathcal{S}$.

Remark 1.25 Proposition 1.24 gives a necessary condition for convexity of a function f. If the gradient of f

$$f' = \begin{bmatrix} \frac{\partial f}{\partial x_1} & \cdots & \frac{\partial f}{\partial x_n} \end{bmatrix}$$

exists and is continuous at $x_0 \in \mathcal{S}$ then $g = f'(x_0)$ is the only subgradient of f at x_0. In that case one can prove that f is convex if and only if the gradient f' is *monotone non-decreasing* in the sense that for all $x_1, x_2 \in \mathcal{S}$ there holds that $[f'(x_1) - f'(x_2)](x_1 - x_2) \geqslant 0$.

Interpretation 1.26 If we consider the right hand side of (1.2.6), then trivially $g^\top(x - x_0) > 0$ implies that $f(x) > f(x_0)$. Thus all points in the half space $\{x \in \mathcal{S} \mid g^\top(x - x_0) > 0\}$ lead to larger values of f than $f(x_0)$. In particular, in searching the global minimum of f we can disregard this entire half-space.

The observation in interpretation 1.26 leads to a simple and straightforward recursive algorithm for the computation of optimal solutions.

Algorithm 1.27 (Ellipsoid algorithm (conceptual)) Let $x_0 \in \mathbb{R}^n$ and $P_0 \in \mathbb{R}^{n \times n}$ be a positive definite matrix. Consider the problem of minimizing the convex function $f : \mathbb{R}^n \to \mathbb{R}$ over $x \in \mathbb{R}^n$ subject to the constraint

$$(x - x_0)^\top P_0^{-1}(x - x_0) \leqslant 1.$$

Step 0 Set $\mathcal{E}_0 := \{x \in \mathbb{R}^n \mid (x - x_0)^\top P_0^{-1}(x - x_0) \leqslant 1\}$.

Step k For $k \in \mathbb{Z}_+$

- Compute one subgradient $g_k \in \mathbb{R}^n$ for f at x_k and put

$$\mathcal{R}_k := \{x \in \mathbb{R}^n \mid x \in \mathcal{E}_k \text{ and } g_k^\top(x - x_k) \leqslant 0\}.$$

- Compute $x_{k+1} \in \mathbb{R}^n$ and $P_k > 0$ with minimal determinant $\det(P_{k+1})$ such that the ellipsoid

$$\mathcal{E}_{k+1} := \{x \in \mathbb{R}^n \mid (x - x_{k+1})^\top P_{k+1}^{-1}(x - x_{k+1}) \leqslant 1\}$$

  contains $\mathcal{R}_k$.
- Set k to $k + 1$ and return to Step k.

The sequence of ellipsoids $\mathcal{E}_k$ and the sets $\mathcal{R}_k$ have the property that they contain an optimal solution. The subgradients $g_k \in \mathbb{R}^n$ divide $\mathbb{R}^n$ in the two halfspaces $\{x \mid g_k(x - x_k) < 0\}$ and $\{x \mid g_k(x - x_k) > 0\}$ while the *cutting plane* $\{x \mid g_k(x - x_k) = 0\}$ passes through the center of the ellipsid $\mathcal{E}_k$ for each k. In particular $f(x_k)$ converges to a minimizer of f. The algorithm therefore does not calculate a solution but only the minimal value of f. convergence of the algorithm is in 'polynomial time' due to the factr that the volume of the ellipsoids decreases geometrically. However, in practice convergence is rather slow.

1.3   Linear matrix inequalities

1.3.1   What are they?

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

$$F(x) := F_0 + x_1 F_1 + \ldots + x_m F_m > 0 \qquad (1.3.1)$$

where

- $x = (x_1, \ldots, x_m)$ is a vector of real numbers.

- $F_0, \ldots, F_m$ are real symmetric matrices, i.e., $F_i = F_i^\top \in \mathbb{R}^{n \times n}$, $i = 0, \ldots, m$ for some $n \in \mathbb{Z}_+$.

- the inequality > 0 in (1.3.1) means 'positive definite', i.e., $u^\top F(x)u > 0$ for all $u \in \mathbb{R}^n$, $u \neq 0$. Equivalently, the smallest eigenvalue of $F(x)$ is positive.

Stated slightly more general,

Definition 1.28 (Linear Matrix Inequality) A *linear matrix inequality* (LMI) is an inequality

$$F(x) > 0 \qquad (1.3.2)$$

where F is an *affine function* mapping a finite dimensional vector space $\mathbb{V}$ to the set $\mathbb{S}^n := \{M \mid M = M^\top \in \mathbb{R}^{n \times n}\}$, $n > 0$, of real symmetric matrices.

Remark 1.29 The term 'linear matrix inequality' is now common use in the literature on systems and control, but the terminology is not consistent with the expression $F(x) > 0$ as F does not need to be a linear function. 'Affine matrix inequality' would definitely be a better name.

Remark 1.30 Recall from definition 1.23 that an affine mapping $F : \mathbb{V} \to \mathbb{S}^n$ necessarily takes the form $F(x) = F_0 + T(x)$ where $F_0 \in \mathbb{S}^n$ and $T : \mathbb{V} \to \mathbb{S}^n$ is a linear transformation. Thus if $\mathbb{V}$ is finite dimensional, say of dimension m, and $\{e_1, \ldots, e_m\}$ constitutes a basis for $\mathbb{V}$, then we can write

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

where the elements $\{x_1, \ldots, x_m\}$ are such that $x = \sum_{j=1}^{m} x_j e_j$ and $F_j = T(e_j)$ for $j = 1, \ldots, m$. Hence we obtain (1.3.1) as a special case.

Remark 1.31 The same remark applies to mappings $F : \mathbb{R}^{m_1 \times m_2} \to \mathbb{S}^n$ where $m_1, m_2 \in \mathbb{Z}_+$. A simple example where $m_1 = m_2$ is the Lyapunov inequality $F(X) = A^\top X + XA + Q > 0$. Here, $A, Q \in \mathbb{R}^{m \times m}$ are assumed to be given and $X \in \mathbb{R}^{m \times m}$ is the unknown. The

unknown variable is therefore a *matrix*. Note that this defines an LMI only if Q is symmetric. In this case, the domain $\mathbb{V}$ of F in definition 1.28 is equal to $\mathbb{S}^m$. We can view this LMI as a special case of (1.3.1) by defining a basis $E_1, \ldots, E_M$ of $\mathbb{S}^m$ and writing $X = \sum_{j=1}^{M} x_j E_j$. Indeed,

$$F(X) = F\left(\sum_{j=1}^{M} x_j E_j\right) = F_0 + \sum_{j=1}^{M} x_j F(E_j) = F_0 + \sum_{j=1}^{M} x_j F_j$$

which is of the form (1.3.1).

Remark 1.32 A *non-strict LMI* is a linear matrix inequality where $>$ in (1.3.1) and (1.3.2) is replaced by $\geqslant$. The matrix inequalities $F(x) < 0$, and $F(x) > G(x)$ with F and G affine functions are obtained as special cases of definition 1.28 as they can be rewritten as the linear matrix inequality $-F(x) > 0$ and $F(x) - G(x) > 0$.

1.3.2 Why are they interesting?

The linear matrix inequality (1.28) defines a *convex constraint* on x. That is, the set $\mathcal{F} := \{x \mid F(x) > 0\}$ is convex. Indeed, if $x_1, x_2 \in \mathcal{F}$ and $\alpha \in (0,1)$ then

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

where in the first equality we used that F is affine and the last inequality follows from the fact that $\alpha \geqslant 0$ and $(1-\alpha) \geqslant 0$.

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

$$F_1(x) > 0, \ldots, F_k(x) > 0. \tag{1.3.3}$$

It is a simple but essential property that every system of LMI's can be rewritten as one single LMI. Precisely, $F_1(x) > 0, \ldots, F_k(x) > 0$ if and only if

$$F(x) := \begin{pmatrix} F_1(x) & 0 & \ldots & 0 \\ 0 & F_2(x) & \ldots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \ldots & F_k(x) \end{pmatrix} > 0.$$

The last inequality indeed makes sense as $F(x)$ is symmetric. Further, since the set of eigenvalues of $F(x)$ is simply the union of the eigenvalues of $F_1(x), \ldots, F_k(x)$, any x that satisfies $F(x) > 0$ also satisfies the system of LMI's (1.3.3) and vice versa.

A second trivial but 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

$$\begin{cases} F(x) > 0 \\ x = Ay + b \text{ for some } y \end{cases}$$

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

$$\begin{cases} F(x) > 0 \\ x \in \mathcal{M} \end{cases} \tag{1.3.4}$$

where $\mathcal{M}$ is an *affine subset* of $\mathbb{R}^n$, i.e.,

$$\mathcal{M} = x_0 + \mathcal{M}_0 = \{x_0 + m \mid m \in \mathcal{M}_0\}$$

with $x_0 \in \mathbb{R}^n$ and $\mathcal{M}_0$ a linear subspace of $\mathbb{R}^n$, can be rewritten in the form of one single linear matrix inequality $\bar{F}(x) > 0$. To actually do this, let $e_1, \ldots, e_k \in \mathbb{R}^n$ be a basis of $\mathcal{M}_0$ and let $F(x) = F_0 + T(x)$ be decomposed as in remark 1.30. Then (1.3.4) can be rewritten as

$$0 < F(x) = F_0 + T(x_0 + \sum_{j=1}^{k} x_j e_j) = \underbrace{F_0 + T(x_0)}_{\text{constant part}} + \underbrace{\sum_{j=1}^{k} x_j T(e_j)}_{\text{linear part}}$$

$$= \bar{F}_0 + x_1 \bar{F}_1 + \ldots + x_k \bar{F}_k$$

$$=: \bar{F}(\bar{x})$$

where $\bar{F}_0 = F_0 + T(x_0)$, $\bar{F}_j = T(e_j)$ and $\bar{x} = (x_1, \ldots, x_k)$. This implies that $x \in \mathbb{R}^n$ satisfies (1.3.4) if and only if $\bar{F}(\bar{x}) > 0$. Note that the dimension of $\bar{x}$ is smaller than the dimension of x.

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

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

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

$$M > 0 \iff \begin{pmatrix} M_{11} & 0 \\ 0 & S \end{pmatrix} > 0$$

$$\iff \begin{cases} M_{11} > 0 \\ S > 0 \end{cases}$$

For the interested reader, the result is obtained by observing that $M > 0$ if and only if $u^\top M u > 0$ for all non-zero $u \in \mathbb{R}^n$. Let $F \in \mathbb{R}^{r \times (n-r)}$. Then $M > 0$ if and only if for all $u_1 \in \mathbb{R}^r$ and $u_2 \in \mathbb{R}^{n-r}$

$$0 < \begin{bmatrix} u_1 + Fu_2 \\ u_2 \end{bmatrix}^\top \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} u_1 + Fu_2 \\ u_2 \end{bmatrix}$$

$$= \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}^\top \begin{bmatrix} M_{11} & M_{11}F + M_{12} \\ M_{21} + F^\top M_{11} & M_{22} + F^\top M_{11}F + F^\top M_{12} + M_{21}F \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

The result then follows by taking $F = -M_{11}^{-1}M_{12}$.

An immediate consequence of this observation is the following proposition.

Proposition 1.34 (Schur complement) *Let $F : \mathbb{V} \to \mathbb{S}^n$ be an affine function which is partitioned according to*

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

where $F_{11}(x)$ is square. Then $F(x) > 0$ if and only if

$$\begin{cases} F_{11}(x) > 0 \\ F_{22}(x) - F_{12}(x)F_{11}^{-1}(x)F_{21}(x) > 0 \end{cases} \quad . \tag{1.3.5}$$

Note that the second inequality in (1.3.5) is a *non-linear* matrix inequality in x. Using this result, it follows that non-linear matrix inequalities of the form (1.3.5) can be converted to linear matrix inequalities. In particular, it follows that the non-linear inequalities (1.3.5) define a convex constraint on the varable x in the sense that all x satisfying (1.3.5) define a convex set.

1.3.3   What are they good for?

As we will see, many optimization problems in control design, identification and signal processing can be formulated (or reformulated) using linear matrix inequalities. Clearly, it only makes sense to cast these problems in terms of LMI's if these inequalities can be solved efficiently and in a 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 $f : \mathcal{S} \to \mathbb{R}$ with $\mathcal{S} := \{x \mid F(x) > 0\}$ 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 $F, G : \mathbb{V} \to \mathbb{S}^{n_1}$ and $H : \mathbb{V} \to \mathbb{S}^{n_2}$ are affine functions. There are three generic problems related to the study of linear matrix inequalities:

1. **Feasibility:** The test whether or not there exist solutions x of $F(x) > 0$ is called a feasibility problem. The LMI is called *non-feasible* if no solutions exist.

2. **Optimization:** Let $f : \mathcal{S} \to \mathbb{R}$ and suppose that $\mathcal{S} = \{x \mid F(x) > 0\}$. The problem to determine

$$V_{\mathrm{opt}} = \inf_{x \in \mathcal{S}} f(x)$$

   is called an optimization problem with an LMI constraint. This problem involves the determination of the infimum V_{opt} and for arbitrary $\varepsilon > 0$ the calculation of an *almost optimal solution* x which satisfies $x \in \mathcal{S}$ and $V_{\mathrm{opt}} \leqslant f(x) \leqslant V_{\mathrm{opt}} + \varepsilon$.

3. **Generalized eigenvalue problem:** This problem amounts to minimizing a scalar $\lambda \in \mathbb{R}$ subject to

$$\begin{cases} \lambda F(x) - G(x) > 0 \\ F(x) > 0 \\ H(x) > 0 \end{cases}$$

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

Example 1

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

$$\dot{x} = Ax \tag{1.3.6}$$

where $A \in \mathbb{R}^{n \times n}$. By this, we mean the problem to decide whether or not all functions $x : \mathbb{R} \to \mathbb{R}^n$ which satisfy (1.3.6) have the property that $\lim_{t \to \infty} x(t) = 0$. Lyapunov taught us that this system is asymtotically stable if and only if there exists $X \in \mathbb{S}^n$ such that $X > 0$ and $A^\top X + XA < 0$. Thus, asymptotic stability of the system (1.3.6) is equivalent to feasibility of the LMI

$$\begin{pmatrix} X & 0 \\ 0 & -A^\top X - XA \end{pmatrix} > 0.$$

Example 2

Experts in μ-analysis (but other people as well!) regularly face the problem to determine a diagonal matrix D such that $\|DMD^{-1}\| < 1$ where M is some given matrix. Since

$$\|DMD^{-1}\| < 1 \iff D^{-\top}M^{\top}D^{\top}DMD^{-1} < I$$
$$\iff M^{\top}D^{\top}DM < D^{\top}D$$
$$\iff X - M^{\top}XM > 0$$

where $X := D^{\top}D > 0$ we see that the existence of such a matrix is an LMI feasibility problem.

Example 3

Let $F : \mathbb{V} \to \mathbb{S}^n$ be an affine function and consider the problem to minimize $f(x) := \lambda_{\max}(F(x))$ over x. Clearly,

$$\lambda_{\max}(F^{\top}(x)F(x)) < \gamma \iff \gamma I - F^{\top}(x)F(x) > 0$$
$$\iff \begin{pmatrix} \gamma I & F^{\top}(x) \\ F(x) & I \end{pmatrix} > 0$$

where the second inequality follows by taking Schur complements. If we define

$$\bar{x} := \begin{pmatrix} x \\ \gamma \end{pmatrix}, \quad \bar{F}(\bar{x}) := \begin{pmatrix} \gamma I & F^{\top}(x) \\ F(x) & I \end{pmatrix}, \quad \bar{f}(\bar{x}) := \gamma$$

then $\bar{F}$ is an *affine function* of $\bar{x}$ and the problem to minimize the maximum egenvalue of $F(x)$ is equivalent to determining $\inf \bar{f}(\bar{x})$ subject to the LMI $\bar{F}(\bar{x}) > 0$. Hence, this is an optimization problem with a *linear objective function* $\bar{f}$ and an LMI constraint.

Example 4

Consider the linear time-invariant system

$$\dot{x} = A_i x + B_i u$$

where $A_i \in \mathbb{R}^{n \times n}$ and $B_i \in \mathbb{R}^{n \times m}$, $i = 1, \ldots, k$. This represents k linear time-invariant systems with n dimensional state space and m-dimensional input space. The question of *simultaneous stabilization* amounts to finding a state feedback law $u = Fx$ with $F \in \mathbb{R}^{m \times n}$ such that the eigenvalues $\lambda(A_i + B_i F)$ belong to the lef-half complex plane for all $i = 1, \ldots, k$. Using example 1 above, this problem is solved when we can find matrices F and X_i, $i = 1, \ldots, k$, such that for all of these i's

$$\begin{cases} X_i > 0 \\ (A_i + B_i F)^{\top} X_i + X_i(A_i + B_i) < 0 \end{cases} . \tag{1.3.7}$$

Note that this is *not* a system of LMI's in the variables X_i and F. A simplification of this problem is obtained by assuming the existence of a joint Lyapunov function, i.e. $X_1 = \ldots = X_k =: X$. Moreover, if we introduce new variables $Y = X^{-1}$ and $K = FY$ then (1.3.7) reads

$$\begin{cases} Y > 0 \\ A_i Y + Y A_i^\top + B_i K + K^\top B_i^\top < 0 \end{cases}$$

for $i = 1, \ldots, k$. The latter is a system of LMI's in the variables Y and K. The joint stabilization problem therefore has a solution if this system of LMI's is feasible.

Example 5

Consider the linear autonomous system

$$\dot{x} = Ax \tag{1.3.8}$$

together with an arbitrary (but fixed) initial value $x(0) = x_0$ and the criterion function

$$J := \int_0^\infty x^\top(t) Q x(t) \, dt$$

where $Q = Q^\top \in \mathbb{R}^{n \times n}$ is non-negative definite. Assume that the system is asymptotically stable. Then all solutions x of (1.3.8) are square integrable so that $J < \infty$. Now consider the non-strict linear matrix inequality

$$A^\top X + X A + Q \leqslant 0. \tag{1.3.9}$$

For any solution $X = X^\top$ of this LMI we can differentiate the function $x^\top(t) X x(t)$ along solutions x of (1.3.8) to get

$$\frac{d}{dt}[x^\top(t) X x(t)] = x^\top(t)[A^\top X + X A] x(t) \leqslant -x^\top(t) Q x(t)$$

If we assume in addition that $X > 0$ then integrating the latter inequality from $t = 0$ till ∞ yields the upper bound

$$J = \int_0^\infty x^\top(t) Q x(t) \, dt \leqslant x_0^\top X x_0.$$

Here, we used that $\lim_{t \to \infty} x(t) = 0$. Moreover, the smallest upperbound of J is obtained by minimizing $x_0^\top X x_0$ over X subject to the system of LMI's

$$\begin{cases} X > 0 \\ A^\top X + X A + Q \leqslant 0 \end{cases}.$$

1.3.4 How are they solved?

The three problems defined in the previous subsection can be solved in a numerically efficient way. In this section we discuss the basic theoretical ideas behind the 'LMI-solvers'.

Ellipsoid method for LMI's

We first give a solution which is based on the ellipsoidal algorithm as explained in the previous section. This solution is a simple but not a very efficient one. Let $F : \mathcal{S} \to \mathbb{S}^n$ be an affine function with $\mathcal{S} \subset \mathbb{R}^m$. Recall that $F(x) < 0$ if and only if $\lambda_{\max} F(x) < 0$. Define $f(x) := \lambda_{\max}(F(x))$ and consider the problem to minimize f. If $\inf f(x) < 0$ then the LMI $F(x) < 0$ is feasible, if $\inf f(x) \geqslant 0$ then the LMI $F(x) < 0$ is not feasible.

There are a few observations to make to apply Proposition 1.24. The first one is to establish that f is a convex function. Indeed, this we showed in Example 3 of the previous subsection. Secondly, for any x_0 we need to determine a subgradient g on the point $(x_0, f(x_0))$ of the graph of f. To do this, we will use the fact that

$$f(x) = \lambda_{\max}(F(x)) = \max_{u^\top u = 1} u^\top F(x)u.$$

This means that for an arbitrary $x_0 \in \mathcal{S}$ we can determine a vector $u_0 \in \mathbb{R}^n$ with $u_0^\top u_0 = 1$ such that $\lambda_{\max}(F(x_0)) = u_0^\top F(x_0)u_0$. But then

$$\begin{aligned}
f(x) - f(x_0) &= \max_{u^\top u = 1} u^\top F(x)u - u_0^\top F(x_0)u_0 \\
&\geqslant u_0^\top F(x)u_0 - u_0^\top F(x_0)u_0 \\
&= u_0^\top \left(F(x) - F(x_0)\right) u_0.
\end{aligned}$$

The last expression is an affine functional that vanishes in x_0. This means that the right-hand side of this expresson must be of the form $g^\top(x - x_0)$ for some vector $g \in \mathbb{R}^m$. To obtain g, we can write

$$u_0^\top F(x)u_0 = \underbrace{u_0^\top F_0 u_0}_{g_0} + \sum_{j=1}^m x_j \underbrace{u_0^\top F_j u_0}_{g_j}$$

$$= g_0 + g^\top x.$$

Here g is the vector with components $g_1, \ldots, g_m$. In particular, we obtain that

$$f(x) - f(x_0) \geqslant g^\top(x - x_0).$$

The ellipsoid algorithm is now as follows.

Algorithm 1.35 (Ellipsoid algorithm)

Step 0 Let $x_0 \in \mathcal{S}$ and $P_0 \in \mathbb{S}^n$ be a positive definite matrix. Define the ellipsoid

$$\mathcal{E}_0 := \{x \in \mathcal{S} \mid (x - x_0)^\top P_0^{-1}(x - x_0) \leqslant 1\}.$$

from the initialization step of the ellipsoid algorithm.

Step k For $k \in \mathbb{Z}_+$:

1. Compute a subgradient $g_k \in \mathbb{R}^n$ for f at x_k and put

$$\mathcal{R}_k := \{x \in \mathcal{S} \mid x \in \mathcal{E}_k \text{ and } g_k^\top (x - x_k) \leqslant 0\}.$$

2. Compute $x_{k+1} \in \mathcal{S}$ and $P_{k+1} > 0$ such that the ellipsoid

$$\mathcal{E}_{k+1} := \{x \in \mathbb{R}^n \mid (x - x_{k+1})^\top P_{k+1}^{-1} (x - x_{k+1}) \leqslant 1\}$$

entirely contains $\mathcal{R}_k$. One such x_{k+1} and P_{k+1} are given by

$$x_{k+1} := x_k - \frac{P_k g_k}{(m+1)\sqrt{g_k^\top P_k g_k}}$$

$$P_{k+1} := \frac{m^2}{m^2 - 1} \left(P_k - \frac{2}{(m+1)g_k^\top P_k g_k} P_k g_k g_k^\top P_k \right)$$

3. Set k to $k+1$ and repeat Step k.

As noticed earlier, this recursive scheme generates a sequence of ellipsoids that are guaranteed to contain a minimizer of f in $\mathcal{S}$. The algorithm needs an initialization step in which P_0 and x_0 are determined. Note that this is the only 'non-automated' step in the algorithm. If $\mathcal{S}$ is a bounded subset in $\mathbb{R}^m$ then the safest choice of the initial ellipsoid $\mathcal{E}_0$ would be to guarantee that $\mathcal{S} \subseteq \mathcal{E}_0$.

Interior point methods

A major breakthrough in convex optimization lies in the introduction of interior-point methods. These methods were developed in a series of papers [12] and became of true interest in the context of LMI problems in the work of Yrii Nesterov and Arkadii Nemirovskii [20].

The main idea is rather simple. To solve a convex optimization problem

$$\min f(x)$$

over all x which satisfy the linear matrix inequality $F(x) > 0$, it is first necessary to introduce a *barrier function*. This is a smooth function ϕ which is required to

1. be *stricly convex* on the interior of the feasibility set

$$\mathcal{F} := \{x \mid F(x) > 0\},$$

and

2. approach infinity along each sequence of points x_n in the interior of $\mathcal{F}$ that converge to a boundary point of $\mathcal{F}$.

Given such a barrier function ϕ, the constraint optimization problem to minimize $f(x)$ over all $x \in \mathcal{F}$ is replaced by the *unconstrained optimization problem* to minimize the functional

$$f_t(x) := tf(x) + \phi(x) \qquad (1.3.10)$$

where $t > 0$ is a so called *penalty parameter*. The main idea is to determine a minimizer $x(t)$ of f_t and to consider the behavior of $x(t)$ as function of the penalty parameter $t > 0$. In almost all interior point methods the latter unconstrained optimization problem is solved with the classical Newton-Raphson iteration technique to approximate the minimum of f_t. Under mild assumptions and for a suitably defined sequence of penalty parameters t_n with $t_n \to \infty$ as $n \to \infty$, the sequence $x(t_n)$ with $n \in \mathbb{Z}_+$ will then converge to a point x which is a solution of the original convex optimization problem.

A small modification of this theme is obtained by replacing the the original constraint optimization problem by the unconstrained optimization problem to minimize

$$g_t(x) := \phi_0(t - f(x)) + \phi(x) \qquad (1.3.11)$$

where $t > t_0 := \inf_{F(x)>0} f(x)$ and ϕ_0 is a barrier function for the non-negative real half-axis. Again, the idea is to calculate a minimizer $x(t)$ of g_t (typically using the classical Newton algorithm) and to consider the 'path' $x(t)$ as function of the penalty parameter t. The curve given by $x(t)$ with $t > t_0$ is called the *path of centers* for the optimization problem. Under suitable conditions the solutions $x(t)$ are analytic and have a limit as $t \downarrow t_0$, say x_{opt}. The point x_{opt} is optimal since for $t > t_0$, $x(t)$ is feasible and satisfies $f(x(t)) < t$.

Interior point methods can be applied to each of the three problems as defined in the previous section. If we consider the *feasibility problem* associated with the LMI $F(x) > 0$ then (f does not play a role and) one candidate barrier function is the logarithmic function

$$\phi(x) := \begin{cases} \log \det F(x)^{-1} & \text{if } x \in \mathcal{F} \\ \infty & \text{otherwise} \end{cases}.$$

Under the assumption that the feasible set $\mathcal{F}$ is bounded and non-empty, it follows that ϕ is strictly convex and hence it defines a barrier function for the feasibility set $\mathcal{F}$. By invoking proposition 1.16, we know that there exists a uniquely defined $x_0 \in \mathcal{F}$ such that $\phi(x_{\mathrm{opt}})$ is the global minimum of ϕ. This point x_{opt} obviously belongs to $\mathcal{F}$ and is called the *analytic center* of the feasibility set $\mathcal{F}$. It is usually obtained in a very efficient way from the classical Newton iteration

$$x_{k+1} = x_k - (\phi''(x_k))^{-1}\phi'(x_k). \qquad (1.3.12)$$

Here ϕ' and ϕ'' denote the gradient and the Hessian of ϕ, respectively.

The convergence of this algorithm can be analyzed as follows. Since ϕ is strongly convex and sufficiently smooth, there exist numbers L and M such that for all vectors u with norm

$\|u\| = 1$ there holds

$$u^\top \phi''(x)u \geqslant M$$
$$\|\phi''(x)u - \phi''(y)u\| \leqslant L\|x - y\|.$$

In that case,

$$\|\phi'(x_{k+1})\|^2 \leqslant \frac{L}{2M^2}\|\phi'(x_k)\|^2$$

so that whenever the initial value x_0 is such that $\frac{L}{2M^2}\|\phi'(x_0)\| < 1$ the method is guaranteed to converge *quadratically*.

The idea will be to implement this algorithm in such a way that quadratic convergence can be guaranteed for the largest possible set of initial values x_0. For this reason the iteration (1.3.12) is modified as follows

$$x_{k+1} = x_k - \alpha_k(\lambda(x_k))\phi''(x_k)^{-1}\phi(x_k)$$

where

$$\alpha_k(\lambda) := \begin{cases} 1 & \text{if } \lambda < 2 - \sqrt{3} \\ \frac{1}{1+\lambda} & \text{if } \lambda \geqslant 2 - \sqrt{3} \end{cases}.$$

and $\lambda(x) := \sqrt{\phi'(x)^\top \phi''(x)\phi'(x)}$ is the so called *Newton decrement* associated with ϕ. It is this damping factor that guarantees that x_k will converge to the analytic center x_{opt}, the unique minimizer of ϕ. It is important to note that the step-size is variable in magnitude. The algorithm guarantees that x_k is always feasible in the sense that $F(x_k) > 0$ and that x_k converges globally to an optimum x_{opt}. It can be shown that $\phi(x_k) - \phi(x_{\text{opt}}) \leqslant \epsilon$ whenever

$$k \geqslant c_1 + c_2 \log\log(1/\epsilon) + c_3\left(\phi(x_0) - \phi(x_{\text{opt}})\right)$$

where c_1, c_2 and c_3 are constants. The first and second terms on the right-hand side do not dependent on the optimization criterion and the specific LMI constraint. The second term can almost be neglected for small values of ϵ.

The *optimization problem* to minimize $f(x)$ subject to the LMI $F(x) > 0$ can be viewed as a feasibility problem for the LMI

$$\bar{F}_t(x) := \begin{pmatrix} t - f(x) & 0 \\ 0 & F(x) \end{pmatrix} > 0.$$

where $t > t^* := \inf_{F(x)>0} f(x)$ is a penalty parameter. Using the same barrier function for this linear matrix inequality yields the unconstrained optimization problem to minimize

$$g_t(x) := \log\det \bar{F}_t(x)^{-1} = \underbrace{\log\frac{1}{t - f(x)}}_{\phi_0(t-f(x))} + \underbrace{\log\det F(x)^{-1}}_{\phi(x)}$$

which is of the form (1.3.11). Due to the strict convexity of g_t the minimizer $x(t)$ of g_t is unique for all $t > t^*$. It can be shown that the sequence $x(t)$ is feasible for all $t > t^*$ and approaches the infimum $\inf_{F(x)>0} f(x)$ as $t \uparrow t^*$.

1.3.5 How to compute solutions?

The LMI toolbox in MATLAB provides various routines for the computation of solutions to the three generic problems that were formulated in subsection 1.3.3. The manual [6] is a well written introduction for the usage of this software and provides plenty of examples. Since nobody likes to read software manuals we give a 'nutshell summary' of the relevant routines and their purposes in this section.

The best introduction to the LMI toolbox (and in particular the LMI-lab which comprises the routines for LMI solvers) is to run and study the tutorial `lmidem` of the LMI toolbox. We recommend every 'beginner' to try this tutorial at least once.

Specification of LMI's

In MATLAB, the data for the description of a linear matrix inequality is internally represented in one vector. The LMI-lab can handle any system of LMI's of the form

$$N^\top L(X_1, \ldots, X_K)N < M^\top R(X_1, \ldots, X_K)M \qquad (1.3.13)$$

where the X_i are the unknown matrix variables, possibly with some prescribed structure, N and M are given matrices with identical dimensions and $L(\cdot)$ and $R(\cdot)$ are symmetrical block matrices with identical block structures. Each block in $L(\cdot)$ and $R(\cdot)$ defines an affine function of $X_1, \ldots, X_K$ and their transposes.

An LMI of this type is internally specified in MATLAB with help of the routines

- `lmiedit` an interactive graphical interface.
- `lmivar` to specify unknown variables and their structure
- `lmiterm` to specify the term content of an LMI
- `setlmis` used in connection with `lmiterm`
- `getlmis` gets the internal representation of an LMI

The specification of an LMI should begin with the command `setlmis` and should be ended with `getlmis`. The command

```
lmisys = getlmis;
```

returns the internal MATLAB description of an LMI in the variable `lmisys` (Don't forget the semicolon as you do not want to see or understand the entries of this variable). The latter variable subsequently serves as input to the LMI solver routines.

The simplest way to specify an LMI is by typing

```
lmiedit
```

at the MATLAB prompt. This will give you a graphical user interface all data for the specification of an LMI can be entered in symbolic terms. Just try it!

LMI solvers

The basic routines for numerically solving the three generic problems formulated in section 1.3.3 are the following

- `feasp` to compute a solution to the feasibility problem
- `mincx` to compute a solution to the optimization problem
- `gevp` to compute a solution to the generalized eigenvalue problem.

Each of these routines is implemented as a .mex file in MATLAB and takes a variable which represents the data of an LMI as its input.

Information retrieval

The routine `lmiinfo` can be used to interactively retrieve information about a linear matrix inequality. It provides information about the linear matrix inequality (1.3.13) and the specific structure and the number of variables and blocks appearing in the affine functions $L(\cdot)$ and $R(\cdot)$.

Validation

A solution $X_1, \ldots, X_K$ of (1.3.13) can be validated with the routines `evallmi` and `showlmi`. We refer to the corresponding help information of these routines for more details.

1.3.6   When were they invented?

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 Aleksandr 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 1890, Lyapunov made a systematic study of the local expansion and contraction properties of motions of dynamical systems around an attractor. He worked out the idea that an invariant set of a differential equation is stable in the sense that it attracts all solutions if one can find a function that is bounded from below and decreases along all solutions outside the invariant set.

Aleksandr Mikhailovich Lyapunov was born on May 25, 1857 and published in 1892 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 served as his doctoral dissertation and was defended on September 1892 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 positions and velocities of material taking values in a finite dimensional state space $\mathcal{X} = \mathbb{R}^n$. As Theorem I in Chapter 1, section 16 it contains the statement[2] 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 simple and 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 systems 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. Because of the importance of this result we devote the next section to the subject of Lyapunov stability.

1.4 Lyapunov stability

Translated in modern jargon, Lyapunov considered the differential equation

$$\dot{x} = f(x) \tag{1.4.1}$$

with finite dimensional state space $\mathcal{X} = \mathbb{R}^n$ and $f : \mathbb{R}^n \to \mathbb{R}^n$ an analytic function. For system theorists this is an example of an *autonomous dynamical system*. Assume that for all initial conditions $x_0 \in \mathcal{X}$ there exists a *unique* solution $x : \mathbb{R}_+ \to \mathcal{X}$ of (1.4.1) which passes through x_0 at the initial time $t = 0$. With some abuse of notation this solution will

---

[2]Translation by A.T. Fuller as published in the special issue of the International Journal of Control in March 1992 and in [16].

be denoted as $x(t, x_0)$ to explicitly display the dependence of the initial value. In particular, $x(0, x_0) = x_0$.

A set $\mathcal{S} \subset \mathcal{X}$ is called an *invariant set* of (1.4.1) if $x_0 \in \mathcal{S}$ implies that $x(t, x_0) \in \mathcal{S}$ for all $t \in \mathbb{R}$. The idea of an invariant set is therefore that a solution remains in the set once it started there. A point x^* in $\mathcal{X}$ is called an *equilibrium point* of the flow if the singleton $\mathcal{S} = \{x^*\}$ is an invariant set. Obviously, every equilibrium point defines a constant solution $x(t, x^*) = x^*$, $t \geqslant 0$ of the differential equation (1.4.1). In particular, an equilibrium point x^* of (1.4.1) satisfies $0 = f(x^*)$. To investigate the issue of stability, we will be interested in the behavior of solutions $x(t, x_0)$ with $t \geqslant 0$ and initial condition x_0 in the neighborhood of an equilibrium point x^*. To do this, we equip the state space $\mathcal{X}$ with its natural (Euclidean) norm $\| \cdot \|$.

Definition 1.36 (Lyapunov stability) Consider the differential equation (1.4.1).

1. An equilibrium point $x^* \in \mathcal{X}$ is called *stable* (in the sense of Lyapunov) if given any $\epsilon > 0$, there exists $\delta > 0$ (only depending on ϵ and not on t) such that

$$\|x^* - x_0\| \leqslant \delta \quad \Longrightarrow \quad \|x(t, x_0) - x^*\| \leqslant \epsilon \text{ for all } t \geqslant 0$$

2. The equilibrium point $x^* \in \mathcal{X}$ is called an *attractor* if there exists $\epsilon > 0$ with the property that

$$\|x^* - x_0\| \leqslant \epsilon \quad \Longrightarrow \quad \lim_{t \to \infty} x(t, x_0) = x^*$$

3. It is called *asymptotically stable* (in the sense of Lyapunov) if x^* is both stable (in the sense of Lyapunov) and an attractor.

4. The equilibrium point $x^* \in \mathcal{X}$ is said to be *unstable* if it is not stable (in the sense of Lyapunov).

There are many variations to these concepts. The *region of attraction* associated with an equilibrium point x^* is defined to be set of all initial states $x_0 \in \mathcal{X}$ for which $x(t, x_0) \to x^*$ as $t \to \infty$. If this region coincides with $\mathcal{X}$ then x^* is said to be a global attractor. We will say that an equilibrium x^* is *globally asymptotically stable* if it is stable and globally attractive. Lyapunov functions are defined as follows.

Definition 1.37 (Lyapunov functions) A function $V : \mathcal{X} \to \mathbb{R}$ is called a *Lyapunov function* in a neighborhood $\mathcal{N}(x^*) \subset \mathcal{X}$ of an equilibrium point x^* if

1. V is continuous at x^*,

2. V attains a strong local mimimum at x^*, i.e., there exists a function $\alpha : \mathbb{R}_+ \to \mathbb{R}_+$ which is continuous, strictly increasing, with $\alpha(0) = 0$, such that

$$V(x) - V(x^*) \geqslant \alpha(\|x - x^*\|)$$

for all $x \in \mathcal{N}(x^*)$.

3. V is monotone non-increasing along all solutions $x(t, x_0)$ of (1.4.1) with $x_0 \in \mathcal{N}(x^*)$, i.e., $V(x(t, x_0))$ is monotone non-increasing as a function of t for all $x_0 \in \mathcal{N}(x^*)$.

If a Lyapunov function V is differentiable, the last item states that $\dot{V}(x) := \sum_{j=1}^{n} \frac{\partial V}{\partial x_j} f_j(x)$ is less than or equal to zero for solutions $x(t, x_0)$ of (1.4.1) with initial condition x_0 nearby the equilibrium x^*. The main stability results for autonomous systems of the form (1.4.1) are summarized in the following proposition.

Proposition 1.38 (Lyapunov theorem) *Consider the differential equation (1.4.1) and let $x^* \in X$ be an equilibrium point.*

1. *x^* is a stable equilibrium if there exists a Lyapunov function V in a neigborhood $\mathcal{N}(x^*)$ of x^*.*

2. *x^* is an aymptotically stable equilibrium if there exists a Lyapunov function V in a neighborhood $\mathcal{N}(x^*)$ of x^* such that the only solution x of (1.4.1) in $\mathcal{N}(x^*)$ for which $\dot{V}(x(t)) = 0$ is $x(t) = x^*$.*

Proof.   1. Suppose that V is a Lyapunov function. Let $\epsilon > 0$ be given. As V is continuous and $V(0) = 0$, there exists $\delta > 0$ such that $V(x_0) - V(x^*) < \alpha(\epsilon)$ for every $x_0 \in \mathcal{X}$ with $\|x - x^*\| < \delta$. Now, for all $t \geqslant 0$ and $x_0 \in \mathcal{N}(x^*)$ we have that

$$0 \leqslant \alpha(\|x(t, x_0) - x^*\|) \leqslant V(x(t, x_0)) - V(x^*) \leqslant V(x_0) - V(x^*) < \alpha(\epsilon)$$

Since α is strictly increasing, it follows that $\|x(t, x_0) - x^*\| < \epsilon$ for all $t \geqslant 0$.

2. Similarly proven.   ■

Together with the flow (1.4.1) let us also consider the *linear autonomous system*

$$\dot{x} = Ax \tag{1.4.2}$$

where $A : \mathbb{R}^n \to \mathbb{R}^n$ is a linear map obtained as the *linearization* of $f : \mathbb{R}^n \to \mathbb{R}$ around an equilibrium point $x^* \in \mathcal{X}$ of (1.4.1). Precisely, for $x^* \in X$ we write

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

The linearization of f around x^* is defined by the system (1.4.2) with A defined by the real $n \times n$ matrix

$$A := \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x^*).$$

It is well known that the origin of the *linear flow* (1.4.2) is asymptotically stable if and only if all the eigenvalues of A have negative real parts. Equivalently, the origin of the linear flow (1.4.2) is asymptotically stable if and only if there exists an ellipsoid

$$\mathcal{E} = \{x \in \mathcal{X} \mid x^\top X x = 1\}, \quad X > 0$$

with center in the origin such that the *velocity vector* Ax is directed inward at any point x of the ellipsoid $\mathcal{E}$. The positive definite quadratic function $V : \mathcal{X} \to \mathbb{R}$ defined by

$$V(x) = x^\top X x$$

is an example of a Lyapunov function. The derivative of $V(x)$ in the direction of the vector field Ax is given by

$$x^\top [A^\top X + X A] x$$

and should be negative to guarantee that the origin is an asymptotic stable equilibrium point of (1.4.2). We thus obtain the following result

Proposition 1.39 *The following statements are equivalent.*

1. *The origin is an asymptotic stable equilibrium point of* (1.4.2).

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

3. *The linear matrix inequality*
$$A^\top X + X A < 0$$
   admits a positive definite solution $X = X^\top > 0$.

Moreover, if one of these statements hold, then the equilibrium x^* of the flow (1.4.1) *is asymptotically stable.*

The most important conclusion of Proposition 1.39 is that asymptotic stability of the equilibrium x^* of the nonlinear flow (1.4.1) can be concluded from the asymptotic stability of the linearized system. It is evident that this result has important consequences for systems and control.

1.5   Some simple applications and examples

In this section we collected various examples and illustrations of the theory treated in this chapter.

1.5.1 A Leontief economy

A manufacturer may be 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, $i = 1, \ldots, m$. The manufacturer probably wishes to maximize his profit

$$p(x_1, \ldots, x_n) := p_1 x_1 + p_2 x_2 + \ldots + p_n x_n$$

subject to the production constraints

$$a_{11} x_1 + a_{12} x_2 + \ldots + a_{1n} x_n \leqslant a_1$$
$$a_{21} x_1 + a_{22} x_2 + \ldots + a_{2n} x_n \leqslant a_2$$
$$\vdots$$
$$a_{m1} x_1 + a_{m2} x_2 + \ldots + a_{mn} x_n \leqslant a_m$$

and $x_j \geqslant 0$, $j = 1, \ldots, n$. Note that this is an optimization problem subject to a system of non-strict linear matrix inequalities.

(Wassily Leontief was born in 1906 in St. Petersburg and is winner of the 1973 Nobel Prize of Economics. Among many things, he used input-output analysis to study the characteristics of trade flow between the U.S. and other countries).

1.5.2 Noise descriptions for system identification

Suppose that input-output data has been collected for the purpose of system identification. That is, let

$$(u_t, y_t), \qquad t = 1, \ldots, N$$

denote a finite set of scalar valued input-output samples that were obtained from some unknown dynamical system. In many applications, the time domain data is supposed to be generated by a system with transfer function

$$H(z) = \sum_{k=0}^{\infty} h_k z^{-k}$$

and it is assumed that the observed output is corrupted by additive noise. That is, it is assumed that

$$y_t = \sum_{k=0}^{t-1} h_k u_{t-k} + \eta_t$$

where $t = 1, \ldots, N$ and η_t is some noise sequence. To define the *a-priori* noise sets it is sometimes useful to consider LMI regions. That is, the noise set can be definied as the *feasibility region* of a *set of LMI's*:

$$\mathcal{N} := \{\eta \in \mathbb{R}^N \mid F(\eta) := F_0 + \sum_{k=1}^N F_k \eta_k > 0\} \tag{1.5.1}$$

where F_k are given real matrices. for example, the set

$$\ell_\infty(\varepsilon) := \{\eta : [1, N] \to \mathbb{R} \mid \sup_{t \in [1,N]} |\eta_t| < \varepsilon\}$$

of bounded amplitude disturbances is often used in this context and is just obtained as a special case of 1.5.1 by taking $F_k \in \mathbb{R}^{2N \times N}$ as follows

$$F_k := \begin{cases} \operatorname{diag} \begin{pmatrix} \varepsilon & \varepsilon \end{pmatrix}^\top & \text{if } k = 0 \\ \operatorname{diag} \begin{pmatrix} 1 & -1 \end{pmatrix}^\top & \text{if } 1 \leqslant k \leqslant N \end{cases}.$$

An identification problem can now be formalized which amounts to determining the set of *consistent models*

$$\mathcal{M} := \left\{ H(z) = \sum_{k=0}^N h_k z^{-k} \mid Y - HU \in \mathcal{N} \right\}.$$

Here, $Y = \begin{pmatrix} y_1, \ldots, y_N \end{pmatrix}^\top$, $U = \begin{pmatrix} u_1, \ldots, u_N \end{pmatrix}^\top$ represent the data, and

$$H := \begin{pmatrix} h_0 & 0 & \ldots & 0 \\ h_1 & h_0 & \ldots & 0 \\ \vdots & \vdots & \vdots & 0 \\ h_{N-1} & \ldots & h_1 & h_0 \end{pmatrix}.$$

Stated otherwise, the set of consistent models are those linear, time-invariant systems for which the output error belongs to the feasible set of a linear matrix inequality. The data is said to be *consistent* with the noise $\mathcal{N}$ if $\mathcal{M} \neq \varnothing$.

1.5.3   Convex combinations of linear systems

In definition 1.6 we introduced the notion of a convex combination of a finite set of points. This notion gets considerable relevance in the context of dynamical systems if 'points' become systems. Consider a time-varying dynamical system

$$\frac{dx}{dt}(t) = A(t)x(t) + B(t)u(t)$$
$$y(t) = C(t)x(t) + D(t)u(t)$$

with input u, output y and state x. Suppose that its *system matrix*

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

is a time varying object which for any time instant $t \in \mathbb{R}$ can be written as a *convex combination* of the n system matrices $S_1, \ldots, S_n$. This means that for any time instant $t \in \mathbb{R}$ there exist real numbers $\alpha_i(t) \geqslant 0$, (real numbers that are possibly depending on t), with $\sum_{i=1}^n \alpha_i(t) = 1$ such that

$$S(t) = \sum_{i=1}^n \alpha_i(t) S_i$$

Here,

$$S_i = \begin{pmatrix} A_i & B_i \\ C_i & D_i \end{pmatrix}, \quad i = 1, \ldots, n$$

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 $S_1, \ldots, S_n$, i.e.,

$$S(t) \in \mathrm{co}(S_1, \ldots, S_n).$$

Such models are called *polytopic linear differential inclusions* and arise in a wide variety of modeling problems. The LMI toolbox in MATLAB provides interesting software to represent such models and to perform simulations with them. See the routines

ltisys	to convert a state space model to a system matrix
ltiss	to convert a system matrix to a state space model
ltitf	to compute the transfer function (SISO only)
sinfo	to extract inquiries about system matrices
splot	to plot characteristic responses of systems
psys	to define a polytopic model
psinfo	to extract inquiries about polytopic models

See the help information of these routines for more specific details on their usage.

1.5.4 Affine combinations of linear systems

Models of physical systems are often expressed in terms of state space systems in which the components of the state variable represent a physical quantity. In these models uncertainty about specific parameters is therefore often reflected as uncertainty in specific entries of the state space matrices A, B, C, D. Let $p = (p_1, \ldots, p_n)$ denote the parameter vector which expresses the uncertain quantities in the system and suppose that this parameter vector belongs to some subset $\mathcal{P} \subset \mathbb{R}^n$. Then the uncertain model can be thought of as being

parametrized by $p \in \mathcal{P}$ through its state space representation

$$\begin{aligned}\dot{x} &= A(p)x + B(p)u \\ y &= C(p)x + D(p)u.\end{aligned} \tag{1.5.2}$$

One way to think of equations of this sort is to view them as a set of linear time-invariant systems as parametrized by $p \in \mathcal{P}$. However, if p is time, then (1.5.2) defines a linear time-varying dynamical system and it can therefore also be viewed as such. If components of p are time varying and coincide with state components then (1.5.2) is better viewed as a non-linear system.

Of particular interest will be those systems in which the system matrices *affinely depend on* p. This means that

$$\begin{aligned}A(p) &= A_0 + p_1 A_1 + \ldots + p_n A_n \\ B(p) &= B_0 + p_1 B_1 + \ldots + p_n B_n \\ C(p) &= C_0 + p_1 C_1 + \ldots + p_n C_n \\ D(p) &= D_0 + p_1 D_1 + \ldots + p_n D_n\end{aligned}$$

or, written in more compact form

$$S(p) = S_0 + p_1 S_1 + \ldots + p_n S_n$$

where

$$S(p) = \begin{pmatrix} A(p) & B(p) \\ C(p) & D(p) \end{pmatrix}$$

is the system matrix associated with (1.5.2). We call these models *affine parameter dependent models*.

In MATLAB such a system is represented with the routines `psys` and `pvec`. For $n = 2$ and a parameter box

$$\mathcal{P} := \{(p_1, p_2) \mid p_1 \in [p_1^{\min}, p_1^{\max}], \ p_2 \in [p_2^{\min}, p_2^{\max}]\}$$

the syntax is

```
affsys = psys( p, [s0, s1, s2] )
p = pvec( 'box', [p1min p1max ; p2min p2max])
```

where `p` is the parameter vector whose i-th component ranges between $p_i^{\min}$ and $p_i^{\max}$. See also the routines

```
pdsimul   for time simulations of affine parameter models
aff2pol   to convert an affine model to an equivalent polytopic model.
pvinfo    to inquire about the parameter vector.
```

1.5.5 Stability regions for LTI systems

As we have seen, the autonomous dynamical system

$$\dot{x} = Ax$$

is asymptotically stable if and only if all eigenvalues of A lie in $\mathbb{C}_-$, the open left half complex plane. For many applications in control and engineering we may be interested in more general stability regions. Let us define a *stability region* as a subset $\mathbb{C}_{\text{stab}} \subseteq \mathbb{C}$ with the following two properties

$$\begin{cases} \text{Property 1:} & \lambda \in \mathbb{C}_{\text{stab}} \implies \bar{\lambda} \in \mathbb{C}_{\text{stab}} \\ \text{Property 2:} & \mathbb{C}_{\text{stab}} \text{ is convex .} \end{cases}$$

Typical examples of common stability sets include

$$\begin{aligned}
&\mathbb{C}_{\text{stab}} = \mathbb{C}_- && \text{open left half complex plane} \\
&\mathbb{C}_{\text{stab}} = \mathbb{C} && \text{no stability requirement} \\
&\mathbb{C}_{\text{stab}} = \{s \in \mathbb{C} \mid \Re(s) < -\alpha\} && \text{guaranteed damping} \\
&\mathbb{C}_{\text{stab}} = \{s \in \mathbb{C} \mid \Re(s) < -\alpha, \ |s| < r\} && \text{maximal damping and oscillation} \\
&\mathbb{C}_{\text{stab}} = \{s \in \mathbb{C} \mid \alpha_1 < \Re(s) < \alpha_2\} && \text{vertical strip} \\
&\mathbb{C}_{\text{stab}} = \{s \in \mathbb{C} \mid |\Im(s)| < \alpha\} && \text{horizontal strip} \\
&\mathbb{C}_{\text{stab}} = \{s \in \mathbb{C} \mid \Re(s)\tan\theta < -|\Im(s)|\} && \text{conic stability region.}
\end{aligned}$$

Here, $\theta \in (0, \pi/2)$ and $r, \alpha, \alpha_1, \alpha_2$ are real numbers. We consider the question whether we can derive a feasibility test to verify whether the eigenmodes of the system $\dot{x} = Ax$ belong to either of these sets. This can indeed be done in the case of the given examples. To see this, we observe that

$$\begin{aligned}
\Re(s) < 0 \qquad &\Longleftrightarrow \quad s + \bar{s} < 0 \\
\Re(s) < -\alpha \qquad &\Longleftrightarrow \quad s + \bar{s} + 2\alpha < 0 \\
|s| < r \qquad &\Longleftrightarrow \quad \begin{pmatrix} -r & s \\ \bar{s} & -r \end{pmatrix} < 0 \\
\alpha_1 < \Re(s) < \alpha_2 \qquad &\Longleftrightarrow \quad \begin{pmatrix} (s+\bar{s}) - 2\alpha_2 & 0 \\ 0 & -(s+\bar{s}) + 2\alpha_1 \end{pmatrix} < 0 \\
\Re(s)\tan(\theta) < -|\Im(s)| \qquad &\Longleftrightarrow \quad \begin{pmatrix} (s+\bar{s})\sin\theta & (s-\bar{s})\cos\theta \\ (s-\bar{s})\cos\theta & (s+\bar{s})\sin\theta \end{pmatrix} < 0.
\end{aligned}$$

Here we used that $|s| < r$ if and only if $s\bar{s} < r^2$ which in turn is equivalent to $r - sr^{-1}\bar{s} > 0$. The latter expression can then be recognized as a Schur complement of r.

In any case, each of these regions can be expressed in the form

$$\{s \in \mathbb{C} \mid P + Qs + Q^\top \bar{s} < 0\} \tag{1.5.3}$$

where $P = P^\top$ and Q are real matrices. As $P + Qs + Q^\top \bar{s}$ is Hermitian for all $s \in \mathbb{C}$ the set (1.5.3) is convex and coincides with the feasibility set of a linear matrix inequality. The matrix valued function

$$f(s) := P + Qs + Q^\top \bar{s}$$

is called the *characteristic function* of the stability region (1.5.3). This set includes the examples given above and regions bounded by circles, ellipses, strips, parabolas and hyperbolas. Since finite intersections of such regions can be obtained by systems of LMI's one can virtually approximate any convex region in the complex plane by an LMI of the form

$$\begin{pmatrix} P_1 & 0 & \dots & 0 \\ 0 & P_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & P_k \end{pmatrix} + \begin{pmatrix} Q_1 & 0 & \dots & 0 \\ 0 & Q_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & Q_k \end{pmatrix} s + \begin{pmatrix} Q_1 & 0 & \dots & 0 \\ 0 & Q_2 & \dots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & Q_k \end{pmatrix} \bar{s} < 0$$

which is again of the form (1.5.3). Stability regions $\mathbb{C}_{\text{stab}}$ of the form (1.5.3) lead to the following interesting generalization of the Lyapunov inequality. The result can be found as Theorem 2.2 in [3].

Proposition 1.40 (M. Chilali and P. Gahinet) *Let $P = P^\top, Q$ and A be real matrices. Then A has all its eigenvalues in the stability region*

$$\mathbb{C}_{stab} := \{s \in \mathbb{C} \mid P + Qs + Q^\top \bar{s} < 0\}$$

if and only if there exists a real symmetric matrix $X = X^\top > 0$ with

$$\begin{pmatrix} p_{11}X + q_{11}AX + q_{11}XA^\top & \dots & p_{1k}X + q_{1k}AX + q_{k1}XA^\top \\ \vdots & \ddots & \vdots \\ p_{k1}X + q_{k1}AX + q_{1k}XA^\top & \dots & p_{kk}X + q_{kk}AX + q_{kk}XA^\top \end{pmatrix} < 0 \qquad (1.5.4)$$

where p_{ij} and q_{ij} are the ij-th entry of P and Q, respectively.

Stated otherwise, A has all its eigenvalues in the stability region $\mathbb{C}_{\text{stab}}$ with characteristic function $P + Qs + Q^\top \bar{s}$ if and only if there exists a positive definite matrix X such that

$$(p_{ij}X + q_{ij}XA + q_{ji}A^\top X) < 0.$$

for all i, j. Note that this is an LMI in X and that the classical Lyapunov theorem corresponds to the characteristic function $f(s) = s + \bar{s}$. Note also that the condition (1.5.4) is related to the characteristic function of the stability region by the substitution $(A, AX, XA^\top) \leftrightarrow (1, s, \bar{s})$.

1.6   Further reading

Optimization: [15]
Convex function analysis: [25], [43]

Lyapunov theory: [7, 16, 27, 49]
Interior point methods: [20]
Software issues: [6]

1.7 Exercises

Exercise 1

In section 1.2 we defined sublevel sets and related them to the convexity of functions $f :$ $\mathcal{S} \to \mathbb{R}$. Define a suitable notion of *suplevel sets* (yes, this is a "p") and formulate and prove a sufficient condition (in the spirit of proposition 1.13) for suplevel sets to be compact.

Exercise 2

Give an example of a non-convex function $f : \mathcal{S} \to \mathbb{R}$ whose sublevel sets $\mathcal{S}_\alpha$ are convex for all $\alpha \in \mathbb{R}$.

Exercise 3

Let $f : \mathcal{S} \to \mathbb{R}$ be a convex function.

1. Show the so called *Jensen's inequality* which states that for a convex combination $x = \sum_{i=1}^{n} \alpha_i x_i$ of $x_1, \ldots x_n \in \mathcal{S}$ there holds that

$$f(\sum_{i=1}^{n} \alpha_i x_i) \leqslant \sum_{i=1}^{n} \alpha_i f(x_i).$$

   Hint: A proof by induction on n may be the easiest.

2. Show that $\mathrm{co}(\mathcal{S})$ is equal to the set of all convex combinations of $\mathcal{S}$

Exercise 4

Run the MATLAB demo `lmidem`.

Exercise 5

Use a feasibility test of the LMI toolbox to verify the asymptotic stability of the system $\dot{x} = Ax$, where

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & -3 & -4 \end{pmatrix}.$$

To do this, use the routine `ltisys` to convert a state space model to an internal format which is used for the LTI toolbox. Use the routine `feasp` to compute a solution to the corresponding LMI feasibility problem. (See the MATLAB help information for these routines).

Exercise 6

In this exercise we investigate the stability of the linear time-varying system

$$\dot{x} = A(t)x \qquad\qquad\qquad (1.7.1)$$

where for all $t \in \mathbb{R}_+$ the matrix $A(t)$ is a convex combination of the triple

$$A_1 := \begin{pmatrix} -1 & 1 \\ -1 & -0.2 \end{pmatrix}, \quad A_2 := \begin{pmatrix} -1 & 1 \\ -2 & -0.7 \end{pmatrix}, \quad A_3 := \begin{pmatrix} -2 & 1 \\ -1.2 & 0.4 \end{pmatrix}.$$

That is,

$$A(t) \in \mathrm{co}(A_1, A_2, A_3)$$

for all values of $t \in \mathbb{R}_+$. This is referred to as a *polytopic model*. (See subsection 1.5.3). It is an interesting fact that the time-varying system (1.7.1) is asymptotically stable if there exists a $X = X^\top > 0$ such that

$$A_1^\top X + X A_1 < 0$$
$$A_2^\top X + X A_2 < 0$$
$$A_3^\top X + X A_3 < 0.$$

If such an X exists then (1.7.1) is stable *irrespective of how fast the time variations of $A(t)$ take place*!

1. Reformulate the question of asymptotic stability of (1.7.1) as a feasibility problem.

2. The MATLAB function `quadstab` tests the (quadratic) stability of polytopic models. To invoke this routine, first use `ltisys` to represent the state space systems $\dot{x} = A_i x$ for $i = 1, 2, 3$ in internal LMI format. Then define the polytopic model (1.7.1) by using `psys`.

3. Give a Lyapunov function for this model.

A second approach to solve the feasibility problem amounts to using the graphical interface `lmiedit` *to enter the relevant linear matrix inequalities in the unknown* X. *The procedure* `feasp` *of the LMI toolbox solves the feasibility problem and thus checks whether* (1.7.1) *is asymptotically stable. We leave this alternative approach as an option.*

Exercise 7

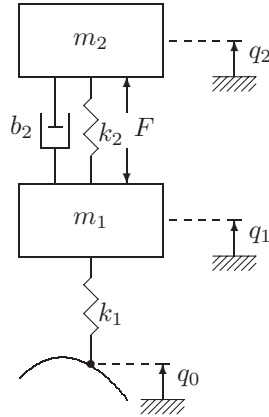Consider the active suspension system of a transport vehicle as depicted in Figure 1.1. The

Figure 1.1: Model for suspension system

system is modeled by the equations

$$m_2\ddot{q}_2 + b_2(\dot{q}_2 - \dot{q}_1) + k_2(q_2 - q_1) - F = 0$$
$$m_1\ddot{q}_1 + b_2(\dot{q}_1 - \dot{q}_2) + k_2(q_1 - q_2) + k_1(q_1 - q_0) + F = 0$$

where F (resp. $-F$) is a force acting on the chassis mass m_2 (the axle mass m_1). Here, $q_2 - q_1$ is the distance between chassis and axle, and $\ddot{q}_2$ denotes the acceleration of the chassis mass m_2. b_2 is a damping coefficient and k_1 and k_2 are spring coefficients. The

m_1	m_2	k_1	k_2	b_2
1.5×10^3	1.0×10^4	5.0×10^6	5.0×10^5	50×10^3

Table 1.1: Physical parameters

variable q_0 represents the road profile. A 'real life' set of system parameters is given in Table 1.1.

1. Consider the case where $F = 0$ and $q_0 = 0$ (thus no active force between chassis and

axle and 'flat' road characteristic) and compute the eigenmodes of this autonomous system.

2. Determine a Lyapunov function $V(q_1, q_2, \dot{q}_1, \dot{q}_2)$ of this system (with $F = 0$ and $q_0 = 0$) and show that its derivative is negative along solutions of the autonomous behavior of the system (i.e. $F = 0$ and $q_0 = 0$).

3. Design your favorite road profile q_0 in MATLAB and simulate the response of the system to this road profile (the force F is kept 0). Plot the variables q_1 and q_2. What are your conclusions?

4. Depending on the load of the truck, the chassis mass m_2 is a major source of uncertainty. Simulate the response of the system to the road profile q_0 when the mass m_2 varies between its minimum value $m_{2,\min} = 1.5 \times 10^3$ (unloaded) and its maximum value $m_{2,\max} = 1.0 \times 10^4$ (fully loaded). Moreover, it is reasonable to assume that the air spring stiffness k_2 is influenced by the load according to

$$k_2 = \frac{m_2 \times k_{2,\max}}{m_{2,\max}}$$

where $k_{2,\max} = 5.0 \times 10^5$. Introduce an arbitrary, non-constant function $m_2(t)$ reflecting the time-varying changes in the chassis mass m_2 and let the air spring stiffness k_2 depend on m_2 as above. Use the routines introduced in Subsection 1.5.4 for a time simulation of this situation and plot the variables q_1 and q_2.

Chapter 2

Dissipativity and robustness analysis

2.1 Dissipative dynamical systems and LMI's

2.1.1 Introduction

The notion of dissipativity is a most important concept in systems theory both for theoretical consideratons as well as from a practical point of view. Especially in the 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, an increase of entropy, mass, electromagnetic 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 section we will formalize the notion of a dissipative dynamical system for a very general class of systems. It will be shown that linear matrix inequalities occur in a very natural way in the study of linear dissipative systems. Perhaps the most appealing framework 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, robustness, and H_∞ and H_2 controller design problems. More details on the theory of this section can be found in [45, 47, 48].

2.1.2   Dissipativity

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

$$\dot{x} = f(x, u) \tag{2.1.1a}$$

$$y = g(x, u) \tag{2.1.1b}$$

Here, x is the state which takes its values in a *state space* X, u is the input taking its values in an *input space* U and y denotes the output of the system which assumes its values in the *output space* Y. Throughout this section, the precise representation of the systems will not be relevant. What we need, though, is that for any initial condition $x(0) = x_0$ of (2.1.1a) and for any input u belonging to an input class $\mathcal{U}$, the state x and the output y are uniquely defined and depend on u in a causal way. The system (2.1.1) therefore generates outputs from inputs and initial conditions. Let

$$s : U \times Y \to \mathbb{R}$$

be a mapping and assume that for all $t_0, t_1 \in \mathbb{R}$ and for all input-output pairs u, y satisfying (2.1.1) the composite function

$$s(t) := s(u(t), y(t))$$

is locally integrable, i.e., $\int_{t_0}^{t_1} |s(t)| dt < \infty$. (We do realize that we abuse notation here). The mapping s will be referred to as the *supply function*.

Definition 2.1 (Dissipativity) The system Σ with supply rate s is said to be *dissipative* if there exists a non-negative function $V : X \to \mathbb{R}$ such that

$$V(x(t_0)) + \int_{t_0}^{t_1} s(u(t), y(t)) dt \ \geqslant \ V(x(t_1)) \tag{2.1.2}$$

for all $t_0 \leqslant t_1$ and all trajectories (u, x, y) which satisfy (2.1.1).

Interpretation 2.2 The supply function (or *supply rate*) s should be interpreted as the supply *delivered to the system*. This means that in a time interval $[0, t]$ work has been done *on* the system whenever $\int_0^t s(\tau) d\tau$ is positive, while work is done *by* the system if this integral is negative. The non-negative function V is called a *storage function* and generalizes the notion of an energy function for a dissipative system. With this interpretation, inequality (2.1.2) formalizes the intuitive 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 $[t_0, t_1]$ will never

exceed the amount of supply that flows into the system (or the 'work done on the system'). This means that part of what is supplied to the system is stored, while the remaining part is dissipated. Inequality (2.1.2) is known as the *dissipation inequality*.

Remark 2.3 Since the system communicates with its environment through the variables u and y, it is logical to assume that the supply function s is a function defined on the external signal space $U \times Y$. Moreover, since storage is a concept related to the status of internal variables of the system, it is logical to define storage functions as *state functions*.

Remark 2.4 Note that whenever the composite function $V(x(\cdot))$ with V a storage function and $x : \mathbb{R} \to X$ a state trajectory satisfying(2.1.1a), is differentiable as a function of time, then (2.1.2) can be equivalently written as

$$\dot{V}(t) \leqslant s(u(t), y(t)). \tag{2.1.3}$$

Remark 2.5 There are a few refinements to Definition 2.1 which are worth mentioning. The system Σ is said to be *conservative* (or *lossless*) if equality holds in (2.1.2) for all $t_0 \leqslant t_1$ and all (u, x, y) which satisfy (2.1.1). Also, Definition 2.1 can be generalized to *time-varying* systems by letting the supply rate s depend on time. We will not need such a generalization for our purposes. Many authors have proposed a definition of dissipativity for discrete time systems, but since we can not think of any physical example of such a system, there seems little point in doing this. Another refinement consists of the idea that a system Σ may be dissipative with respect to more than one supply function. See Example 2.8 below.

The notion of strict dissipativity is a refinement of Definition 2.1 which we will use in the sequel. It is defined as follows.

Definition 2.6 (Strict dissipativity) The system Σ with supply rate s is said to be *strictly dissipative* if there exists an $\epsilon > 0$ and a non-negative function $V : X \to \mathbb{R}$ which attains a strong global minimum (see Chapter 1) such that the dissipation inequality

$$V(x(t_0)) + \int_{t_0}^{t_1} \left(s(u(t), y(t)) - \epsilon^2 \|u(t)\|^2 \right) dt \; \geqslant \; V(x(t_1)) \tag{2.1.4}$$

holds for all $t_0 \leqslant t_1$ and all trajectories (u, x, y) which satisfy (2.1.1).

Clearly, a system is strictly dissipative only if a strict inequality holds in (2.1.2). Note that storage functions of strictly dissipative systems are assumed to have a strong global minimum at some point $x^* \in X$. This means that for any state x in a neighborhood of x^* the storage $V(x)$ is strictly larger than $V(x^*)$.

Example 2.7 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 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 input and output variables are taken to be effort and flow variables and the supply function s is invariably taken to be the product of these two variables. A bond-graph is therefore a special case of a dissipative system (and not the other way around!).

Example 2.8 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 or chemical principles or through experimentation– the mathematical model of the thermodynamic system has been decided upon and is given by the time invariant system (2.1.1). The first and second law of thermodynamics may then be formulated in the sense of Definition 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 two basic laws of thermodynamics state that for all system trajectories (T, Q, W) and all time instants $t_0 \leqslant t_1$

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

(which is conservation of thermodynamical energy) and the second law of thermodynamics states that the system trajectories satisfy

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

for a storage function S. 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. Note that thermodynamical systems are dissipative with respect to more than one supply function!

Example 2.9 Typical examples of supply functions $s : U \times Y \to \mathbb{R}$ are

$$s(u, y) = u^\top y,$$
$$s(u, y) = \|y\|^2 - \|u\|^2$$
$$s(u, y) = \|y\|^2 + \|u\|^2$$
$$s(u, y) = \|y\|^2$$

which arise in network theory, bondgraph theory, scattering theory, H_∞ theory, game theory and LQ-optimal control and H_2-optimal control theory.

If Σ is dissipative with storage function V then we will assume that there exists a reference point $x^* \in X$ of *minimal storage*, i.e. x^* is such that $V(x^*) = \min_{x \in X} V(x)$. Given a storage

function V, its *normalization* (with respect to x^*) is defined as $\bar{V}(x) := V(x) - V(x^*)$. Obviously, $\bar{V}$ is a storage function of Σ whenever V is and $\bar{V}(x^*) = 0$.

Instead of considering the set of all possible storage functions associated with a dynamical system Σ we will restrict attention to the set of *normalized storage functions*. Formally, the set of normalized storage functions (associated with (Σ, s)) is defined by

$$\mathcal{V}(x^*) := \{V : X \to \mathbb{R}_+ \mid V(x^*) = 0 \text{ and } (2.1.2) \text{ holds}\}.$$

The existence of a reference point x^* of minimal storage implies that for a dissipative system

$$\int_0^{t_1} s(u(t), y(t)) \, dt \geqslant 0$$

for any $t_1 \geqslant 0$ and any (u, x, y) satisfying (2.1.1) with $x(0) = x^*$. Stated otherwise, any trajectory of the system which emanates from x^* has the property that the net flow of supply is *into* the system. In many treatments of dissipativity this property is often taken as definition of *passivity*. We introduce two mappings $V_{\mathrm{av}} : X \to \mathbb{R}_+ \cup \infty$ and $V_{\mathrm{req}} : X \to \mathbb{R} \cup \{-\infty\}$ which will play a crucial role in the sequel. They are defined by

$$V_{\mathrm{av}}(x_0) := \sup\left\{ -\int_0^{t_1} s(t) \, dt \mid t_1 \geqslant 0; \ (u, x, y) \text{ satisfy } (2.1.1) \text{ with } x(0) = x_0 \right\} \quad (2.1.5a)$$

$$V_{\mathrm{req}}(x_0) := \inf\left\{ \int_{t_{-1}}^0 s(t) \, dt \mid t_{-1} \leqslant 0; \ (u, x, y) \text{ satisfy } (2.1.1) \text{ with} \right. \quad (2.1.5b)$$

$$x(0) = x_0 \text{ and } x(t_{-1}) = x^*\} \quad (2.1.5c)$$

Then $V_{\mathrm{av}}(x)$ denotes the maximal amount of internal storage that may be recovered from the system over all state trajectories starting from x. Similarly, $V_{\mathrm{req}}(x)$ reflects the minimal supply the environment has to deliver to the system in order to excite the state x via any trajectory in the state space originating in x^*. We refer to V_{av} and V_{req} as the *available storage* and the *required supply*, respectively. Note that in (2.1.5b) it is assumed that the point $x_0 \in X$ is reachable from the reference pont x^*, i.e. it is assumed that there exist a control input u which brings the state trajectory x from x^* at time $t = t_{-1}$ to x_0 at time $t = 0$. This is the case when the system Σ is controllable.

Proposition 2.10 (Willems) *Let the system Σ be described by (2.1.1) and let s be a supply function. Then*

1. *Σ is dissipative if and only if $V_{av}(x)$ is finite for all $x \in X$.*

2. *If Σ is dissipative and controllable then*

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

   (b) *$\{V \in \mathcal{V}(x^*)\} \ \Rightarrow \ \{0 \leqslant V_{av} \leqslant V \leqslant V_{req}\}$.*

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

Interpretation 2.11 Proposition 2.10 gives a necessary and sufficient condition for a system to be dissipative. It shows that both the available storage and the required supply are possible storage functions. Moreover, statement (b) shows that the available storage and the required supply are the *extremal storage functions* in $\mathcal{V}(x^*)$. In particular, for any state of a dissipative system, the available storage is at most equal to the required supply. In addition, convex combinations of the available storage and the required supply are candidate storage functions for a dissipative system.

Proof. 1. Let Σ be dissipative, V a storage function and $x_0 \in X$. From (2.1.2) it then follows that for all $t_1 \geqslant 0$ and all (u,x,y) satisfing (2.1.1) with $x(0) = x_0$,

$$-\int_0^{t_1} s(u(t), y(t))dt \leqslant V(x_0) < infty.$$

Taking the supremum over all $t_1 \geqslant 0$ and all such trajectories (u,x,y) (with $x(0) = x_0$) yields that $V_{av}(x_0) \leqslant V(x_0) < \infty$. To prove the converse implication it suffices to show that V_{av} is a storage function. To see this, first note that $V_{av}(x) \geqslant 0$ for all $x \in X$ (take $t_1 = 0$ in (2.1.5a)). To prove that V_{av} satisfies (2.1.2), let $t_0 \leqslant t_1 \leqslant t_2$ and (u,x,y) satisfy (2.1.1). Then

$$V_{av}(x(t_0)) \geqslant -\int_{t_0}^{t_1} s(u(t), y(t))dt - \int_{t_1}^{t_2} s(u(t), y(t))dt.$$

Since the second term in the right hand side of this inequality holds for arbitrary $t_2 \geqslant t_1$ and arbitrary $(u,x,y)|_{[t_1,t_2]}$ (with $x(t_1)$ fixed), we can take the supremum over all such trajectories to conclude that

$$V_{av}(x(t_0)) \geqslant -\int_{t_0}^{t_1} s(u(t), y(t))dt - V_{av}(x(t_1)).$$

which shows that V_{av} satisfies (2.1.2).

2a. Suppose that Σ is dissipative and let V be a storage function. Then $\bar{V}(x) := V(x) - V(x^*) \in \mathcal{V}(x^*)$ so that $\mathcal{V}(x^*) \neq \varnothing$. Observe that $V_{av}(x^*) \geqslant 0$ and $V_{req}(x^*) \leqslant 0$ (take $t_1 = t_{-1} = 0$ in (2.1.5)). Suppose that the latter inequalities are strict. Then, using controllability of the system, there exists $t_{-1} \leqslant 0 \leqslant t_1$ and a state trajectory x with $x(t_{-1}) = x(0) = x(t_1) = x^*$ such that $-\int_0^{t_1} s(t)dt > 0$ and $\int_{t_{-1}}^0 s(t)dt < 0$. But this yields a contradiction with (2.1.2) as both $\int_0^{t_1} s(t)dt \geqslant 0$ and $\int_{t_{-1}}^0 s(t)dt \geqslant 0$. Thus, $V_{av}(x^*) = V_{req}(x^*) = 0$. We already proved that V_{av} is a storage function so that $V_{av} \in \mathcal{V}(x^*)$. Along the same lines one shows that also $V_{req} \in \mathcal{V}(x^*)$.

2b. If $V \in \mathcal{V}(x^*)$ then

$$-\int_0^{t_1} s(u(t), y(t))dt \leqslant V(x_0) \leqslant \int_{t_{-1}}^0 s(u(t), y(t))dt$$

for all $t_{-1} \leqslant 0 \leqslant t_1$ and (u, x, y) satisfying (2.1.1) with $x(t_{-1}) = x^*$ and $x(0) = x_0$. Now take the supremum and infimum over all such trajectories to obtain that $V_{\text{av}} \leqslant V \leqslant V_{\text{req}}$.

2c. Follows trivially from (2.1.2). ∎

If the system Σ is dissipative with respect to the supply function s then

$$V(x(t_0)) + \int_{t_0}^{t_1} s(u(t), y(t)) \, dt - V(x(t_1))$$

is a non-negative quantity that can be interpreted as the amount of supply that is dissipated in the system in the time interval $[t_0, t_1]$.

Definition 2.12 (Dissipation functions) A function $d : X \times U \to \mathbb{R}$ is called a *dissipation function* for (Σ, s) if there exists a storage function $V : X \to \mathbb{R}_+$ such that

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

hold for all $t_0 \leqslant t_1$ and all trajectories (u, x, y) which satisfy (2.1.1).

Obviously, the system Σ is dissipative with respect to the supply function s if and only if the dissipation function d is non-negative in the sense that $d(x, u) \geqslant 0$ for all x, u. Note that if d is a dissipation function then the system is conservative with respect to the supply function $s - d$.

2.1.3 Storage functions and Lyapunov functions

Storage functions and Lyapunov functions (introduced in the previous chapter) are closely related. Indeed, if $u(t) = u^*$ with $u^* \in U$ is taken as a constant input in (2.1.1) then we obtain the *autonomous system*

$$\dot{x} = f(x, u^*)$$
$$y = g(x, u^*).$$

Let x^* be an equilibrium point of this system and suppose that the system defined by (2.1.1) is dissipative with supply

$$s(u^*, y) = s(u^*, g(x, u^*)) \leqslant 0$$

for all x in a neigborhood of x^*. From Remark 2.4 we then infer that any (differentiable) storage function V of this system is non-negative and monotone non-increasing along solutions in a neighborhood of x^*. Consequently, by Lyapunov's theorem, x^* is a stable equilibrium if the storage function V attains a strong local minimum at x^*. In that case, the storage function V is nothing else than a *Lyapunov function* defined in a neighborhood of x^*. (See the corresponding definitions in the previous chapter).

2.2   Linear dissipative systems with quadratic supply rates

In this section we will apply the above theory of dissipative dynamical systems to linear input-output systems Σ described by the equations

$$\dot{x} = Ax + Bu$$
$$y = Cx + Du \tag{2.2.1}$$

with state space $X = \mathbb{R}^n$, input space $U = \mathbb{R}^m$ and output space $Y = \mathbb{R}^p$.

Consider a general *quadratic supply function* $s : U \times Y \to \mathbb{R}$ defined by

$$s(u, y) = \begin{pmatrix} y \\ u \end{pmatrix}^\top \begin{pmatrix} Q_{yy} & Q_{yu} \\ Q_{uy} & Q_{uu} \end{pmatrix} \begin{pmatrix} y \\ u \end{pmatrix} \tag{2.2.2}$$

Here,

$$Q := \begin{pmatrix} Q_{yy} & Q_{yu} \\ Q_{uy} & Q_{uu} \end{pmatrix}$$

is a real symmetric matrix which is partitioned conformally with u and y. We emphasize that no *a priori* definiteness assumptions are made on Q.

Note that the supply functions given in Example 2.9 can all be written in the form (2.2.2) for an appropriate partitioned real symmetric matrix Q.

Remark 2.13 Substituting the output equation $y = Cx + Du$ in the supply function (2.2.2) shows that (2.2.2) can equivalently be viewed as a quadratic function in the variables u and x. Indeed,

$$s(u, y) = s(u, Cx + Du) = \begin{pmatrix} x \\ u \end{pmatrix}^\top \begin{pmatrix} Q_{xx} & Q_{xu} \\ Q_{ux} & Q_{uu} \end{pmatrix} \begin{pmatrix} x \\ u \end{pmatrix}$$

where

$$\begin{pmatrix} Q_{xx} & Q_{xu} \\ Q_{ux} & Q_{uu} \end{pmatrix} = \begin{pmatrix} C & D \\ 0 & I \end{pmatrix}^\top \begin{pmatrix} Q_{yy} & Q_{yu} \\ Q_{uy} & Q_{uu} \end{pmatrix} \begin{pmatrix} C & D \\ 0 & I \end{pmatrix}.$$

The following proposition is the main result of this section. It provides necessary and sufficient conditions for dissipativeness of the pair (Σ, s) and it shows that linear dissipative systems with quadratic supply functions have at least one quadratic storage function.

Proposition 2.14 *Suppose that the system Σ described by (2.2.1) is controllable and let the supply function s be defined by (2.2.2). Then the following statements are equivalent.*

    1. (Σ, s) is dissipative.

2. *There exists a quadratic storage function $V(x) := x^\top K x$ with $K = K^\top \geqslant 0$ such that (2.1.2) holds for all $t_0 \leqslant t_1$ and (u, x, y) satisfying (2.2.1).*

3. *There exists $K = K^\top \geqslant 0$ such that*

$$F(K) := -\begin{pmatrix} A^\top K + KA & KB \\ B^\top K & 0 \end{pmatrix} + \begin{pmatrix} C & D \\ 0 & I \end{pmatrix}^\top \begin{pmatrix} Q_{yy} & Q_{yu} \\ Q_{uy} & Q_{uu} \end{pmatrix} \begin{pmatrix} C & D \\ 0 & I \end{pmatrix} \geqslant 0 \quad (2.2.3)$$

4. *There exists $K_- = K_-^\top \geqslant 0$ such that $V_{av}(x) = x^\top K_- x$.*

5. *There exists $K_+ = K_+^\top \geqslant 0$ such that $V_{req}(x) = x^\top K_+ x$.*

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

Proof. (1⇒2,4). If (Σ, s) is dissipative then we infer from Proposition 2.10 that the available storage $V_{av}(x)$ is finite for any $x \in \mathbb{R}^n$. We claim that $V_{av}(x)$ is a quadratic function of x. This follows from [46] upon noting that the supply function s is quadratic and that

$$V_{av}(x) = \sup - \int_0^{t_1} s(t)dt = -\inf \int_0^{t_1} s(t)dt$$

denotes the optimal cost of a linear quadratic optimization problem. It is well known that this infimum is a quadratic form in x.

(4⇒1). Obvious from Proposition (2.10).

(2⇒3). If $V(x) = x^\top K x$ with $K \geqslant 0$ is a storage function then the dissipation inequality can be rewritten as

$$\int_{t_0}^{t_1} \left(-\frac{d}{dt} x(t)^\top K x(t) + s(u(t), y(t)) \right) dt \geqslant 0.$$

Substituting the system equations (2.2.1), this is equivalent to

$$\int_{t_0}^{t_1} \begin{pmatrix} x(t) \\ u(t) \end{pmatrix}^\top \underbrace{-\begin{pmatrix} A^\top K + KA & KB \\ B^\top K & 0 \end{pmatrix} + \begin{pmatrix} C & D \\ 0 & I \end{pmatrix}^\top \begin{pmatrix} Q_{yy} & Q_{yu} \\ Q_{uy} & Q_{uu} \end{pmatrix} \begin{pmatrix} C & D \\ 0 & I \end{pmatrix}}_{F(K)} \begin{pmatrix} x(t) \\ u(t) \end{pmatrix} dt \geqslant 0.$$

$$(2.2.4)$$

Since (2.2.4) holds for all $t_0 \leqslant t_1$ and all inputs u this reduces to the requirement that $K \geqslant 0$ satisfies the LMI $F(K) \geqslant 0$.

(3⇒2). Conversely, if there exist $K \geqslant 0$ such that $F(K) \geqslant 0$ then (2.2.4) holds and it follows that $V(x) = x^\top K x$ is a storage function which satisfies the dissipation inequality.

(1⇔5). If (Σ, s) is dissipative then by Proposition (2.10), V_{req} is a storage function. Since V_{req} is defined as an optimal cost corresponding to a linear quadratic optimization problem,

V_{req} is quadratic. Hence, if the reference point $x^* = 0$, $V_{\text{req}}(x)$ is of the form $x^\top K_+ x$ for some $K_+ \geqslant 0$. Conversely, if $V_{\text{req}} = x^\top K_+ x$, $K_+ \geqslant 0$, then it is easily seen that V_{req} satisfies the dissipation inequality (2.1.2) which implies that (Σ, s) is dissipative.   ■

We recognize in (2.2.3) a non-strict linear matrix inequality. The matrix $F(K)$ is usually called the *dissipation matrix*. Observe that in the above proposition the set of *quadratic storage functions* in $\mathcal{V}(0)$ is completely characterized by the inequalities $K \geqslant 0$ and $F(K) \geqslant 0$. In other words, the set of normalized quadratic storage functions associated with (Σ, s) coincides with the feasibility set of the system of LMI's $K = K^\top \geqslant 0$ and $F(K) \geqslant 0$. In particular, the available storage and the required supply are quadratic storage functions and hence K_- and K_+ also satisfy $F(K_-) \geqslant 0$ and $F(K_+) \geqslant 0$. Using Proposition 2.10 it moreover follows that any solution $K = K^\top \geqslant 0$ of $F(K) \geqslant 0$ has the property that

$$0 \leqslant K_- \leqslant K \leqslant K_+.$$

In other words, among the set of positive semi-definite solutions K of the LMI $F(K) \geqslant 0$ there exists a smallest and a largest element.

For conservative and strictly dissipative systems with quadratic supply functions a similar characterization can be given. The precize formulation for conservative systems is evident from Proposition 2.14 and is left to the reader. For strictly dissipative system the result is worth mentioning.

Proposition 2.15 *Suppose that the system Σ described by (2.2.1) is controllable and let the supply function s be defined by (2.2.2). Then the following statements are equivalent.*

1. *(Σ, s) is strictly dissipative.*

2. *There exists $K = K^\top > 0$ such that*

$$F(K) := -\begin{pmatrix} A^\top K + KA & KB \\ B^\top K & 0 \end{pmatrix} + \begin{pmatrix} C & D \\ 0 & I \end{pmatrix}^\top \begin{pmatrix} Q_{yy} & Q_{yu} \\ Q_{uy} & Q_{uu} \end{pmatrix} \begin{pmatrix} C & D \\ 0 & I \end{pmatrix} > 0 \quad (2.2.5)$$

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

Proof.   The proof proceeds along the same lines as the proof of Proposition 2.14. We only prove the implication $(2 \Rightarrow 1)$ here. Let $K > 0$ be such that $F(K) > 0$. Then obviously, there exists $\epsilon > 0$ such that

$$F(K) - \begin{pmatrix} 0 & 0 \\ 0 & \epsilon^2 I \end{pmatrix} \geqslant 0.$$

But then

$$\int_{t_0}^{t_1} \begin{pmatrix} x(t) \\ u(t) \end{pmatrix}^\top F(K) \begin{pmatrix} x(t) \\ u(t) \end{pmatrix} dt \geqslant \int_{t_0}^{t_1} \begin{pmatrix} x(t) \\ u(t) \end{pmatrix}^\top \begin{pmatrix} 0 & 0 \\ 0 & \epsilon^2 I \end{pmatrix} \begin{pmatrix} x(t) \\ u(t) \end{pmatrix} dt = \int_{t_0}^{t_1} \epsilon^2 |u(t)|^2 dt$$

from which it follows that $V(x) = x^\top K x$ is a storage function with a strong global minimum at $x = 0$ which satisfies the dissipation inequality (2.1.4). Hence, the pair (Σ, s) is strictly dissipative. ∎

There is a simple relation between the dissipation matrix $F(K)$ and dissipation functions. Indeed, if $K = K^\top \geqslant 0$ (or > 0) is such that $F(K) \geqslant 0$ then the dissipation matrix can be factorized as

$$F(K) = \begin{pmatrix} M_K & N_K \end{pmatrix}^\top \begin{pmatrix} M_K & N_K \end{pmatrix}.$$

where $\begin{pmatrix} M_K & N_K \end{pmatrix}$ is a real matrix with $n + m$ columns and $r \geqslant r_K := \operatorname{rank}(F(K))$ rows. For any triple (u, x, y) satisfying (2.2.1) we then obtain that

$$\int_{t_0}^{t_1} \|M_K x(t) + N_K u(t)\|^2 \, dt = \int_{t_0}^{t_1} \begin{pmatrix} x \\ u \end{pmatrix}^\top F(K) \begin{pmatrix} x \\ u \end{pmatrix} dt$$

$$= x(t_0)^\top K x(t_0) - x(t_1)^\top K x(t_1) + \int_{t_0}^{t_1} s(t) \, dt.$$

In other words the function

$$d(x, u) := \|M_K x + N_K u\|^2$$

is a dissipation function of the system (2.2.1).

We conclude this subsection with a frequency domain characterization of dissipativeness. Let $G(s) := C(Is - A)^{-1} B + D$ denote the *transfer function* associated with (2.2.1).

Proposition 2.16 (Yakubovich-Kalman) *Under the same conditions as Proposition 2.14, the following statements are equivalent.*

1. (Σ, s) *is dissipative.*

2. *For all $\omega \in \mathbb{R}$ with $\det(i\omega I - A) \neq 0$, there holds*

$$\begin{pmatrix} G(i\omega) \\ I \end{pmatrix}^* \begin{pmatrix} Q_{yy} & Q_{yu} \\ Q_{uy} & Q_{uu} \end{pmatrix} \begin{pmatrix} G(i\omega) \\ I \end{pmatrix} \geqslant 0 \tag{2.2.6}$$

Proof.   Because of the simplicity of the argument, we only prove the implication (1⇒2) here. Let $\omega \in \mathbb{R}$ be such that $\det(i\omega I - A) \neq 0$ and consider the (complex) input $u(t) = \exp(i\omega t) u_0$ with $u_0 \in \mathbb{R}^m$. Define $x(t) := \exp(i\omega t)(i\omega I - A)^{-1} B u_0$ and $y(t) := C x(t) + D u(t)$. Then $y(t) = \exp(i\omega t) G(i\omega) u_0$ and the triple (u, x, y) satisfies (2.2.1). Moreover,

$$s(u(t), y(t)) = \bar{u}_0^* \begin{pmatrix} G(i\omega) \\ I \end{pmatrix}^* \begin{pmatrix} Q_{yy} & Q_{yu} \\ Q_{uy} & Q_{uu} \end{pmatrix} \begin{pmatrix} G(i\omega) \\ I \end{pmatrix} u_0$$

which is a constant for all time $t \in \mathbb{R}$. Now suppose that (Σ, s) is dissipative. For non-zero frequencies ω the triple (u, x, y) is periodic with period $P = 2\pi/\omega$. In particular, there

must exist a time instant t_0 such that $x(t_0) = x(t_0 + kP) = 0$, $k \in \mathbb{Z}$. Since $V(0) = 0$, the dissipation inequality (2.1.2) reads

$$\int_{t_0}^{t_1} s(u(t), y(t)) \, dt = \int_{t_0}^{t_1} \bar{u}_0^* \begin{pmatrix} G(i\omega) \\ I \end{pmatrix}^* \begin{pmatrix} Q_{yy} & Q_{yu} \\ Q_{uy} & Q_{uu} \end{pmatrix} \begin{pmatrix} G(i\omega) \\ I \end{pmatrix} u_0$$

$$= (t_1 - t_0) \bar{u}_0^* \begin{pmatrix} G(i\omega) \\ I \end{pmatrix}^* \begin{pmatrix} Q_{yy} & Q_{yu} \\ Q_{uy} & Q_{uu} \end{pmatrix} \begin{pmatrix} G(i\omega) \\ I \end{pmatrix} u_0 \; \geqslant \; 0$$

for all $t_1 > t_0$. Since u_0 and $t_1 > t_0$ are arbitrary this yields that statement 2 holds.

The converse implication is much more involved and will be omitted here.   ■

The frequency domain characterization of strictly dissipative systems is similar and involves a strict inequality in (2.2.6). Proposition 2.16 has a long history in system theory. The result goes back to V.A. Yakubovich (1962) and R. Kalman (1963).

2.2.1   The positive real lemma

Consider the system (2.2.1) together with the quadratic supply function $s(u, y) = y^\top u$ as defined in 2.2.2. The following is then an immediate consequence of Proposition 2.14.

Corollary 2.17 *Suppose that the system Σ described by (2.2.1) is controllable and has transfer function G. Let $s(u, y) = y^\top u$ be a supply function. Then equivalent statements are*

1. *(Σ, s) is dissipative.*

2. *the system of LMI's*

$$K = K^\top \geqslant 0$$

$$\begin{pmatrix} -A^\top K - KA & -KB + C^\top \\ -B^\top K + C & D + D^\top \end{pmatrix} \; \geqslant \; 0$$

   is feasible

3. *For all $\omega \in \mathbb{R}$ with $\det(i\omega I - A) \neq 0$ $G(i\omega)^* + G(i\omega) \geqslant 0$.*

Moreover, $V(x) = x^\top K x$ defines a quadratic storage function if and only if K satisfies the above system of LMI's.

Remark 2.18 Corollary 2.17 is known as the Kalman-Yacubovich-Popov or the *positive real lemma* and 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*.

2.2.2   The bounded real lemma

Consider the quadratic supply function

$$s(u, y) = \gamma^2 u^\top u - y^\top y \tag{2.2.7}$$

where $\gamma \geqslant 0$. In a similar fashion we obtain the following result as an immediate consequence of Proposition 2.14.

Corollary 2.19 *Suppose that the system Σ described by (2.2.1) is controllable and has transfer function G. Let $s(u, y) = \gamma^2 u^\top u - y^\top y$ be a supply function. Then equivalent statements are*

1. *(Σ, s) is dissipative.*

2. *The system of LMI's*

$$K = K^\top \geqslant 0$$
$$\begin{pmatrix} A^\top K + KA + C^\top C & KB + C^\top D \\ B^\top K + D^\top C & D^\top D - \gamma^2 I \end{pmatrix} \leqslant 0$$

   is feasible.

3. *For all $\omega \in \mathbb{R}$ with $\det(i\omega I - A) \neq 0$ $G(i\omega)^* G(i\omega) \leqslant \gamma^2 I$.*

Moreover, $V(x) = x^\top K x$ defines a quadratic storage function if and only if K satisfies the above system of LMI's.

Let us analyze the importance of this result for control. If Σ is dissipative with respect to the supply function (2.2.7) then we infer from Remark 2.4 that for any quadratic storage function $V(x) = x^\top K x$,

$$\dot{V} \leqslant \gamma^2 u^\top u - y^\top y \tag{2.2.8}$$

along solutions (u, x, y) of (2.2.1). Suppose that $x(0) = 0$, A has all its eigenvalues in the open left-half complex plane and the input u is taken from the set $\mathcal{L}_2$ of square integrable functions, i.e., u is such that

$$\|u\|_2^2 := \int_0^\infty u^\top(t) u(t) \, dt \ < \ \infty.$$

Then both the state x and the output y of (2.2.1) are square integrable functions and $\lim_{t\to\infty} x(t) = 0$. We can therefore integrate (2.2.8) from $t = 0$ till ∞ to obtain that for all $u \in \mathcal{L}_2$

$$\gamma^2 \|u\|_2^2 - \|y\|_2^2 \geqslant 0.$$

Equivalently,

$$\sup_{u \in \mathcal{L}_2} \frac{\|y\|_2}{\|u\|_2} \leqslant \gamma. \qquad (2.2.9)$$

The left-hand side of (2.2.9) is the so called $\mathcal{L}_2$-*induced norm* or $\mathcal{L}_2$-*gain* of the system (2.2.1). The $\mathcal{L}_2$-gain is therefore obtained as the smallest $\gamma \geqslant 0$ for which (2.2.9) holds.

Translated in terms of LMI's, we therefore obtain that the upperbound (2.2.9) holds if and only if there exists K that satisfies the linear matrix inequalities of Corollary (2.19). This provides a feasibility test, parametrized in $\gamma > 0$, to determine $\mathcal{L}_2$-gain of the system.

2.3   Nominal performance and LMI's

In this section we will use the previous results on dissipative systems to characterize a number of relevant performance criteria for dynamical systems. In view of forthcoming chapters we consider the system (2.2.1) and assume throughout this section that the system is asymptotically stable (i.e. the eigenvalues of A are in the open left-half complex plane). The following lemma is a simple but useful variation on one of the implications of Proposition 2.15.

Lemma 2.20 *Consider the following statements.*

1. *(Σ, s) is strictly dissipative*

2. *There exists $K = K^\top > 0$ such that $F(K) > 0$*

3. *There exists $\epsilon > 0$ such that*

$$\int_0^\infty \begin{pmatrix} y(t) \\ u(t) \end{pmatrix}^\top \begin{pmatrix} Q_{yy} & Q_{yu} \\ Q_{uy} & Q_{uu} \end{pmatrix} \begin{pmatrix} y(t) \\ u(t) \end{pmatrix} dt \geqslant \epsilon \int_0^\infty u(t)^\top u(t) dt - x(0)^\top K x(0). \quad (2.3.1)$$

Then statements 1. and 2. are equivalent and each of these statements implies 3.

Proof.   If $K = K^\top > 0$ is such that $K(K) > 0$ then we can choose $\epsilon > 0$ such that

$$G(K) := F(K) - \begin{pmatrix} 0 & 0 \\ 0 & \epsilon^2 I \end{pmatrix} \geqslant 0.$$

Viewing this as a dissipation inequality for a dissipative system yields that for all $u \in \mathcal{L}_2$

$$\int_0^\infty \begin{pmatrix} x(t) \\ u(t) \end{pmatrix}^\top G(K) \begin{pmatrix} x(t) \\ u(t) \end{pmatrix} dt$$

$$\int_0^\infty -\left(\frac{d}{dt} x(t)^\top K x(t) - \epsilon^2 u(t)^\top u(t) + s(u(t), y(t)) \right) dt$$

where $s(u, y)$ is the quadratic supply function given in (2.2.2). Stability of the system yields that for all $u \in \mathcal{L}_2$ the state $x(t)$ vanishes for $t \to \infty$. The inequality (2.3.1) then follows. ∎

2.3.1 H_∞ nominal performance

A popular performance measure of a stable linear time-invariant system is the H_∞ norm of its transfer function. It is defined as follows. The *transfer function* associated with (2.2.1) is defined by

$$G(s) := C(Is - A)^{-1}B + D$$

where $s \in \mathbb{C}$. Since the system is assumed to be stable, $G(s)$ is bounded for all $s \in \mathbb{C}$ with positive real part. By this, we mean that the largest singular value $\sigma_{\max}(G(s))$ is finite for all $s \in \mathbb{C}$ with $\Re s > 0$. The space H_∞ consists of all complex valued functions which are analytic in $\Re s > 0$ and for which

$$\|G\|_\infty := \sup_{\Re s > 0} \sigma_{\max}(G(s)) \ < \ \infty.$$

The left hand side of this expression satisfies the axioms of a norm and defines the H_∞ norm of G. It can be shown that each function in H_∞ has a unique extension to the imaginary axis and that in fact the H_∞ norm is given by

$$\|G\|_\infty = \sup_{\omega \in \mathbb{R}} \sigma_{\max}(G(j\omega)).$$

In words, the H_∞ norm of a transfer function is the supremum of the maximum singular value of the frequency response of the system.

Remark 2.21 Various graphical representations of frequency responses are illustrative to investigate system properties like bandwidth, gains, etc. Probably the most important one is a plot of the singular values $\sigma_i(G(j\omega))$ $(i = 1, \ldots, \min(m, p))$ viewed as function of frequency $\omega \in \mathbb{R}$. For single-input single-output systems there is only one singular value and $\sigma(G(j\omega)) = |G(j\omega)|$. A *Bode diagram* is a plot of the latter and provides information to what extent the system amplifies purely harmonic input signals with frequencies $\omega \in \mathbb{R}$. In order to interpret these diagrams one usually takes logarithmic scales on the ω axis and plots $20^{10} \log(G(j\omega))$ to get units in dB. The H_∞ norm of a transfer function is then nothing else than the highest peak value occuring in the Bode plot.

The connection between time-domain and frequency-domain norms is most clearly expressed in the following standard result.

Proposition 2.22 *Let $G \in H_\infty$ be the transfer function of the asymptotically stable system (2.2.1). Then for all initial conditions $x(0) = x_0$, $u \in \mathcal{L}_2$ implies $y \in \mathcal{L}_2$ and with $x(0) = 0$ there holds that*

$$\|G\|_\infty = \sup_{u \in \mathcal{L}_2} \frac{\|y\|_2}{\|u\|_2}.$$

Interpretation 2.23 For a stable system, the H_∞ norm of its transfer function is the $\mathcal{L}_2$-induced norm of the input-output operator associated with the system.

This yields a simple relation between the Bounded Real Lemma, Lemma 2.20 and bounds on the H_∞ norm of transfer functions:

Proposition 2.24 *If the system* (2.2.1) *is asymptotically stable then* $\|G\|_\infty < \gamma$ *whenever there exists a solution* $K = K^\top > 0$ *to the LMI*

$$\begin{pmatrix} A^\top K + KA + C^\top C & KB + C^\top D \\ B^\top K + D^\top C & D^\top D - \gamma^2 I \end{pmatrix} < 0. \tag{2.3.2}$$

Proof.   Apply Lemma 2.20 with $x_0 = 0$, $Q_{yy} = I$, $Q_{yu} = 0$, $Q_{uy} = 0$ and $Q_{uu} = -\gamma^2 I$ to infer that the existence of a solution $K > 0$ to the LMI implies that for all $u \in \mathcal{L}_2$

$$\|y\|_2^2 = \int_0^\infty y(t)^\top y(t) dt \;\leqslant\; (\gamma^2 - \epsilon) \int_0^\infty u(t)^\top u(t) dt$$
$$< \gamma^2 \int_0^\infty u(t)^\top u(t) dt = \gamma^2 \|u\|_2^2.$$

Proposition 2.22 then yields that $\|G\|_\infty < \gamma$ as desired.   ∎

Interpretation 2.25 We can compute the smallest possible upperbound of the L_2-induced gain of the system (which is the H_∞ norm of the transfer function) by minimizing $\gamma > 0$ over all variables γ and $K > 0$ that satisfy the LMI (2.3.2).

2.3.2 H_2 nominal performance

Consider the system (2.2.1) and suppose that we are interested only in the impulse responses of this system. This means, that we take impulsive inputs[1] of the form

$$u(t) = \delta(t) e_i$$

with e_i the $i - th$ basis vector in the standard basis of the input space $\mathbb{R}^m$. (i runs from 1 till m). With zero initial conditions, the corresponding output y^i belongs to $\mathcal{L}_2$ and is given by

$$y^i(t) = \begin{cases} C \exp(At) B e_i & \text{for } t > 0 \\ D e_i \delta(t) & \text{for } t = 0 \\ 0 & \text{for } t < 0 \end{cases}.$$

---

[1] Formally, the impulse δ is not a function and for this reason it is neither a signal. It requires a complete introduction to distribution theory to make these statements more precise, but we will not do this at this place.

Only if $D = 0$, the sum of the squared norms of all such impulse responses is well defined and given by

$$\sum_{i=1}^{m} \|y^i\|_2^2 = \text{trace} \int_0^\infty B^\top \exp(A^\top t) C^\top C \exp(At) B \; dt$$

$$= \text{trace} \int_0^\infty C \exp(At) BB^\top \exp(A^\top t) C^\top \; dt$$

Long ago, Parseval taught us that the latter is equal to

$$\frac{1}{2\pi} \text{trace} \int_{-\infty}^\infty G(j\omega) G(j\omega)^* \; d\omega \qquad (2.3.3)$$

where G is the transfer function of the system. The square root of this expression satisfies the axioms of a norm of the transfer function G and we will refer to the square root of (2.3.3) as the H_2 norm of G.

Precisely, H_2 is the class of complex valued functions which are analytic at $\Re s > 0$ and for which

$$\|G\|_2 := \sqrt{\frac{1}{2\pi} \sup_{\sigma > 0} \text{trace} \int_{-\infty}^\infty G(\sigma + j\omega)[G(\sigma + j\omega)]^* \; d\omega}$$

is finite. This defines the H_2 norm of G. As in H_∞, it can be shown that each function in H_2 has a unique extension to the imaginary axis and that in fact the H_2 norm is given by the square root of (2.3.3). We have the following characterization of the H_2 norm a transfer function.

Proposition 2.26 *Suppose that the system* (2.2.1) *is asymptotically stable and let* $G(s) = C(Is - A)^{-1} B + D$ *denote its transfer function. Then*

1. $\|G\|_2 < \infty$ *if and only if* $D = 0$.

2. *The following statements are equivalent*

   (a) $\|G\|_2 < \gamma$

   (b) there exists $K = K^\top > 0$ *and* Z *such that*

   $$\begin{pmatrix} A^\top K + KA & KB \\ B^\top K & -I \end{pmatrix} < 0; \quad \begin{pmatrix} K & C^\top \\ C & Z \end{pmatrix} > 0; \quad \text{trace}(Z) < \gamma^2 \qquad (2.3.4)$$

   (c) there exists $K = K^\top > 0$ *and* Z *such that*

   $$\begin{pmatrix} AK + KA^\top & KC^\top \\ CK & -I \end{pmatrix} < 0; \quad \begin{pmatrix} K & B \\ B^\top & Z \end{pmatrix} > 0; \quad \text{trace}(Z) < \gamma^2. \qquad (2.3.5)$$

Proof.   The first claim is immediate from the definition of the H_2 norm. To prove the second part, note that $\|G\|_2 < \gamma$ is equivalent to requiring that the controllability gramian $W := \int_0^\infty \exp(At) BB^\top \exp(A^\top t)\, dt$ satisfies $\mathrm{trace}(CWC^\top) < \gamma^2$. Since the controllability gramian is the unique positive definite solution to the Lyapunov equation $AW + WA^\top + BB^\top = 0$ this is equivalent to saying that there exists $X > 0$ such that

$$AX + XA^\top + BB^\top < 0; \qquad \mathrm{trace}(CXC^\top) < \gamma^2.$$

With a change of variables $K := X^{-1}$, this is equivalent to the existence of $K > 0$ and Z such that

$$A^\top K + KA + KBB^\top K < 0; \quad CK^{-1}C^\top < Z; \quad \mathrm{trace}(Z) < \gamma^2.$$

Now, using Schur complements for the first two inequalities yields that $\|G\|_2 < \gamma$ is equivalent to the existence of $K > 0$ and Z such that

$$\begin{pmatrix} A^\top K + KA & KB \\ B^\top K & -I \end{pmatrix} < 0; \quad \begin{pmatrix} K & C^\top \\ C & Z \end{pmatrix} > 0; \quad \mathrm{trace}(Z) < \gamma^2$$

which is (2.3.4) as desired. The equivalence with (2.3.5) is obtained by a direct dualization and the observation that $\|G\|_2 = \|G^\top\|_2$. ∎

Interpretation 2.27 The smallest possible upperbound of the H_2-norm of the transfer function can be calculated by minimizing the criterion $\mathrm{trace}(Z)$ over the variables $K > 0$ and Z that satisfy the LMI's defined by the first two inequalities in (2.3.4) or (2.3.5).

2.3.3   Generalized H_2 nominal performance

Consider again the system (2.2.1) and suppose that $D = 0$. Assuming that the system is stable, any input $u \in \mathcal{L}_2$ yields an output y which has finite infinity (or *amplitude*) norm

$$\|y\|_\infty := \max_{i=1,\dots,p} \sup_{t \geqslant 0} \|y_i(t)\|.$$

For fixed initial condition $x(0) = 0$ this system defines a mapping from inputs $u \in \mathcal{L}_2$ to outputs $y \in \mathcal{L}_\infty$ and we will be interested in the induced (or 'energy to peak' or 'generalized H_2') norm

$$\sup_{u \in \mathcal{L}_2} \frac{\|y\|_\infty}{\|u\|_2}.$$

The following result characterizes an upperbound on this quantity.

Proposition 2.28 *Suppose that the system* (2.2.1) *is asymptotically stable and* $D = 0$. *Then the* $\mathcal{L}_2$-$\mathcal{L}_\infty$ *induced norm (or generalized H_2 norm) is smaller than* γ *if and only if there exists a solution* $K = K^\top > 0$ *to the LMI's*

$$\begin{pmatrix} A^\top K + KA & KB \\ B^\top K & -I \end{pmatrix} < 0; \qquad \begin{pmatrix} K & C^\top \\ C & \gamma^2 I \end{pmatrix} > 0 \qquad\qquad (2.3.6)$$

Proof. Firstly, infer from Proposition 2.14 that the existence of $K > 0$ with

$$\begin{pmatrix} A^\top K + KA & KB \\ B^\top K & -I \end{pmatrix} < 0$$

is equivalent to the dissipativity of the system (2.2.1) with respect to the supply function $s(u, y) = u^\top u$. Equivalently, for all $u \in \mathcal{L}_2$ there holds

$$x(t)^\top K x(t) \leqslant \int_0^t u(\tau)^\top u(\tau) \, d\tau.$$

Secondly, using Schur complements, the LMI

$$\begin{pmatrix} K & C^\top \\ C & \gamma^2 I \end{pmatrix} > 0$$

is equivalent to the existence of an $\epsilon > 0$ such that $C^\top C < (\gamma^2 - \epsilon^2)K$. Together, this yields that for all $t \geqslant 0$

$$y(t)^\top y(t) = x(t)^\top C^\top C x(t) \leqslant (\gamma - \epsilon)x(t)^\top K x(t) \leqslant (\gamma^2 - \epsilon^2) \int_0^t u(\tau)^\top u(\tau) \, d\tau.$$

Taking the supremum over $t \geqslant 0$ yields the existence of $\epsilon > 0$ such that for all $u \in \mathcal{L}_2$

$$\|z\|_\infty^2 \; \leqslant \; (\gamma^2 - \epsilon^2)\|u\|_2^2.$$

Dividing by $\|u\|_2^2$ and taking the supremum over all $u \in \mathcal{L}_2$ then yields the result. ∎

Interpretation 2.29 The smallest possible upperbound of the $\mathcal{L}_2$-$\mathcal{L}_\infty$ gain of a system can be calculated by minimizing γ over all variables γ and $K > 0$ for which the LMI's (2.3.6) are feasible.

2.3.4 L_1 or peak-to-peak nominal performance

Consider the system (2.2.1) and assume again that the system is stable. For fixed initial condition $x(0) = 0$ this system defines a mapping from bounded amplitude inputs $u \in \mathcal{L}_\infty$ to bounded amplitude outputs $y \in \mathcal{L}_\infty$ and a relevant performance criterion is the 'peak-to-peak' or $\mathcal{L}_\infty$-induced norm of this mapping

$$\sup_{u \in \mathcal{L}_\infty} \frac{\|y\|_\infty}{\|u\|_\infty}.$$

We just remark that this induced norm is equal to the L_1 norm of the impulse response of the system. The following result gives a sufficient condition for an upperbound γ of the peak-to-peak gain of the system.

Proposition 2.30 *If there exists $K > 0$, $\lambda > 0$ and $\mu > 0$ such that*

$$\begin{pmatrix} A^\top K + KA + \lambda K & KB \\ B^\top K & -\mu I \end{pmatrix} < 0; \qquad \begin{pmatrix} \lambda K & 0 & C^\top \\ 0 & (\gamma - \mu)I & D^\top \\ C & D & \gamma I \end{pmatrix} > 0 \qquad (2.3.7)$$

then the peak-to-peak (or $\mathcal{L}_\infty$ induced) norm of the system is smaller than γ.

Proof. The first inequality in (2.3.7) implies that

$$\frac{d}{dt} x(t)^\top K x(t) + \lambda x(t) K x(t) - \mu u(t)^\top u(t) \;<\; 0.$$

for all u and x for which $\dot{x} = Ax + Bu$. Now assume that $x(0) = 0$ and $u \in \mathcal{L}_\infty$ with $\|u\|_\infty \leqslant 1$. Then, since $K > 0$ we obtain (pointwise in $t \geqslant 0$) that

$$x(t)^\top K x(t) \leqslant \frac{\mu}{\lambda}.$$

Taking a Schur complement of the second inequality in (2.3.7) yields that

$$\begin{pmatrix} \lambda K & 0 \\ 0 & (\gamma - \mu)I \end{pmatrix} - \frac{1}{\gamma - \epsilon} \begin{pmatrix} C^\top \\ D^\top \end{pmatrix} \begin{pmatrix} C & D \end{pmatrix} \;>\; 0$$

so that, pointwise in $t \geqslant 0$ and for all $\|u\|_\infty \leqslant 1$ we can write

$$y(t)^\top y(t) \leqslant (\gamma - \epsilon)[\lambda x(t)^\top K x(t) + (\gamma - \mu)u(t)^\top u(t)]$$
$$\leqslant \gamma(\gamma - \epsilon)$$

Consequently, the peak-to-peak gain of the system is smaller than γ. ∎

Remark 2.31 We emphasize that Proposition 2.30 gives only a sufficient condition for an upperbound γ of the peak-to-peak gain of the system. The minimal $\gamma \geqslant 0$ for which the there exist $K > 0$, $\lambda > 0$ and $\mu \geqslant 0$ such that (2.3.7) is satisfied is usually only an upperbound of the real peak-to-peak gain of the system.

2.4 Quadratic Lyapunov functions

An important issue in the design of control systems involves the question to what extent 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 system defined by

$$\dot{x} = A(\delta)x \qquad (2.4.1)$$

where the state matrix $A(\delta)$ is a function of a real valued parameter vector $\delta = (\delta_1, \ldots, \delta_k) \in \mathbb{R}^k$. Let $\mathcal{X} = \mathbb{R}^n$ be the state space of this system. If you like, you can think of this

(autonomous) system as a feedback interconnection of a plant and a control system. We will analyze the *robust stability* of the equilibrium point $x^* = 0$ of this system. More precisely, we analyze to what extent the equilibrium point $x^* = 0$ is asymptotically stable when δ varies in a prescribed set, say Δ, of uncertain parameters.

There are two particular cases of this robust stability problem that are of special interest.

1. the parameter vector δ is a fixed but unknown element of a parameter set $\Delta \subseteq \mathbb{R}^k$.

2. the parameter vector δ is a time varying function $\delta : \mathbb{R} \to \mathbb{R}^k$ which belongs to some set Δ of functions in $(\mathbb{R}^k)^{\mathbb{R}}$. The differential equation (2.4.1) is then to be interpreted as $\frac{dx}{dt}(t) = A(\delta(t))x(t)$.

The first case typically appears in models in which the physical parameters are fixed but only approximately known up to some accuracy. Note that for these parameters (2.4.1) defines a linear time-invariant system. The second case involves time-varying uncertain parameters. For this case one can in addition distinguish between the situations where Δ consists of one element only (known time varying perturbations) and the situation where Δ is a higher dimensional set of time functions (arbitrary time varying perturbations). Robust stability against time-varying perturbations is generally a more demanding requirement for the system than robust stability against time-invariant parameter uncertainties. This, because 2 is obviously a special case of 1. We will mainly consider the general case of time-varying parametric uncertainties in the sequel.

Remark 2.32 We emphasize that in the case of constant uncertain parameters, the system $\dot{x} = A(\delta)x$ is asymptotically stable if and only if the eigenvalues of $A(\delta)$ lie in the open left-half complex plane for all admissible perturbations. However, we emphasize that such a test of stability does not hold for time varying systems. In particular, for time-varying perturbations it is not true that the asymptotic stability of $\dot{x}(t) = A(\delta(t))x(t)$ is equivalent to the condition that the (time-varying) eigenvalues $\lambda(A(\delta(t)))$ belong to the stability region $\mathbb{C}_-$ for all admissible perturbations $\delta(\cdot)$.

2.4.1   Quadratic stability

A sufficient condition for $x^* = 0$ to be an asymptotically stable equilibrium point of (2.4.1) is the existence of a quadratic Lyapunov function

$$V(x) = x^{\top} K x$$

with $K = K^{\top} > 0$ such that

$$\dot{V} = \frac{dV(x(t))}{dt} \leqslant 0$$

along state trajectories x of (2.4.1) that originate in a neighborhood of the equilibrium $x^* = 0$.

Definition 2.33 (Quadratic stability) The system (2.4.1) is said to be *quadratically stable for perturbations* Δ if there exists a matrix $K = K^\top > 0$ such that

$$A(\delta(t))^\top K + KA(\delta(t)) < 0$$

for all perturbations $\delta \in \Delta$.

Interpretation 2.34 If the system (2.4.1) is quadratically stable for perturbations Δ then $V(x) = x^\top Kx$ is a quadratic Lyapunov function for (2.4.1) for all $\delta \in \Delta$. By Proposition 1.38 the existence of a quadratic Lyapunov function implies that the equilibrium point $x^* = 0$ is asymptotically stable. Quadratic stability for perturbations Δ is therefore equivalent to the existence of a quadratic Lyapunov function $V(x) = x^\top Kx$, $K > 0$ such that

$$\frac{dV(x(t))}{dt} = x^\top [A(\delta(t))^\top K + KA(\delta(t))]x < 0$$

for all $\delta \in \Delta$.

Note that in general quadratic stability of the system for an uncertainty class Δ places an infinite number of constraints on the symmetric matrix K. It is the purpose of this section to make additional assumptions on the way the uncertainty enters the system, so as to convert the robust stability problem into a numerically tractable problem.

2.4.2 Quadratic stability of affine models

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

$$A(\delta(t)) = A_0 + \delta_1(t)A_1 + \ldots + \delta_k(t)A_k$$

for all $\delta \in \Delta$. This is refered to as an *affine parameter dependent model*. Note that these do not impose restrictions on the rate of changes in the parameters, i.e., arbitrary fast time variations in the parameters are allowed.

Suppose that the uncertain parameters $\delta_j(t)$, $j = 1, \ldots, k$, $t \in \mathbb{R}$ assume their values in an interval $[\underline{\delta}_j, \bar{\delta}_j]$, i.e.,

$$\delta_j(t) \in [\underline{\delta}_j, \bar{\delta}_j].$$

This means that the uncertainty of each independent parameter is assumed to be bounded between two extremal values. Define the set of *corners* of the uncertainty region as

$$\Delta_0 := \{\delta = (\delta_1, \ldots, \delta_k) \mid \delta_j \in \{\underline{\delta}_j, \bar{\delta}_j\}\} \tag{2.4.2}$$

and observe that this set has a *finite number of elements*.

Proposition 2.35 *If* (2.4.1) *is an affine parameter dependent model then it is quadratically stable if and only if there exists $K > I$ such that*

$$A(\delta)^\top K + K A(\delta) < 0$$

for all $\delta \in \Delta_0$.

Proof.   The proof of this result is an application of Proposition 1.19 in Chapter 1. Indeed, fix $x \in \mathbb{R}^n$ and consider the mapping $f_x : \Delta \to \mathbb{R}$ defined by

$$f_x(\delta) := x^\top [A(\delta(t))^\top K + K A(\delta(t))]x.$$

The domain Δ of this mapping is a convex set and by definition of Δ_0, it is the convex hull of Δ_0, i.e., $\Delta = \mathrm{co}(\Delta_0)$. Further, since $A(\delta)$ is an affine function of δ it follows that $f_x(\delta)$ is a *convex* function of δ. In particular, Proposition 1.19 yields that $f_x(\delta) < 0$ for all $\delta \in \Delta$ if and only if $f_x(\delta) < 0$ for all $\delta \in \Delta_0$. Since x is arbitrary it follows that

$$A(\delta(t))^\top K + K A(\delta(t)) < 0, \qquad \delta \in \Delta$$

if and only if

$$A(\delta(t))^\top K + K A(\delta(t)) < 0, \qquad \delta \in \Delta_0$$

which yields the result.   ∎

Obviously, the importance of this result lies in the fact that quadratic stability can be concluded from a *finite test* of matrix inequalities. Note that the condition stated in Proposition 2.35 is a feasibility test of a (finite) system of LMI's.

2.4.3   Quadratic stability of polytopic models

A second case of interest amounts to considering uncertainty defined by convex combinations of the form

$$A(\delta(t)) = \alpha_1(t)A_1 + \alpha_k(t)A_k \qquad (2.4.3)$$

where $\alpha_j(t) \geqslant 0$ and $\sum_{j=1}^{k} \alpha_j(t) = 1$ for all $t \in \mathbb{R}$. Here, $A_1, \ldots, A_k$ are fixed real matrices of dimension $n \times n$ and the α_j are to be interpreted as coefficients of a convex decomposition of the uncertain time-varying matrix $A(\delta(t))$ over the set of vertices $(A_1, \ldots, A_k)$, that is for all $\delta \in \Delta$ and $t \in \mathbb{R}$ we assume the existence of coefficients $\alpha_j(t) \geqslant 0$ with $\sum_{j=1}^{k} \alpha_j(t) = 1$ such that (2.4.3) holds. We refer to such a model as a *polytopic model*. The state matrix of a polytopic model is therefore equivalently specified as

$$A(\delta(t)) \in \mathrm{co}(A_1, \ldots, A_k)$$

for all time-varying perturbations $\delta \in \Delta$. In particular, these polytopic models do not impose restrictions on the rate of changes in the parameters, i.e., arbitrarily fast time variations in the parameters are allowed. The main result concerning quadratic stabilization of a class of uncertain polytopic models is as follows.

Proposition 2.36 *If* (2.4.1) *is a polytopic parameter dependent model where* $A(\delta(t)) \in$ co$(A_1, \ldots, A_k)$ *for all* $\delta \in \Delta$ *then it is quadratically stable if and only if there exists* $K > I$ *such that*

$$A_i^\top K + K A_i < 0$$

for all $i = 1, \ldots, k$.

Instead of proving Proposition 2.36 (See exercise 6 below) it is useful to understand the relation between Proposition 2.35 and Proposition 2.36. In fact, the class of affine models as introduced in the previous subsection can be converted to a class of polytopic models. To see this, suppose that the map $A : \Delta \to \mathbb{R}^{n \times n}$ with

$$\Delta := \{(\delta_1, \ldots, \delta_k) \mid \delta_j \in [\underline{\delta}_j, \bar{\delta}_j], \; j = 1, \ldots, k\} \tag{2.4.4}$$

is affine. Let Δ_0 be the set of corners as defined in (2.4.2). Then Δ_0 has dimension 2^k and since $\Delta = \mathrm{co}(\Delta_0)$ we have

$$A(\mathrm{co}(\Delta_0)) = A(\Delta) = \{A(\delta) \mid \delta \in \Delta\} = \mathrm{co}\{A(\delta) \mid \delta \in \Delta_0\} = \mathrm{co}(A(\Delta_0)).$$

The 'corner elements' $\delta \in \Delta_0$ in the parameter space are mapped by A onto a set of vertices $A(\delta)$, $\delta \in \Delta_0$ of a polytopic model. In other words, the affine model $\dot{x} = A(\delta)x$ with $\delta \in \Delta = \mathrm{co}(\Delta_0)$ can equivalently be viewed as a polytopic model where $A(\delta) \in \mathrm{co}(A(\Delta_0))$.

2.5 Parameter dependent Lyapunov functions

The main disadvantage in searching 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 test of quadratic stability does not discriminate between systems that have slow time-varying parameters and systems whose dynamical characteristics quickly vary in time. To reduce conservatism of the quadratic stability test we will consider quadratic Lyapunov functions for the system (2.4.1) which are parameter dependent, i.e., Lyapunov functions of the form

$$V(x) = x^\top K(\delta) x$$

where the Lyapunov matrix $K(\delta)$ is allowed to dependent on the uncertain parameter δ. More specifically, we will be interested in Lyapunov functions that are *affine* in the parameter δ, i.e.,

$$K(\delta) = K_0 + \delta_1 K_1 + \ldots + \delta_k K_k$$

where $K_0, \ldots, K_k$ are real matrices of dimension $n \times n$ and $\delta = (\delta_1 \; \ldots \; \delta_k)$. Clearly, with

$$K_1 = K_2 = \ldots = K_k = 0$$

we are back to the case of parameter *independent* quadratic Lyapunov functions as discussed in the previous section.

Definition 2.37 (Affine quadratic stability) The system (2.4.1) is called *affinely quadratically stable* if there exists matrices $K_0, \ldots, K_k$ such that

$$K(\delta) := K_0 + \delta_1 K_1 + \ldots \delta_k K_k > I \qquad (2.5.1a)$$

$$A(\delta)^\top K(\delta) + K(\delta)A(\delta) + \frac{dK(\delta)}{dt} < 0 \qquad (2.5.1b)$$

for all $\delta \in \Delta$.

Interpretation 2.38 The affine function $K(\delta)$ which satisfies (2.5.1) for all $\delta \in \Delta$ defines a quadratic Lyapunov function

$$V(x, \delta) = x^\top K(\delta)x$$

for the system (2.4.1). Indeed, from (2.5.1) we infer that $V(x, \delta) > 0$ for all $x \neq 0$ and

$$\frac{dV(x(t))}{dt} = x^\top \left(A(\delta)^\top K(\delta) + K(\delta)A(\delta)\frac{dK(\delta)}{dt} \right) x < 0$$

for all non-zero x so that the equilibrium point $x^* = 0$ is (globally) asymptotically stable if the conditions (2.5.1) are satisfied.

Remark 2.39 As in the previous section, we emphasize that the conditions (2.5.1) impose in general an infinite number of constraints on the matrices $K_0, \ldots, K_k$.

Given a system (2.4.1) with a set Δ of uncertain parameters the affine quadratic stabilization problem therefore amounts to finding matrices $K_0, \ldots K_k$ such that the conditions (2.5.1) are satisfied. In solving this problem we will distinguish between the two cases of time-invariant and time-varying uncertainty descriptions as introduced in section 2.4.

2.5.1   Time-invariant uncertainty

If the uncertainty set $\Delta \subset \mathbb{R}^k$ contains constant uncertain parameters then obviously the Lyapunov matrix $K(\delta)$ does not vary in time, so that for any $\delta \in \Delta$ we have that

$$\frac{dK(\delta)}{dt} = 0$$

in (2.5.1b). We can therefore guarantee affine quadratic stability of the system $\dot{x} = A(\delta)x$ with $\delta \in \Delta$ if we can find matrices $K_0, \ldots, K_k$ such that (2.5.1a) and

$$A(\delta)^\top K(\delta) + K(\delta)A(\delta) < 0$$

hold for all $\delta \in \Delta$. Let the uncertainty set Δ again be a convex set as defined in (2.4.4) and let Δ_0 be the corresponding set of vertices as defined in (2.4.2). Note that the expression

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

is in general not affine in δ not even when A is an affine mapping. As a consequence, for fixed $x \in \mathbb{R}^n$, the function $f_x : \Delta \to \mathbb{R}$ defined by

$$f_x(\delta) := x^\top L(\delta) x \qquad\qquad (2.5.2)$$

is in general not convex so that the negativity of the function f_x at Δ is not equivalent to its negativity at the vertices Δ_0 of Δ.

To achieve that f_x is a convex function (for any $x \in \mathbb{R}^n$) we will impose additional constraints. Suppose that both the system matrix $A(\delta)$ and the Lyapunov matrix $K(\delta)$ are affine in δ. Expanding $L(\delta)$ then yields

$$L(\delta) = [A_0 + \sum_{j=1}^{k} \delta_j A_j]^\top [K_0 + \sum_{j=1}^{k} \delta_j K_j] + [K_0 + \sum_{j=1}^{k} \delta_j K_j][A_0 + \sum_{j=1}^{k} \delta_j A_j]$$

$$= A_0^\top K_0 + K_0 A_0 + \sum_{j=1}^{k} \delta_j [A_j^\top K_0 + K_0 A_j + A_0^\top K_j + K_j A_0]$$

$$+ \sum_{j=1}^{k} \sum_{i=1}^{j-1} \delta_i \delta_j [A_i^\top K_j + K_j A_i + A_j^\top K_i + K_i A_j] + \sum_{j=1}^{k} \delta_j^2 [A_j^\top K_j + K_j A_j].$$

Now, let $x \in \mathbb{R}^n$ be fixed and consider the function f_x as defined in (2.5.2). Then for any $\delta \in \Delta$ this function takes the form

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

where c_0, c_j, c_{ij} and d_j are constants. A sufficient condition for the implication

$$\{f_x(\delta) < 0 \text{ for all } \delta \in \Delta_0\} \Rightarrow \{f_x(\delta) < 0 \text{ for all } \delta \in \Delta\}$$

is that $f_x(\delta_1, \ldots, \delta_j, \ldots, \delta_k)$ is convex in each of its arguments δ_j, $j = 1, \ldots, k$. This is the case when

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

for $j = 1, \ldots, k$. Since x is arbitrary, we obtain that

$$A_j^\top K_j + K_j A_j \geqslant 0, \qquad j = 1, \ldots, k$$

is a sufficient condition to infer the negativity of f_x over the uncertainty set Δ from the negativity of f_x at the vertices Δ_0 of Δ. This leads to the following main result.

Proposition 2.40 *If (2.4.1) is an affine parameter dependent model and $\Delta \subset \mathbb{R}^k$ is the uncertainty set defined in (2.4.4) then the system $\dot{x} = A(\delta)x$, $\delta \in \Delta$ is affinely quadratically*

stable if there exist real matrices $K_0, \ldots, K_k$ such that

$$A(\delta)^\top K(\delta) + K(\delta)A(\delta) < 0 \text{ for all } \delta \in \Delta_0 \qquad (2.5.3\text{a})$$

$$K(\delta) > I \text{ for all } \delta \in \Delta_0 \qquad (2.5.3\text{b})$$

$$A_i^\top K + K A_i < 0 \text{ for } i = 1, \ldots, k \qquad (2.5.3\text{c})$$

Here, $A(\delta) = A_0 + \sum_{j=1}^{k} \delta_j A_j$ and $K(\delta) = K_0 + \sum_{j=1}^{k} \delta_j K_j$. Moreover, in that case $V(x, \delta) := x^\top K(\delta)x$ is a quadratic parameter-dependent Lyapunov function of the system.

Proof.   It suffices to prove that (2.5.3) implies (2.5.1) for all $\delta \in \Delta$. Let x be a non-zero fixed but arbitrary element of $\mathbb{R}^n$. Since $K(\delta)$ is affine in δ, the mapping

$$\delta \to x^\top K(\delta)x$$

with $\delta \in \Delta$ is convex. Consequently, $x^\top [K(\delta) - I]x$ is larger than zero for all $\delta \in \Delta$ if it is larger than zero for all $\delta \in \Delta_0$. As x is arbitrary, this yields that (2.5.1a) holds for all $\delta \in \Delta$ whenever (2.5.3b) holds. Since for time-invariant perturbations $dK(\delta)/dt = 0$ it now suffices to prove that (2.5.3a) and (2.5.3c) imply that (2.5.1b) holds for all $\delta \in \Delta$. This however, we showed in the arguments preceding the proposition. ∎

Interpretation 2.41 Proposition 2.40 reduces the problem to verify affine quadratic stability of the system (2.4.1) to a feasibility problem of a (finite) set of linear matrix inequalities. The latter problem is a numerically tractable one and is readily implemented in the LMI toolbox.

2.5.2   Time varying uncertainty

We conclude this section with a result on robust stability of the system (2.4.1) for time-varying parameters. Let the uncertainty set Δ be defined as

$$\Delta := \big\{ (\delta_1, \ldots, \delta_k) : \mathbb{R} \to \mathbb{R}^k \mid \delta_j \text{ is differentiable and for all } t \in \mathbb{R},$$
$$\delta_j(t) \in [\underline{\delta}_j, \bar{\delta}_j], \text{ and}$$
$$\dot{\delta}_j(t) \in [\underline{\lambda}_j, \bar{\lambda}_j] \big\}. \qquad (2.5.4)$$

This means that we assume the uncertain parameters to have bounded variation and bounded rate of variation. We further introduce the vertex sets

$$\Delta_0 := \big\{ (\delta_1, \ldots, \delta_k) : \mathbb{R} \to \mathbb{R}^k \mid \delta_j \in \{\underline{\delta}_j, \bar{\delta}_j\} \big\}$$
$$\Lambda_0 := \big\{ (\lambda_1, \ldots, \lambda_k) : \mathbb{R} \to \mathbb{R}^k \mid \lambda_j \in \{\underline{\lambda}_j, \bar{\lambda}_j\} \big\}.$$

Thus Δ_0 represents the vertices of the convex hull in which the uncertain parameters take their values, whereas Λ_0 is a set of vertices whose convex hull represents the admissible rates of variation of the parameters.

Remark 2.42 There are two extreme cases of uncertainty sets (2.5.4) worth mentioning. Firstly, if the rate of variation of the uncertain parameter δ_j is set to zero, δ_j represents a constant, time-invariant perturbation as treated in the previous section. If all perturbations are known to be time-invariant then $\underline{\lambda}_j = \bar{\lambda}_j = 0$ for $j = 1, \ldots, k$ in which case Λ_0 becomes a singleton. Secondly, arbitrarily fast perturbations of δ_j are obtained by putting $\underline{\lambda}_j = -\infty$ and $\bar{\lambda}_j = \infty$.

] The reason for considering this type of uncertainty sets lies in the fact that the last term in the left-hand side of (2.5.1b) can be evaluated exactly whenever the Lyapunov matrix $K(\delta)$ is affine in δ. Specifically, if $K(\delta)$ is affine in δ we infer that

$$\frac{dK(\delta)}{dt} = \dot{\delta}_1 K_1 + \ldots + \dot{\delta}_k K_k = K(\dot{\delta}) - K_0$$

which is an affine function in $\dot{\delta}$. The main result is now as follows.

Proposition 2.43 *If (2.4.1) is an affine parameter dependent model and Δ is the uncertainty set as defined in (2.5.4) then the system $\dot{x} = A(\delta)x$, $\delta \in \Delta$ is affinely quadratically stable if there exist real matrices $K_0, \ldots, K_k$ such that*

$$A(\delta)^\top K(\delta) + K(\delta)A(\delta) + K(\lambda) < K_0 \text{ for all } \delta \in \Delta_0 \text{ and } \lambda \in \Lambda_0 \qquad (2.5.5a)$$

$$K(\delta) > I \text{ for all } \delta \in \Delta_0 \qquad (2.5.5b)$$

$$A_i^\top K + K A_i < 0 \text{ for } i = 1, \ldots, k \qquad (2.5.5c)$$

Here, $A(\delta) = A_0 + \sum_{j=1}^k \delta_j A_j$ and $K(\delta) = K_0 + \sum_{j=1}^k \delta_j K_j$. Moreover, in that case $V(x, \delta) := x^\top K(\delta)x$ is a quadratic parameter-dependent Lyapunov function of the system.

Proof.   The proof of this proposition is basically a generalization of the proof of Proposition 2.40 to the time-varying case. First fix $\lambda \in \mathbb{R}^k$. Then a similar argument as in the proof of Proposition 2.40 yields that (2.5.5) implies

$$K(\delta) > I$$
$$A(\delta)^\top K(\delta) + K(\delta)A(\delta) + K(\lambda) - K_0 < 0$$

for all $\delta \in \Delta$. Since $K(\lambda)$ is affine in λ and this last inequality holds for any $\lambda \in \Lambda_0$ we conclude from the definition of Δ that (2.5.5) implies

$$K(\delta) > I$$
$$A(\delta)^\top K(\delta) + K(\delta)A(\delta) + K(\dot{\delta}) - K_0 < 0$$

for all $\delta \in \Delta$. Now use that $K(\dot{\delta}) - K_0 = dK(\delta)/dt$ to conclude that (2.5.1) holds for all $\delta \in \Delta$ which implies the affine quadratic stability of the system. ∎

Interpretation 2.44 Proposition 2.43 reduces the problem to verify affine quadratic stability of the system $\dot{x}(t) = A(\delta(t))x(t)$ with time-varying perturbations (2.5.4) to a feasibility test of a *finite* set of LMI's.

Remark 2.45 It is interesting to compare the numerical complexity of the conditions of Proposition 2.40 with the conditions mentioned in Proposition 2.43. If the uncertainty vector δ is k-dimensional then the vertex set Δ_0 has dimension 2^k so that the verification of conditions (2.5.3) requires a feasibility test of

$$2^k + 2^k + k$$

linear matrix inequalities. In this case, also the vertex set Λ_0 has dimension 2^k which implies that (2.5.5) requires a feasibility test of

$$2^{2k} + 2^k + k + 1 = 4^k + 2^k + k$$

linear matrix inequalities.

2.5.3   Some software aspects

In MATLAB both affine and polytopic models are implemented with the routine `psys`. With $k = 2$ the syntax is

- `affmodel = psys( pv, [S0, S1, S2] )` for the specification of affine models and

- `polmodel = psys( [S1, S2] )` for the specification of polytopic models.

Here, `pv` is a *parameter vector* which is supposed to be specified by the routine `pvec`, i.e.,

```
pv = pvec('box', [ d1min d1max; d2min d2max]);
```

which implements the ranges of the uncertain parameters. `S0`, `S1` and `S2` are *system matrices* which are supposed to be defined by expressions of the form

```
S0 = ltisys(A0, B0, C0, D0, E0);
S1 = ltisys(A1, B1, C1, D1, E1);
S2 = ltisys(A2, B2, C2, D2, E2);
```

where $(A_i, B_i, C_i, D_i, E_i)$ define the state space parameters of the model[2]

$$E_i \dot{x} = A_i x + B_i u; \qquad y = C_i x + D_i u. \tag{2.5.6}$$

---

[2]The matrix E_i in the state space descriptions was introduced to handle descriptor systems in the LMI toolbox. By default the E-entries are put to the identity matrix if you omit the last argument.

Remark 2.46 The presence of the E matrix in the system representations of the LMI toolbox can be pretty disturbing. In particular, an affine combination of models of the form (2.5.6) with $i = 0, \ldots, k$ results in a model

$$E\dot{x} = Ax + Bu; \qquad y = Cx + Du$$

where

$$E = E_0 + \sum_{i=1}^{k} \delta_i E_i;$$

$$A = A_0 + \sum_{i=1}^{k} \delta_i A_i; \qquad B = B_0 + \sum_{i=1}^{k} \delta_i B_i$$

$$C = C_0 + \sum_{i=1}^{k} \delta_i C_i; \qquad D = D_0 + \sum_{i=1}^{k} \delta_i D_i.$$

If E is supposed to be independent of parameter variations then you need to explicitly set $E_1 = \ldots = E_k = 0$.

Information concerning the implemented models and parameter vectors can be retrieved via the routines `psinfo` and `pvinfo`, respectively. The routine

```
aff2pol
```

converts affine models to polytopic ones and the routine

```
quadstab
```

tests the quadratic stability of the affine or polytopic models. With `model` denoting either the affine system `affmodel` or the polytopic system `polmodel`, the syntax is

- `quadstab(model)` to verify the quadratic stability of `model`

- `[t,K] = quadstab(model)` to calculate a quadratic Lyapunov function $V(x) = x^\top K x$ for the class of models specified in `model`.

Parameter dependent Lyapunov functions can be calculated with the procedure `pdlstab`. For $k = 2$ the syntax is

```
[t,Q0,Q1,Q2]=pdlstab(affmodel)
```

where `affmodel` is as an affine model as specified above. This results in the parameter dependent function

$$Q(\delta) := Q_0 + \delta_1 Q_1 + \delta_2 Q_2$$

which defines the Lyapunov function

$$V(x, \delta) := x^\top Q(\delta)^{-1} x$$

for the uncertain system `affmodel`. Note that the LMI toolbox therefore computes the *inverses* of our Lyapunov function $K(\delta)$. For more details on the software we refer to the corresponding help files in the LMI toolbox.

2.6 Further reading

The material on dissipative systems originates from [47, 48] and has been further developed in [42] and [44, 45]. The material on quadratic stability and parameter dependent Lyapunov functions has been documented in the papers [5, 6].

2.7 Exercises

Exercise 1

Show that for conservative controllable systems the set of normalized storage functions $\mathcal{V}(x^*)$ consist of one element only. (*Consequently, storage functions of conservative or lossless systems are unique up to normalization!*).

Exercise 2

Show that the set of dissipation functions associated with a dissipative system is convex.

Exercise 3

Suppose that

$$\dot{x} = Ax + Bu$$
$$y = Cx + Du$$

is a minimal (i.e. controllable and observable) representation of a linear time-invariant dynamical system Σ. Show that Σ is stable (in the sense that $\dot{x} = Ax$ is Lyapunov stable) whenever Σ is dissipative with respect to the supply function $s(u, y) = y^\top u$.

Exercise 4

Consider the suspension system Σ of Exercise 4 in Chapter 1 with the nominal system parameters as given in Table 1.1.

1. Derive a state space model of the form 2.2.1 of the system which assumes

$$u = \begin{pmatrix} q_0 \\ F \end{pmatrix}; \qquad y = \begin{pmatrix} q_1 \\ \dot{q}_1 \\ q_2 \\ \dot{q}_2 \end{pmatrix}$$

   as its input and output, respectively.

2. Define a supply function $s : U \times Y \to \mathbb{R}$ such that (Σ, s) is dissipative. (Base your definition on physical insight).

3. Characterize the set of all quadratic storage functions of the system as the feasibility set of a linear matrix inequality.

4. Use the MATLAB routine `feasp` to compute a quadratic storage function $V(x) = x^\top K x$ of this system. (Use the routine `lmiedit` to implement the linear matrix inequality in MATLAB.

5. Use MATLAB to determine a dissipation function $d : X \times U \to \mathbb{R}$ for this system.

Exercise 5

Consider the transfer functions

1. $G_1(s) = 1/(s+1)$

2. $G_2(s) = (s-1)/(s+1)$

3. $G_3(s) = \begin{pmatrix} (s+2)(s-1)/(s+1)^2 & (s+3)/(s+4) \\ (s-1)/(s+0.5) & (s+1)/(s+2) \end{pmatrix}$.

Determine for each of these transfer functions

1. whether or not they are positive real

2. their $\mathcal{L}_2$-induced norm

Reformulate this problem as a feasibility test involving a suitably defined LMI (See Corollary 2.17 and Corollary 2.19 of this chapter).

Exercise 6

Give a proof of Proposition 2.36.

Exercise 7

Time-invariant perturbations and arbitrary fast perturbations can be viewed as two extreme cases of time-varying uncertainty sets of the type (2.5.4). (See Remark 2.42). These two extreme manifestations of time-varying perturbations reduce Proposition 2.43 to two special cases.

1. Show that the result of Proposition 2.40 is obtained as a special case of Proposition 2.43 if

$$\underline{\lambda}_j = \bar{\lambda}_j = 0; \qquad j = 1, \ldots, k.$$

2. Show that if

$$\underline{\lambda}_j = -\infty, \quad \bar{\lambda}_j = \infty; \qquad j = 1, \ldots, k.$$

   then the matrices $K_0, \ldots, K_k$ satisfying the conditions of Proposition 2.43 necessarily satisfy $K_1 = \ldots = K_k = 0$.

The latter property implies that with arbitrary fast perturbations the only affine parameter-dependent Lyapunov matrices $K(\delta) = K_0 + \sum_{j=1}^{k} \delta_j K_j$ are the constant (parameter-independent) ones. It is in this sense that Propostion 2.43 reduces to Proposition 2.35 for arbitrarily fast perturbations.

Exercise 8

Consider the suspension system Σ of Figure 2.1. and suppose that the road profile $q_0 = 0$
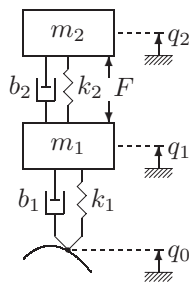

Figure 2.1: Model for suspension system

and the active suspension force $F = 0$. Here, b_1 and b_2 denote the tire damping and the

suspension damping and k_1 and k_2 denote the tire stiffness and the suspension stiffness, respectively. This system is modeled by the equations

$$m_2\ddot{q}_2 + b_2(\dot{q}_2 - \dot{q}_1) + k_2(q_2 - q_1) = 0$$
$$m_1\ddot{q}_1 + b_2(\dot{q}_1 - \dot{q}_2) + k_2(q_1 - q_2) + k_1 q_1 + b_1\dot{q}_1 = 0$$

The nominal values of the parameters defining this system are given in Table 2.1. The tire

m_1 [kg]	m_2 [kg]	k_1 N/m	k_2 N/m	b_1 Ns/m	b_2 Ns/m
1.5×10^3	1.0×10^4	5.0×10^6	$7.5 \times 10^4 - 50 \times 10^4$	$0.5 \times 10^3 - 1.0 \times 10^4$	5.0×10^4

Table 2.1: Physical parameters

damping b_1 and the suspension stiffness k_2 are supposed to be uncertain parameters with ranges indicated in Table 2.1. Let

$$\delta = \begin{pmatrix} b_1 \\ k_2 \end{pmatrix}$$

be the vector containing the uncertain physical parameters.

1. Let $x = \begin{pmatrix} q_1 & q_2 & \dot{q}_1 & \dot{q}_2 \end{pmatrix}^\top$ denote the state of this system and write this system in the form (2.4.1). Verify whether $A(\delta)$ is affine in the uncertainty parameter δ.

2. Use Proposition 2.35 to verify whether this system is quadratically stable. If so, give a quadratic Lyapunov function for this system.

3. Calculate vertex matrices $A_1, \ldots, A_k$ such that

$$A(\delta) \in \mathrm{co}(A_1, \ldots, A_k)$$

   for all δ satisfying the specifications of Table 2.1.

4. Suppose that b_1 and k_2 are time-varying and that their rates of variations satisfy

$$|\dot{b}_1| \leqslant \beta \tag{2.7.1a}$$
$$|\dot{k}_2| \leqslant \kappa \tag{2.7.1b}$$

   where $\beta = 1$ and $\kappa = 3.7$. Use Proposition 2.43 to verify whether there exists a parameter dependent Lyapunov function that proves affine quadratic stability of the uncertain system. If so, give such a Lyapunov function.

5. Suppose that

$$b_1(t) = b_1^{\mathrm{nominal}}(1 + \delta_0(t)) \tag{2.7.2a}$$
$$k_2(t) = k_2^{\mathrm{nominal}}(1 + \delta_0(t)) \tag{2.7.2b}$$

where $b_1^{\text{nominal}} = 5250$ Ns/m and $k_2^{\text{nominal}} = 28.75 \times 10^4$ N/m (the average values of b_1 and k_2, respectively).

The *parameter margin* is defined as the largest number $\delta_{\text{margin}} > 0$ such that the system $\dot{x}(t) = A(\delta(t))x(t)$ is affinely quadratically stable for all perturbations

$$\delta \in \Delta := \{(b_1, k_2) : \mathbb{R} \to \mathbb{R}^2 \mid (2.7.1) \text{ and } (2.7.2) \text{ hold for all } |\delta_0(t)| \leqslant \delta_{\text{margin}}\}.$$

Our aim is to evaluate the parameter margin δ_{margin} as function of β and κ. Suppose that $\beta = 0$ and make a plot of the function

$$\kappa \longrightarrow \delta_{\text{margin}}$$

where the rate of variation of the stiffness κ ranges from 10^{-2} till 10^2.

6. Similarly, assume that $\kappa = 0$ and make a plot of the function

$$\beta \longrightarrow \delta_{\text{margin}}$$

where the rate of variation of the tire damping β ranges from 10^{-1} till 10. What are your conclusions?

Chapter 3

Analysis of Input-Output Behavior

3.1   Basic Notions

The main concern in control theory is to study how signals are processed by dynamical systems and how this processing can be influenced to achieve a certain desired behavior.

For that purpose one has to specify the signals (time series, trajectories) which are of interest. This is done by deciding on the set of values which the signals can take (such as $\mathbb{R}^n$) and on the time set on which they are considered (such as the full time axis $\mathbb{R}$, the half axis $[0, \infty)$ or the corresponding discrete time versions $\mathbb{Z}$ and $\mathbb{N}$).

A dynamical system is then nothing but a mapping that assigns to a certain input signal an output signal (sorry, Jan!). Very often, such a mapping is defined by a differential equation with a fixed initial condition or by an integral equation, such that one considers systems or mappings with a specific description or representation.

The first purpose of this section is to discuss stability properties of systems in the general setting. In a second step we specify the system representations and investigate in how far one can obtain refined results which are amenable to computational techniques.

Note that we take a specialized point of view by considering a system as a mapping of signals. It is not very difficult to extend our results to a more general setting by viewing a system as a subset of signals or, in the modern language, as a behavior. Note that this point of view has been adopted also in the older literature on input-output stability where systems are relations instead of mappings. Obviously, if there is no clear cause-effect between signals,

this latter viewpoint is more appropriate.

To be concrete let us now specify the signal class. We denote by L^n the set of all mappings $x : [0, \infty) \to \mathbb{R}^n$ that are Lebesgue-measurable. Without bothering too much about the exact definition, one should recall that all piece-wise continuous signals are contained in L^n.

For any $x \in L^n$ one can calculate the integral in

$$\|x\|_2 := \sqrt[2]{\int_0^\infty \|x(t)\|^2 \, dt}$$

of the signal x; $\|x\|_2$ is either finite or infinite. If we collect all signals with a finite value, we arrive at the space

$$L_2^n := \{x \in L^n : \|x\|_2 < \infty\}.$$

It can be shown that L_2^n is a linear space, that $\|.\|_2$ is a norm on L_2^n, and that L_2^n is complete. Mathematically, L_2^n is a Banach space. $\|x\|_2$ is often called the *energy* of the signal x.

Remark. If the number of components n of the underlying signals is understood from the context or irrelevant, we simply write L_2 instead of L_2^n.

There is an additional structure. Indeed, define the bilinear form

$$\langle x, y \rangle = \int_0^\infty x(t)^T y(t) \, dt$$

on $L_2^n \times L_2^n$. Bilinearity just means that $\langle ., y \rangle$ is linear for each $y \in L_2^n$ and $\langle x, . \rangle$ is linear for each $x \in L_2^n$. It is not difficult to see that $\langle ., . \rangle$ defines an inner product. Moreover, the norm on L_2^n results from this inner product as $\|x\|_2^2 = \langle x, x \rangle$. Therefore, L_2^n is in fact a Hilbert space.

For any $x \in L_2^n$ one can calculate the Fourier transform $\widehat{x}$ which is a function mapping the imaginary axis $\mathbb{C}^0$ into $\mathbb{C}^n$ such that

$$\int_{-\infty}^\infty \widehat{x}(i\omega)^* \widehat{x}(i\omega) \, d\omega \quad \text{is finite.}$$

Indeed, a fundamental results in the theory of the Fourier transformation on L_2-spaces, the so-called Parseval theorem, states that

$$\int_0^\infty x(t)^T y(t) \, dt = \frac{1}{2\pi} \int_{-\infty}^\infty \widehat{x}(i\omega)^* \widehat{y}(i\omega) \, d\omega.$$

(Note that the Fourier transform $\widehat{x}$ has, in fact, a unique continuation into $\mathbb{C}^0 \cup \mathbb{C}^+$ that is analytic in $\mathbb{C}^+$. Hence, $\widehat{x}$ is not just an element of $L_2(\mathbb{C}^0, \mathbb{C}^n)$ but even of the subspace $H_2(\mathbb{C}^+, \mathbb{C}^n)$, one of the Hardy spaces. Indeed, one has $L_2(\mathbb{C}^0, \mathbb{C}^n) = H_2(\mathbb{C}^-, \mathbb{C}^n) + H_2(\mathbb{C}^+, \mathbb{C}^n)$, the sum on the right is direct, and the two spaces are orthogonal to each other.

This corresponds via the Payley-Wiener theorem to the orthogonal direct sum decomposition $L_2(\mathbb{R}, \mathbb{C}^n) = L_2((-\infty, 0], \mathbb{C}^n) + L_2([0, \infty), \mathbb{C}^n)$. This is only mentioned for clarification and not required later. The beautiful mathematical background is excellently exposed in [28].)

Stability of systems will be, roughly speaking, related to the property that it maps any signal in L_2 into a signal that is also contained in L_2. Since we also need to deal with unstable systems, we cannot confine ourselves to signals with finite L_2-norm. Hence we introduce a larger class of signals that have finite energy over finite intervals only.

For that purpose it is convenient to introduce for each $T \geqslant 0$ the so-called *truncation operator* P_T: It assigns to any signal $x \in L^n$ the signal $P_T x$ that is identical to x on $[0, T]$ and that vanishes identically on (T, ∞):

$$P_T : L^n \to L^n, \ (P_T x)(t) := \begin{cases} x(t) & \text{for} & t \in [0, T] \\ 0 & \text{for} & t \in (T, \infty) \end{cases}$$

Note that L^n is a linear space and that P_T is a linear operator on that space with the property $P_T P_T = P_T$. Hence P_T is a projection.

Now it is easy to define the space L_{2e}^n. It just consists of all signals $x \in L^n$ such that $P_T x$ has finite energy for all $T \geqslant 0$:

$$L_{2e}^n := \{x \in L^n : \ P_T x \in L_2^n \ \text{for all} \ T \geqslant 0\}.$$

Hence any $x \in L_{2e}^n$ has the property that

$$\|P_T x\|_2 = \int_0^T \|x(t)\|^2 \, dt$$

is finite for every T. (This is nothing but an integrability condition.) Note that $\|P_T x\|_2$ does not decrease if T increases. Therefore, $\|P_T x\|_2$ viewed as a function of T either stays bounded for $T \to \infty$, such that it converges, or it is unbounded, such that it diverges to ∞. We conclude for $x \in L_{2e}$: $\|P_T x\|_2$ is bounded (i.e. there exists a K such that $\|P_T x\|_2 \leqslant K$ for all $T \geqslant 0$) iff x is contained in L_2. Moreover,

$$x \in L_2^n \ \text{implies} \ \|x\|_2 = \lim_{T \to \infty} \|P_T x\|_2.$$

A *dynamical system* S is a mapping

$$S : L_{2e}^k \to L_{2e}^l.$$

The system S is *causal* if

$$P_T S(u) = P_T S(P_T u) \ \text{for all} \ T \geqslant 0, \ u \in L_{2e}^k.$$

It is easily seen that $P_T S = P_T S P_T$ is equivalent to the following more intuitive fact: If u_1 and u_2 are two input signals that are identical on $[0, T]$, $P_T u_1 = P_T u_2$, then Su_1 and

Su_2 are also identical on $[0, T]$, $P_T S(u_1) = P_T S(u_2)$. In other words, the future values of the inputs do not have an effect on the past outputs. This matches the intuitive notion of causality.

Our main interest in this abstract setting is to characterize invertibility and stability of a system. Among the many possibilities to define stability of a system, two concepts have turned out to be of prominent importance: the finite gain and finite incremental gain property of a system.

Definition 3.1 The L_2-*gain* of the system $S : L_{2e}^k \to L_{2e}^l$ is defined as

$$\|S\|_2 := \sup\{\frac{\|P_T S(u)\|_2}{\|P_T u\|_2} \mid u \in L_{2e}^k,\ T \geqslant 0,\ \|P_T u\|_2 \neq 0\} =$$
$$= \inf\{\gamma \in \mathbb{R} \mid \forall u \in L_{2e}^k, T \geqslant 0 :\ \|P_T S(u)\|_2 \leqslant \gamma \|P_T u\|_2\}.$$

If $\|S\|_2 < \infty$, S is said to have *finite L_2-gain*.

(Why does the equality in the definition of $\|S\|_2$ hold?) Clearly, S has finite L_2-gain iff there exists a real $\gamma > 0$ such that

$$\|P_T S(u)\|_2 \leqslant \gamma \|P_T u\|_2 \quad \text{for all}\ \ T \geqslant 0,\ u \in L_{2e}^k. \tag{3.1.1}$$

If S has finite L_2-gain, we conclude

$$\|S(u)\|_2 \leqslant \gamma \|u\|_2 \quad \text{for all}\ \ u \in L_2^k. \tag{3.1.2}$$

Hence, if the input has finite energy, then the output has finite energy and the output energy is bounded by a constant times the input energy. If S is *causal*, the converse is true: then (3.1.2) implies (3.1.1). Indeed, causality implies $\|P_T S(u)\|_2 = \|P_T S(P_T u)\|_2 \leqslant \|S(P_T u)\|_2$ and (3.1.2) shows (since $P_T u$ is in L_2^k) that $\|S(P_T u)\|_2 \leqslant \gamma \|P_T u\|_2$. Combining both inequalities gives (3.1.1).

If S is causal, we hence infer

$$\|S\|_2 = \sup_{u \in L_2,\, \|u\|_2 > 0} \frac{\|S(u)\|_2}{\|u\|_2}.$$

The L_2-gain of S is the worst amplification of the system if measuring the size of the input- and output-signals by their L_2-norm.

For nonlinear systems it is often more relevant to investigate how the increment $S(u_1) - S(u_2)$ relates to the increment of the input signals $u_1 - u_2$. Indeed, one wishes to have the outputs close to each other if the inputs are close what amounts to a certain continuity property. We arrive at the notion of incremental L_2-gain.

Definition 3.2 The *incremental L_2-gain* of the system $S : L_{2e}^k \to L_{2e}^l$ is defined as

$$\|S\|_{2i} := \sup\{\frac{\|P_T S(u_1) - P_T S(u_2)\|_2}{\|P_T u_1 - P_T u_2\|_2} \mid u_1, u_2 \in L_{2e}^k,\ T \geqslant 0,\ \|P_T u_1 - P_T u_2\|_2 \neq 0\} =$$
$$= \inf\{\gamma \in \mathbb{R} \mid \forall u_1, u_2 \in L_{2e}^k, T \geqslant 0 :\ \|P_T S(u_1) - P_T S(u_2)\|_2 \leqslant \gamma \|P_T u_1 - P_T u_2\|_2\}.$$

If $\|S\|_{2i} < \infty$, S is said to have *finite incremental L_2-gain.*

Similarly as before, the system S has finite incremental L_2-gain if there exists a real $\gamma > 0$ such that

$$\|P_T S(u_1) - P_T S(u_2)\|_2 \leqslant \gamma \|P_T u_1 - P_T u_2\|_2 \quad \text{for all} \;\; T \geqslant 0, \; u_1, u_2 \in L_{2e}^k. \qquad (3.1.3)$$

This reveals

$$\|S(u_1) - S(u_2)\|_2 \leqslant \gamma \|u_1 - u_2\|_2 \quad \text{for all} \;\; u_1, u_2 \in L_2^k. \qquad (3.1.4)$$

If S is causal, (3.1.4) implies (3.1.3). Moreover, for causal S, we have

$$\|S\|_{2i} = \sup_{u_1, u_2 \in L_2, \, \|u_1 - u_2\|_2 > 0} \frac{\|S(u_1) - S(u_2)\|_2}{\|u_1 - u_2\|_2}.$$

If S is linear, it is obvious that

$$\|S\|_2 = \|S\|_{2i}$$

and, hence, S has finite L_2-gain iff it has finite incremental L_2-gain. Only for nonlinear systems, the two notions are different. Even for nonlinear S, on can related both concepts if $S(0) = 0$; then one has

$$\|S\|_2 \leqslant \|S\|_{2i};$$

hence S has finite L_2-gain if it has finite incremental L_2-gain. The converse is, in general, not true. (Construct an example!)

In the remainder, stability of S will be mostly related to the property that S has finite L_2-gain or finite incremental L_2-gain.

All these concepts can be extended in literally the same fashion to all L_p-spaces for $1 \leqslant p \leqslant \infty$. Let us briefly comment on $p = \infty$ since it will emerge later. The space L_∞^n is constructed on the basis of

$$\|x\|_\infty = \text{ess} \sup_{t \geqslant 0} \|x(t)\| \quad \text{for} \;\; x \in L^n.$$

(We don't discuss the exact definition of the essential supremum. For piece-wise continuous and right-continuous signals $x \in L^n$, the essential supremum is nothing but $\sup_{t \geqslant 0} \|x(t)\|$.) Contrary to what is often done in the literature, we use the Euclidean norm $\|x(t)\|^2 = x(t)^T x(t)$ to measure the size of the real vector $x(t) \in \mathbb{R}^n$. Now L_∞^n is defined as

$$L_\infty^n := \{x \in L^n \mid \|x\|_\infty < \infty.\}$$

It is well-known that L_∞^n with the norm $\|.\|_\infty$ defines a Banach space. The space $L_{\infty e}^n$ and the L_∞-gain or incremental L_∞-gain $\|S\|_\infty$ or $\|S\|_{\infty i}$ for a system $S : L_{\infty e}^k \rightarrow L_{\infty e}^l$ are defined literally as before, and similar properties hold true.

Let us mention a few modifications or generalizations in various respects. As said before, the extension to all L_p-spaces for $1 \leqslant p \leqslant \infty$ is straightforward. The time-axis can be chosen as all nonnegative integers to discuss discrete-time systems. A mixture of continuous- and

discrete-time allows to consider hybrid systems or systems with jumps. Finally, the set on which the signals take their values can be an arbitrary normed space; this allows to discuss infinite dimensional systems.

In addition, the stability concept used here is one out of a multitude of possibilities. Often, it is only required that S maps L_2 into L_2 without necessarily having a finite L_2-gain. Indeed, the theory can be extended to general designer chosen stability properties that only obey certain (technical) axiomatic hypotheses. We just include these remarks to stress that we somehow artificially confine ourselves to pretty specific cases for reasons of space.

3.2   Robust Input-Output Stability

3.2.1   A Specific Feedback Interconnection

In robust control, one encounters systems that are affected by uncertainties (parametric variations, unmodeled dynamics). In order to study the effects of uncertainties, one has to start with a structure that captures how variations in the uncertainties affect the system to be investigated. Although one could think of a broad range of such structures, the technique of 'pulling out the uncertainties' allows to reduce many of these variants to one common setup that is represented in Figure 3.1.

Here, M is viewed as the nominal model and Δ captures the (varying) uncertainties. Both the nominal system and the uncertainty are interconnected via feedback. M is usually viewed as a fixed system, whereas Δ is allowed to vary in a certain class $\mathbf{\Delta}$.

Typical examples include the case of linear time-invariant systems that are affected by additive or multiplicative uncertainties. If looking at a large interconnection of such small uncertain components, one arrives at structured uncertainties as they are considered in μ-theory. The set $\mathbf{\Delta}$ should hence be seen as capturing both the nature of the uncertainty (linear/nonlinear, time-invariant/time-varying, static/dynamic), their size (bounds on norm, gain or incremental gain) and their structure (block-diagonal, full-block). We will not repeat how to pull out the uncertainties in specific feedback interconnections what should have been presented in a basic course on robust control.

Having specified M and the class $\mathbf{\Delta}$, one of the central tasks is to characterize whether the feedback interconnection of the stable systems M and Δ remains stable for all $\Delta \in \mathbf{\Delta}$. Let us now be more specific by introducing the mathematical setup.

Here are the standing hypotheses.

Assumption 3.3 The mapping

$$M : L_{2e}^k \to L_{2e}^l \quad \text{is causal, of finite } L_2\text{-gain, and linear.}$$
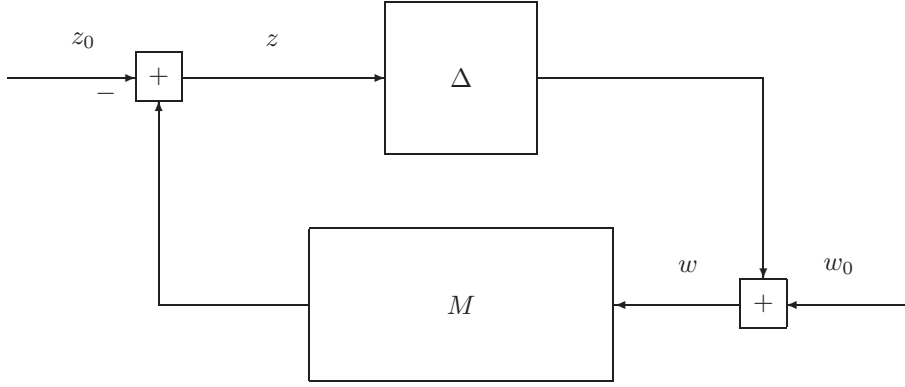
Figure 3.1: Uncertainty Feedback Configuration

The uncertainty set $\boldsymbol{\Delta}$ is a set of systems

$$\Delta : L_{2e}^l \to L_{2e}^k \quad \text{that are causal and of finite } L_2\text{-gain.}$$

Moreover, it is star-shaped with star center 0:

$$\Delta \in \boldsymbol{\Delta} \implies \tau\Delta \in \boldsymbol{\Delta} \quad \text{for all} \quad \tau \in [0,1].$$

Note that the third property implies $0 \in \boldsymbol{\Delta}$; this is consistent with viewing Δ as an uncertainty where $\Delta = 0$ is related to the unperturbed or nominal system. Recall that $\tau \to \tau\Delta$ just defines a line in the set of all causal and stable systems connecting 0 with Δ what justifies the terminology that $\boldsymbol{\Delta}$ is star-shaped with center 0.

Let us finally stress that linearity of M is not crucial at all for the results to follow; they can be easily extended with minor variations to nonlinear systems M.

For any $\Delta \in \boldsymbol{\Delta}$, we investigate the feedback interconnection in Figure 3.1 that is defined by the relations

$$\begin{pmatrix} I & -\Delta \\ M & -I \end{pmatrix} \begin{pmatrix} w \\ z \end{pmatrix} := \begin{pmatrix} w \\ Mw \end{pmatrix} - \begin{pmatrix} \Delta(z) \\ z \end{pmatrix} = \begin{pmatrix} w_0 \\ z_0 \end{pmatrix}. \tag{3.2.1}$$

Here, the signals $\begin{pmatrix} w_0 \\ z_0 \end{pmatrix}$ are viewed as external inputs or disturbances, and $\begin{pmatrix} w \\ z \end{pmatrix}$ constitutes the response of the interconnection.

A first important property of the feedback interconnection is *well-posedness*: Does there exists for each w_0, z_0 a unique response w, z satisfying (3.2.1) such that the mapping

$$\begin{pmatrix} w_0 \\ z_0 \end{pmatrix} \to \begin{pmatrix} w \\ z \end{pmatrix} \tag{3.2.2}$$

is causal?

Secondly, one is interested in the stability of the feedback interconnection: If the interconnection is well-posed, does the mapping (3.2.2) have finite L_2-gain or finite incremental L_2-gain? If the interconnection is stable for all $\Delta \in \mathbf{\Delta}$, it is said to be robustly stable.

Thirdly, one might look at uniform robust stability: The interconnection is well-posed and robustly stable, and the L_2-gain or the incremental L_2-gain of the mapping (3.2.2) is bounded by a constant for all $\Delta \in \mathbf{\Delta}$. (The latter property just means that the (incremental) L_2-gain is uniformly bounded in $\Delta \in \mathbf{\Delta}$.)

It simplifies notations if we introduce the abbreviations

$$X_e = L_{2e}^{k+l}, \quad X = L_2^{k+l}, \quad \mathcal{I}_M(\Delta) := \begin{pmatrix} I & -\Delta \\ M & -I \end{pmatrix}.$$

Consequently, any signal $x \in X_e$ is partitioned as $x = \begin{pmatrix} w \\ z \end{pmatrix}$ and we note that the system $\mathcal{I}_M(\Delta)$ captures both the subsystems M and Δ and the specific interconnection structure that we are considering. (M and Δ are asymmetrically entering the notation $\mathcal{I}_M(\Delta)$ in order to stress the fact M is fixed whereas Δ is allowed to vary in $\mathbf{\Delta}$.)

Recall that the loop is required to have a a unique response x to any external input x_0: For each $x_0 \in X_e$ there exists a unique $x \in X_e$ with $\mathcal{I}_M(\Delta)(x) = x_0$. Mathematically, this simply amounts to the mapping $\mathcal{I}_M(\Delta) : X_e \to X_e$ having an inverse $\mathcal{I}_M(\Delta)^{-1}$. If this inverse is, in addition, causal, $\mathcal{I}_M(\Delta)$ is said to be well-posed.

Definition 3.4 $\mathcal{I}_M(\Delta) : X_e \to X_e$ is *well-posed* if it has a causal inverse.

Well-posedness can be often assured by standard results about the existence of solutions of differential equations. However, we will also provide simple explicit conditions on M and $\mathbf{\Delta}$ that imply well-posedness.

If we impose a certain stability property on M and Δ, the composed system $\mathcal{I}_M(\Delta)$ shares this stability property, whereas this is usually *not true for the inverse* $\mathcal{I}_M(\Delta)^{-1}$. In stability theory, it is hence of fundamental interest to find additional conditions to guarantee that this inverse indeed shares its stability properties with M, Δ, and $\mathcal{I}_M(\Delta)$.

In these notes we obtain sufficient conditions for the following facts:

- Under the hypothesis that $\mathcal{I}_M(\Delta)^{-1}$ exists and is causal, we characterize that there exists a c with $\|\mathcal{I}_M(\Delta)^{-1}\| \leqslant c$ for all $\Delta \in \mathbf{\Delta}$; in particular, any $\mathcal{I}_M(\Delta)^{-1}$ then has finite L_2-gain.

- If all systems in $\mathbf{\Delta}$ have finite incremental L_2-gain, we characterize that $\mathcal{I}_M(\Delta)$ does have a causal inverse and that there exists a c with $\|\mathcal{I}_M(\Delta)^{-1}\|_{2i} \leqslant c$ for all $\Delta \in \mathbf{\Delta}$; in particular, any $\mathcal{I}_M(\Delta)^{-1}$ has a finite incremental L_2-gain.

It is important to observe the difference in these two characterizations: In the first scenario one has to *assume* that the interconnection is well-posed, whereas in the second situation one can *conclude* this property.

Our main goal is to summarize many of the results that are available in the literature in two basic theorems that are related to the two differing hypotheses and conclusions sketched above. Let us first turn to a very simple auxiliary observation that simplifies the proofs.

3.2.2 An Elementary Auxiliary Result

Suppose W and Z are arbitrary subset of the normed space X. Our goal is to characterize that
$$\frac{\|w\|^2 + \|z\|^2}{\|w - z\|^2}$$
remains bounded for all $w \in W$, $z \in Z$ such that $w \neq z$. This implies that W and Z can have at most the vector 0 in common, and that the squared distance of any two vectors is at least as large as a constant times the sum of the squared norms of these vectors. If W and Z are subspaces, there are very close relations to the gap or angle of these subspaces.

The desired characterization is provided in terms of a mapping $\Sigma : X \to \mathbb{R}$ that is quadratically continuous [24].

Definition 3.5 $\Sigma : X \to \mathbb{R}$ is *quadratically continuous* if for every $\delta > 0$ there exists a $\sigma_\delta > 0$ with
$$|\Sigma(x_1) - \Sigma(x_2)| \leqslant \sigma_\delta \|x_1 - x_2\|^2 + \delta \|x_2\|^2 \quad \text{for all} \quad x_1, x_2 \in X.$$

As a typical important example, let $\langle ., . \rangle : X \times X \to \mathbb{R}$ be any biadditive form that is bounded: there exists a $\sigma > 0$ such that
$$\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle, \quad \langle x, y + z \rangle = \langle x, y \rangle + \langle x, z \rangle, \quad |\langle x, y \rangle| \leqslant \sigma \|x\| \|y\|. \tag{3.2.3}$$

Then $\Sigma(x) := \langle x, x \rangle$ is quadratically continuous. Indeed, we have
$$|\Sigma(x + y) - \Sigma(x)| - \delta \|x\|^2 = |\langle y, y \rangle + \langle x, y \rangle + \langle y, x \rangle| - \delta \|x\|^2 \leqslant$$
$$\leqslant \sigma \|y\|^2 + 2\sigma \|x\| \|y\| - \delta \|x\|^2 \leqslant \sigma_\delta \|y\|^2$$
with $\sigma_\delta := \max\{\sigma + 2\sigma t - \delta t^2 : t \in \mathbb{R}\}$ such that $\sigma + 2\sigma \frac{\|x\|}{\|y\|} - \delta \frac{\|x\|^2}{\|y\|^2} \leqslant \sigma_\delta$ for $\|y\| \neq 0$.

Lemma 3.6 *Suppose* $\Sigma : X \to \mathbb{R}$ *is quadratically continuous. If*
$$\Sigma(z) \geqslant 0 \quad \text{for all} \quad z \in Z$$

and if there exists an $\epsilon > 0$ such that

$$\Sigma(w) \leqslant -\epsilon \|w\|^2 \quad for\ all\ \ w \in W,$$

then one can find a $c > 0$ (that only depends on Σ and ϵ) with

$$\|w\|^2 + \|z\|^2 \leqslant c^2 \|w - z\|^2 \quad for\ all\ \ w \in W,\ z \in Z. \tag{3.2.4}$$

Proof. The proof is trivial: We have $\frac{1}{2}\epsilon\|w\|^2 \leqslant \Sigma(z) - \Sigma(w) - \frac{1}{2}\epsilon\|w\|^2 \leqslant \sigma_{\epsilon/2}\|z - w\|^2$ and hence

$$\|w\|^2 \leqslant \frac{\sigma_{\epsilon/2}}{\epsilon/2}\|z - w\|^2.$$

In addition, we infer

$$\|z\|^2 \leqslant (\|z - w\| + \|w\|)^2 \leqslant 2(\|z - w\|^2 + \|w\|^2) \leqslant 2(1 + \frac{\sigma_{\epsilon/2}}{\epsilon/2})\|z - w\|^2.$$

Hence (3.2.4) holds with $c = \sqrt{2 + 3\frac{\sigma_{\epsilon/2}}{\epsilon/2}}$. ■

Note that the choice of the square of the norms in all these definitions is completely arbitrary and only made for later, more concrete, applications. One can replace $\|.\|^2$ everywhere by $\alpha(\|.\|)$ where $\alpha : [0, \infty) \to [0, \infty)$ is any function that is monotone and for which there exists a constant $\alpha_0 > 0$ with $\alpha(t_1 + t_2) \leqslant \alpha_0(\alpha(t_1) + \alpha(t_2))$. The proof remains unchanged.

3.2.3   An Abstract Stability Characterization

In this section we assume that $\mathcal{I}_M(\Delta)$ is well-posed for all $\Delta \in \mathbf{\Delta}$. Hence, $\mathcal{I}_M(\Delta)^{-1}$ exists and is causal. We intend to get criteria that this inverse admits a uniform bound on its L_2-gain. It is not difficult to provide an abstract criterion on the basis of Lemma 3.6.

Theorem 3.7 *Let $\Sigma : L_2^{k+l} \to \mathbb{R}$ be quadratically continuous. Suppose that for all $\Delta \in \mathbf{\Delta}$*

$$\mathcal{I}_M(\Delta) \ \ is\ well\mbox{-}posed\ and\ \ \Sigma \left(\begin{array}{c} \Delta(z) \\ z \end{array} \right) \geqslant 0 \ \ for\ all\ \ z \in L_2^l. \tag{3.2.5}$$

If there exists an $\epsilon > 0$ with

$$\Sigma \left(\begin{array}{c} w \\ Mw \end{array} \right) \leqslant -\epsilon\|w\|_2^2 \ \ for\ all\ \ w \in L_2^k, \tag{3.2.6}$$

there exists a constant c such that $\|\mathcal{I}_M(\Delta)^{-1}\|_2 \leqslant c$ for all $\Delta \in \mathbf{\Delta}$.

Proof. Fix any $\Delta \in \mathbf{\Delta}$. Recall that all $\tau\Delta$ are also contained in $\mathbf{\Delta}$ if $\tau \in [0, 1]$. For any such τ define

$$W := \{ \left(\begin{array}{c} w \\ Mw \end{array} \right) :\ w \in L_2^k \}, \ \ Z := \{ \left(\begin{array}{c} \tau\Delta(z) \\ z \end{array} \right) :\ z \in L_2^l \}.$$

Due to $\left\|\begin{pmatrix} w \\ Mw \end{pmatrix}\right\|_2^2 \leqslant (1+\|M\|_2^2)\|w\|_2^2$ we have $-\epsilon\|w\|_2^2 \leqslant -\frac{\epsilon}{1+\|M\|_2^2}\left\|\begin{pmatrix} w \\ Mw \end{pmatrix}\right\|_2^2$ what allows to apply Lemma 3.6. Then (3.2.4) implies

$$\left\|\begin{pmatrix} w \\ Mw \end{pmatrix}\right\|_2^2 + \left\|\begin{pmatrix} \tau\Delta(z) \\ z \end{pmatrix}\right\|_2^2 \leqslant c^2\left\|\begin{pmatrix} w \\ Mw \end{pmatrix} - \begin{pmatrix} \tau\Delta(z) \\ z \end{pmatrix}\right\|_2^2$$

for all $w \in L_2^k$ and $z \in L_2^l$. Since the left-hand side bounds $\left\|\begin{pmatrix} w \\ z \end{pmatrix}\right\|_2^2$ from above, we arrive at

$$\|x\|_2 \leqslant c\|\mathcal{I}_M(\tau\Delta)(x)\|_2 \quad \text{for all} \quad x \in X. \tag{3.2.7}$$

If we pick $x_0 \in X$ and we *know* that $x = \mathcal{I}_M(\tau\Delta)^{-1}(x_0)$ is contained in X (and not only in X_e), we can conclude

$$\|\mathcal{I}_M(\tau\Delta)^{-1}(x_0)\|_2 \leqslant c\|x_0\|_2. \tag{3.2.8}$$

Consequently, if $\mathcal{I}_M(\tau\Delta)^{-1}$ has finite L_2-gain (such that it maps any vector $x_0 \in X$ into a vector $x \in X$), we infer (3.2.8) for all $x_0 \in X$. This implies that the L_2-gain of $\mathcal{I}_M(\tau\Delta)^{-1}$ is bounded by the constant c:

$$\|\mathcal{I}_M(\tau\Delta)^{-1}\|_2 \leqslant c.$$

Note that c only depends on Σ, ϵ, $\|M\|_2$; it is independent of the particular Δ or τ!

Let us now prove the following statement:

$$\|\mathcal{I}_M(\tau_0\Delta)^{-1}\|_2 < \infty \implies \|\mathcal{I}_M(\tau\Delta)^{-1}\|_2 \leqslant c \quad \text{for} \quad |\tau - \tau_0| < \frac{1}{c\|\Delta\|_2}, \ \tau \in [0,1]. \tag{3.2.9}$$

Indeed, if $\mathcal{I}_M(\tau_0\Delta)^{-1}$ has finite L_2-gain, we get $\|\mathcal{I}_M(\tau_0\Delta)^{-1}\|_2 \leqslant c$. Let us now take any $x_0 \in X$ and set $x = \mathcal{I}_M(\tau\Delta)^{-1}(x_0)$. We infer

$$\mathcal{I}_M(\tau_0\Delta)(x) = \begin{pmatrix} 0 & (\tau-\tau_0)\Delta \\ 0 & 0 \end{pmatrix}(x) + \mathcal{I}_M(\tau\Delta)(x) = \begin{pmatrix} 0 & (\tau-\tau_0)\Delta \\ 0 & 0 \end{pmatrix}(x) - x_0.$$

This leads to

$$\|P_T x\|_2 = \|P_T\mathcal{I}_M(\tau_0\Delta)^{-1}[\begin{pmatrix} 0 & (\tau-\tau_0)\Delta \\ 0 & 0 \end{pmatrix}(P_T x) - P_T x_0]\|_2 \leqslant$$

$$\leqslant c\|\begin{pmatrix} 0 & (\tau-\tau_0)\Delta \\ 0 & 0 \end{pmatrix}(P_T x)\|_2 + c\|P_T x_0\|_2 \leqslant$$

$$\leqslant c|\tau-\tau_0|\,\|\Delta\|\,\|P_T x\|_2 + c\|P_T x_0\|_2$$

and hence

$$(1 - c|\tau-\tau_0|\|\Delta\|_2)\|P_T x\|_2 \leqslant c\|P_T x_0\|_2 \leqslant c\|x_0\|_2.$$

Since the factor on the left is positive for all τ that satisfy the hypothesis in (3.2.9), we infer that $\|P_T x\|_2$ is bounded what implies $x \in X$. As argued above, we obtain (3.2.8), and since x_0 was arbitrary, we get $\|\mathcal{I}_M(\tau\Delta)^{-1}\|_2 \leqslant c$.

Let us now set $\delta_j := \min\{1, j/(2c\|\Delta\|_2)\}$ such that $\delta_j \in [0, 1]$ and $|\delta_j - \delta_{j-1}| < 1/(c\|\Delta\|_2)$. For $\delta_0 = 0$, we have $\|\mathcal{I}_M(\delta_0\Delta)^{-1}\|_2 < \infty$ such that $\|\mathcal{I}_M(\delta_1\Delta)^{-1}\|_2 \leqslant c$. In this way we can successively conclude that $\|\mathcal{I}_M(\delta_j\Delta)^{-1}\|_2 \leqslant c$ for all j. Since $\delta_j = 1$ for some j, the proof is finished.   ∎

Theorem 3.7 is applied as follows. For the given system M and for the set of uncertainties Δ, one tries to find a quadratically continuous Σ that guarantees (3.2.5) and (3.2.6). Then we can conclude that $\mathcal{I}_M(\Delta)^{-1}$ admits a uniform bound on its L_2-gain. In more concrete situation, we will later see how the search of such a mapping Σ can be performed by solving an LMI problem.

However, one can also change the viewpoint: Given a quadratically continuous mapping Σ, *define* the class of uncertainties Δ as those that satisfy (3.2.5). Then all systems that have the property (3.2.6) cannot be destabilized by this class of uncertainties. Classical small-gain and passivity theorems fall in this class as will be discussed Section 3.3.

Remarks.

- To characterize that $\mathcal{I}_M(\Delta)^{-1}$ has finite L_2-gain for a *fixed* Δ requires the stronger hypotheses

$$\Sigma(P_T \begin{pmatrix} \Delta(z) \\ z \end{pmatrix}) \geqslant 0 \quad \text{for all} \ \ z \in L_{2e}^l, \ T \geqslant 0$$

  and, with some $\epsilon > 0$,

$$\Sigma(P_T \begin{pmatrix} w \\ Mw \end{pmatrix}) \leqslant -\epsilon \|P_T w\|^2 \quad \text{for all} \ \ w \in L_{2e}^k, \ T \geqslant 0.$$

  This result is very easy to prove and, in fact, closer to what is usually found in the literature.

- The proof of the theorem proceeds via a homotopy argument: The L_2-gain of $\mathcal{I}_M(0)^{-1}$ is finite. Then one proves that $\mathcal{I}_M(\Delta)^{-1}$ has finite L_2-gain by showing that the gain $\|\mathcal{I}_M(\tau\Delta)^{-1}\|_2$ stays below a constant c and, hence, does not blow up if τ varies from 0 to 1.

  The line $[0, 1] \ni \tau \to \tau\Delta$ can be replaced without difficulty by any continuous curve $\gamma : [0, 1] \to \Delta$ connecting 0 and Δ as $\gamma(0) = 0$, $\gamma(1) = \Delta$. Hence, instead of being star-shaped, it suffices to require that Δ contains 0 and is path-wise connected.

  Note that all this is very similar to proving the usual Nyquist-based stability results for LTI systems.

3.2.4 A Characterization of Well-Posedness and Stability

In the last section we required $\mathcal{I}_M(\Delta)$ to have a causal inverse. In this section we intend to get rid of this hypothesis. As a price to be paid, we have to assume that all Δ have finite

incremental L_2-gain, and we have to replace the second condition in (3.2.5) on $\begin{pmatrix} \Delta(z) \\ z \end{pmatrix}$ by

the same on the increment $\begin{pmatrix} \Delta(z_1) \\ z_2 \end{pmatrix} - \begin{pmatrix} \Delta(z_2) \\ z_2 \end{pmatrix}$. (If M is nonlinear, the same holds for

(3.2.6) what is not pursued here.) Instead of a bound on the L_2-gain of $\mathcal{I}_M(\Delta)^{-1}$ we then obtain a bound on the incremental L_2-gain of this mapping.

If both M and Δ are linear but otherwise still general, both criteria coincide. Hence, only for nonlinear uncertainties Δ we require a stronger hypotheses to get to the desired stronger conclusions.

Technically, we exploit the fact that $X = L_2$ is a Banach space and we apply Banach's fixed point theorem to derive well-posedness.

Theorem 3.8 *Let $\Sigma : L_2^{k+l} \to \mathbb{R}$ be quadratically continuous. Suppose that all $\Delta \in \mathbf{\Delta}$ have finite incremental L_2-gain and satisfy*

$$\Sigma \begin{pmatrix} \Delta(z_1) - \Delta(z_2) \\ z_1 - z_2 \end{pmatrix} \geqslant 0 \quad \text{for all} \quad z_1, z_2 \in L_2^l. \tag{3.2.10}$$

If there exists an $\epsilon > 0$ with (3.2.6), then $\mathcal{I}_M(\Delta)$ does have a causal inverse and there exists a constant c with $\|\mathcal{I}_M(\Delta)^{-1}\|_{2i} \leqslant c$ for all $\Delta \in \mathbf{\Delta}$.

Proof. Similarly as in the proof of Theorem 3.7 we pick $\Delta \in \mathbf{\Delta}$, $\tau \in [0,1]$ and define $W := \left\{ \begin{pmatrix} w \\ Mw \end{pmatrix} : w \in L_2^k \right\}$, $Z := \left\{ \begin{pmatrix} \tau\Delta(z_1) - \tau\Delta(z_2) \\ z_1 - z_2 \end{pmatrix} : z_1, z_2 \in L_2^l \right\}$. Again, we can apply Lemma 3.6 and to infer from (3.2.4) that

$$\left\| \begin{pmatrix} w \\ Mw \end{pmatrix} \right\|_2^2 + \left\| \begin{pmatrix} \tau\Delta(z_1) - \tau\Delta(z_2) \\ z_1 - z_2 \end{pmatrix} \right\|_2^2 \leqslant c^2 \left\| \begin{pmatrix} w \\ Mw \end{pmatrix} - \begin{pmatrix} \tau\Delta(z_1) - \tau\Delta(z_2) \\ z_1 - z_2 \end{pmatrix} \right\|_2^2$$

for all $w \in L_2^k$ and $z_1, z_2 \in L_2^l$. This relation leads to

$$\|x_1 - x_2\|_2^2 \leqslant c^2 \|\mathcal{I}_M(\tau\Delta)(x_1) - \mathcal{I}_M(\tau\Delta)(x_2)\|_2^2 \quad \text{for all} \quad x_1, x_2 \in X.$$

Let us temporarily assume that $\mathcal{I}_M(\tau_0\Delta) : X \to X$ with $\tau_0 \in [0,1]$ has an inverse. Then we conclude

$$\|\mathcal{I}_M(\tau_0\Delta)^{-1}(y_1) - \mathcal{I}_M(\tau_0\Delta)^{-1}(y_2)\|_2^2 \leqslant c^2 \|y_1 - y_2\|_2^2 \quad \text{for all} \quad y_1, y_2 \in X. \tag{3.2.11}$$

It is essential to observe that c on the right does does neither depend on Δ nor on τ!

Now we take another $\tau \in [0,1]$ that is close (we will specify how close) to τ_0. To verify that $\mathcal{I}_M(\tau\Delta) : X \to X$ has an inverse amounts to checking that for all $y \in X$ there is a unique $x \in X$ satisfying $\mathcal{I}_M(\tau\Delta)(x) = y$. Let us now bring in the mapping of which we know that it has an inverse; we rewrite the equation to

$$\mathcal{I}_M(\tau\Delta)(x) - \mathcal{I}_M(\tau_0\Delta)(x) + \mathcal{I}_M(\tau_0\Delta)(x) = y.$$

This is easily rearranged to the fixed-point equation

$$x = \mathcal{I}_M(\tau_0 \Delta)^{-1}(y - \mathcal{I}_M(\tau \Delta)(x) + \mathcal{I}_M(\tau_0 \Delta)(x)). \tag{3.2.12}$$

Let us abbreviate

$$F(x) := \mathcal{I}_M(\tau_0 \Delta)^{-1}(y - \mathcal{I}_M(\tau \Delta)(x) + \mathcal{I}_M(\tau_0 \Delta)(x)).$$

Note that F maps X into X. We have reduced the original problem to the question of whether there exists a unique $x \in X$ with $F(x) = x$, i.e., whether F has a unique fixed point in X. If X is a Banach space, and if there exists a $f < 1$ with

$$\|F(x_1) - F(x_2)\|_2 \leqslant f\|x_1 - x_2\| \quad \text{for all} \quad x_1, x_2 \in X, \tag{3.2.13}$$

Banach's fixed point theorem leads to the desired conclusion; it is then guaranteed that F indeed has exactly one fixed point. In our case, X is Banach. The only thing to be assured is (3.2.13). Using (3.2.11), we arrive at

$$\|F(x_1) - F(x_2)\|_2^2 \leqslant$$
$$\leqslant c^2 \|[y - \mathcal{I}_M(\tau \Delta)(x_1) + \mathcal{I}_M(\tau_0 \Delta)(x_1)] - [y - \mathcal{I}_M(\tau \Delta)(x_2) + \mathcal{I}_M(\tau_0 \Delta)(x_2)]\|_2^2 \leqslant$$
$$\leqslant c^2 \|[\mathcal{I}_M(\tau_0 \Delta)(x_1) - \mathcal{I}_M(\tau \Delta)(x_1)] + [\mathcal{I}_M(\tau_0 \Delta)(x_2) - \mathcal{I}_M(\tau \Delta)(x_2)]\|_2^2 \leqslant$$
$$\leqslant c^2 \| \begin{pmatrix} 0 & (\tau - \tau_0)\Delta \\ 0 & 0 \end{pmatrix}(x_1) + \begin{pmatrix} 0 & (\tau - \tau_0)\Delta \\ 0 & 0 \end{pmatrix}(x_2)\|_2^2 \leqslant c^2 |\tau - \tau_0|^2 \|\Delta(z_1) - \Delta(z_2)\|_2^2 \leqslant$$
$$\leqslant c^2 |\tau - \tau_0|^2 \|\Delta\|_{2i}^2 \|z_1 - z_2\|_2^2 \leqslant c^2 |\tau - \tau_0|^2 \|\Delta\|_{2i}^2 \|x_1 - x_2\|_2^2.$$

Hence

$$|\tau_0 - \tau| < \frac{1}{c\|\Delta\|_{2i}}$$

implies (3.2.13) with $f = c^2 |\tau - \tau_0|^2 \|\Delta\|_{2i}^2 < 1$. Then F has a unique fixed point in X. Therefore, $\mathcal{I}_M(\tau \Delta) : X \to X$ has an inverse that satisfies, a posteriori, (3.2.11).

Obviously, $\mathcal{I}_M(0) : X \to X$ does have an inverse. Therefore, we can successively conclude that $\mathcal{I}_M(\tau_j \Delta) : X \to X$ has an inverse satisfying (3.2.8) for $\tau_j = \min\{1, j/(2c\|\Delta\|_{2i})\}$, $j = 0, 1, 2, \ldots$, and hence also for $\tau = 1$.

So far, we have shown that $\mathcal{I}_M(\Delta) : X \to X$ has the inverse $\mathcal{I}_M(\Delta)^{-1} : X \to X$ with bound c on its incremental L_2-gain. It is a simple exercise to prove that the causal mapping $\mathcal{I}_M(\Delta) : X_e \to X_e$ then also has a causal inverse $\mathcal{I}_M(\Delta)^{-1} : X_e \to X_e$ with the same bound c on its incremental L_2-gain.   ■

Even if dealing with nonlinear uncertainties Δ, they often have the property $\Delta(0) = 0$. Then we infer that $\mathcal{I}_M(\Delta)(0) = 0$ such that the same must hold for its inverse. Therefore, we have

$$\|\mathcal{I}_M(\Delta)^{-1}\|_2 \leqslant \|\mathcal{I}_M(\Delta)^{-1}\|_{2i}$$

and Theorem 3.8 also provides a bound on the L_2-gain of the inverse.

Since Theorem 3.8 also guarantees the existence of the inverse of $\mathcal{I}_M(\Delta)$ and, therefore (Exercise 1), also of $I - M\Delta$, we will mainly build in the sequel on this result under under the additional

Assumption 3.9 All $\Delta \in \mathbf{\Delta}$ have finite incremental L_2-gain and satisfy $\Delta(0) = 0$.

We stress again that these properties are trivially satisfied if the uncertainties are linear.

3.3 Small-Gain and Passivity Tests

As an illustration, let us consider for a symmetric matrix Π the bilinear mapping

$$\langle x, y \rangle = \int_0^\infty x(t)^T \Pi y(t)\, dt \tag{3.3.1}$$

on L_2^{k+l}. Since this mapping satisfies (3.2.3), $\Sigma(x) := \langle x, x \rangle$ is quadratically continuous. We assume P to be partitioned as

$$\Pi = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} \in \mathbb{R}^{(k+l) \times (k+l)} \tag{3.3.2}$$

according to the size of the signals w and z.

Let us make the specific choice

$$\Pi = \begin{pmatrix} -I & 0 \\ 0 & I \end{pmatrix}, \quad \Sigma \begin{pmatrix} w \\ z \end{pmatrix} = \int_0^\infty z(t)^T z(t) - w(t)^T w(t)\, dt.$$

Obviously, the set of all causal uncertainties that satisfy (3.2.5) is given by all

$$\text{causal } \Delta : L_{2e} \to L_{2e} \text{ with } \|\Delta\|_2 \leqslant 1. \tag{3.3.3}$$

We immediately arrive at the following classical small L_2-gain result.

Corollary 3.10 *Suppose that, for all Δ with (3.3.3), $\mathcal{I}_M(\Delta)$ has a causal inverse. If $\|M\|_2 < 1$, there exists a c such that $\|\mathcal{I}_M(\Delta)^{-1}\|_2 \leqslant c$ for all Δ as in (3.3.3).*

Similarly, the set of all causal uncertainties with (3.2.10) is nothing but all

$$\text{causal } \Delta : L_{2e} \to L_{2e} \text{ with } \|\Delta\|_{2i} \leqslant 1. \tag{3.3.4}$$

This leads us to the following standard small incremental L_2-gain result.

Corollary 3.11 *Suppose that $\|M\|_2 < 1$. Then $\mathcal{I}_M(\Delta)$ has a causal inverse, and there exists a c with $\|\mathcal{I}_M(\Delta)^{-1}\|_{2i} \leqslant c$ for all Δ satisfying (3.3.4).*

The choice

$$\Pi = \begin{pmatrix} 0 & \frac{1}{2}I \\ \frac{1}{2}I & 0 \end{pmatrix}, \quad \Sigma\begin{pmatrix} w \\ z \end{pmatrix} = \int_0^\infty z(t)^T w(t)\, dt \qquad (3.3.5)$$

leads to the standard passivity tests for robust stability.

Corollary 3.12 *Suppose there exists an $\epsilon > 0$ with*

$$\int_0^\infty w(t)^T (Mw)(t)\, dt \leqslant -\epsilon\|w\|_2^2.$$

- *Let $\mathcal{I}_M(\Delta)$ have a causal inverse for all causal $\Delta : L_{2e} \to L_{2e}$ with*

$$\int_0^\infty z(t)^T \Delta(z)(t)\, dt \geqslant 0.$$

  Then there exists a c such that $\|\mathcal{I}_M(\Delta)^{-1}\|_2 \leqslant c$ holds for all these uncertainties.

- *The mapping $\mathcal{I}_M(\Delta)$ has a causal inverse, and there exists a c with $\|\mathcal{I}_M(\Delta)^{-1}\|_{2i} \leqslant c$ for all causal uncertainties $\Delta : L_{2e} \to L_{2e}$ that satisfy*

$$\int_0^\infty [z_1(t) - z_2(t)]^T [\Delta(z_1)(t) - \Delta(z_2)(t)]\, dt \geqslant 0.$$

We have obtain with ease four classical results that are usually formulated and proved independently. Even better, the approach taken here allows much further reaching generalizations that can only be partially addressed in these notes.

Remark. The small-gain theorem for L_p spaces can be recovered with the mapping

$$\Sigma\begin{pmatrix} w \\ z \end{pmatrix} = \int_0^\infty \|z(t)\|^p - \|w(t)\|^p\, dt$$

and with $\alpha(t) = t^p$ instead of $\alpha(t) = t^2$ in Section 3.2.2. The details are only variations of what has been presented and can be left as an exercise.

3.4   Integral Quadratic Constraints

3.4.1   Stability Tests with Integral Quadratic Constraints

In this section we assume that the uncertainties Δ are general but that $z = Mw$ is defined with a proper stable rational matrix $\widehat{M}$ as

$$\widehat{z}(i\omega) = \widehat{M}(i\omega)\widehat{w}(i\omega), \quad i\omega \in \mathbb{C}^0.$$

Recall that $\widehat{x}$ denotes the Fourier transform of the signal $x \in L_2$.

Instead of general quadratically continuous mappings, let us consider so-called integral quadratic forms. Suppose $\Pi : i\omega \rightarrow \Pi(i\omega)$ is any (measurable) mapping that assigns to every $i\omega \in \mathbb{C}^0$ a Hermitian matrix $\Pi(i\omega)$ of dimension $(k + l) \times (k + l)$ that is bounded:

$$\|\Pi(i\omega)\| \leqslant p \text{ for all } i\omega \in \mathbb{C}^0.$$

(Note that we will consider in most cases mappings that are defined with a rational matrix valued function $\Pi(s)$; then it is just required that this rational matrix is Hermitian on the imaginary axis, and that it has neither a pole in $\mathbb{C}^0$ nor at infinity such that it is proper.)

For any $x, y \in L_2^{k+l}$ we can define with their Fourier transforms $\widehat{x}, \widehat{y}$ the mapping

$$\langle x, y \rangle := \int_{-\infty}^{\infty} \widehat{x}(i\omega)^* \Pi(i\omega) \widehat{y}(i\omega)\, d\omega$$

which satisfies (3.2.3) with bound $\sigma = p$.

Condition (3.2.6) then amounts to

$$\int_{-\infty}^{\infty} \widehat{w}(i\omega)^* \left(\begin{matrix} I \\ \widehat{M}(i\omega) \end{matrix} \right)^* \Pi(i\omega) \left(\begin{matrix} I \\ \widehat{M}(i\omega) \end{matrix} \right) \widehat{w}(i\omega)\, d\omega \leqslant -\frac{\epsilon}{2\pi} \int_{-\infty}^{\infty} \widehat{w}(i\omega)^* \widehat{w}(i\omega)\, d\omega$$

for all $w \in L_2^k$. This is obviously implied by the frequency domain inequality (FDI)

$$\left(\begin{matrix} I \\ \widehat{M}(i\omega) \end{matrix} \right)^* \Pi(i\omega) \left(\begin{matrix} I \\ \widehat{M}(i\omega) \end{matrix} \right) \leqslant -\frac{\epsilon}{2\pi} I \text{ for all } i\omega \in \mathbb{C}^0.$$

It is not required for our arguments and not difficult to see that the converse holds as well; both characterization are in fact equivalent.

Lemma 3.13 *Suppose Π is a (measurable) bounded Hermitian valued mapping on $\mathbb{C}^0$. Then the following two statements are equivalent:*

- $\int_{-\infty}^{\infty} \widehat{x}(i\omega)^* \Pi(i\omega) \widehat{x}(i\omega)\, d\omega \leqslant -\alpha \int_{-\infty}^{\infty} \widehat{x}(i\omega)^* \widehat{x}(i\omega)\, d\omega$ *for all $x \in L_2$.*

- $\Pi(i\omega) \leqslant -\alpha I$ *for all $i\omega \in \mathbb{C}^0$.*

The reason for this re-formulation: the frequency domain inequality is easier to check.

Now we get as immediate corollaries to Theorems 3.7 and 3.8 the following stability results using *integral quadratic constraints* (IQC's).

Theorem 3.14 *Suppose that, for all $\Delta \in \mathbf{\Delta}$, $\mathcal{I}_M(\Delta)$ is well-posed and that*

$$\int_{-\infty}^{\infty} \left(\begin{matrix} \widehat{\Delta(z)}(i\omega) \\ \widehat{z}(i\omega) \end{matrix} \right)^* \Pi(i\omega) \left(\begin{matrix} \widehat{\Delta(z)}(i\omega) \\ \widehat{z}(i\omega) \end{matrix} \right) d\omega \geqslant 0 \text{ for all } z \in L_2^l. \qquad (3.4.1)$$

If there exists an $\epsilon > 0$ with

$$\left(\begin{array}{c} I \\ \widehat{M}(i\omega) \end{array} \right)^{*} \Pi(i\omega) \left(\begin{array}{c} I \\ \widehat{M}(i\omega) \end{array} \right) \leqslant -\epsilon I \ \ for \ all \ \ \omega \in \mathbb{R}, \qquad (3.4.2)$$

then the L_2-gain of $\mathcal{I}_M^{-1}(\Delta)$ is bounded uniformly in $\Delta \in \mathbf{\Delta}$.

Theorem 3.15 *Suppose that any $\Delta \in \mathbf{\Delta}$ has finite incremental L_2-gain and satisfies*

$$\int_{-\infty}^{\infty} \left(\begin{array}{c} \widehat{\Delta(z_1)}(i\omega) - \widehat{\Delta(z_2)}(i\omega) \\ \widehat{z_1}(i\omega) - \widehat{z_2}(i\omega) \end{array} \right)^{*} \Pi(i\omega) \left(\begin{array}{c} \widehat{\Delta(z_1)}(i\omega) - \widehat{\Delta(z_2)}(i\omega) \\ \widehat{z_1}(i\omega) - \widehat{z_2}(i\omega) \end{array} \right) d\omega \geqslant 0 \qquad (3.4.3)$$

for all $z_1, z_2 \in L_2^l$. If there exists an $\epsilon > 0$ with (3.4.2), then $\mathcal{I}_M(\Delta)$ is well-posed, and the incremental L_2-gain of its inverse is uniformly bounded in $\Delta \in \mathbf{\Delta}$.

Remarks.

- One should read $\widehat{\Delta(z)}(i\omega)$ correctly: Take z, let it pass through Δ to get the signal $\Delta(z)$, take its Fourier transform $\widehat{\Delta(z)}$, and evaluate this Fourier transform at $i\omega$ to obtain $\widehat{\Delta(z)}(i\omega)$. Therefore, the signal z with power distribution $\widehat{z}$ is mapped into the signal $\Delta(z)$ with power distribution $\widehat{\Delta(z)}$. (In general, of course, there is no nice operation - such as the multiplication by a transfer matrix - that maps $\widehat{z}$ directly into $\widehat{\Delta(z)}$. However, since we only transform signals, no complication arises.)

  The inequality (3.4.1) defined via Π hence restricts how the power distribution of z can and cannot be rearranged in $\widehat{\Delta(z)}$; (3.4.1) could be called a power distribution constraint. The constraint (3.4.3) admits the same interpretation for increments.

- In principal, the inequality (3.4.2) is easy to verify: one just needs to plot the largest eigenvalue of the left-hand side over frequency and read of the maximum that this curve takes. This could be viewed as a generalization of plotting the largest singular values of a certain transfer matrix to apply the small-gain theorem.

  If Π is real-rational and proper, the Kalman-Yakubovich-Popov Lemma allows to reduce this condition to the solvability of a linear matrix inequality; this is the reason why IQC's play such prominent role in the LMI approach to robust control. We will elaborate on these points in Section 3.6.

3.4.2   The Philosophy for Applying IQC's

So far we have considered *one* quadratically continuous mapping Σ and *one* IQC to characterize stability. For small-gain and passivity conditions, this was sufficient to arrive at standard stability results. However, if one has a more detailed picture about the uncertainty, one can often find more than one IQC that are satisfied by the uncertainties.

For the purpose of illustration let us look at a simple example. Consider the structured nonlinear uncertainties $\Delta : L_{2e}^l \to L_{2e}^k$ that are defined for fixed partitions

$$z = \begin{pmatrix} z^1 \\ \vdots \\ z^m \end{pmatrix}, \quad w = \begin{pmatrix} w^1 \\ \vdots \\ w^m \end{pmatrix}$$

(where the signals z^j and w^j can have different sizes) with the causal mappings $\Delta_j : L_{2e} \to L_{2e}$, $\Delta_j(0) = 0$, as

$$w = \Delta(z), \quad \Delta\left(\begin{pmatrix} z^1 \\ \vdots \\ z^m \end{pmatrix} \right) = \begin{pmatrix} \Delta_1(z^1) \\ \vdots \\ \Delta_m(z^m) \end{pmatrix}.$$

Furthermore, it is assumed that $\|\Delta_j\|_{2i} \leqslant 1$ such that, as well, $\|\Delta_j\|_2 \leqslant 1$.

Note that the set of all these uncertainties is star-shaped. Due to $\|\Delta\|_{2i} \leqslant 1$, the incremental small-gain theorem applies. Then $\|\widehat{M}\|_\infty < 1$ implies that $\mathcal{I}_M(\Delta)^{-1}$ exists, is causal, and uniformly incrementally bounded. However, this also holds for the much larger class of all uncertainties Δ with $\|\Delta\|_{2i} \leqslant 1$, even if they do not have the specific structure considered here.

Hence we should find more IQC's that provide a way to capture this structure. Motivated by μ-theory, we consider the IQC's defined with the constant matrices Π given as

$$\Pi = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix}, \quad S = 0, \quad Q = \mathrm{diag}(-r_1 I, \dots, -r_m I), \quad R = \mathrm{diag}(r_1 I, \dots, r_m I), \quad r_j > 0$$

(3.4.4)

where the sizes of the identity blocks in Q and R correspond to the sizes of the signals w^j and z^j respectively. We infer

$$\int_0^\infty \begin{pmatrix} \Delta(z_1)(t) - \Delta(z_2)(t) \\ z_1(t) - z_2(t) \end{pmatrix}^T \Pi \begin{pmatrix} \Delta(z_1)(t) - \Delta(z_2)(t) \\ z_1(t) - z_2(t) \end{pmatrix} dt =$$

$$= \int_0^\infty \sum_{j=1}^m -r_j \|\Delta_j(z_1)(t) - \Delta_j(z_2)(t)\|^2 + r_j \|z_1(t) - z_2(t)\|^2 \, dt =$$

$$= \sum_{j=1}^m \int_0^\infty r_j [\|z_1(t) - z_2(t)\|^2 - \|\Delta_j(z_1)(t) - \Delta_j(z_2)(t)\|^2] \, dt \geqslant 0$$

such that the incremental IQC's (3.4.3) hold for all uncertainties and for all Π.

We have found a whole family of IQC's for our class of uncertainties, parameterized by the numbers r_j. If we just find among these infinitely many IQC's one for which, in addition, the FDI (3.4.2) holds, we conclude exactly the same stability properties of $\mathcal{I}_M(\Delta)$ as before.

Again, we stress that *all* IQC's (3.4.3) must be satisfied by the uncertainties, but only for *one* IQC we need to assure (3.4.2)! Hence, the more IQC's we find for the uncertainties, the

more freedom we have if trying to fulfill the FDI and the better the chances are to verify robust stability.

Let us now have a more detailed look at (3.4.2) for the specific scalings (3.4.4). The inequality simply reads as

$$\widehat{M}(i\omega)^* R \widehat{M}(i\omega) - R < 0 \quad \text{for all} \quad \omega \in \mathbb{R} \cup \{\infty\}. \tag{3.4.5}$$

(Since we have replaced $\leqslant -\epsilon I$ by < 0, we have to include $\omega = \infty$ in the condition. Why?) The goal is to find some R (structured as in (3.4.4)) that satisfies this FDI. It will turn out that the search for R can be cast into an LMI problem.

In order to relate to μ-theory, re-parameterize

$$R = D^T D$$

with D in the same class as R. Then $\widehat{M}(i\omega)^* D^T D \widehat{M}(i\omega) - D^T D < 0$ is equivalent to $\|D\widehat{M}(i\omega)D^{-1}\| < 1$ if $\|.\|$ denotes the maximal singular value for complex matrices. Therefore, (3.4.5) is nothing but

$$\|D\widehat{M}D^{-1}\|_\infty < 1 \tag{3.4.6}$$

which is a scaled H_∞ condition. Such conditions - possibly with frequency dependent scalings D - appear in μ-theory. Note, however, that the conclusions made in μ-theory are usually only valid for *linear time-invariant uncertainties that admit a Fourier transform with suitable properties*. Our conclusions hold for a much larger class of uncertainties since our proof was not based on a Nyquist type argument in the frequency domain.

We have shown that we can replace $\|\widehat{M}\|_\infty < 1$ by the scaled small-gain condition (3.4.6) to come to the same robust stability conclusions. The scalings D capture the knowledge about the structure of the uncertainties and provide us extra freedom to satisfy (3.4.5). Hence, the scalings reduce the conservatism that is involved in the simple but rough condition $\|\widehat{M}\|_\infty < 1$.

Let us introduce a terminology: We will call the matrices Π that define the IQC's *scalings* or *multipliers*. The first name is motivated by the above mentioned relation to μ-theory. The second name reminds of the relation to classical multipliers that have been used in loop transformation arguments.

The example reveals the philosophy in applying the robust stability results discussed here: Try to find as many multipliers Π as possible such that the IQC's (3.4.1) (or (3.4.3)) hold for the considered class of uncertainties. Then find, among all these multipliers, one that also satisfies the FDI (3.4.2). If this is possible, one can conclude (existence and) uniform boundedness of the (incremental) L_2-gain of $\mathcal{I}_M(\Delta)^{-1}$.

A simple trick often allows to increase the number of multipliers. Indeed, if $\Pi_1, \dots, \Pi_k$ are multipliers that satisfy (3.4.1) (or (3.4.3)), the same is true of all

$$\sum_{j=1}^{k} \tau_j \Pi_j \quad \text{if} \quad \tau_j \geqslant 0. \tag{3.4.7}$$

One can hence easily construct out of finitely many multipliers an infinite family of multipliers parameterized by τ_j. The same trick applies to an infinite set of multipliers. (Those familiar with the corresponding concepts will recognize that we just need to take the convex conic hull; any set of multipliers can, therefore, always assumed to be a convex cone.)

Finding multipliers such that a specific class of uncertainties satisfies the corresponding IQC is not really supported by theory; this is indeed the hard part in concrete applications. For suitable parameterizations of the family of considered multipliers (such as (3.4.7) or more general versions), the second step of finding one multiplier that also renders the FDI (3.4.2) satisfied will turn out to be an LMI problem.

3.4.3 Examples of IQC's

In what follows we provide a non-exhaustive list of uncertainties and suitable multipliers. We recall that one needs to always verify Assumption 3.3, in particular star-shapeness with center 0, in order to apply Theorem 3.7, Theorem 3.8 or their IQC counterparts.

- The structured nonlinear uncertainties

$$\Delta\left(\begin{pmatrix} z^1 \\ \vdots \\ z^m \end{pmatrix}\right) = \begin{pmatrix} \Delta_1(z^1) \\ \vdots \\ \Delta_m(z^m) \end{pmatrix} \tag{3.4.8}$$

with causal Δ_j that satisfy $\|\Delta_j\|_2 \leqslant 1$ or $\|\Delta_j\|_{2i} \leqslant 1$ fulfill (3.4.1) or (3.4.3) for the class of multipliers

$$\mathbf{\Pi} := \{\begin{pmatrix} Q & 0 \\ 0 & R \end{pmatrix}, \ Q = \mathrm{diag}(-r_1 I, \ldots, -r_m I), R = \mathrm{diag}(r_1 I, \ldots, r_m I) > 0\}. \tag{3.4.9}$$

- In (3.4.8) we can confine the attention to *linear* causal mappings Δ_j only. Beautiful results by Shamma and Megretsky [19, 40] show that, then, the resulting scaled H_∞-condition (3.4.2) is not only *sufficient* for robust stability (as we have proved) but even *necessary* (what is harder to show).

- We can specialize further and use (3.4.9) also for the block-diagonal time-varying parametric uncertainties
$$w^j(t) = \Delta_j(t) z^j(t)$$
with (measurable) matrix valued functions satisfying

$$\|\Delta_j(t)\| \leqslant 1 \ \text{ for all } \ t \geqslant 0.$$

Equivalently, we have

$$w(t) = \Delta(t) z(t), \quad \Delta(t) = \mathrm{diag}(\Delta_1(t), \ldots, \Delta_m(t)), \quad \|\Delta(t)\| \leqslant 1 \ \text{ for } \ t \geqslant 0.$$

In this case, for any Π in the class (3.4.9), the uncertainties even satisfy the quadratic constraint

$$\left(\begin{array}{c} \Delta(t) \\ I \end{array} \right)^T \Pi \left(\begin{array}{c} \Delta(t) \\ I \end{array} \right) \geqslant 0. \qquad (3.4.10)$$

(We will see in Section 3.7 that this implies exponential stability.) The quadratic constraint still holds if using a time-varying multiplier. Let $P : [0, \infty) \to \Pi$ be (measurable and essentially) bounded. Note that $P(t)$ admits exactly the same structure as the constant multipliers above. For any such time-varying scaling we infer

$$\left(\begin{array}{c} \Delta(t) \\ I \end{array} \right)^T P(t) \left(\begin{array}{c} \Delta(t) \\ I \end{array} \right) \geqslant 0 \quad \text{for all} \ \ t \geqslant 0.$$

With the quadratically continuous mapping (why?)

$$\Sigma(x) := \int_0^\infty x(t)^T P(t) x(t) \, dt$$

on L_2^{k+l}, we infer (3.2.10) by linearity. Hence, if there exists an $\epsilon > 0$ with

$$\int_0^\infty \left(\begin{array}{c} w(t) \\ M(w)(t) \end{array} \right)^T P(t) \left(\begin{array}{c} w(t) \\ M(w)(t) \end{array} \right) dt \leqslant -\epsilon \|w\|_2^2, \qquad (3.4.11)$$

we can apply the more abstract Theorem 3.8 to infer that $\mathcal{I}_M(\Delta)^{-1}$ exists and has a uniformly bounded L_2-gain. Again, (3.4.11) amounts to a scaled small-gain condition with time-varying scalings. If M can be described by

$$\dot{x} = A(t)x + B(t)w, \quad z = C(t)x + D(t)w, \quad x(0) = 0,$$

where $\dot{x} = A(t)x$ is exponentially stable, the validity of (3.4.11) can be characterized by a differential linear matrix inequality.

- The so-called repeated structured uncertainties are defined as

$$w(t) = \Delta(t)z(t), \quad \Delta(t) = \mathrm{diag}(\delta_1(t)I, \ldots, \delta_m(t)I), \quad |\delta_j(t)| \leqslant 1 \ \ \text{for} \ \ t \geqslant 0.$$

Here the blocks on the diagonal of $\Delta(t)$ are repeated scalar valued functions. If we choose the multipliers Π in the partition (3.3.2) as

$$R = \mathrm{diag}(R_1, \ldots, R_m) > 0, \ Q = -R, \ S = \mathrm{diag}(S_1, \ldots, S_m), \ S_j + S_j^T = 0,$$

we infer

$$\left(\begin{array}{c} \Delta(t) \\ I \end{array} \right)^T \Pi \left(\begin{array}{c} \Delta(t) \\ I \end{array} \right) = \sum_{j=1}^m -\delta_j(t)^2 R_j + \delta_j(t)(S_j + S_j^T) + R_j = \sum_{j=1}^m (1 - \delta_j(t)^2) R_j \geqslant 0.$$

Again, we have found a class of multipliers for which (3.4.1) or (3.4.3) hold, and we could generalize to time-varying scalings.

- Note that there exists a $\delta(t)$ with $|\delta(t)| \leqslant 1$ and $w(t) = \delta(t)z(t)$ iff $w(t)w(t)^T \leqslant z(t)z(t)^T$. This leads to the notion of structured repeated nonlinear uncertainties. They are defined as in (3.4.8) where the diagonal maps Δ_j satisfy

$$\int_0^\infty \Delta_j(z^j)(t)\Delta_j(z^j)(t)^T \, dt \leqslant \int_0^\infty z^j(t)z^j(t)^T \, dt.$$

  The same scalings as in the previous item can be used to infer (3.4.1). If we ask the property to hold for the increments, we obtain (3.4.3). Then we arrive at robust stability results against repeated nonlinear uncertainties.

- In the above examples we have used for parametric uncertainties and the corresponding nonlinear uncertainties the same class of scalings. However, for parametric uncertainties, one can work with a class of scalings that is only indirectly described but larger than that considered so far; since this results in more IQC's, it reduces the conservatism in the stability results.

  Let us assume that the uncertainty is defined as

$$w(t) = \Delta(t)z(t)$$

  where (the measurable) $\Delta : [0, \infty) \to \mathbb{R}^{k \times l}$ satisfies

$$\Delta(t) \in \mathrm{co}\{\Delta_1, ..., \Delta_N\} \quad \text{for all} \quad t \geqslant 0.$$

  Here, Δ_j are fixed matrices which generate the convex hull that defines the set of values which can be taken by the time-varying uncertainties; these generators capture the structure and the size of the parametric uncertainty. Obviously, the repeated diagonal structure is a special case of this more general setup (Why?). The goal is to define the scalings in order to assure (3.4.10). We just ask the condition (3.4.10) to hold at the generators of the convex hull:

$$\begin{pmatrix} \Delta_j \\ I \end{pmatrix}^T \Pi \begin{pmatrix} \Delta_j \\ I \end{pmatrix} \geqslant 0 \quad \text{for all} \quad j = 1, \ldots, N. \tag{3.4.12}$$

  If we impose an additional constraint on Π such that these finitely many inequalities imply (3.4.10), we are done. The simplest possible condition is to require the left-upper block of Π being negative definite; this leads to the class

$$\mathbf{\Pi} := \{\Pi = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} : Q < 0, \ \begin{pmatrix} \Delta_j \\ I \end{pmatrix}^T \Pi \begin{pmatrix} \Delta_j \\ I \end{pmatrix} > \quad \text{for} \quad j = 1, \ldots, N\}.$$

  A very simple convexity argument reveals that any $\Pi \in \mathbf{\Pi}$ indeed satisfies (3.4.10). Instead of what we have done previously, the multipliers are now only indirectly described. Since we have strengthened the non-strict inequality to a strict inequality, however, one can easily implement this indirect description as constraints in an LMI solver. That allows to reduce, again, the search for a multiplier satisfying (3.4.1) to an LMI problem. As a considerable advantage of this latter technique we observe that we do not need to bother at all about the specific structure of the uncertainties and theoretically derive the corresponding structure of the multipliers - the numerical algorithm does the job for us.
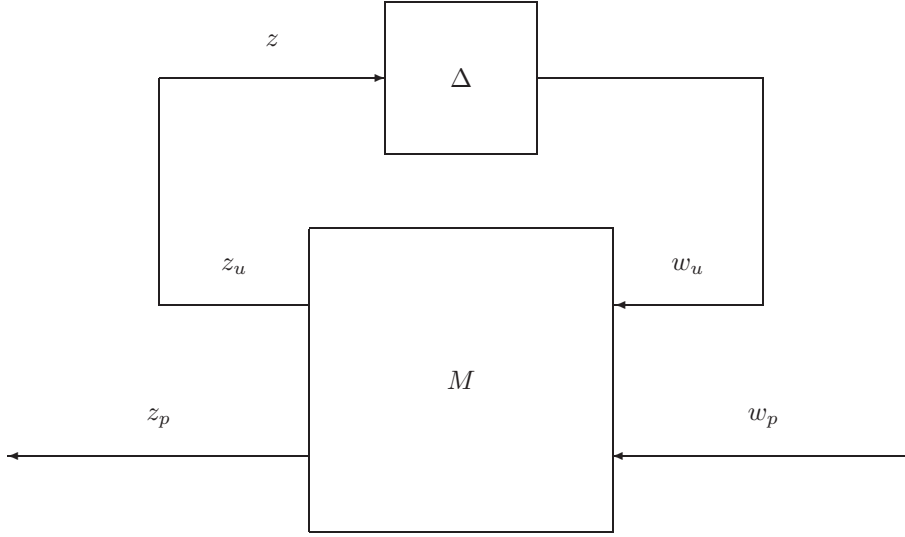
Figure 3.2: Uncertain System with Performance Channel

3.5　Guaranteeing Robust Performance

So far we have considered robust stability. However, the techniques presented so far allow a simple extension to provide sufficient conditions for robust performance.

3.5.1　An Abstract Condition for Robust Performance

Let us consider the uncertain system in Figure 3.2 where the blocks M and $\Delta \in \mathbf{\Delta}$ satisfy the Assumptions 3.3 and 3.9.

As indicated we assume that M is partitioned as

$$\begin{pmatrix} z_u \\ z_p \end{pmatrix} = \begin{pmatrix} M_{uu} & M_{up} \\ M_{pu} & M_{pp} \end{pmatrix} \begin{pmatrix} w_u \\ w_p \end{pmatrix}. \tag{3.5.1}$$

The signals w_u and z_u form the *uncertainty channel* $w_u \to z_u$ and $w_p \to z_p$ denotes the *performance channel*. The reason for this terminology is simple. The uncertain system is described by (3.5.1) and by closing the upper loop with any uncertainty $\Delta \in \mathbf{\Delta}$ as

$$w_u = \Delta(z_u). \tag{3.5.2}$$

This leads to $z_u = M_{uu}\Delta(z_u) + M_{up}w_p$. If $I - M_{uu}\Delta$ has a causal inverse, we arrive at $z_u = (I - M_{uu}\Delta)^{-1}(M_{up}w_p)$. The perturbed system hence admits the description

$$z_p = M_{pp}w_p + M_{pu}\Delta(I - M_{uu}\Delta)^{-1}(M_{up}w_p). \tag{3.5.3}$$

The performance specification under considerations is then specified in terms of the channel $w_p \to z_p$.

As a typical performance specification, the L_2-gain from w_p to z_p should not exceed one (or any other number what can be always brought back to one by scaling); this specification amounts to

$$\int_0^\infty z_p(t)^T z_p(t)\, dt \leqslant \int_0^\infty w_p(t)^T w_p(t)\, dt \quad \text{for all} \ \ w_p \in L_2.$$

For technical reasons (to arrive at necessary and sufficient conditions) one often tries to characterize that the L_2-gain is *strictly* smaller than one. Contrary to what is often stated in the literature, we cannot just replace the $\leqslant$ by $<$ in the above inequality! (Why?) The correct formulation is as follows: there exists an $\epsilon > 0$ such that

$$\int_0^\infty z_p(t)^T z_p(t) - w_p(t)^T w_p(t)\, dt \leqslant -\epsilon \int_0^\infty w_p(t)^T w_p(t)\, dt \quad \text{for all} \ \ w_p \in L_2.$$

Similarly, alternative performance specifications are passivity

$$\int_0^\infty z_p(t)^T w_p(t)\, dt \leqslant 0 \quad \text{for all} \ \ w_p \in L_2$$

or strict passivity: there exists an $\epsilon > 0$ with

$$\int_0^\infty z_p(t)^T w_p(t)\, dt \leqslant -\epsilon \int_0^\infty w_p(t)^T w_p(t)\, dt \quad \text{for all} \ \ w_p \in L_2.$$

These cases are easily seen to be specialization of the following general *performance specification*: there exists an $\epsilon > 0$ such that

$$\Sigma_p \begin{pmatrix} w_p \\ z_p \end{pmatrix} \leqslant -\epsilon \|w_p\|_2^2 \quad \text{for all} \ \ w_p \in L_2. \tag{3.5.4}$$

Here, Σ_p is an arbitrary mapping

$$\Sigma_p : L_2 \ni \begin{pmatrix} w_p \\ z_p \end{pmatrix} \to \Sigma_p \begin{pmatrix} w_p \\ z_p \end{pmatrix} \in \mathbb{R} \ \ \text{satisfying} \ \ \Sigma_p \begin{pmatrix} 0 \\ z_p \end{pmatrix} \geqslant 0. \tag{3.5.5}$$

(The second condition is of technical nature - it is required in the proof of the next theorem.)

The goal is to characterize robust stability and robust performance. For the precise definition of robust stability, we need to introduce (as earlier) the auxiliary signals w_0 and z_0 as in Figure 3.3.

The interconnection (3.2) is said to be *uniformly robustly stable* if the relations

$$\begin{pmatrix} z_u \\ z_p \end{pmatrix} = M \begin{pmatrix} w_u \\ w_p \end{pmatrix}, \quad w_u = \Delta(z) + w_0, \quad z = z_u - z_0$$
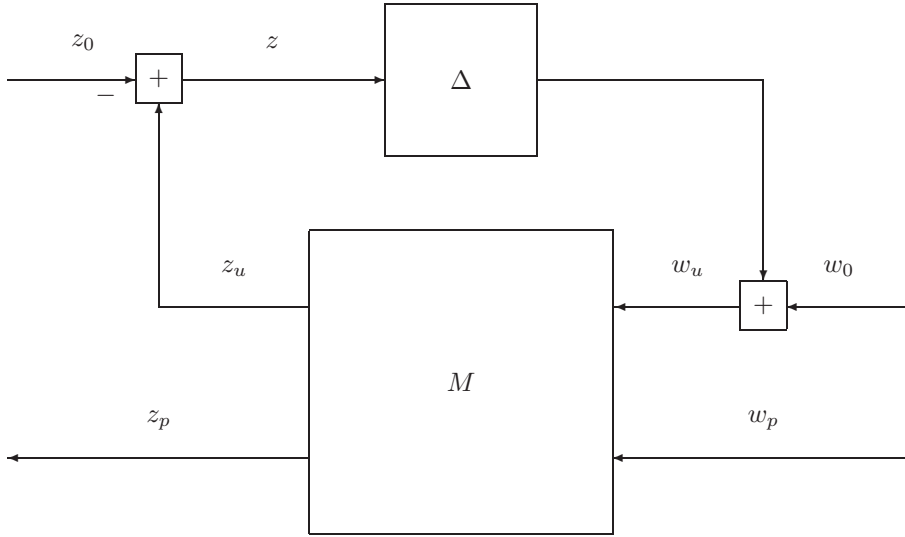
Figure 3.3: Setup for Robust Performance

that correspond to Figure 3.3 define, for each $\Delta \in \boldsymbol{\Delta}$, a causal mapping

$$
L_{2e} \ni \begin{pmatrix} w_0 \\ z_0 \\ w_p \end{pmatrix} \rightarrow \begin{pmatrix} w_u \\ z \\ z_p \end{pmatrix} \in L_{2e},
$$

and if the incremental L_2-gain of this mapping is bounded uniformly in $\Delta \in \boldsymbol{\Delta}$.

Since M and Δ have finite (incremental) L_2-gain, it is very simple to verify that robust stability is equivalent to $\mathcal{I}_{M_{uu}}(\Delta)$ or (Exercise 1) $I - M_{uu}\Delta$ having a causal inverse whose incremental L_2-gain is bounded uniformly in $\Delta \in \boldsymbol{\Delta}$.

Theorem 3.16 *Suppose* $\Sigma : L_2 \rightarrow L_2$ *is quadratically continuous and that all* $\Delta \in \boldsymbol{\Delta}$ *satisfy (3.2.10). Moreover, suppose there exists an* $\epsilon > 0$ *such that*

$$
\Sigma \begin{pmatrix} w_u \\ M_{uu}w_u + M_{up}w_p \end{pmatrix} + \Sigma_p \begin{pmatrix} w_p \\ M_{pu}w_u + M_{pp}w_p \end{pmatrix} \leqslant -\epsilon(\|w_u\|_2^2 + \|w_p\|_2^2) \qquad (3.5.6)
$$

for all $w_u \in L_2$, $w_p \in L_2$. *Then* $I - M_{uu}\Delta$ *has a causal inverse whose incremental* L_2-gain *is bounded uniformly in* $\Delta \in \boldsymbol{\Delta}$, *and the uncertain system (3.5.3) satisfies (3.5.4).*

Proof. The proof is extremely simple. We can set $w_p = 0$ to infer by the second property in (3.5.5) from (3.5.6) that

$$
\Sigma \begin{pmatrix} w_u \\ M_{uu}w_u \end{pmatrix} \leqslant -\epsilon\|w_u\|_2^2
$$

for all $w_u \in L_2$. Hence we can apply Theorem 3.8 to conclude that $\mathcal{I}_{M_{uu}}(\Delta)$ or, equivalently (Exercise 1), $I - M_{uu}\Delta$ have causal inverses with uniformly bounded incremental L_2-gain.

Note that, in particular, the uncertain system system (3.5.3) then defines a mapping with uniformly bounded incremental L_2-gain. Even more, for any $w_p \in L_2$, we use $w_u = \Delta(z_u)$ and (3.2.10) to infer

$$\Sigma \left(\begin{array}{c} w_u \\ M_{uu}w_u + M_{up}w_p \end{array} \right) = \Sigma \left(\begin{array}{c} w_u \\ z_u \end{array} \right) = \Sigma \left(\begin{array}{c} \Delta(z_u) \\ z_u \end{array} \right) \geqslant 0$$

(where we require Assumption 3.9). Then (3.5.6) leads to (3.5.4). ∎

It is straightforward to generalize Theorem 3.7 along the same lines. One needs to *assume* that $\mathcal{I}_{M_{uu}}(\Delta)$ or $I - M_{uu}\Delta$ have a causal inverses, and one can work with the weaker hypothesis (3.2.5) to conclude in a similar fashion robust stability and robust performance. The details can be left to the reader and are omitted.

3.5.2   Guaranteeing Robust Quadratic Performance with IQC's

After this abstract motivating introduction, we turn our attention to the case that (3.5.1) is described by a finite dimensional stable LTI system; the corresponding transfer matrix is again denoted as $\widehat{M}$.

Let us look at the *quadratic performance* index

$$\Sigma_p \left(\begin{array}{c} w_p \\ z_p \end{array} \right) := \int_0^\infty \left(\begin{array}{c} w_p(t) \\ z_p(t) \end{array} \right)^T P_p \left(\begin{array}{c} w_p(t) \\ z_p(t) \end{array} \right) dt$$

where P_p is a fixed symmetric matrix that satisfies

$$P_p = \left(\begin{array}{cc} Q_p & S_p \\ S_p^T & R_p \end{array} \right), \quad R_p \geqslant 0 \tag{3.5.7}$$

to guarantee (3.5.5).

If Σ is defined by an IQC based on Π, (3.5.6) amounts to

$$\int_{-\infty}^\infty \left(\begin{array}{c} * \\ * \end{array} \right)^* \Pi(i\omega) \left(\begin{array}{c} \widehat{w}_u(i\omega) \\ \widehat{M}_{uu}(i\omega)\widehat{w}_u(i\omega) + \widehat{M}_{up}(i\omega)\widehat{w}_p(i\omega) \end{array} \right) d\omega +$$

$$+ \int_{-\infty}^\infty \left(\begin{array}{c} * \\ * \end{array} \right) P_p \left(\begin{array}{c} \widehat{w}_p(i\omega) \\ \widehat{M}_{pu}(i\omega)\widehat{w}_u(i\omega) + \widehat{M}_{pp}(i\omega)\widehat{w}_p(i\omega) \end{array} \right) d\omega \leqslant$$

$$\leqslant -\frac{\epsilon}{2\pi} \int_{-\infty}^\infty \left(\begin{array}{c} \widehat{w}_u(i\omega) \\ \widehat{w}_p(i\omega) \end{array} \right)^* \left(\begin{array}{c} \widehat{w}_u(i\omega) \\ \widehat{w}_p(i\omega) \end{array} \right) d\omega.$$

We can apply Lemma 3.13 to arrive at the equivalent condition

$$
\left(\begin{array}{cc} I & 0 \\ \widehat{M}_{uu} & \widehat{M}_{up} \end{array} \right)^* \Pi \left(\begin{array}{cc} I & 0 \\ \widehat{M}_{uu} & \widehat{M}_{up} \end{array} \right) + \left(\begin{array}{cc} 0 & I \\ \widehat{M}_{pu} & \widehat{M}_{pp} \end{array} \right)^* P_p \left(\begin{array}{cc} 0 & I \\ \widehat{M}_{pu} & \widehat{M}_{pp} \end{array} \right) \leqslant -\frac{\epsilon}{2\pi} I \quad (3.5.8)
$$

on $\mathbb{C}^0$. Again, this amounts to a simple frequency domain condition for which we provide an alternative formula in the following result.

Theorem 3.17 *Suppose that any $\Delta \in \mathbf{\Delta}$ satisfies the incremental IQC (3.4.3) for the multiplier $\Pi = \left(\begin{array}{cc} Q & S \\ S^* & R \end{array} \right)$. Moreover, suppose there exists an $\epsilon > 0$ with*

$$
\left(\begin{array}{cc} I & 0 \\ 0 & I \\ \widehat{M}_{uu} & \widehat{M}_{up} \\ \widehat{M}_{pu} & \widehat{M}_{pp} \end{array} \right)^* \left(\begin{array}{cc|cc} Q & 0 & S & 0 \\ 0 & Q_p & 0 & S_p \\ \hline S^* & 0 & R & 0 \\ 0 & S_p^T & 0 & R_p \end{array} \right) \left(\begin{array}{cc} I & 0 \\ 0 & I \\ \widehat{M}_{uu} & \widehat{M}_{up} \\ \widehat{M}_{pu} & \widehat{M}_{pp} \end{array} \right) \leqslant -\frac{\epsilon}{2\pi} I \quad on \ \mathbb{C}^0. \quad (3.5.9)
$$

Then the interconnection (3.2) is robustly stable, and for any $w_p \in L_2$ one has

$$
\int_0^\infty \left(\begin{array}{c} w_p(t) \\ z_p(t) \end{array} \right)^T P_p \left(\begin{array}{c} w_p(t) \\ z_p(t) \end{array} \right) dt \leqslant -\epsilon \|w_p\|_2^2.
$$

3.5.3   Guaranteeing Robust H_2 Performance with IQC's

For a strictly proper stable rational matrix $\widehat{M}$, the squared H_2-norm is defined as

$$
\|\widehat{M}\|_2^2 = \frac{1}{2\pi} \int_{-\infty}^\infty \text{trace}(\widehat{M}(i\omega)^* \widehat{M}(i\omega)) \, d\omega.
$$

We have mentioned that there are many interpretations and motivations why to consider this norm in design problems. Therefore, there are different manners to define the robust H_2 analysis problem.

In these notes we concentrate on one of these possible generalizations. For that purpose we characterize $\|\widehat{M}\|_2 < \gamma$ by requiring the existence of a symmetric matrix Q such that

$$
\frac{1}{2\pi} \int_{-\infty}^\infty \widehat{M}(i\omega)^* \widehat{M}(i\omega) \, d\omega < Q, \quad \text{trace}(Q) < \gamma^2.
$$

The first of these two inequalities admits a simple signal based interpretation; indeed it is equivalent to

$$
\int_0^\infty z(t)^T z(t) \, dt - w^T Q w \leqslant -\epsilon \|w\|^2
$$

for all real vectors w and $\widehat{z}(i\omega) = \widehat{M}(i\omega)w$. (Why?)

We intend to generalize this characterization to uncertain systems that are again described using a linear mapping

$$\left(\begin{array}{c} z_u \\ z_p \end{array} \right) = \left(\begin{array}{cc} M_{uu} & M_{up} \\ M_{pu} & M_{pp} \end{array} \right) \left(\begin{array}{c} w_u \\ w_p \end{array} \right) \tag{3.5.10}$$

that takes

$$w_u \in L_{2e}^{k_u}, \;\; w_p \in \mathbb{R}^{k_p} \;\; \text{into} \;\; z_u \in L_{2e}^{l_u}, \;\; z_p \in L_{2e}^{l_p}.$$

We assume that M_{uu}, M_{up} are causal and of finite L_2-gain, and that M_{up}, M_{pp} have finite gain as mappings from $\mathbb{R}^{k_p}$ to L_2; this means that there exists a constant m with

$$\| \left(\begin{array}{c} M_{up} w_p \\ M_{pp} w_p \end{array} \right) \|_2 \leqslant m \| w_p \| \;\; \text{for all} \;\; w_p \in \mathbb{R}^{k_p}.$$

(The theory presented so far does not directly encompass this case since we considered time-signals as inputs. Although one could easily extend the setup - by simply admitting the time-set $\{0\}$ for some components of the signals, - we view, instead, $M_{up} w_p$ and $M_{pp} w_p$ as L_2 disturbances and directly apply the techniques developed up to now.)

Robust stability is defined analogously as earlier, and the uncertainty system is said to have a robust H_2-level $\gamma > 0$ if there exists a symmetric Q with

$$\text{trace}(Q) < \gamma^2 \tag{3.5.11}$$

such that, for some $\epsilon > 0$,

$$-w_p^T Q w_p + \int_0^\infty z_p(t)^T z_p(t) \, dt \leqslant -\epsilon \| w_p \|^2 \;\; \text{for all} \;\; w_p \in \mathbb{R}^{k_p}. \tag{3.5.12}$$

It is now straightforward to arrive at the following IQC test for robust H_2 performance.

Theorem 3.18 *Suppose that every* $\Delta \in \mathbf{\Delta}$ *satisfies (3.4.3) and that* Q *is a symmetric matrix. Moreover, suppose there exists an* $\epsilon > 0$ *such that*

$$\int_{-\infty}^\infty \left(\begin{array}{c} \widehat{w}_u(i\omega) \\ \widehat{z}_u(i\omega) \end{array} \right)^* \Pi(i\omega) \left(\begin{array}{c} \widehat{w}_u(i\omega) \\ \widehat{z}_u(i\omega) \end{array} \right) + \widehat{z}_p(i\omega)^* \widehat{z}_p(i\omega) \, d\omega - w_p^T Q w_p \leqslant -\frac{\epsilon}{2\pi} (\| w_u \|_2^2 + \| w_p \|^2)$$

$$\tag{3.5.13}$$

for all $w_u \in L_2^{k_u}$, $w_p \in \mathbb{R}^{k_p}$ *and the corresponding outputs as defined in (3.5.10). Then* $I - M_{uu}\Delta$ *has a causal inverse with uniformly bounded incremental* L_2-*gain, and the uncertain system (3.5.3) satisfies (3.5.12).*

Proof. As for robust quadratic performance, the proof is a simple exercise and left to the reader. ∎

In a concrete test we have to view Q as an extra variable (besides Π) with (3.5.11) in order to satisfy (3.5.13). Hence Q can be viewed as a *varying* performance multiplier.

All this can be applied if M is an LTI system; then it is defined with real rational proper stable matrices $\widehat{M}_{uu}, \widehat{M}_{pu}, \widehat{M}_{up}, \widehat{M}_{pp}$ where $\widehat{M}_{up}(\infty) = 0$, $\widehat{M}_{pp}(\infty) = 0$, as

$$\begin{pmatrix} \widehat{z}_u(i\omega) \\ \widehat{z}_p(i\omega) \end{pmatrix} = \begin{pmatrix} \widehat{M}_{uu}(i\omega) & \widehat{M}_{up}(i\omega) \\ \widehat{M}_{pu}(i\omega) & \widehat{M}_{pp}(i\omega) \end{pmatrix} \begin{pmatrix} \widehat{w}_u(i\omega) \\ w_p \end{pmatrix}.$$

Remarks.

- Obviously, our robust H_2 property is an immediate extension of robust quadratic performance if equipping the space $\mathbb{R}^{k_p} \times L_2^{k_u}$ with the norm $\sqrt{\|w_p\|^2 + \|w_u\|_2^2}$ and defining the quadratic form $\Sigma_p : \mathbb{R}^{k_p} \times L_2^{l_p} \to \mathbb{R}$ as

$$\Sigma_p \begin{pmatrix} w_p \\ z_p \end{pmatrix} = -w_p^T Q w_p + \int_0^\infty z_p(t)^T z_p(t)\, dt.$$

  As earlier, one can also consider a general mapping Σ_p that satisfies $\Sigma_p \begin{pmatrix} 0 \\ z_p \end{pmatrix} \geqslant 0$.

- Suppose M is LTI. Paganini [22] has observed that the H_2-norm can be approximated by the gain of the underlying system for a certain class of finite power signals that approximate white noise. He shows that this concept can be then extended to uncertain systems similarly as done here, and derives necessary and sufficient conditions for robust H_2-performance against arbitrarily fast and arbitrarily slow time-varying uncertainties (similarly as Shamma [40], Megretsky [19], Poolla, Tikku [23] did for the robust L_2-gain problem). The conditions of [22] can be extended to the more general setting considered here and then read as follows: There exists an $\epsilon > 0$ and a measurable Hermitian valued Φ on $\mathbb{C}^0$

$$\frac{1}{2\pi} \int_{-\infty}^\infty \mathrm{trace}(\Phi(i\omega))\, d\omega < \gamma^2$$

  such that

$$\begin{pmatrix} I & 0 \\ 0 & I \\ \hline M_{uu} & M_{up} \\ M_{pu} & M_{pp} \end{pmatrix}^* \begin{pmatrix} Q & 0 & S & 0 \\ 0 & -\Phi & 0 & 0 \\ \hline S^* & 0 & R & 0 \\ 0 & 0 & 0 & I \end{pmatrix} \begin{pmatrix} I & 0 \\ 0 & I \\ \hline M_{uu} & M_{up} \\ M_{pu} & M_{pp} \end{pmatrix} \leqslant \begin{pmatrix} -\frac{\epsilon}{2\pi} I & 0 \\ 0 & 0 \end{pmatrix} \quad \text{on } \mathbb{C}^0,$$

  where Π is again partitioned as in Theorem 3.17. (Note that the latter inequality implies $M_{pp}^* M_{pp} \leqslant \Phi$ such that Φ is positive semi-definite. If the (2,2) block of the matrix on the right-hand side was negative definite, the $L_2(\mathbb{C}^0)$-norm of Φ could not be finite!) Again, the proof of sufficiency of these conditions is straightforward and left to the reader.

- In the robust quadratic performance problem, one can view P_p as an extra parameter that varies in a certain given class of matrices. D'Andrea [4] has worked out interesting variations of the H_∞-performance criterion that are amenable to this technique.

3.6   IQC's in the State-Space

Let us now look at multipliers Π that are real-rational. Then we can assume Π to be given in a specific form.

Lemma 3.19 *If Π is real rational and bounded on $\mathbb{C}^0$, there exists a P and a real rational proper stable Ψ with*

$$\Pi(i\omega) = \Psi(i\omega)^* P \Psi(i\omega).$$

Proof. Since Π is bounded, it is proper and there exists some $\alpha > 0$ such that $\Pi(i\omega) + \alpha I > 0$ for all $\omega \in \mathbb{R} \cup \{\infty\}$. Hence there exists a proper stable rational matrix T with $\Pi(i\omega) + \alpha I = T(i\omega)^* T(i\omega)$. This implies $\Pi(i\omega) = \begin{pmatrix} T(i\omega) \\ I \end{pmatrix}^* \begin{pmatrix} I & 0 \\ 0 & -\alpha I \end{pmatrix} \begin{pmatrix} T(i\omega) \\ I \end{pmatrix}$. ∎

The IQC

$$\int_{-\infty}^{\infty} \begin{pmatrix} \widehat{w}(i\omega) \\ \widehat{z}(i\omega) \end{pmatrix}^* \Pi(i\omega) \begin{pmatrix} \widehat{w}(i\omega) \\ \widehat{z}(i\omega) \end{pmatrix} d\omega = \int_{-\infty}^{\infty} \begin{pmatrix} \widehat{w}(i\omega) \\ \widehat{z}(i\omega) \end{pmatrix}^* \Psi(i\omega)^* P \Psi(i\omega) \begin{pmatrix} \widehat{w}(i\omega) \\ \widehat{z}(i\omega) \end{pmatrix} d\omega \geqslant 0$$

can, therefore, be rewritten as a static quadratic constraint

$$\int_0^{\infty} z_\Psi(t)^T P z_\Psi(t)\, dt \geqslant 0 \tag{3.6.1}$$

on the output of

$$z_\Psi = \Psi \begin{pmatrix} w \\ z \end{pmatrix}. \tag{3.6.2}$$

The system Ψ can be interpreted as a filter which encompasses the dynamics in the multiplier Π. Non-dynamic multipliers are simply obtained with $\Psi = I$ such that the only possibly freedom is left in P. With a minimal realization

$$\Psi = \left[\begin{array}{c|c} A_\Psi & B_\Psi \\ \hline C_\Psi & D_\Psi \end{array} \right],$$

we arrive at yet another parameterization of the multiplier Π in the state-space. Indeed, the system (3.6.2) now admits the description

$$\begin{pmatrix} \dot{x}_\Psi \\ z_\Psi \end{pmatrix} = \begin{pmatrix} A_\Psi \\ C_\Psi \end{pmatrix} x_\Psi + \begin{pmatrix} B_\Psi \\ D_\Psi \end{pmatrix} \begin{pmatrix} w \\ z \end{pmatrix}, \quad x_\Psi(0) = 0.$$

3.6.1   Robust Stability

In order to apply Theorems 3.15 or 3.14 to guarantee robust stability, we need to check the FDI

$$\begin{pmatrix} I \\ \widehat{M} \end{pmatrix}^* \Pi(i\omega) \begin{pmatrix} I \\ \widehat{M} \end{pmatrix} = [\Psi \begin{pmatrix} I \\ \widehat{M} \end{pmatrix}]^* P [\Psi \begin{pmatrix} I \\ \widehat{M} \end{pmatrix}] < 0 \quad \text{on } \mathbb{C}^0 \cup \{\infty\}. \tag{3.6.3}$$

For a state-space characterization, we choose a realization of

$$\begin{pmatrix} w \\ z \end{pmatrix} = \begin{pmatrix} I \\ \widehat{M} \end{pmatrix} w$$

as

$$\dot{x} = Ax + Bw, \quad \begin{pmatrix} w \\ z \end{pmatrix} = Cx + Dw$$

such that A is Hurwitz. (Note that C and D have a specific structure due to the fact that the first component of the output equals the input w.) Then we infer

$$\Psi \begin{pmatrix} I \\ \widehat{M} \end{pmatrix} = \left[\begin{array}{cc|c} A & 0 & B \\ B_\Psi C & A_\Psi & B_\Psi D \\ \hline D_\Psi C & C_\Psi & D_\Psi D \end{array} \right] = \left[\begin{array}{c|c} \tilde{A} & \tilde{B} \\ \hline \tilde{C} & \tilde{D} \end{array} \right].$$

The Kalman-Yakubovich-Popov Lemma reveals that the FDI (3.6.3) is equivalent to the solvability of a linear matrix inequality.

Lemma 3.20 *The FDI (3.6.3) holds iff there exists a symmetric solution X of the LMI*

$$\begin{pmatrix} \tilde{A}^T X + X\tilde{A} & X\tilde{B} \\ \tilde{B}^T X & 0 \end{pmatrix} + \begin{pmatrix} \tilde{C} & \tilde{D} \end{pmatrix}^T P \begin{pmatrix} \tilde{C} & \tilde{D} \end{pmatrix} < 0.$$

3.6.2 Robust Quadratic Performance

The FDI (3.5.8) that characterizes robust quadratic performance is treated similarly. Indeed, it is equivalent to

$$\begin{pmatrix} 0 & I \\ \widehat{M}_{uu} & \widehat{M}_{up} \end{pmatrix}^* \Psi^* P \Psi \begin{pmatrix} 0 & I \\ \widehat{M}_{uu} & \widehat{M}_{up} \end{pmatrix} + \begin{pmatrix} 0 & I \\ \widehat{M}_{pu} & \widehat{M}_{pp} \end{pmatrix}^* P_p \begin{pmatrix} 0 & I \\ \widehat{M}_{pu} & \widehat{M}_{pp} \end{pmatrix} < 0 \quad (3.6.4)$$

on $\mathbb{C}^0 \cup \{\infty\}$.

Introduce the minimal realization

$$\left(\begin{pmatrix} I & 0 \\ \widehat{M}_{uu} & \widehat{M}_{up} \end{pmatrix} \\ \begin{pmatrix} 0 & I \\ \widehat{M}_{pu} & \widehat{M}_{pp} \end{pmatrix} \right) = \left[\begin{array}{c|cc} A & B_u & B_p \\ \hline C_u & D_{uu} & D_{up} \\ C_p & D_{pu} & D_{pp} \end{array} \right] \qquad (3.6.5)$$

with an A that is Hurwitz to arrive at

$$\left(\Psi \begin{pmatrix} I & 0 \\ \widehat{M}_{uu} & \widehat{M}_{up} \end{pmatrix} \\ \begin{pmatrix} 0 & I \\ \widehat{M}_{pu} & \widehat{M}_{pp} \end{pmatrix} \right) = \left[\begin{array}{cc|cc} A & 0 & B_u & B_p \\ B_\Psi C_u & A_\Psi & B_\Psi D_{uu} & B_\Psi D_{up} \\ \hline D_\Psi C_u & C_\Psi & D_\Psi D_{uu} & D_\Psi D_{up} \\ C_p & 0 & D_{pu} & D_{pp} \end{array} \right] = \left[\begin{array}{c|cc} \tilde{A} & \tilde{B}_u & \tilde{B}_p \\ \hline \tilde{C}_\Psi & \tilde{D}_{\Psi u} & \tilde{D}_{\Psi p} \\ \tilde{C}_p & \tilde{D}_{pu} & \tilde{D}_{pp} \end{array} \right].$$

$$(3.6.6)$$

Again, the Kalman-Yakubovich-Popov Lemma allows to characterize this FDI in terms of the solvability of an LMI.

Lemma 3.21 *The FDI (3.6.4) holds iff there exits an X satisfying*

$$\begin{pmatrix} \tilde{A}^T X + X \tilde{A} & X \tilde{B}_u & X \tilde{B}_p \\ \tilde{B}_u^T X & 0 & 0 \\ \tilde{B}_p^T X & 0 & 0 \end{pmatrix} + \begin{pmatrix} \tilde{C}_\Psi & \tilde{D}_{\Psi u} & \tilde{D}_{\Psi p} \\ \tilde{C}_p & \tilde{D}_{pu} & \tilde{D}_{pp} \end{pmatrix}^T \begin{pmatrix} P & 0 \\ 0 & P_p \end{pmatrix} \begin{pmatrix} \tilde{C}_\Psi & \tilde{D}_{\Psi u} & \tilde{D}_{\Psi p} \\ \tilde{C}_p & \tilde{D}_{pu} & \tilde{D}_{pp} \end{pmatrix} < 0.$$

3.6.3 Robust H_2 Performance

For robust H_2 performance we have to guarantee (3.5.13). We introduce the realization (3.6.5) (with stable A) and recall that $\widehat{M}_{up}$ and $\widehat{M}_{pp}$ are strictly proper such that $D_{up} = 0$ and $D_{pp} = 0$. This implies $\tilde{D}_{up} = 0$ and $\tilde{D}_{pp} = 0$ for the realization (3.6.6) that incorporates the dynamics of the multiplier. Let us introduce the abbreviation

$$P_p := \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix}$$

such that

$$z_p^T z_p = \begin{pmatrix} w_p \\ z_p \end{pmatrix}^T P_p \begin{pmatrix} w_p \\ z_p \end{pmatrix}.$$

Then (3.5.13) is equivalent to the existence of some $\epsilon > 0$ with

$$\int_0^\infty \begin{pmatrix} \tilde{z}_u(t) \\ \tilde{z}_p(t) \end{pmatrix}^T \begin{pmatrix} P & 0 \\ 0 & P_p \end{pmatrix} \begin{pmatrix} \tilde{z}_u(t) \\ \tilde{z}_p(t) \end{pmatrix} dt - w_p^T Q w_p \leqslant -\epsilon(\|w_u\|_2^2 + \|w_p\|^2) \qquad (3.6.7)$$

for all $w_u \in L_2^{k_u}$, $w_p \in \mathbb{R}^{k_p}$ and for the output of

$$\begin{pmatrix} \dot{x} \\ \tilde{z}_u \\ \tilde{z}_p \end{pmatrix} = \begin{pmatrix} \tilde{A} & \tilde{B}_u \\ \tilde{C}_\Psi & \tilde{D}_{\Psi u} \\ \tilde{C}_p & \tilde{D}_{pu} \end{pmatrix} \begin{pmatrix} x \\ w_u \end{pmatrix}, \quad x(0) = \tilde{B}_p w_p. \qquad (3.6.8)$$

Again, by the Kalman-Yakubovich-Popov Lemma, this condition turns out to be equivalent to the solvability of an LMI.

Lemma 3.22 *The conditions (3.5.13) for robust H_2-performance holds iff there exists an X satisfying*

$$\tilde{B}_P^T X \tilde{B}_p < Q, \quad \begin{pmatrix} \tilde{A}^T X + X \tilde{A} & X \tilde{B}_u \\ \tilde{B}_u^T X & 0 \end{pmatrix} + \begin{pmatrix} \tilde{C}_\Psi & \tilde{D}_{\Psi u} \\ \tilde{C}_p & \tilde{D}_{pu} \end{pmatrix}^T \begin{pmatrix} P & 0 \\ 0 & P_p \end{pmatrix} \begin{pmatrix} \tilde{C}_\Psi & \tilde{D}_{\Psi u} \\ \tilde{C}_p & \tilde{D}_{pu} \end{pmatrix} < 0.$$

One should compare with robust quadratic performance.

3.7 A Summary and Extensions

In this section we intend to merge robust quadratic performance and the robust H_2 specification into one result. Furthermore, we summarize the required hypotheses, we provide alternative proofs based on Lyapunov arguments whenever possible, and we discuss the consequences of strengthening the IQC hypotheses on the uncertainties.

With a set $\mathbf{\Delta}$ of systems $\Delta : L_{2e}^l \to L_{2e}^k$ that are causal and of finite L_2-gain, we consider the uncertain system

$$\begin{pmatrix} \dot{x} \\ \hline z_u \\ z_p \end{pmatrix} = \left(\begin{array}{c|cc} A & B_u & B_p \\ \hline C_u & D_{uu} & D_{up} \\ C_p & D_{pu} & D_{pp} \end{array} \right) \begin{pmatrix} x \\ \hline w_u \\ w_p \end{pmatrix}, \quad x(0) = x_0, \quad w_u = \Delta(z_u), \quad \Delta \in \mathbf{\Delta}. \quad (3.7.1)$$

P_p is the performance index matrix that satisfies

$$\begin{pmatrix} 0 \\ I \end{pmatrix}^T P_p \begin{pmatrix} 0 \\ I \end{pmatrix} \geqslant 0. \quad (3.7.2)$$

The dynamics of the considered IQC is given as

$$\begin{pmatrix} \dot{x}_\Psi \\ z_\Psi \end{pmatrix} = \begin{pmatrix} A_\Psi \\ C_\Psi \end{pmatrix} x_\Psi + \begin{pmatrix} B_\Psi \\ D_\Psi \end{pmatrix} \begin{pmatrix} w \\ z \end{pmatrix}, \quad x_\Psi(0) = 0 \quad (3.7.3)$$

and P is the index matrix of the IQC.

Let us now interconnect (3.7.1) with (3.7.2) as

$$\begin{pmatrix} w \\ z \end{pmatrix} = \begin{pmatrix} w_u \\ z_u \end{pmatrix}. \quad (3.7.4)$$

The dynamics of the resulting system admits the state-space description

$$\begin{pmatrix} \dot{x} \\ \dot{x}_\Psi \\ z_\Psi \\ \begin{pmatrix} w_p \\ z_p \end{pmatrix} \end{pmatrix} = \underbrace{\left(\begin{array}{cc|cc} A & 0 & B_u & B_p \\ B_\Psi \begin{pmatrix} 0 \\ C_u \end{pmatrix} & A_\Psi & B_\Psi \begin{pmatrix} I \\ D_{uu} \end{pmatrix} & B_\Psi \begin{pmatrix} 0 \\ D_{up} \end{pmatrix} \\ \hline D_\Psi \begin{pmatrix} 0 \\ C_u \end{pmatrix} & C_\Psi & D_\Psi \begin{pmatrix} I \\ D_{uu} \end{pmatrix} & D_\Psi \begin{pmatrix} 0 \\ D_{up} \end{pmatrix} \\ \begin{pmatrix} 0 \\ C_p \end{pmatrix} & 0 & \begin{pmatrix} 0 \\ D_{pu} \end{pmatrix} & \begin{pmatrix} I \\ D_{pp} \end{pmatrix} \end{array} \right)}_{\left(\begin{array}{c|cc} \tilde{A} & \tilde{B}_u & \tilde{B}_p \\ \hline \tilde{C}_\Psi & \tilde{D}_{\Psi u} & \tilde{D}_{\Psi p} \\ \tilde{C}_p & \tilde{D}_{pu} & \tilde{D}_{pp} \end{array} \right)} \begin{pmatrix} x \\ x_\Psi \\ \hline w_u \\ w_p \end{pmatrix}. \quad (3.7.5)$$

Finally, suppose that the symmetric matrix X satisfies the LMI

$$\begin{pmatrix} I & 0 & 0 \\ \tilde{A} & \tilde{B}_u & \tilde{B}_p \end{pmatrix}^T \begin{pmatrix} 0 & X \\ X & 0 \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ \tilde{A} & \tilde{B}_u & \tilde{B}_p \end{pmatrix} +$$

$$+ \begin{pmatrix} \tilde{C}_\Psi & \tilde{D}_{\Psi u} & \tilde{D}_{\Psi p} \\ \tilde{C}_p & \tilde{D}_{pu} & \tilde{D}_{pp} \end{pmatrix}^T \begin{pmatrix} P & 0 \\ 0 & P_p \end{pmatrix} \begin{pmatrix} \tilde{C}_\Psi & \tilde{D}_{\Psi u} & \tilde{D}_{\Psi p} \\ \tilde{C}_p & \tilde{D}_{pu} & \tilde{D}_{pp} \end{pmatrix} < 0. \quad (3.7.6)$$

Note that the first term in this LMI just equals

$$\begin{pmatrix} \tilde{A}^T X + X\tilde{A} & X\tilde{B}_u & X\tilde{B}_p \\ \tilde{B}_u^T X & 0 & 0 \\ \tilde{B}_p^T X & 0 & 0 \end{pmatrix}$$

but the formula given above is slightly more illustrative for our purposes.

This list of ingredients is motivated by the discussion in Section 3.6. At this point we have not yet specified the exact relation of the uncertainties and the IQC dynamics what will be done in the next subsections. We first proceed with some preparatory remarks.

For some small $\epsilon > 0$, we can replace the matrix P_p in (3.7.6) by $P_p + \epsilon I$ and the right-hand side 0 by $-\epsilon I$ without violating the LMI (3.7.6). For any trajectory of (3.7.5), we right-multiply the resulting inequality with

$$\left(\begin{pmatrix} x(t) \\ x_\Psi(t) \end{pmatrix} \\ w_u(t) \\ w_p(t) \right)$$

and left-multiply with its transpose. We obtain

$$\frac{d}{dt} \begin{pmatrix} x(t) \\ x_\Psi(t) \end{pmatrix}^T X \begin{pmatrix} x(t) \\ x_\Psi(t) \end{pmatrix} + z_\Psi(t)^T P z_\Psi(t) + \begin{pmatrix} w_p(t) \\ z_p(t) \end{pmatrix}^T (P_p + \epsilon I) \begin{pmatrix} w_p(t) \\ z_p(t) \end{pmatrix} \leqslant$$

$$\leqslant -\epsilon \| \begin{pmatrix} x(t) \\ x_\Psi(t) \end{pmatrix} \|^2 - \epsilon \| \begin{pmatrix} w_u(t) \\ w_p(t) \end{pmatrix} \|^2 \quad \text{for all } t \geqslant 0 \quad (3.7.7)$$

and, after integration on $[0, T]$,

$$\begin{pmatrix} x(T) \\ x_\Psi(T) \end{pmatrix}^T X \begin{pmatrix} x(T) \\ x_\Psi(T) \end{pmatrix} - \begin{pmatrix} x(0) \\ x_\Psi(0) \end{pmatrix}^T X \begin{pmatrix} x(0) \\ x_\Psi(0) \end{pmatrix} +$$

$$+ \int_0^T z_\Psi(t)^T P z_\Psi(t)\, dt + \int_0^T \begin{pmatrix} w_p(t) \\ z_p(t) \end{pmatrix}^T (P_p + \epsilon I) \begin{pmatrix} w_p(t) \\ z_p(t) \end{pmatrix} dt \leqslant$$

$$\leqslant -\epsilon \int_0^T \| \begin{pmatrix} x(t) \\ x_\Psi(t) \end{pmatrix} \|^2 + \| \begin{pmatrix} w_u(t) \\ w_p(t) \end{pmatrix} \|^2 dt \quad \text{for all } T \geqslant 0. \quad (3.7.8)$$

We further exploit (3.7.2) to find a $\delta > 0$ with

$$\begin{pmatrix} \delta I & 0 \\ 0 & 0 \end{pmatrix} + (P_p + \epsilon I) > 0$$

(why?) such that we can conclude

$$- \begin{pmatrix} w_p(t) \\ z_p(t) \end{pmatrix}^T (P_p + \epsilon I) \begin{pmatrix} w_p(t) \\ z_p(t) \end{pmatrix} \leqslant \delta \|w_p(t)\|^2. \tag{3.7.9}$$

All new (Lyapunov based) arguments in this section are based on these three relations.

3.7.1   Well-Posedness with Soft Incremental IQC's

We assume that A and A_Ψ are Hurwitz and that $\mathbf{\Delta}$ is star-shaped with star center 0. Moreover, for each $\Delta \in \mathbf{\Delta}$ and $z_1, z_2 \in L_2$, the output of (3.7.3) for

$$\begin{pmatrix} w \\ z \end{pmatrix} = \begin{pmatrix} \Delta(z_1) - \Delta(z_2) \\ z_1 - z_2 \end{pmatrix}$$

satisfies

$$\int_0^\infty z_\Psi(t)^T P z_\Psi(t)\, dt \geqslant 0. \tag{3.7.10}$$

Then a suitable adaption of the proof of Theorem 3.8 reveals that the solvability of (3.7.6) implies

Well-posedness: For each $x_0 \in \mathbb{R}^n$ and $w_p \in L_{2e}$, the system (3.7.1) admits a unique response $x, w_u \in L_{2e}$.

Recall that this proof heavily relies on the fact that L_2 is a Banach space. However, for all the remaining statements in this section, this property will not be exploited. Hence, one could e.g. guarantee well-posedness by standard results on the existence of solutions of differential equations (such as Lipschitz conditions plus linear boundedness properties) that are derived by other techniques. It is then no problem to adapt the considered class of signals - such as to the set of piece-wise continuous (and continuous/piecewise continuously differentiable x) or continuous signals of finite energy - to the result that has been applied. Hence, the 'smoothness' properties of the trajectories is mainly dictated by those results that are available to show well-posedness. Once well-posedness has been established (by whatsoever technique), one can often restrict the attention to signal subspaces that are technically easier to handle.

Let us mention a situation in which well-posedness is easy to verify. Suppose that the uncertainty $w = \Delta(z)$ is described by an LTI system

$$\dot{x}_\Delta = A_\Delta x_\Delta + B_\Delta z, \quad \begin{pmatrix} w \\ z \end{pmatrix} = \begin{pmatrix} C_\Delta \\ 0 \end{pmatrix} x_\Delta + \begin{pmatrix} D_\Delta \\ I \end{pmatrix} z, \quad x_\Delta(0) = 0.$$

Then (3.7.1) is well-posed if

$$\begin{pmatrix} I & D_\Delta \\ D_{uu} & I \end{pmatrix} \quad \text{is nonsingular.} \tag{3.7.11}$$

(Why?)

The soft IQC condition on Δ now reads as (3.7.10) on the output of

$$\begin{pmatrix} \dot{x}_\Delta \\ \dot{x}_\Psi \\ z_\Psi \end{pmatrix} = \left(\begin{array}{cc|c} A_\Delta & 0 & B_\Delta \\ B_\Psi \begin{pmatrix} C_\Delta \\ 0 \end{pmatrix} & A_\Psi & B_\Psi \begin{pmatrix} D_\Delta \\ I \end{pmatrix} \\ \hline D_\Psi \begin{pmatrix} C_\Delta \\ 0 \end{pmatrix} & C_\Psi & D_\Psi \begin{pmatrix} D_\Delta \\ I \end{pmatrix} \end{array} \right) \begin{pmatrix} x_\Delta \\ x_\Psi \\ z \end{pmatrix}, \quad \begin{pmatrix} x_\Delta \\ x_\Psi \end{pmatrix}(0) = 0.$$

On the one hand, taking the resulting FDI at $\omega = \infty$ reveals

$$[D_\Psi \begin{pmatrix} D_\Delta \\ I \end{pmatrix}]^T P D_\Psi \begin{pmatrix} D_\Delta \\ I \end{pmatrix} \geqslant 0. \tag{3.7.12}$$

On the other hand, (3.7.6) implies

$$0 > \begin{pmatrix} \tilde{D}_{\Psi u} \\ \tilde{D}_{pu} \end{pmatrix}^T \begin{pmatrix} P & 0 \\ 0 & P_p \end{pmatrix} \begin{pmatrix} \tilde{D}_{\Psi u} \\ \tilde{D}_{pu} \end{pmatrix} =$$

$$= [D_\Psi \begin{pmatrix} I \\ D_{uu} \end{pmatrix}]^T P [D_\Psi \begin{pmatrix} I \\ D_{uu} \end{pmatrix}] + \begin{pmatrix} 0 \\ D_{pu} \end{pmatrix}^T P_p \begin{pmatrix} 0 \\ D_{pu} \end{pmatrix}.$$

Due to (3.7.2), we infer

$$[D_\Psi \begin{pmatrix} I \\ D_{uu} \end{pmatrix}]^T P [D_\Psi \begin{pmatrix} I \\ D_{uu} \end{pmatrix}] < 0. \tag{3.7.13}$$

The two inequalities (3.7.12) and (3.7.13) imply that

$$\begin{pmatrix} D_\Psi \begin{pmatrix} I \\ D_{uu} \end{pmatrix} & D_\Psi \begin{pmatrix} D_\Delta \\ I \end{pmatrix} \end{pmatrix}$$

has full column rank what leads to well-posedness (3.7.11).

Similar arguments apply for time-varying uncertainties if one can assure (3.7.12) to hold. Note that we used a frequency domain argument to *infer* (3.7.12) from the soft IQC; this argument breaks down for time-varying uncertainties.

From now on we **assume that well-posedness has been verified**.

3.7.2   Soft IQC's

Again, suppose A and A_Ψ are Hurwitz, that the LMI (3.7.6) holds, and that $\boldsymbol{\Delta}$ is star-shaped with center 0. For each $\Delta \in \boldsymbol{\Delta}$ and for any $z, w \in L_2$ such that $w = \Delta(z)$, the output of (3.7.3) satisfies (3.7.10). Similarly as Theorem 3.7 one proves

Robust Stability: For all $x_0 \in \mathbb{R}^n$ and $w_p \in L_2$, the unique system response of (3.7.1) satisfies $x, w_u \in L_2$.

For the remaining properties of uniform robust stability and robust performance we provide independent (very elementary) proofs.

Uniform Robust Stability: There exist constants K_1, K_2 such that for every $x_0 \in \mathbb{R}^n$ and every $w_p \in L_2$

$$\|x\|_2^2 + \|w_u\|_2^2 \leqslant K_1 \|w_p\|_2^2 + K_2 \|x_0\|^2 \quad \text{and} \quad \lim_{t \to \infty} x(t) = 0.$$

We could add (after possibly modifying K_1, K_2) $\|z_u\|_2^2$, $\|z_p\|_2^2$ and $\|x_\Psi\|_2^2$, $\|z_\Psi\|_2^2$ on the left-hand side and the inequality still remains true. (Why?)

Proof. For $x_0 \in \mathbb{R}^n$, $w_p \in L_2$, we conclude for the unique system response that $w_u, x \in L_2$ and hence (with the system's differential equation and since L_2 is a linear space) $\dot{x}, z_u, z_p \in L_2$; this implies $\lim_{t \to \infty} x(t) = 0$. For the interconnection of (3.7.1) and (3.7.3) according to (3.7.4), we infer

$$\begin{pmatrix} \dot{x}_\Psi \\ z_\Psi \end{pmatrix} = \begin{pmatrix} A_\Psi \\ C_\Psi \end{pmatrix} x_\Psi + \begin{pmatrix} B_\Psi \\ D_\Psi \end{pmatrix} \begin{pmatrix} \Delta(z_u) \\ z_u \end{pmatrix}, \quad x_\Psi(0) = 0.$$

Since A_Ψ is stable, we infer $x_\Psi, \dot{x}_\Psi, z_\Psi \in L_2$ and $\lim_{t \to \infty} x_\Psi(t) = 0$. Moreover, (3.7.10) holds true. If we take the limit $T \to \infty$ in (3.7.8), and if we combine with (3.7.10), we get

$$\int_0^\infty \begin{pmatrix} w_p(t) \\ z_p(t) \end{pmatrix}^T (P_p + \epsilon I) \begin{pmatrix} w_p(t) \\ z_p(t) \end{pmatrix} dt \leqslant$$

$$\leqslant \begin{pmatrix} x_0 \\ 0 \end{pmatrix}^T X \begin{pmatrix} x_0 \\ 0 \end{pmatrix} - \epsilon \int_0^\infty \| \begin{pmatrix} x(t) \\ x_\Psi(t) \end{pmatrix} \|^2 + \| \begin{pmatrix} w_u(t) \\ w_p(t) \end{pmatrix} \|^2 dt. \quad (3.7.14)$$

Exploiting (3.7.9) leads, as required, to

$$\epsilon \int_0^\infty \| \begin{pmatrix} x(t) \\ x_\Psi(t) \end{pmatrix} \|^2 + \| \begin{pmatrix} w_u(t) \\ w_p(t) \end{pmatrix} \|^2 dt \leqslant \delta \int_0^\infty \|w_p(t)\|^2 dt + \begin{pmatrix} x(0) \\ 0 \end{pmatrix}^T X \begin{pmatrix} x(0) \\ 0 \end{pmatrix}.$$

∎

Robust Performance: The system is (uniformly) robustly stable, and for all $x_0 \in \mathbb{R}$ and $w_p \in L_2$ one has

$$\int_0^\infty \begin{pmatrix} w_p(t) \\ z_p(t) \end{pmatrix}^T P_p \begin{pmatrix} w_p(t) \\ z_p(t) \end{pmatrix} dt \leqslant \begin{pmatrix} x_0 \\ 0 \end{pmatrix}^T X \begin{pmatrix} x_0 \\ 0 \end{pmatrix}.$$

Proof. This immediately follows from (3.7.14). ■

For some given symmetric $Q \in \mathbb{R}^{n \times n}$ and some subspace $\mathcal{X}$ of $\mathbb{R}^n$, we can impose the extra constraint

$$\begin{pmatrix} I \\ 0 \end{pmatrix}^T X \begin{pmatrix} I \\ 0 \end{pmatrix} < Q \quad \text{on} \quad \mathcal{X}$$

to infer

$$\int_0^\infty \begin{pmatrix} w_p(t) \\ z_p(t) \end{pmatrix}^T P_p \begin{pmatrix} w_p(t) \\ z_p(t) \end{pmatrix} dt \leqslant x_0^T Q x_0$$

for every $w_p \in L_2$ and every $x_0 \in \mathcal{X}$; this reveals that we have indeed merged the previous robust H_2 and robust quadratic performance specification into one result.

3.7.3 Hard IQC's

Recall that we assume well-posedness of (3.7.1). Suppose that, for each $\Delta \in \mathbf{\Delta}$ and $z, w \in L_{2e}$ with $w = \Delta(z)$, the output of (3.7.3) satisfies

$$\int_0^T z_\Psi(t)^T P z_\Psi(t) \, dt \geqslant 0 \quad \text{for all} \quad T \geqslant 0. \tag{3.7.15}$$

We say that Δ satisfies a *hard IQC*. In addition, we include the hypothesis that the LMI (3.7.6) has a solution

$$X > 0.$$

Remark. Note that the set $\mathbf{\Delta}$ is not necessarily star-shaped; it can even be just a singleton. Moreover, note that A and A_Ψ are not necessarily stable. Whether positivity of X implies or is implied by the stability of A, A_Ψ depends on the specific matrices P and P_p, and no general claim is possible.

Now we can provide a direct proof of robust stability and we can strengthen the uniform robust stability and robust performance conclusions.

Robust Stability: For $x_0 \in \mathbb{R}^n$, $w_p \in L_2$, the unique response of (3.7.1) satisfies $x, w_u \in L_2$.

Proof. The proof is directly based on (3.7.8) and (3.7.9). Indeed, we infer from (3.7.8) that

$$\begin{pmatrix} x(T) \\ x_\Psi(T) \end{pmatrix}^T X \begin{pmatrix} x(T) \\ x_\Psi(T) \end{pmatrix} + \int_0^T \begin{pmatrix} w_p(t) \\ z_p(t) \end{pmatrix}^T (P_p + \epsilon I) \begin{pmatrix} w_p(t) \\ z_p(t) \end{pmatrix} dt \leqslant$$

$$\leqslant \begin{pmatrix} x_0 \\ 0 \end{pmatrix}^T X \begin{pmatrix} x_0 \\ 0 \end{pmatrix} - \epsilon \int_0^T \left\| \begin{pmatrix} x(t) \\ x_\Psi(t) \end{pmatrix} \right\|^2 + \left\| \begin{pmatrix} w_u(t) \\ w_p(t) \end{pmatrix} \right\|^2 dt \tag{3.7.16}$$

and combining with (3.7.9) gives

$$\begin{pmatrix} x(T) \\ x_\Psi(T) \end{pmatrix}^T X \begin{pmatrix} x(T) \\ x_\Psi(T) \end{pmatrix} + \epsilon \int_0^T \| \begin{pmatrix} x(t) \\ x_\Psi(t) \end{pmatrix} \|^2 + \| \begin{pmatrix} w_u(t) \\ w_p(t) \end{pmatrix} \|^2 \, dt \leqslant$$

$$\leqslant \begin{pmatrix} x_0 \\ 0 \end{pmatrix}^T X \begin{pmatrix} x_0 \\ 0 \end{pmatrix} + \delta \int_0^T \|w_p(t)\|^2 \, dt \quad (3.7.17)$$

for all $T \geqslant 0$. Since the right-hand side is bounded for $T \to \infty$, we infer the same (due to $X > 0$) for $\int_0^T \|x(t)\|^2 \, dt$, $\int_0^T \|x_\Psi(t)\|^2 \, dt$, and $\int_0^T \|w_u(t)\|^2 \, dt$; this shows $x, x_\Psi, w_u \in L_2$.

Uniform Robust Stability: The system is robustly stable, and there exist constants K_1, K_2 such that for every $x_0 \in \mathbb{R}^n$ and every $w_p \in L_2$

$$\|x(T)\|^2 + \int_0^T \|x(t)\|^2 + \|w_u(t)\|^2 \, dt \leqslant K_1 \int_0^T \|w_p(t)\|^2 \, dt + K_2 \|x_0\|^2 \quad \text{for all } T \geqslant 0$$

and $\lim_{t \to \infty} x(t) = 0$. Again we could add (for possibly other constants) $\int_0^T \|z_u(t)\|^2 + \|z_p(t)\|^2 + \|x_\Psi(t)\|^2 + \|z_\Psi(t)\|^2 \, dt$ and $\|x_\Psi(T)\|$ on the left-hand side. (Why?)

Proof. This is immediate from (3.7.17). ∎

Robust Performance: The system is (uniformly) robustly stable, and for all $x_0 \in \mathbb{R}$ and $w_p \in L_2$ one has

$$\begin{pmatrix} x(T) \\ x_\Psi(T) \end{pmatrix}^T X \begin{pmatrix} x(T) \\ x_\Psi(T) \end{pmatrix} + \int_0^T \begin{pmatrix} w_p(t) \\ z_p(t) \end{pmatrix}^T P_p \begin{pmatrix} w_p(t) \\ z_p(t) \end{pmatrix} \, dt \leqslant$$

$$\leqslant \begin{pmatrix} x_0 \\ 0 \end{pmatrix}^T X \begin{pmatrix} x_0 \\ 0 \end{pmatrix} \quad \text{for all } T \geqslant 0.$$

Proof. Immediate with (3.7.16). ∎

3.7.4 QC's

Again, we assume well-posedness of (3.7.1). Furthermore, suppose that for each $\Delta \in \mathbf{\Delta}$ and $w, z \in L_{2e}$ with $w = \Delta(z)$, the output of (3.7.3) satisfies

$$z_\Psi(t)^T P z_\Psi(t) \geqslant 0 \quad \text{for all } t \geqslant 0. \quad (3.7.18)$$

This is a *quadratic constraint* (QC) in time. Moreover, let us suppose that the LMI (3.7.6) has the solution $X > 0$.

Note that this QC implies the hard IQC (3.7.15) such that all conclusions from the previous subsection are still true. However, due to the constraint point-wise in time, we can now argue directly with (3.7.7); this leads to uniform exponential stability.

Uniform Exponential Stability: There exist constants $\alpha > 0$, K_3, K_4 such that

$$\|x(T)\|^2 \leqslant K_3 \|x(t_0)\|^2 e^{-\alpha(T-t_0)} + K_4 \int_{t_0}^T \|w_p(t)\|_2^2 \, dt \quad \text{for all} \ \ T \geqslant t_0 \geqslant 0.$$

If $w_p = 0$, this implies that the system state $x(T)$ converges exponentially to zero for $T \to \infty$. Since the constants K_3 and α do not depend on the specific Δ, the exponential stability is uniform in the uncertainty.

Proof. Let us introduce the abbreviations

$$v(t) := \left(\begin{array}{c} x(t) \\ x_\Psi(t) \end{array} \right)^T X \left(\begin{array}{c} x(t) \\ x_\Psi(t) \end{array} \right), \quad \alpha := \frac{\epsilon}{\lambda_{\min}(X)} > 0.$$

Then (3.7.7) clearly implies

$$\frac{d}{dt}v(t) + \alpha v(t) + \left(\begin{array}{c} w_p(t) \\ z_p(t) \end{array} \right)^T (P_p + \epsilon I) \left(\begin{array}{c} w_p(t) \\ z_p(t) \end{array} \right) \leqslant 0 \quad \text{for all} \ \ t \geqslant 0.$$

Combining with (3.7.9) gives

$$\frac{d}{dt}v(t) + \alpha v(t) \leqslant \delta \|w_p(t)\|^2 \quad \text{for all} \ \ t \geqslant 0.$$

For any $t_0 \geqslant 0$ we conclude

$$v(T) \leqslant v(t_0)e^{-\alpha(T-t_0)} + \delta \int_{t_0}^T e^{-\alpha(T-t)} \|w_p(t)\|^2 \, dt.$$

We conclude that there exist constants K_3, K_4 such that

$$\|x(T)\|^2 \leqslant K_3 \|x(t_0)\|^2 e^{-\alpha(T-t_0)} + K_4 \int_0^T \|w_p(t)\|_2^2 \, dt.$$

Remarks.

- Typical examples for the latter results to apply are time-varying parametric uncertainties, static non-linearities, or non-linearities defined by differential equations. Often, mild Lipschitz-type conditions suffice to guarantee that (3.7.1) has a solution for each initial condition and for each disturbance. Usually, such solutions are only given locally in time. (Recall that nonlinear differential equations can have a finite escape time.) The stability results for hard IQC's or QC's might allow, however, to exclude e.g. a finite escape time such that solutions can be extended to the whole interval $[0, \infty)$. Hence, apart from stability, also certain aspects of well-posedness could be shown with the arguments that we presented in this section.

- **Further weakening the IQC's.** We have assumed the hard IQC's or the QC's to hold for all $w, z \in L_{2e}$ with $w = \Delta(z)$. However, it is obvious that all proofs given in this section only required them to be satisfied if, in addition, (3.7.4) holds for some trajectory of the uncertain system (3.7.1).

- We include two exercises about the multi-variable circle criterion and the Popov criterion that reveal the subtleties discussed here and, nevertheless, allow to re-prove classical results in a straightforward manner without the need for technical hypotheses that often occur in the literature [14].

3.8   Other Performance Specifications in the State-Space

In this section we want to clarify how to extend the results to other performance criteria different from robust quadratic or H_2-performance. As a rule, all those criteria that can be formulated in terms of a quadratic Lyapunov function on the system extended with the IQC dynamics can be considered. In [2, Chapter 5 and 6] one finds many variations of these criteria which are not listed since, after an understanding of the basic IQC principle, all these results can be easily derived not only for polytopic and LFT parametric uncertainties, but they can be effectively extended to much larger classes of uncertainties described by dynamic IQC's.

We just confine ourselves to the generalized H_2 and peak-to-peak upper bound specifications; we employ the same setup as in Section 3.7.

3.8.1   Robust Generalized H_2-Performance

Let us define
$$P_{p1} = \begin{pmatrix} -I & 0 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad P_{p2} = \begin{pmatrix} 0 & 0 \\ 0 & I \end{pmatrix}$$

with a partition according to those of the rows of $\tilde{C}_p$ in (3.7.5). (This choice will be motivated by the considerations to follow). Suppose that $X > 0$ satisfies

$$\begin{pmatrix} I & 0 & 0 \\ \tilde{A} & \tilde{B}_u & \tilde{B}_p \end{pmatrix}^T \begin{pmatrix} 0 & X \\ X & 0 \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ \tilde{A} & \tilde{B}_u & \tilde{B}_p \end{pmatrix} +$$
$$+ \begin{pmatrix} \tilde{C}_\Psi & \tilde{D}_{\Psi u} & \tilde{D}_{\Psi p} \\ \tilde{C}_p & \tilde{D}_{pu} & \tilde{D}_{pp} \end{pmatrix}^T \begin{pmatrix} P & 0 \\ 0 & \gamma P_{p1} \end{pmatrix} \begin{pmatrix} \tilde{C}_\Psi & \tilde{D}_{\Psi u} & \tilde{D}_{\Psi p} \\ \tilde{C}_p & \tilde{D}_{pu} & \tilde{D}_{pp} \end{pmatrix} < 0 \quad (3.8.1)$$

and

$$\begin{pmatrix} \tilde{C}_p & \tilde{D}_{pu} & \tilde{D}_{pp} \end{pmatrix}^T \frac{1}{\gamma} P_{p2} \begin{pmatrix} \tilde{C}_p & \tilde{D}_{pu} & \tilde{D}_{pp} \end{pmatrix} \leqslant \begin{pmatrix} I & 0 & 0 \end{pmatrix}^T X \begin{pmatrix} I & 0 & 0 \end{pmatrix}. \quad (3.8.2)$$

Then the conclusions about well-posedness and robust stability remain unchanged.

If the uncertainty satisfies the hard IQC as in Section 3.7.3, we infer for $x(0) = 0$ and $x_\Psi(0) = 0$ from (3.7.8) by

$$\int_0^T \left(\begin{array}{c} w_p(t) \\ z_p(t) \end{array} \right)^T P_{p1} \left(\begin{array}{c} w_p(t) \\ z_p(t) \end{array} \right) dt = \int_0^T w_p(t)^T w_p(t) \, dt$$

(what motivates the definition of P_{p1}) that

$$\left(\begin{array}{c} x(T) \\ x_\Psi(T) \end{array} \right)^T X \left(\begin{array}{c} x(T) \\ x_\Psi(T) \end{array} \right) \leqslant (\gamma - \epsilon) \int_0^T w_p(t)^T w_p(t) \, dt \ \text{ for all } \ T \geqslant 0.$$

Hence, we have

$$\left(\begin{array}{c} x(T) \\ x_\Psi(T) \end{array} \right)^T X \left(\begin{array}{c} x(T) \\ x_\Psi(T) \end{array} \right) \leqslant (\gamma - \epsilon) \|w_p\|_2^2 \ \text{ for all } \ T \geqslant 0$$

such that the state-trajectory is caught in an ellipsoid defined by X. The second inequality implies

$$\frac{1}{\gamma} \left(\begin{array}{c} w_p(t) \\ z_p(t) \end{array} \right)^T P_{p2} \left(\begin{array}{c} w_p(t) \\ z_p(t) \end{array} \right) = \frac{1}{\gamma} z_p(t)^T z_p(t) \leqslant \left(\begin{array}{c} x(t) \\ x_\Psi(t) \end{array} \right)^T X \left(\begin{array}{c} x(t) \\ x_\Psi(t) \end{array} \right) \ \text{ for all } \ t \geqslant 0$$

(what motivates the choice of P_{p2}.) With the L_∞-norm definition $\|z_p\|_\infty = \sup_{t \geqslant 0} \|z_p(t)\|$, we infer by combining both inequalities that

$$\|z_p\|_\infty^2 \leqslant \gamma(\gamma - \epsilon) \|w_p\|_2^2.$$

Hence, the gain of $w_p \ni L_2 \to z_p \in L_\infty$ is robustly strictly smaller than γ.

Due to $P_{p2} \geqslant 0$, we note that (3.8.2) is equivalent to the relations

$$P_{p2} \tilde{D}_{pu} = 0, \quad P_{p2} \tilde{D}_{pp} = 0, \quad \frac{1}{\gamma} \tilde{C}_p^T P_{p2} \tilde{C}_p < X.$$

Hence, the non-strict inequality (3.8.2) indeed comprises certain strict properness conditions that are required to render the gain $w_p \ni L_2 \to z_p \in L_\infty$ *finite*.

Suppose we know (as for parametric perturbations) that the uncertainties even satisfy a QC as in Section 3.7.4 for the indices P_1 and P_2. Then we can replace (3.8.2) by

$$\left(\begin{array}{ccc} \tilde{C}_\Psi & \tilde{D}_{\Psi u} & \tilde{D}_{\Psi p} \\ \tilde{C}_p & \tilde{D}_{pu} & \tilde{D}_{pp} \end{array} \right)^T \left(\begin{array}{cc} P_2 & 0 \\ 0 & \frac{1}{\gamma} P_{p2} \end{array} \right) \left(\begin{array}{ccc} \tilde{C}_\Psi & \tilde{D}_{\Psi u} & \tilde{D}_{\Psi p} \\ \tilde{C}_p & \tilde{D}_{pu} & \tilde{D}_{pp} \end{array} \right) \leqslant \left(\begin{array}{ccc} I & 0 & 0 \end{array} \right)^T X \left(\begin{array}{ccc} I & 0 & 0 \end{array} \right)$$
$$(3.8.3)$$

and, still, infer that the generalized H_2-gain is robustly smaller than γ.

Note that we have used different multipliers P_1 and P_2 in both inequalities (what will be relevant if searching for suitable multipliers in a whole family as discussed in Section 3.9.)

The extra multiplier P_2 leads to extra freedom to render the inequality satisfied; hence it reduces conservatism. Contrary to what we could conclude previously, the inequality does not necessarily lead to $\tilde{D}_{pu} = 0$, $\tilde{D}_{pp} = 0$ but, still, it implies that the gain $w_p \ni L_2 \rightarrow z_p \in L_\infty$ is finite. Hence, (3.8.3) allows a more complicated dependence on the uncertainties at the expense of a non-strict inequality that cannot be easily re-formulated to a strict one.

Remark. The presented techniques do not apply for soft IQC's. In addition, it might be more suitable to look at uncertainties Δ that map L_∞ causally into L_2; the abstract theory developed earlier can be easily extended to cope with such situations.

3.8.2   Robust Bound on Peak-to-Peak Gain

Let us assume that the class of uncertainties $\boldsymbol{\Delta}$ consists of causal mappings $\Delta : L_{\infty e} \rightarrow L_{\infty e}$ of finite L_∞-gain. We assume that, for each $\Delta \in \boldsymbol{\Delta}$, the perturbed system (3.7.1) is

well-posed: For each $x_0 \in \mathbb{R}^n$ and $w_p \in L_{\infty e}$, the system (3.7.1) admits a unique response $x, w_u \in L_{\infty e}$.

Let us now assume that for any $z_u \in L_{\infty e}$ and any $\Delta \in \boldsymbol{\Delta}$, the QC (3.7.18) holds for $w_u = \Delta(z_u)$ and for $P = P_1$, $P = P_2$ (similarly as in Section 3.7.4).

Let X, $\lambda > 0$, $\mu \in \mathbb{R}$ satisfy the inequalities

$$X > 0, \quad \begin{pmatrix} I & 0 & 0 \\ \tilde{A} & \tilde{B}_u & \tilde{B}_p \end{pmatrix}^T \begin{pmatrix} \lambda X & X \\ X & 0 \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ \tilde{A} & \tilde{B}_u & \tilde{B}_p \end{pmatrix} +$$
$$+ \begin{pmatrix} \tilde{C}_\Psi & \tilde{D}_{\Psi u} & \tilde{D}_{\Psi p} \\ \tilde{C}_p & \tilde{D}_{pu} & \tilde{D}_{pp} \end{pmatrix}^T \begin{pmatrix} P & 0 \\ 0 & \mu P_{p1} \end{pmatrix} \begin{pmatrix} \tilde{C}_\Psi & \tilde{D}_{\Psi u} & \tilde{D}_{\Psi p} \\ \tilde{C}_p & \tilde{D}_{pu} & \tilde{D}_{pp} \end{pmatrix} < 0$$

and

$$\begin{pmatrix} \tilde{C}_\Psi & \tilde{D}_{\Psi u} & \tilde{D}_{\Psi p} \\ \tilde{C}_p & \tilde{D}_{pu} & \tilde{D}_{pp} \end{pmatrix}^T \begin{pmatrix} P_2 & 0 \\ 0 & \frac{1}{\gamma} P_{p2} \end{pmatrix} \begin{pmatrix} \tilde{C}_\Psi & \tilde{D}_{\Psi u} & \tilde{D}_{\Psi p} \\ \tilde{C}_p & \tilde{D}_{pu} & \tilde{D}_{pp} \end{pmatrix} \leqslant$$
$$\leqslant \begin{pmatrix} I & 0 & 0 \\ 0 & 0 & I \end{pmatrix}^T \begin{pmatrix} \lambda X & 0 \\ 0 & (\gamma - \mu) I \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ 0 & 0 & I \end{pmatrix}. \quad (3.8.4)$$

This leads to **uniform exponential stability**, for every $w_p \in L_\infty$

$$x(T) \leqslant \frac{\lambda_{\max}(X)}{\lambda_{\min}(X)} e^{-\lambda T} x(0) + \frac{\mu}{\lambda_{\min}(X)\lambda} \sup_{t \in [0,T]} w_p(t)^T w_p(t),$$

and to the **robust peak-to-peak norm bound** γ: for every $w_p \in L_\infty$

$$\|z_p\|_\infty \leqslant \gamma \|w_p\|_\infty.$$

Proof. Again, set

$$v(t) := \left(\begin{array}{c} x(t) \\ x_\Psi(t) \end{array} \right)^T X \left(\begin{array}{c} x(t) \\ x_\Psi(t) \end{array} \right).$$

Then the first inequality shows

$$\frac{d}{dt} v(t) + \lambda v(t) \leqslant \mu w_p(t)^T w_p(t) \text{ for all } t \geqslant 0$$

and hence

$$v(T) \leqslant e^{-\lambda T)} v(0) + \frac{\mu}{\lambda} \sup_{t \in [0,T]} w_p(t)^T w_p(t) \text{ for all } T \geqslant 0.$$

This clearly implies the statement on robust stability. Moreover, $x(0) = 0$ reveals

$$\lambda v(T) \leqslant \mu \|w_p\|_\infty^2 \text{ for all } T \geqslant 0.$$

Now we exploit (3.8.4) to infer

$$\frac{1}{\gamma} \|z_p(t)\|^2 \leqslant \lambda v(t) + (\gamma - \mu)\|w_p(t)\|^2 \text{ for all } t \geqslant 0.$$

Combining both inequalities implies

$$z_p(t)^T z_p(t) \leqslant \gamma^2 \|w_p\|_\infty^2$$

what reveals that γ is a bound on the peak-to-peak gain of $L_\infty \ni w_p \to z_p \in L_\infty$. ∎

Remark. The above inequalities imply well-posedness if A, A_Ψ are Hurwitz, if all $\Delta \in \mathbf{\Delta}$ are of finite incremental L_∞-gain, if $\mathbf{\Delta}$ is star-shaped with star center 0, and if each $\Delta \in \mathbf{\Delta}$ satisfies (3.7.10) for all $z_1, z_2 \in L_\infty$ and $z = z_1 - z_2$, $w = \Delta(z_1) - \Delta(z_2)$. The proof is based Banach's fixed point theorem that exploits the completeness of L_∞ and just requires a slight modification of that of Theorem 3.8.

Remark. As for the generalized H_2 norm, we can reduce (3.8.4) to a strict inequality with algebraic constraints if setting $P_2 = 0$; we can replace (3.8.4) equivalently by

$$P_{p2}\tilde{D}_{pu} = 0, \quad \frac{1}{\gamma} \left(\tilde{C}_p \ \ \tilde{D}_{pp} \right)^T P_{p2} \left(\tilde{C}_p \ \ \tilde{D}_{pp} \right) < \left(\begin{array}{cc} \lambda X & 0 \\ 0 & (\gamma - \mu)I \end{array} \right).$$

3.9 Multiple IQC's in the State-Space

Although most of the results provided so far have been given for one IQC, it is always important to keep in mind that one usually considers a whole family of IQC's that is parameterized in a suitable fashion.

In the latter sections we have considered IQC's that are described by a dynamical part (3.7.3) defined through A_Ψ, B_Ψ, C_Ψ, D_Ψ, and an index matrix P. In principle, all of these

parameters could be varied in specific classes to describe a set of IQC's. However, the final task is to find one of these parameters and an X (possibly with $X > 0$) such that the LMI (3.7.6) is satisfied. Hence we will prefer those parameterizations for which the search for X and a suitable IQC turns out to be a standard LMI problem.

Obvious cases include those where the dynamic part of the IQC is fixed and only the quadratic index varies in a class $\boldsymbol{P}$. If this class $\boldsymbol{P}$ is described by infinitely many LMI's, a semi-infinite LMI problem results; although convex, such problems cannot be directly handled with existing software. However, if the set $\boldsymbol{P}$ is described by finitely LMI's, standard software can be used to search for X and a suitable multiplier.

As a simple example, suppose that

$$\Pi_j(s) = C_j(sI - A_j)^{-1}B_j + D_j$$

(A_j Hurwitz) are finitely many rational multipliers. Searching in the set of all multipliers parameterized as (3.4.7) with $\tau_j > 0$ has indeed the form as just described. (Derive the dynamics and the class of indices P.)

3.10   Parametric Uncertainties

In this final section we would like to briefly return to the situation that Δ in (3.7.1) is time-varying parametric and can vary arbitrarily fast.

3.10.1   LFT Representations

The set of values of the parametric uncertainties is assumed to be given by finitely many generating points

$$\boldsymbol{\Delta}_g = \{\Delta_1, ..., \Delta_N\}$$

as

$$\boldsymbol{\Delta}_c = \mathrm{co}\,\boldsymbol{\Delta}_g = \mathrm{co}\{\Delta_1, ..., \Delta_N\}.$$

We assume $0 \in \boldsymbol{\Delta}_c$. The set of uncertainties is given by all curves $\Delta(t)$ defined on $[0, \infty)$ and taking their values in $\boldsymbol{\Delta}_c$:

$$\boldsymbol{\Delta} := \{\Delta \mid \Delta : [0, \infty) \to \boldsymbol{\Delta}_c \text{ is continuous}\}.$$

Note that any $\Delta \in \boldsymbol{\Delta}$ acts in (3.7.1) as a multiplication operator

$$w(t) = \Delta(t)z(t).$$

If, for $\Delta \in \boldsymbol{\Delta}$,

$$\begin{pmatrix} I & \Delta(t) \\ D_{uu} & I \end{pmatrix} \quad \text{is nonsingular for all } t \geqslant 0,$$

the uncertain system (3.7.1) admits the alternative representation

$$\begin{pmatrix} \dot{x} \\ z_p \end{pmatrix} = \left(\begin{pmatrix} A & B_p \\ C_p & D_{pp} \end{pmatrix} + \begin{pmatrix} B_u \\ D_{pu} \end{pmatrix} \Delta(t)(I - D_{uu}\Delta(t))^{-1} \begin{pmatrix} C_u & D_{up} \end{pmatrix} \right) \begin{pmatrix} x \\ w_p \end{pmatrix}, \quad x(0) = x_0.$$

This motivates to define the functions

$$\begin{pmatrix} A_\Delta & B_\Delta \\ C_\Delta & D_\Delta \end{pmatrix} := \begin{pmatrix} A + B_u\Delta(I - D_{uu}\Delta)^{-1}C_u & B_p + B_p\Delta(I - D_{uu}\Delta)^{-1}D_{up} \\ C_p + D_{pu}\Delta(I - D_{uu}\Delta)^{-1}C_u & D_{pp} + D_{pu}\Delta(I - D_{uu}\Delta)^{-1}D_{up} \end{pmatrix}.$$

Note that these functions are rational in the elements of Δ, and they are affine if $D_{uu} = 0$.
However, we are interested in this function only on the set $\boldsymbol{\Delta}_c$. Depending on the structure
of the matrices in $\boldsymbol{\Delta}_c$ and on D_{uu}, it might happen that $\det(I - D_{uu}\Delta)$ vanishes identically
on $\boldsymbol{\Delta}_c$. Even if not vanishing identically, this function can have zeros on the set $\boldsymbol{\Delta}_c$. We
call the LFT well-posed, if

$$\det(I - D_{uu}\Delta) \neq 0 \quad \text{for all} \quad \Delta \in \boldsymbol{\Delta}_c. \tag{3.10.1}$$

Then $(I - D_{uu}\Delta)^{-1}$ and A_Δ, B_Δ, C_Δ, D_Δ are well-defined rational functions that are
continuous on the set $\boldsymbol{\Delta}_c$. (Continuity even implies that these functions are smooth. In
fact, the essence is that they don't have poles on this set.)

If the well-posedness condition (3.10.1) holds, we arrive at the alternative LFT description

$$\begin{pmatrix} \dot{x} \\ z_p \end{pmatrix} = \begin{pmatrix} A_{\Delta(t)} & B_{\Delta(t)} \\ C_{\Delta(t)} & D_{\Delta(t)} \end{pmatrix} \begin{pmatrix} x \\ w_p \end{pmatrix}, \quad x(0) = x_0$$

of the uncertain system (3.7.1).

Let us recall the following converse fact from μ-theory: Suppose $F(\delta)$, $G(\delta)$, $H(\delta)$, $J(\delta)$ are
rational functions of $\delta = \begin{pmatrix} \delta_1 & \cdots & \delta_m \end{pmatrix}$ that are continuous on the parameter box

$$\boldsymbol{\delta}_c := \{\delta = (\delta_1, ..., \delta_m) \mid \delta_j \in [-1, 1]\}.$$

(As usual, we can shift and re-scale the uncertainties to such they take their values in $[-1, 1]$
without without loss of generality.) Then there exist matrices

$$\begin{pmatrix} A & B_u & B_p \\ C_u & D_{uu} & D_{up} \\ C_p & D_{pu} & D_{pp} \end{pmatrix}$$

and nonnegative integers

$$d_1, ..., d_m$$

such that, with

$$\Delta(\delta) = \operatorname{diag}(\delta_1 I_{d_1}, ..., \delta_m I_{d_m}), \tag{3.10.2}$$

we have the following two properties:

$$\det(I - D_{uu}\Delta(\delta)) \neq 0 \quad \text{and} \quad \begin{pmatrix} A_{\Delta(\delta)} & B_{\Delta(\delta)} \\ C_{\Delta(\delta)} & D_{\Delta(\delta)} \end{pmatrix} = \begin{pmatrix} F(\delta) & G(\delta) \\ H(\delta) & J(\delta) \end{pmatrix} \quad \text{on} \quad \boldsymbol{\delta}_c.$$

Consequently, we can summarize:

- On every set $\boldsymbol{\Delta}_c$ where $\det(I - D_{uu}\Delta)$ does not vanish, A_Δ, B_Δ, C_Δ, D_Δ define continuous rational functions of the elements of Δ.

- Arbitrary continuous rational functions $F(\delta)$, $G(\delta)$, $H(\delta)$, $J(\delta)$ without poles on $\boldsymbol{\delta}_c$ admit a well-posed LFT representation; they can be written as $A_{\Delta(\delta)}$, $B_{\Delta(\delta)}$, $C_{\Delta(\delta)}$, $D_{\Delta(\delta)}$ where $\Delta(\delta)$ is a block-diagonal matrix whose blocks admit the form $\delta_j I$.

Remarks. This result shows that choice of a block-diagonal structure (as usually made in μ-theory) is a specific case of the general formulation we started out with. In concrete applications, the extra freedom of not being forced to use block-diagonal matrices should be exploited to arrive at more efficient LFT representations, in particular with respect to the size of Δ. Note also that one is not bound to parameter boxes of the form $\boldsymbol{\delta}_c$ but one can also choose for more general sets that are described as $\boldsymbol{\Delta}_c$.

3.10.2 Robust Quadratic Performance

Let us now return to robust quadratic performance analysis with an index P_p that satisfies, as earlier, the condition (3.5.7). We have given two seemingly different sufficient conditions to guarantee robust performance; one is related to finding a common quadratic Lyapunov function, and the other proceeds via scalings and quadratic constraints.

Indeed, robust quadratic performance is implied by the existence of an X such that

$$X > 0, \quad \begin{pmatrix} A_\Delta^T X + X A_\Delta & X B_\Delta \\ B_\Delta^T X & 0 \end{pmatrix} + \begin{pmatrix} 0 & I \\ C_\Delta & D_\Delta \end{pmatrix}^T P_p \begin{pmatrix} 0 & I \\ C_\Delta & D_\Delta \end{pmatrix} < 0 \ \text{ for all } \ \Delta \in \boldsymbol{\Delta}_v. \tag{3.10.3}$$

Moreover, it is also guaranteed if there exists an X and a scaling

$$P = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} \ \text{ satisfying } \ \begin{pmatrix} \Delta \\ I \end{pmatrix}^T P \begin{pmatrix} \Delta \\ I \end{pmatrix} > 0 \ \text{ for all } \ \Delta \in \boldsymbol{\Delta}_c \tag{3.10.4}$$

such that

$$\begin{pmatrix} I & 0 & 0 \\ A & B_u & B_p \end{pmatrix}^T \begin{pmatrix} 0 & X \\ X & 0 \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ A & B_u & B_p \end{pmatrix} +$$
$$+ \begin{pmatrix} 0 & I & 0 \\ C_u & D_{uu} & D_{up} \end{pmatrix}^T P \begin{pmatrix} 0 & I & 0 \\ C_u & D_{uu} & D_{up} \end{pmatrix} +$$
$$+ \begin{pmatrix} 0 & 0 & I \\ C_p & D_{pu} & D_{pp} \end{pmatrix}^T P_p \begin{pmatrix} 0 & 0 & I \\ C_p & D_{pu} & D_{pp} \end{pmatrix} < 0. \tag{3.10.5}$$

A simple exercise reveals that (3.10.4)-(3.10.5) imply the non-singularity of $I - D_{uu}\Delta$ for all $\Delta \in \boldsymbol{\Delta}_c$ and the validity of (3.10.3) with the same X. (Why?) It is not so trivial to see that the converse holds true as well; in fact, both conditions for robust quadratic performance are equivalent.

Theorem 3.23 *The matrix $I - D_{uu}\Delta$ is nonsingular for all $\Delta \in \mathbf{\Delta}_c$ and there exists an X with (3.10.3) iff there exist symmetric X and P satisfying (3.10.4)-(3.10.5).*

The proof of this theorem is found in [33]. Since based on a more general result about quadratic forms, similar statements can be obtained with ease for all the other performance criteria (such as H_2, generalized H_2 and peak-to-peak upper bound performance) that have been considered in these notes.

Testing the first condition amounts to reducing (3.10.3) to finitely many LMI's that can be based on convexity arguments or gridding techniques. In the second characterization, (3.10.5) poses no problem and only the parameterization of the scalings as in (3.10.4) has to be given with a finite number of LMI's.

If D_{uu} vanishes, the picture becomes clear. Then the functions A_Δ, B_Δ, C_Δ, D_Δ are *affine* in Δ, and one just has to solve the inequalities in (3.10.3) for the finitely many generators in $\mathbf{\Delta}_g$. Similarly, $D_{uu} = 0$ implies that any multiplier P that satisfies (3.10.5) has, in fact, the property

$$Q < 0.$$

(Why? The alert reader recalls that the required argument appeared earlier in the text.) Hence we can introduce this extra constraint *without conservatism*. Under this constraint, however, (3.10.4) holds iff it holds already for the generators $\mathbf{\Delta}_g$. Hence, both robust performance characterizations reduce without conservatism to standard LMI problems.

The situation is more complicated for $D_{uu} \neq 0$ such that A_Δ, B_Δ, C_Δ, D_Δ could (depending on the structure of Δ) be genuinely nonlinear. Then (3.10.3) offers no clear hint how to perform the reduction to a finite number of LMI's. The second characterization, however, allows to give - possibly at the expense of conservatism - certain reduction recipes. We provide three possibilities:

- Just introduce the extra constraint $Q < 0$ and replace $\mathbf{\Delta}_c$ by $\mathbf{\Delta}_g$ in (3.10.4). If the resulting LMI's are feasible, robust performance is guaranteed. Hence this is a sufficient condition for the two tests (3.10.3) or (3.10.4)-(3.10.5) that is, generally, stronger due to the extra constraint for the scaling.

- One can take the specific structure of the set $\mathbf{\Delta}_c$ and of the function

$$\Delta \to \begin{pmatrix} \Delta \\ I \end{pmatrix}^T P \begin{pmatrix} \Delta \\ I \end{pmatrix} \tag{3.10.6}$$

  into account to refine the pretty rough extra constraint $Q < 0$. Just as an example that admits immediate extensions, suppose that

  $$\mathbf{\Delta}_c = \{\Delta = \begin{pmatrix} \delta_1 I & & 0 \\ & \ddots & \\ 0 & & \delta_m I \end{pmatrix} \mid \delta_j \in [-1, 1]\}.$$

Then it suffices to restrict Q only as

$$Q = \begin{pmatrix} Q_{11} & \cdots & Q_{1m} \\ \vdots & \ddots & \vdots \\ Q_{m1} & \cdots & Q_{mm} \end{pmatrix}, \quad Q_{jj} < 0, \quad j = 1, \dots, m$$

(in the same partition as that of Δ.) This renders the function (3.10.6) in $\delta = \begin{pmatrix} \delta_1 & \cdots & \delta_m \end{pmatrix}$ partially convex on the parameter box $\boldsymbol{\delta}_c$, and it suffices to describe the scalings only through inequalities on the extreme points

$$\boldsymbol{\delta}_g := \{\delta = \begin{pmatrix} \delta_1 & \cdots & \delta_m \end{pmatrix} \mid \delta_j \in \{-1, 1\}\}$$

of this box. Clearly, this extra constraint on Q is less stringent than $Q < 0$ what reduces conservatism.

• Finally, we mention a variant that is even more conservative than $Q < 0$. One just uses, for the class of uncertainties as in the previous item, the scalings adjusted to the uncertainty structure as defined in (3.4.3). This provides a subset of all scalings satisfying (3.10.4) that admit an explicit (implementable) description. If the diagonal blocks are not repeated, $Q = -R$ are diagonal and S vanishes; this is the case that reappears throughout the book [2] by applying the so-called S-procedure. Although the restriction to a smaller class of scalings introduces conservatism, it also reduces the number of variables in the LMI test what speeds up the calculations.

This discussion reveals that the introduction of scalings allows to provide guarantees for robust quadratic performance even if the parameters enter in a rational fashion. Moreover, for the least conservatism, one should employ scalings that are full and can only be indirectly described. Using the usual structured scalings as appearing in μ-theory introduces extra conservatism that should be avoided.

There is, however, one point that seems not sufficiently stressed in the literature: No matter by which technique, solving (3.10.4)-(3.10.5) or (3.10.3) amounts to guaranteeing robust performance not only for the systems described with matrices in

$$\left\{ \begin{pmatrix} A_\Delta & B_\Delta \\ C_\Delta & D_\Delta \end{pmatrix} \mid \Delta \in \boldsymbol{\Delta}_c \right\}$$

but even for those described with

$$\mathrm{co}\left\{ \begin{pmatrix} A_\Delta & B_\Delta \\ C_\Delta & D_\Delta \end{pmatrix} \mid \Delta \in \boldsymbol{\Delta}_c \right\}. \tag{3.10.7}$$

(Why?) If A_Δ, B_Δ, C_Δ, D_Δ are affine on $\boldsymbol{\Delta}_c$, both sets coincide since $\boldsymbol{\Delta}_c$ is convex. However, if these functions are nonlinear, taking the convex hull might increase the set considerably such that the desired spec is guaranteed for systems which are not included in the original description. Note that this observations suggests another approach to guarantee

robust performance: Try to find a description of (3.10.7) as the convex hull of finitely many generators. The computation of such a re-parameterization, however, might be very expensive.

Remarks.

- In general, Δ enters the constraints (3.10.3) in a rational fashion, whereas (3.10.5) is independent of Δ and (3.10.4) is quadratic in Δ. Through the auxiliary variable P, the dependence on Δ has been simplified. The variable P is closely related to Lagrange-multipliers as appearing in constraint optimization theory.

- In [2] the authors apply the so-called S-procedure to derive (3.10.5) from (3.10.3) for the specific class of scalings described above; the constraint (3.10.4) does not appear explicitly since the corresponding multipliers satisfy it automatically. In the present notes we have provided two version of robust performance tests: One directly based on the parameter dependent system description, and one based on QC's. We avoided to refer to the S-procedure since the QC results allow powerful generalizations and since our approach provides a better insight in the choice of various classes of scalings and the resulting conservatism. Therefore, the S-procedure only plays a minor role in our notes.

- The discussion in this section is particularly important for the synthesis of robust controllers. Controller synthesis seems not directly possible on the basis of (3.10.3), but one can easily provide variations of the standard D/K-iteration from μ-synthesis if characterizing robust performance by (3.10.4)-(3.10.5).

3.11 Exercises

Exercise 1

a) Suppose that M (linear) and Δ are systems that have finite gain. Show that $\mathcal{I}_M(\Delta)$ has a causal inverse with finite gain iff the same is true of the mapping $I - M\Delta$. Show that the same holds if 'finite gain' is replaced with 'finite incremental gain'.

b) Suppose $S : L_{2e} \rightarrow L_{2e}$ is causal and has finite incremental gain. Moreover, assume that the restriction $S : L_2 \rightarrow L_2$ has an inverse whose incremental gain is finite. Then $S : L_{2e} \rightarrow L_{2e}$ itself has an inverse with finite incremental gain.

Exercise 2

Suppose w and z are two vectors in $\mathbb{R}^n$. Prove:

There exists a $\Delta \in \mathbb{R}^{n \times n}$ with $\|\Delta\| \leqslant 1$ and $w = \Delta z$ iff $w^T w \leqslant z^T z$.

There exists a $\delta \in \mathbb{R}$ with $|\delta| \leqslant 1$ and $w = \delta z$ iff $ww^T \leqslant zz^T$.

Exercise 3

For given $\Delta_j \in \mathbb{R}^{k \times l}$, define the set $\boldsymbol{\Delta} := \mathrm{co}\{\Delta_1, \ldots, \Delta_N\}$. With fixed $Q = Q^T$, S, $R = R^T$ consider the function

$$f(\Delta) := \begin{pmatrix} \Delta \\ I \end{pmatrix}^T \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} \begin{pmatrix} \Delta \\ I \end{pmatrix}.$$

Prove: $Q \leqslant 0$ implies that $\Delta \rightarrow f(\Delta)$ is concave.

Prove: If f is concave then

$$f(\Delta_j) > 0 \quad \text{for all} \quad j = 1, \ldots, N \implies f(\Delta) > 0 \quad \text{for all} \quad \Delta \in \boldsymbol{\Delta}. \tag{3.11.1}$$

Find weaker conditions on Q that lead to the the same implication (3.11.1).

Exercise 4 (MIMO Circle criterion)

Consider the system

$$\dot{x} = Ax + B\Delta(t, Cx)$$

where $\Delta : \mathbb{R} \times \mathbb{R}^l \to \mathbb{R}^k$ is any continuously differentiable function that satisfies, for two matrices K, L, the multi-variable sector condition

$$[\Delta(t, z) - Kz]^T [\Delta(t, z) - Lz] \leqslant 0 \quad \text{for all} \quad (t, z) \in \mathbb{R} \times \mathbb{R}^l. \tag{3.11.2}$$

Note that this is nothing but a static QC. Find a multiplier and the corresponding LMI that proves exponential stability. With $M(s) = C(sI - A)^{-1}B$ define $G(s) = (I - LM(s))(I - KM(s))^{-1}$. Show that the LMI you found has a solution iff

$$A + BKC \quad \text{is stable and} \quad G(i\omega)^* + G(i\omega) > 0 \quad \text{for all} \quad \omega \in \mathbb{R}.$$

(Hence G is strictly positive real. Note that this terminology is often used in the literature for a different property!) Is stability of A required for your arguments?

Exercise 5 (Popov criterion)

Consider the system

$$\dot{x} = Ax + B\Delta(Cx) \tag{3.11.3}$$

with A Hurwitz and Δ a continuously differentiable nonlinearity $\Delta : \mathbb{R} \to \mathbb{R}$ that satisfies

$$0 \leqslant z\Delta(z) \leqslant z^2 \quad \text{for all} \quad z \in \mathbb{R}.$$

Prove the following statements:

1. Δ satisfies the static quadratic constraints

$$\tau \begin{pmatrix} \Delta(z) \\ z \end{pmatrix}^T \begin{pmatrix} -2 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \Delta(z) \\ z \end{pmatrix} \geqslant 0$$

   for all $z \in \mathbb{R}$ and $\tau \geqslant 0$.

2. For any $z \in \mathbb{R}$

$$\int_0^z \Delta(\zeta)\, d\zeta \geqslant 0, \quad \int_0^z \zeta - \Delta(\zeta)\, d\zeta \geqslant 0.$$

   If $z : [0, \infty) \to \mathbb{R}$ is continuously differentiable, then

$$\tau_1 \int_0^T z(t)\dot{z}(t) - \Delta(z(t))\dot{z}(t)\, dt \geqslant -\tau_1 \frac{1}{2} z(0)^2, \quad \tau_2 \int_0^T \Delta(z(t))\dot{z}(t)\, dt \geqslant -\tau_2 \frac{1}{2} z(0)^2$$

   for $\tau_1, \tau_2 \geqslant 0$. (Substitution rule!)

3. Suppose there exist X and $\tau, \tau_1, \tau_2 \geqslant 0$ such that

$$\begin{pmatrix} I & 0 \\ A & B \end{pmatrix}^T \begin{pmatrix} 0 & X \\ X & 0 \end{pmatrix} \begin{pmatrix} I & 0 \\ A & B \end{pmatrix} + \begin{pmatrix} 0 & I \\ C & 0 \end{pmatrix}^T \begin{pmatrix} -2\tau & \tau \\ \tau & 0 \end{pmatrix} \begin{pmatrix} 0 & I \\ C & 0 \end{pmatrix} +$$

$$+ \begin{pmatrix} C & 0 \\ CA & CB \end{pmatrix}^T \begin{pmatrix} 0 & \tau_1 \\ \tau_1 & 0 \end{pmatrix} \begin{pmatrix} C & 0 \\ CA & CB \end{pmatrix} + \begin{pmatrix} 0 & I \\ CA & CB \end{pmatrix}^T \begin{pmatrix} 0 & -\tau_1 \\ \tau_1 & 0 \end{pmatrix} \begin{pmatrix} 0 & I \\ CA & CB \end{pmatrix} +$$

$$+ \begin{pmatrix} 0 & I \\ CA & CB \end{pmatrix}^T \begin{pmatrix} 0 & \tau_2 \\ \tau_2 & 0 \end{pmatrix} \begin{pmatrix} 0 & I \\ CA & CB \end{pmatrix} < 0.$$

Then the equilibrium $x_0 = 0$ of (3.11.3) is globally asymptotically Lyapunov stable. What can you say about exponential stability? Hint: Use a simple hard IQC argument. Note that, along trajectories of (3.11.3), one has $z = Cx$, $\dot{z} = CAx + CBw$ and $w = \Delta(z)$.

4. Show that the condition in the previous exercise is equivalent to the existence of a symmetric K, $\nu \in \mathbb{R}$, $\tau > 0$ with

$$
\begin{pmatrix} A^T K + KA & KB \\ B^T K & 0 \end{pmatrix} + \begin{pmatrix} 0 & I \\ CA & CB \end{pmatrix}^T \begin{pmatrix} 0 & \nu \\ \nu & 0 \end{pmatrix} \begin{pmatrix} 0 & I \\ CA & CB \end{pmatrix} +
$$
$$
+ \begin{pmatrix} 0 & I \\ C & 0 \end{pmatrix}^T \begin{pmatrix} -2\tau & \tau \\ \tau & 0 \end{pmatrix} \begin{pmatrix} 0 & I \\ C & 0 \end{pmatrix} < 0.
$$

5. With $G(s) = C(sI - A)^{-1}B$, the LMI in the last exercise is solvable iff there exists a $q \in \mathbb{R}$ with
$$
\mathrm{Re}((1 + qi\omega)G(i\omega)) < 1 \quad \text{for all} \ \ \omega \in \mathbb{R} \cup \{\infty\}.
$$

This reveals the relation to the classical Popov criterion. Note that q is often assumed to be nonnegative what is, actually, a redundant hypothesis. Show with an example that the extra constraint $q \geqslant 0$ (or $\nu \geqslant 0$ in the LMI) introduces conservatism. (Think of a smart test using LMI-Lab to find an example.)

6. Find an LMI condition for global asymptotic stability of

$$
\dot{x} = Ax + \sum_{j=1}^{k} B_j \Delta_j (C_j x)
$$

where the continuously differentiable $\Delta_j : \mathbb{R} \to \mathbb{R}$ satisfy the sector conditions $\beta_j z^2 \leqslant z\Delta_j(z) \leqslant \alpha_j z^2$ for all $z \in \mathbb{R}$.

Chapter 4

Controller Synthesis

In this chapter we intend to provide a powerful result that allows to step in a straightforward manner from analysis conditions formulated in terms of matrix inequalities to the corresponding matrix inequalities for controller synthesis. We will observe that this is achieved by a nonlinear and essentially bijective transformation of the controller parameters.

4.1 The Setup

Suppose an LTI system is described as

$$
\begin{pmatrix} \dot{x} \\ z_1 \\ \vdots \\ z_q \\ y \end{pmatrix} = \left(\begin{array}{c|cccc} A & B_1 & \cdots & B_q & B \\ \hline C_1 & D_1 & \cdots & D_{1q} & E_1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ C_q & D_{q1} & \cdots & D_q & E_q \\ C & F_1 & \cdots & F_q & 0 \end{array} \right) \begin{pmatrix} x \\ w_1 \\ \vdots \\ w_q \\ u \end{pmatrix}.
\tag{4.1.1}
$$

We denote by u the control input, by y the measured output available for control, and by $w_j \to z_j$ the channels on which we want to impose certain robustness and/or performance objectives. Since we want to extend the design technique to mixed problems with various performance specifications on various channels, we already start at this point with a multi-channel system description. Sometimes we collect the signals as

$$
z = \begin{pmatrix} z_1 \\ \vdots \\ z_q \end{pmatrix}, \quad w = \begin{pmatrix} w_1 \\ \vdots \\ w_q \end{pmatrix}.
$$

Remark. Note that we do not exclude the situation that some of the signals w_j or z_j are identical. Therefore, we only need to consider an equal number of input- and output-signals. Moreover, it might seem restrictive to only consider the diagonal channels and neglect the channels $w_j \to z_k$ for $j \neq k$. This is not the case. As a typical example, suppose we intend to impose for $z = Tw$ specifications on $L_j T R_j$ where L_j, R_j are arbitrary matrices that pick out certain linear combinations of the signals z, w (or of the rows/columns of the transfer matrix if T is described by an LTI system). If we set $w = R_j w_j$, $z_j = L_j z$, we are hence interested in specifications on the diagonal channels of

$$\begin{pmatrix} z_1 \\ z_2 \\ \vdots \end{pmatrix} = \begin{pmatrix} L_1 \\ L_2 \\ \vdots \end{pmatrix} T \begin{pmatrix} R_1 & R_2 & \cdots \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \\ \vdots \end{pmatrix}.$$

If T is LTI, the selection matrices L_j and R_j can be easily incorporated into the realization to arrive at the description (4.1.1).

A *controller* is any finite dimensional linear time invariant system described as

$$\begin{pmatrix} \dot{x}_c \\ u \end{pmatrix} = \begin{pmatrix} A_c & B_c \\ C_c & D_c \end{pmatrix} \begin{pmatrix} x_c \\ y \end{pmatrix} \tag{4.1.2}$$

that has u as its input and y as its output. Controllers are hence simply parameterized by the matrices A_c, B_c, C_c, D_c.

The controlled or closed-loop system then admits the description

$$\begin{pmatrix} \dot{\xi} \\ z \end{pmatrix} = \begin{pmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{pmatrix} \begin{pmatrix} \xi \\ w \end{pmatrix} \quad \text{or} \quad \begin{pmatrix} \dot{\xi} \\ z_1 \\ \vdots \\ z_q \end{pmatrix} = \left(\begin{array}{c|ccc} \mathcal{A} & \mathcal{B}_1 & \cdots & \mathcal{B}_q \\ \hline \mathcal{C}_1 & \mathcal{D}_1 & \cdots & \mathcal{D}_{1q} \\ \vdots & \vdots & \ddots & \vdots \\ \mathcal{C}_q & \mathcal{D}_{q1} & \cdots & \mathcal{D}_q \end{array} \right) \begin{pmatrix} \xi \\ w_1 \\ \vdots \\ w_q \end{pmatrix}. \tag{4.1.3}$$

The corresponding input-output mappings (or transfer matrices) are denoted as

$$w = \mathcal{T}z \quad \text{or} \quad \begin{pmatrix} z_1 \\ \vdots \\ z_q \end{pmatrix} = \begin{pmatrix} \mathcal{T}_1 & & * \\ & \ddots & \\ * & & \mathcal{T}_q \end{pmatrix} \begin{pmatrix} w_1 \\ \vdots \\ w_q \end{pmatrix}.$$

respectively.

One can easily calculate a realization of $\mathcal{T}_j$ as

$$\begin{pmatrix} \dot{\xi} \\ z_j \end{pmatrix} = \begin{pmatrix} \mathcal{A} & \mathcal{B}_j \\ \mathcal{C}_j & \mathcal{D}_j \end{pmatrix} \begin{pmatrix} \xi \\ w_j \end{pmatrix} \tag{4.1.4}$$

where

$$\left(\begin{array}{c|c} \mathcal{A} & \mathcal{B}_j \\ \hline \mathcal{C}_j & \mathcal{D}_j \end{array} \right) = \left(\begin{array}{cc|c} A + BD_cC & BC_c & B_j + BD_cF_j \\ B_cC & A_c & B_cF_j \\ \hline C_j + E_jD_cC & E_jC_c & D_j + E_jD_cF_j \end{array} \right).$$
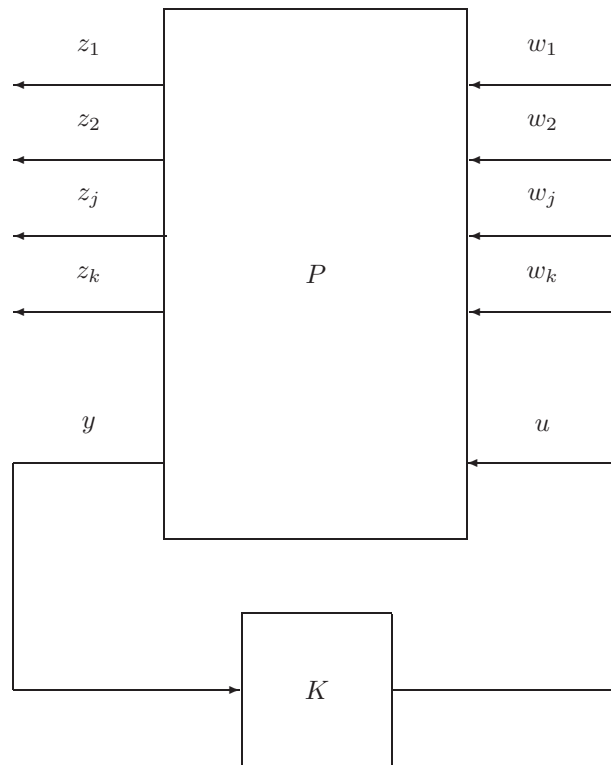
Figure 4.1: Multi-channel Closed-Loop System

It simplifies some calculations if we use the equivalent alternative formula

$$\left(\begin{array}{cc} \mathcal{A} & \mathcal{B}_j \\ \mathcal{C}_j & \mathcal{D}_j \end{array} \right) = \left(\begin{array}{cc|c} A & 0 & B_j \\ 0 & 0 & 0 \\ \hline C_j & 0 & D_j \end{array} \right) + \left(\begin{array}{cc} 0 & B \\ I & 0 \\ 0 & E_j \end{array} \right) \left(\begin{array}{cc} A_c & B_c \\ C_c & D_c \end{array} \right) \left(\begin{array}{cc|c} 0 & I & 0 \\ C & 0 & F_j \end{array} \right). \qquad (4.1.5)$$

Note that the left-hand side depends affinely on the controller parameters what is not made explicit by our notation.

4.2 From Analysis to Synthesis - A General Procedure

Let us consider as a paradigm example the design of a controller that achieves stability and quadratic performance in the channel $w_j \rightarrow z_j$. For that purpose we suppose that we have given a performance index

$$P_j = \left(\begin{array}{cc} Q_j & S_j \\ S_j^T & R_j \end{array} \right) \quad \text{with} \quad R_j \geqslant 0.$$

In Chapter 2 we have revealed that the following conditions are equivalent: The controller (4.1.2) renders (4.1.4) internally stable and leads to

$$\int_0^\infty \left(\begin{array}{c} w_j(t) \\ z_j(t) \end{array} \right)^T P_j \left(\begin{array}{c} w_j(t) \\ z_j(t) \end{array} \right) dt \leqslant -\epsilon \int_0^\infty w_j(t)^T w_j(t)\, dt$$

for some $\epsilon > 0$ if and only if

$$\sigma(\mathcal{A}) \subset \mathbb{C}^- \quad \text{and} \quad \left(\begin{array}{c} I \\ \mathcal{T}_j(i\omega) \end{array} \right)^* P_j \left(\begin{array}{c} I \\ \mathcal{T}_j(i\omega) \end{array} \right) < 0 \quad \text{for all} \quad \omega \in \mathbb{R} \cup \{\infty\}$$

if and only if there exists a symmetric $\mathcal{X}$ satisfying

$$\mathcal{X} > 0, \quad \left(\begin{array}{cc} \mathcal{A}^T \mathcal{X} + \mathcal{X} \mathcal{A} & \mathcal{X} \mathcal{B}_j \\ \mathcal{B}_j^T \mathcal{X} & 0 \end{array} \right) + \left(\begin{array}{cc} 0 & I \\ \mathcal{C}_j & \mathcal{D}_j \end{array} \right)^T P_j \left(\begin{array}{cc} 0 & I \\ \mathcal{C}_j & \mathcal{D}_j \end{array} \right) < 0. \qquad (4.2.1)$$

The corresponding *quadratic performance synthesis problem* reads as follows: Search controller parameters $\left(\begin{array}{cc} A_c & B_c \\ C_c & D_c \end{array} \right)$ and an $\mathcal{X} > 0$ that render (4.2.1) satisfied.

Recall that $\mathcal{A}$ depends on the controller parameters; since $\mathcal{X}$ is also a variable, we observe that $\mathcal{X}\mathcal{A}$ depends non-linearly on the variables to be found.

It has been observed only quite recently [18, 38] that one can find a nonlinear transformation

$$\left(\mathcal{X}, \ \left(\begin{array}{cc} A_c & B_c \\ C_c & D_c \end{array} \right) \right) \rightarrow v = \left(X, \ Y, \ \left(\begin{array}{cc} K & L \\ M & N \end{array} \right) \right) \qquad (4.2.2)$$

and a $\mathcal{Y}$ such that, with the functions

$$
\left.
\begin{array}{l}
\boldsymbol{X}(v) \; := \; \begin{pmatrix} Y & I \\ I & X \end{pmatrix} \\[2ex]
\begin{pmatrix} \boldsymbol{A}(v) & \boldsymbol{B}_j(v) \\ \boldsymbol{C}_j(v) & \boldsymbol{D}_j(v) \end{pmatrix} \; := \;
\left(\begin{array}{cc|c} AY + BM & A + BNC & B_j + BNF_j \\ K & AX + LC & XB_j + LF_j \\ \hline C_jY + E_jM & C_j + E_jNC & D_j + E_jNF_j \end{array} \right)
\end{array}
\right\}
\tag{4.2.3}
$$

one has

$$
\left.
\begin{array}{l}
\mathcal{Y}^T \mathcal{X} \mathcal{Y} \; = \; \boldsymbol{X}(v) \\[2ex]
\begin{pmatrix} \mathcal{Y}^T \mathcal{X} \mathcal{A} \mathcal{Y} & \mathcal{Y}^T \mathcal{X} \mathcal{B}_j \\ \mathcal{C}_j \mathcal{Y} & \mathcal{D}_j \end{pmatrix} \; = \; \begin{pmatrix} \boldsymbol{A}(v) & \boldsymbol{B}_j(v) \\ \boldsymbol{C}_j(v) & \boldsymbol{D}_j(v) \end{pmatrix}
\end{array}
\right\}
\tag{4.2.4}
$$

Hence, under congruence transformations with the matrices

$$
\mathcal{Y} \quad \text{and} \quad \begin{pmatrix} \mathcal{Y} & 0 \\ 0 & I \end{pmatrix},
\tag{4.2.5}
$$

the blocks transform as

$$
\mathcal{X} \rightarrow \boldsymbol{X}(v), \quad \begin{pmatrix} \mathcal{X}\mathcal{A} & \mathcal{X}\mathcal{B}_j \\ \mathcal{C}_j & \mathcal{D}_j \end{pmatrix} \rightarrow \begin{pmatrix} \boldsymbol{A}(v) & \boldsymbol{B}_j(v) \\ \boldsymbol{C}_j(v) & \boldsymbol{D}_j(v) \end{pmatrix}.
$$

Therefore, the original blocks that depend non-linearly on the decision variables $\mathcal{X}$ and $\begin{pmatrix} A_c & B_c \\ C_c & D_c \end{pmatrix}$ are transformed into blocks that are *affine* functions of the new variables v.

If $\mathcal{Y}$ is nonsingular, we can perform a congruence transformation on the two inequalities in (4.2.1) with the nonsingular matrices (4.2.5) to obtain

$$
\mathcal{Y}^T \mathcal{X} \mathcal{Y} > 0, \quad \begin{pmatrix} \mathcal{Y}^T [\mathcal{A}^T \mathcal{X} + \mathcal{X}\mathcal{A}] \mathcal{Y} & \mathcal{Y}^T \mathcal{X} \mathcal{B}_j \\ \mathcal{B}_j^T \mathcal{X} \mathcal{Y} & 0 \end{pmatrix} + \begin{pmatrix} 0 & I \\ \mathcal{C}_j \mathcal{Y} & \mathcal{D}_j \end{pmatrix}^T P_j \begin{pmatrix} 0 & I \\ \mathcal{C}_j \mathcal{Y} & \mathcal{D}_j \end{pmatrix} < 0
\tag{4.2.6}
$$

what is nothing but

$$
\boldsymbol{X}(v) > 0, \quad \begin{pmatrix} \boldsymbol{A}(v)^T + \boldsymbol{A}(v) & \boldsymbol{B}_j(v) \\ \boldsymbol{B}_j(v)^T & 0 \end{pmatrix} + \begin{pmatrix} 0 & I \\ \boldsymbol{C}_j(v) & \boldsymbol{D}_j(v) \end{pmatrix}^T P_j \begin{pmatrix} 0 & I \\ \boldsymbol{C}_j(v) & \boldsymbol{D}_j(v) \end{pmatrix} < 0.
\tag{4.2.7}
$$

For $R_j = 0$ (as it happens in the positive real performance index), we infer $P_j = \begin{pmatrix} Q_j & S_j \\ S_j^T & 0 \end{pmatrix}$ what implies that the inequalities (4.2.7) are affine in v. For a general performance index with $R_j \geqslant 0$, the second inequality in (4.2.7) is non-linear but convex in v. It is straightforward to transform it to a genuine LMI with a Schur complement argument. Since it is more convenient to stay with the inequalities in the form (4.2.7), we rather formulate a general auxiliary result that displays how to perform the linearization whenever it is required for computational purposes.

Lemma 4.1 (Linearization Lemma) *Suppose that A, S are constant matrices, that $B(v)$, $Q(v) = Q(v)^T$ depend affinely on some parameter v, and that $R(v)$ can be decomposed as $TU(v)^{-1}T^T$ with $U(v)$ being affine. Then the non-linear matrix inequalities*

$$U(v) > 0, \quad \begin{pmatrix} A \\ B(v) \end{pmatrix}^T \begin{pmatrix} Q(v) & S \\ S' & R(v) \end{pmatrix} \begin{pmatrix} A \\ B(v) \end{pmatrix} < 0$$

are equivalent to the linear matrix inequality

$$\begin{pmatrix} A^T Q(v) A + A^T S B(v) + B(v)^T S^T A & B(v)^T T \\ T^T B(v) & -U(v) \end{pmatrix} < 0.$$

In order to apply this lemma we rewrite the second inequality of (4.2.7) as

$$\begin{pmatrix} I & 0 \\ \boldsymbol{A}(v) & \boldsymbol{B}_j(v) \\ \hline 0 & I \\ \boldsymbol{C}_j(v) & \boldsymbol{D}_j(v) \end{pmatrix}^T \left(\begin{array}{cc|cc} 0 & I & 0 & 0 \\ I & 0 & 0 & 0 \\ \hline 0 & 0 & Q_j & S_j \\ 0 & 0 & S_j^T & R_j \end{array} \right) \begin{pmatrix} I & 0 \\ \boldsymbol{A}(v) & \boldsymbol{B}_j(v) \\ \hline 0 & I \\ \boldsymbol{C}_j(v) & \boldsymbol{D}_j(v) \end{pmatrix} < 0 \qquad (4.2.8)$$

what is, after a simple permutation, nothing but

$$\begin{pmatrix} I & 0 \\ 0 & I \\ \hline \boldsymbol{A}(v) & \boldsymbol{B}_j(v) \\ \boldsymbol{C}_j(v) & \boldsymbol{D}_j(v) \end{pmatrix}^T \left(\begin{array}{cc|cc} 0 & 0 & I & 0 \\ 0 & Q_j & 0 & S_j \\ \hline I & 0 & 0 & 0 \\ 0 & S_j^T & 0 & R_j \end{array} \right) \begin{pmatrix} I & 0 \\ 0 & I \\ \hline \boldsymbol{A}(v) & \boldsymbol{B}_j(v) \\ \boldsymbol{C}_j(v) & \boldsymbol{D}_j(v) \end{pmatrix} < 0. \qquad (4.2.9)$$

This inequality can be linearized according to Lemma 4.1 with an arbitrary factorization

$$R_j = T_j T_j^T \ \ \text{leading to} \ \ \begin{pmatrix} 0 & 0 \\ 0 & R_j \end{pmatrix} = \begin{pmatrix} 0 \\ T_j \end{pmatrix} \begin{pmatrix} 0 & T_j^T \end{pmatrix}.$$

So far we have discussed how to derive the synthesis inequalities (4.2.7). Let us now suppose that we have verified that these inequalities do have a solution, and that we have computed some solution v. If we can find a preimage $\left(\mathcal{X}, \ \begin{pmatrix} A_c & B_c \\ C_c & D_c \end{pmatrix} \right)$ of v under the transformation (4.2.2) and a nonsingular $\mathcal{Y}$ for which (4.2.4) holds, then we can simply reverse all the steps performed above to reveal that (4.2.7) is equivalent to (4.2.1). Therefore, the controller defined by $\begin{pmatrix} A_c & B_c \\ C_c & D_c \end{pmatrix}$ renders (4.2.1) satisfied and, hence, leads to the desired quadratic performance specification for the controlled system.

Before we comment on the resulting design procedure, let us first provide a proof of the following result that summarizes the discussion.

Theorem 4.2 *There exists a controller $\begin{pmatrix} A_c & B_c \\ C_c & D_c \end{pmatrix}$ and a $\mathcal{X}$ satisfying (4.2.1) iff there exists an v that solves the inequalities (4.2.7). If v satisfies (4.2.7), then $I - XY$ is nonsingular*

and there exist nonsingular U, V with $I - XY = UV^T$. The unique $\mathcal{X}$ and $\begin{pmatrix} A_c & B_c \\ C_c & D_c \end{pmatrix}$ with

$$\begin{pmatrix} Y & V \\ I & 0 \end{pmatrix} \mathcal{X} = \begin{pmatrix} I & 0 \\ X & U \end{pmatrix} \quad \text{and}$$

$$\begin{pmatrix} K & L \\ M & N \end{pmatrix} = \begin{pmatrix} U & XB \\ 0 & I \end{pmatrix} \begin{pmatrix} A_c & B_c \\ C_c & D_c \end{pmatrix} \begin{pmatrix} V^T & 0 \\ CY & I \end{pmatrix} + \begin{pmatrix} XAY & 0 \\ 0 & 0 \end{pmatrix} \qquad (4.2.10)$$

satisfy the LMI's (4.2.1).

Note that U and V are square and nonsingular such that (4.2.10) leads to the formulas

$$\mathcal{X} = \begin{pmatrix} Y & V \\ I & 0 \end{pmatrix}^{-1} \begin{pmatrix} I & 0 \\ X & U \end{pmatrix} \quad \text{and}$$

$$\begin{pmatrix} A_c & B_c \\ C_c & D_c \end{pmatrix} = \begin{pmatrix} U & XB \\ 0 & I \end{pmatrix}^{-1} \begin{pmatrix} K - XAY & L \\ M & N \end{pmatrix} \begin{pmatrix} V^T & 0 \\ CY & I \end{pmatrix}^{-1}.$$

Due to the zero blocks in the inverses, the formulas can be rendered even more explicit. Of course, numerically it is better to directly solve the equations (4.2.10) by a stable technique.

Proof. Suppose a controller and some $\mathcal{X}$ satisfy (4.2.1). Let us partition

$$\mathcal{X} = \begin{pmatrix} X & U \\ U^T & * \end{pmatrix} \quad \text{and} \quad \mathcal{X}^{-1} = \begin{pmatrix} Y & V \\ V^T & * \end{pmatrix}$$

according to $\mathcal{A}$. Define

$$\mathcal{Y} = \begin{pmatrix} Y & I \\ V^T & 0 \end{pmatrix} \quad \text{and} \quad \mathcal{Z} = \begin{pmatrix} I & 0 \\ X & U \end{pmatrix} \quad \text{to get} \quad \mathcal{Y}^T \mathcal{X} = \mathcal{Z}. \qquad (4.2.11)$$

Without loss of generality we can assume that the dimension of A_c is larger than that of A (why?). Hence, U has more columns than rows, and we can perturb this block (since we work with strict inequalities) such that it has full row rank. Then $\mathcal{Z}$ has full row rank and, hence, $\mathcal{Y}$ has full column rank.

Due to $XY + UV^T = I$, we infer

$$\mathcal{Y}^T \mathcal{X} \mathcal{Y} = \begin{pmatrix} Y & I \\ I & X \end{pmatrix} = \mathbf{X}(v)$$

what leads to the first relation in (4.2.4). Let us now consider

$$\begin{pmatrix} \mathcal{Y} & 0 \\ 0 & I \end{pmatrix}^T \begin{pmatrix} \mathcal{X} \mathcal{A} & \mathcal{X} \mathcal{B}_j \\ \mathcal{C}_j & \mathcal{D}_j \end{pmatrix} \begin{pmatrix} \mathcal{Y} & 0 \\ 0 & I \end{pmatrix} = \begin{pmatrix} \mathcal{Y}^T \mathcal{X} \mathcal{A} \mathcal{Y} & \mathcal{Y}^T \mathcal{X} \mathcal{B}_j \\ \mathcal{C}_j \mathcal{Y} & \mathcal{D}_j \end{pmatrix}.$$

Using (4.1.5), a very brief calculation (do it!) reveals that

$$
\begin{pmatrix} \mathcal{Y}^T \mathcal{X} \mathcal{A} \mathcal{Y} & \mathcal{Y}^T \mathcal{X} \mathcal{B}_j \\ \mathcal{C}_j \mathcal{Y} & \mathcal{D}_j \end{pmatrix} = \begin{pmatrix} \mathcal{Z} \mathcal{A} \mathcal{Y} & \mathcal{Z} \mathcal{B}_j \\ \mathcal{C}_j \mathcal{Y} & \mathcal{D}_j \end{pmatrix} = \left(\begin{array}{cc|c} AY & A & B_j \\ 0 & XA & XB_j \\ \hline C_j Y & C_j & D_j \end{array} \right) +
$$

$$
+ \begin{pmatrix} 0 & B \\ I & 0 \\ 0 & E_j \end{pmatrix} \left[\begin{pmatrix} U & XB \\ 0 & I \end{pmatrix} \begin{pmatrix} A_c & B_c \\ C_c & D_c \end{pmatrix} \begin{pmatrix} V^T & 0 \\ CY & I \end{pmatrix} + \begin{pmatrix} XAY & 0 \\ 0 & 0 \end{pmatrix} \right] \left(\begin{array}{cc|c} I & 0 & 0 \\ 0 & C & F_j \end{array} \right).
$$

If we introduce the new parameters $\begin{pmatrix} K & L \\ M & N \end{pmatrix}$ as in (4.2.10), we infer

$$
\begin{pmatrix} \mathcal{Y}^T \mathcal{X} \mathcal{A} \mathcal{Y} & \mathcal{Y}^T \mathcal{X} \mathcal{B}_j \\ \mathcal{C}_j \mathcal{Y} & \mathcal{D}_j \end{pmatrix} =
$$

$$
= \left(\begin{array}{cc|c} AY & A & B_j \\ 0 & XA & XB_j \\ \hline C_j Y & C_j & D_j \end{array} \right) + \begin{pmatrix} 0 & B \\ I & 0 \\ 0 & E_j \end{pmatrix} \begin{pmatrix} K & L \\ M & N \end{pmatrix} \left(\begin{array}{cc|c} I & 0 & 0 \\ 0 & C & F_j \end{array} \right) =
$$

$$
= \left(\begin{array}{cc|c} AY + BM & A + BNC & B_j + BNF_j \\ K & AX + LC & XB_j + LF_j \\ \hline C_j Y + E_j M & C_j + E_j NC & D_j + E_j NF_j \end{array} \right) = \begin{pmatrix} \boldsymbol{A}(v) & \boldsymbol{B}_j(v) \\ \boldsymbol{C}_j(v) & \boldsymbol{D}_j(v) \end{pmatrix}.
$$

Hence the relations (4.2.4) are valid. Since $\mathcal{Y}$ has full column rank, (4.2.1) implies (4.2.6), and by (4.2.4), (4.2.6) is identical to (4.2.7). This proves necessity.

To reverse the arguments we assume that v is a solution of (4.2.7). Due to $\boldsymbol{X}(v) > 0$, we infer that $I - XY$ is nonsingular. Hence we can factorize $I - XY = UV^T$ with square and nonsingular U, V. Then $\mathcal{Y}$ and $\mathcal{Z}$ defined in (4.2.11) are, as well, square and nonsingular. Hence we can choose $\mathcal{X}$, $\begin{pmatrix} A_c & B_c \\ C_c & D_c \end{pmatrix}$ such that (4.2.10) hold true; this implies that, again, the relations (4.2.4) are valid. Therefore, (4.2.7) and (4.2.6) are identical. Since $\mathcal{Y}$ is nonsingular, a congruence transformation with $\mathcal{Y}^{-1}$ and $\mathrm{diag}(\mathcal{Y}^{-1}, I)$ leads from (4.2.6) back to (4.2.1) and the proof is finished. ∎

We have obtained a general procedure for deriving from analysis inequalities the corresponding synthesis inequalities and for construction corresponding controllers as follows:

- Rewrite the analysis inequalities in the blocks $\mathcal{X}$, $\mathcal{X} \mathcal{A}$, $\mathcal{X} \mathcal{B}_j$, $\mathcal{C}_j$, $\mathcal{D}_j$ in order to be able to find a (formal) congruence transformation involving $\mathcal{Y}$ which leads to inequalities in the blocks $\mathcal{Y}^T \mathcal{X} \mathcal{Y}$, $\mathcal{Y}^T \mathcal{X} \mathcal{A} \mathcal{Y}$, $\mathcal{Y}^T \mathcal{X} \mathcal{B}_j$, $\mathcal{C}_j \mathcal{Y}$, $\mathcal{D}_j$.

- Perform the substitution (4.2.4) to arrive at matrix inequalities in the variables v.

- After having solved the synthesis inequalities for v, one factorizes $I - XY$ into nonsingular blocks UV^T and solves the equations (4.2.10) to obtain the controller parameters A_c, B_c, C_c, D_c and a Lyapunov matrix $\mathcal{X}$ which render the analysis inequalities satisfied.

The power of this procedure lies in its simplicity and its generality. Virtually all controller design methods that are based on matrix inequality analysis results can be converted with ease into the corresponding synthesis result. In the subsequent section we will include an extensive discussion of how to apply this technique to the various analysis results that have been obtained in the present notes.

Remark on the controller order. In Theorem 4.2 we have not restricted the order of the controller. In proving necessity of the solvability of the synthesis inequalities, the size of A_c was arbitrary. The specific construction of a controller in proving sufficiency leads to an A_c that has *the same size* as A. Hence Theorem 4.2 also include the side result that controllers of order larger than that of the plant offer no advantage over controllers that have the same order as the plant. The story is very different in reduced order control: Then the intention is to include a constraint $\dim(A_c) \leqslant k$ for some k that is smaller than the dimension of A. It is not very difficult to derive the corresponding synthesis inequalities; however, they include *rank constraints* that are hard if not impossible to treat by current optimization techniques. We will only briefly comment on a concrete result later.

Remark on strictly proper controllers. Note that the direct feed-through of the controller D_c is actually not transformed; we simply have $D_c = N$. If we intend to design a strictly proper controller (i.e. $D_c = 0$), we can just set $N = 0$ to arrive at the corresponding synthesis inequalities. The construction of the other controller parameters remains the same. Clearly, the same holds if one wishes to impose an arbitrary more refined structural constraint on the direct feed-through term as long as it can be expressed in terms of LMI's.

Remarks on numerical aspects. After having verified the solvability of the synthesis inequalities, we recommend to take some precautions to improve the conditioning of the calculations to reconstruct the controller out of the decision variable v. In particular, one should avoid that the parameters v get to large, and that $I - XY$ is close to singular what might render the controller computation ill-conditioned.

We have observed good results with the following two-step procedure:

- Add to the feasibility inequalities the bounds

$$\|X\| < \alpha, \ \|Y\| < \alpha, \ \left\| \begin{pmatrix} K & L \\ M & N \end{pmatrix} \right\| < \alpha$$

  as extra constraints and minimize α. Note that these bounds are equivalently rewritten in LMI form as

$$X < \alpha I, \ Y < \alpha I, \ \left(\begin{array}{cc|cc} \alpha I & 0 & K & L \\ 0 & \alpha I & M & N \\ \hline K^T & M^T & \alpha I & 0 \\ L^T & N^T & 0 & \alpha I \end{array} \right) > 0.$$

  Hence they can be easily included in the feasibility test, and one can directly minimize α to compute the smallest bound α_*.

- In a second step, one adds to the feasibility inequalities and to the bounding inequalities for some enlarged but fixed $\alpha > \alpha_*$ the extra constraint

$$\begin{pmatrix} Y & \beta I \\ \beta I & X \end{pmatrix} > 0.$$

Of course, the resulting LMI system is feasible for $\beta = 1$. One can hence maximize β to obtain a supremal $\beta_* > 1$. The value β_* gives an indication of the conditioning of the controller reconstruction procedure. In fact, the extra inequality is equivalent to $X - \beta^2 Y^{-1} > 0$. Hence, maximizing β amounts to 'pushing X away from Y^{-1}'. Therefore, this step is expected to push the smallest singular value of $I - XY$ away from zero. The larger the smaller singular value of $I - XY$, the larger one can choose the smallest singular values of both U and V in the factorization $I - XY = UV^T$. This improves the conditioning of U and V, and renders the calculation of the controller parameters more reliable.

4.3 Other Performance Specifications

4.3.1 H_∞ Design

The optimal value of the H_∞ problem is defined as

$$\gamma_j^* = \inf_{A_c, B_c, C_c, D_c \text{ such that } \sigma(\mathcal{A}) \subset \mathbb{C}^-} \|\mathcal{T}_j\|_\infty.$$

Clearly, the number γ_j is larger than γ_j^* iff there exists a controller which renders

$$\sigma(\mathcal{A}) \subset \mathbb{C}^- \quad \text{and} \quad \|\mathcal{T}_j\|_\infty < \gamma_j$$

satisfied. These two properties are equivalent to stability and quadratic performance for the index

$$P_j = \begin{pmatrix} Q_j & S_j \\ S_j^T & R_j \end{pmatrix} = \begin{pmatrix} -\gamma_j I & 0 \\ 0 & (\gamma_j I)^{-1} \end{pmatrix}.$$

The corresponding synthesis inequalities (4.2.7) are rewritten with Lemma 4.1 to

$$\boldsymbol{X}(v) > 0, \quad \begin{pmatrix} \boldsymbol{A}(v)^T + \boldsymbol{A}(v) & \boldsymbol{B}_j(v) & \boldsymbol{C}_j(v)^T \\ \boldsymbol{B}_j(v)^T & -\gamma_j I & \boldsymbol{D}_j(v)^T \\ \boldsymbol{C}_j(v) & \boldsymbol{D}_j(v) & -\gamma_j I \end{pmatrix} < 0.$$

Note that the the optimal H_∞ value γ_j^* is then just given by the minimal γ_j for which these inequalities are feasible; one can directly compute γ_l^* by a standard LMI algorithm.

For the controller reconstruction, one should improve the conditioning (as described in the previous section) by an additional LMI optimization. We recommend not to perform this

step with the optimal value γ_j^* itself but with a slightly increased value $\gamma_j > \gamma_j^*$. This is motivated by the observation that, at optimality, the matrix $\boldsymbol{X}(v)$ is often (but not always!) close to singular; then $I - XY$ is close to singular and it is expected to be difficult to render it better conditioned if γ_j is too close to the optimal value γ_j^*.

4.3.2 Positive Real Design

In this problem the goal is to test whether there exists a controller which renders the following two conditions satisfied:

$$\sigma(\mathcal{A}) \subset \mathbb{C}^-, \quad \mathcal{T}_j(i\omega)^* + \mathcal{T}_j(i\omega) > 0 \ \text{ for all } \ \omega \in \mathbb{R} \cup \{\infty\}.$$

This is equivalent to stability and quadratic performance for

$$P_j = \left(\begin{array}{cc} Q_j & S_j \\ S_j^T & R_j \end{array} \right) = \left(\begin{array}{cc} 0 & -I \\ -I & 0 \end{array} \right),$$

and the corresponding synthesis inequalities read as

$$\boldsymbol{X}(v) > 0, \quad \left(\begin{array}{cc} \boldsymbol{A}(v)^T + \boldsymbol{A}(v) & \boldsymbol{B}_j(v) - \boldsymbol{C}_j(v)^T \\ \boldsymbol{B}_j(v)^T - \boldsymbol{C}_j(v) & -\boldsymbol{D}_j(v) - \boldsymbol{D}_j(v)^T \end{array} \right) < 0.$$

4.3.3 H_2-Problems

Let us define the linear functional

$$f_j(Z) := \mathrm{trace}(Z).$$

Then we recall that $\mathcal{A}$ is stable and $\|\mathcal{T}_j\|_2 < \gamma_j$ iff there exists a symmetric $\mathcal{X}$ with

$$\mathcal{D}_j = 0, \quad \mathcal{X} > 0, \quad \left(\begin{array}{cc} \mathcal{A}^T \mathcal{X} + \mathcal{X} \mathcal{A} & \mathcal{X} \mathcal{B}_j \\ \mathcal{B}_j^T \mathcal{X} & -\gamma_j I \end{array} \right) < 0, \quad f_j(\mathcal{C}_j \mathcal{X}^{-1} \mathcal{C}_j^T) < \gamma_j. \tag{4.3.1}$$

The latter inequality is rendered affine in $\mathcal{X}$ and $\mathcal{C}_j$ by introducing the auxiliary variable (or slack variable) Z_j. Indeed, the analysis test is equivalent to

$$\mathcal{D}_j = 0, \quad \left(\begin{array}{cc} \mathcal{A}^T \mathcal{X} + \mathcal{X} \mathcal{A} & \mathcal{X} \mathcal{B}_j \\ \mathcal{B}_j^T \mathcal{X} & -\gamma_j I \end{array} \right) < 0, \quad \left(\begin{array}{cc} \mathcal{X} & \mathcal{C}_j^T \\ \mathcal{C}_j & Z_j \end{array} \right) > 0, \quad f_j(Z_j) < \gamma_j. \tag{4.3.2}$$

(Why?) This version of the inequalities is suited to simply read-off the corresponding synthesis inequalities.

Corollary 4.3 *There exists a controller that renders (4.3.2) for some $\mathcal{X}$, Z_j satisfied iff there exist v and Z_j with*

$$\boldsymbol{D}_j(v) = 0, \quad \left(\begin{array}{cc} \boldsymbol{A}(v)^T + \boldsymbol{A}(v) & \boldsymbol{B}_j(v) \\ \boldsymbol{B}_j(v)^T & -\gamma_j I \end{array} \right) < 0, \quad \left(\begin{array}{cc} \boldsymbol{X}(v) & \boldsymbol{C}_j(v)^T \\ \boldsymbol{C}_j(v) & Z_j \end{array} \right) > 0, \quad f_j(Z_j) < \gamma_j.$$

$$\tag{4.3.3}$$

The proof of this statement and the controller construction are literally the same as for quadratic performance.

For the generalized H_2-norm $\|\mathcal{T}_j\|_{2g}$, we recall that $\mathcal{A}$ is stable and $\|\mathcal{T}_j\|_{2g} < \gamma_j$ iff

$$\mathcal{D}_j = 0, \quad \mathcal{X} > 0, \quad \begin{pmatrix} \mathcal{A}^T \mathcal{X} + \mathcal{X} \mathcal{A} & \mathcal{X} \mathcal{B}_j \\ \mathcal{B}_j^T \mathcal{X} & -\gamma_j I \end{pmatrix} < 0, \quad \mathcal{C}_j \mathcal{X}^{-1} \mathcal{C}_j^T < \gamma_j I.$$

These conditions are nothing but

$$\mathcal{D}_j = 0, \quad \begin{pmatrix} \mathcal{A}^T \mathcal{X} + \mathcal{X} \mathcal{A} & \mathcal{X} \mathcal{B}_j \\ \mathcal{B}_j^T \mathcal{X} & -\gamma_j I \end{pmatrix} < 0, \quad \begin{pmatrix} \mathcal{X} & \mathcal{C}_j^T \\ \mathcal{C}_j & \gamma_j I \end{pmatrix} > 0$$

and it is straightforward to derive the synthesis LMI's.

Note that the corresponding inequalities are equivalent to (4.3.3) for the function

$$f_j(Z) = Z.$$

In contrast to the genuine H_2-problem, there is no need for the extra variable Z_j to render the inequalities affine.

Remarks.

- If f assigns to Z its diagonal $\operatorname{diag}(z_1, \ldots, z_m)$ (where m is the dimension of Z), one characterizes a bound on the gain of $L_2 \ni w_j \rightarrow z_j \in L_\infty$ if equipping L_∞ with the norm $\|x\|_\infty := \operatorname{ess\,sup}_{t \geqslant 0} \max_k |x_k(t)|$ [26, 30]. Note that the three concrete H_2-like analysis results for $f_j(Z) = \operatorname{trace}(Z)$, $f_j(Z) = Z$, $f_j(Z) = \operatorname{diag}(z_1, \ldots, z_m)$ are exact characterizations, and that the corresponding synthesis results do not involve any conservatism.

- In fact, Corollary 4.3 holds for any affine function f that maps symmetric matrices into symmetric matrices (of possibly different dimension) and that has the property $Z \geqslant 0 \Rightarrow f(Z) \geqslant 0$. (Why?) Hence, Corollary 4.3 admits many other specializations.

- Similarly as in the H_∞ problem, we can directly minimize the bound γ_j to find the optimal H_2-value or the optimal generalized H_2-value that can be achieved by stabilizing controllers.

- We observe that it causes no trouble in our general procedure to derive the synthesis inequalities if the underlying analysis inequalities involve certain auxiliary parameters (such as Z_j) as extra decision variables.

- It is instructive to equivalently rewrite (4.3.2) as $\mathcal{X} > 0$, $Z_j > 0$, $f_j(Z_j) < \gamma_j$ and

$$\begin{pmatrix} I & 0 \\ \mathcal{X}\mathcal{A} & \mathcal{X}\mathcal{B}_j \\ 0 & I \end{pmatrix}^T \left(\begin{array}{cc|c} 0 & I & 0 \\ I & 0 & 0 \\ \hline 0 & 0 & -\gamma_j I \end{array} \right) \begin{pmatrix} I & 0 \\ \mathcal{X}\mathcal{A} & \mathcal{X}\mathcal{B}_j \\ 0 & I \end{pmatrix} < 0,$$

$$\begin{pmatrix} I & 0 \\ 0 & I \\ \mathcal{C}_j & \mathcal{D}_j \end{pmatrix}^T \left(\begin{array}{c|cc} -\mathcal{X} & 0 & 0 \\ \hline 0 & 0 & 0 \\ 0 & 0 & Z_j^{-1} \end{array} \right) \begin{pmatrix} I & 0 \\ 0 & I \\ \mathcal{C}_j & \mathcal{D}_j \end{pmatrix} \leqslant 0.$$

Note that the last inequality is non-strict and includes the algebraic constraint $\mathcal{D}_j = 0$. It can be equivalently replaced by

$$\begin{pmatrix} I \\ \mathcal{C}_j \end{pmatrix}^T \begin{pmatrix} -\mathcal{X} & 0 \\ 0 & Z_j^{-1} \end{pmatrix} \begin{pmatrix} I \\ \mathcal{C}_j \end{pmatrix} < 0, \quad \mathcal{D}_j = 0.$$

The synthesis relations then read as $\boldsymbol{X}(v) > 0, \quad Z_j > 0, \quad f_j(Z_j) < \gamma_j$ and

$$\left(\begin{array}{cc} I & 0 \\ \boldsymbol{A}(v) & \boldsymbol{B}_j(v) \\ \hline 0 & I \end{array} \right)^T \left(\begin{array}{cc|c} 0 & I & 0 \\ I & 0 & 0 \\ \hline 0 & 0 & -\gamma_j I \end{array} \right) \left(\begin{array}{cc} I & 0 \\ \boldsymbol{A}(v) & \boldsymbol{B}_j(v) \\ \hline 0 & I \end{array} \right) < 0, \tag{4.3.4}$$

$$\begin{pmatrix} I \\ \boldsymbol{C}_j(v) \end{pmatrix}^T \begin{pmatrix} -\boldsymbol{X}(v) & 0 \\ 0 & Z_j^{-1} \end{pmatrix} \begin{pmatrix} I \\ \boldsymbol{C}_j(v) \end{pmatrix} < 0, \quad \boldsymbol{D}_j(v) = 0. \tag{4.3.5}$$

The first inequality is affine in v, whereas the second one can be rendered affine in v and Z_j with Lemma 4.1.

4.3.4   Upper Bound on Peak-to-Peak Norm

The controller (4.1.2) renders $\mathcal{A}$ stable and the bound

$$\|w_j\|_\infty \leqslant \gamma_j \|z_j\|_\infty \quad \text{for all} \quad z_j \in L_\infty$$

satisfied **if** there exist a symmetric $\mathcal{X}$ and real parameters λ, μ with

$$\lambda > 0, \quad \begin{pmatrix} \mathcal{A}^T\mathcal{X} + \mathcal{X}\mathcal{A} + \lambda\mathcal{X} & \mathcal{X}\mathcal{B}_j \\ \mathcal{B}_j^T\mathcal{X} & 0 \end{pmatrix} + \begin{pmatrix} 0 & I \\ \mathcal{C}_j & \mathcal{D}_j \end{pmatrix}^T \begin{pmatrix} -\mu I & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & I \\ \mathcal{C}_j & \mathcal{D}_j \end{pmatrix} < 0$$

$$\begin{pmatrix} 0 & I \\ \mathcal{C}_j & \mathcal{D}_j \end{pmatrix}^T \begin{pmatrix} 0 & 0 \\ 0 & \frac{1}{\gamma_j}I \end{pmatrix} \begin{pmatrix} 0 & I \\ \mathcal{C}_j & \mathcal{D}_j \end{pmatrix} < \begin{pmatrix} \lambda\mathcal{X} & 0 \\ 0 & (\gamma_j - \mu)I \end{pmatrix}.$$

(Note that $\mathcal{X} > 0$ is built in. Where?) The inequalities are obviously equivalent to

$$\lambda > 0, \quad \begin{pmatrix} \mathcal{A}^T\mathcal{X} + \mathcal{X}\mathcal{A} + \lambda\mathcal{X} & \mathcal{X}\mathcal{B}_j \\ \mathcal{B}_j^T\mathcal{X} & -\mu I \end{pmatrix} < 0, \quad \begin{pmatrix} \lambda\mathcal{X} & 0 & \mathcal{C}_j^T \\ 0 & (\gamma_j - \mu)I & \mathcal{D}_j^T \\ \mathcal{C}_j & \mathcal{D}_j & \gamma_j I \end{pmatrix} > 0,$$

and the corresponding synthesis inequalities thus read as

$$\lambda > 0, \quad \begin{pmatrix} \boldsymbol{A}(v)^T + \boldsymbol{A}(v) + \lambda\boldsymbol{X}(v) & \boldsymbol{B}_j(v) \\ \boldsymbol{B}_j(v)^T & -\mu I \end{pmatrix} < 0, \quad \begin{pmatrix} \lambda\boldsymbol{X}(v) & 0 & \boldsymbol{C}_j(v)^T \\ 0 & (\gamma_j - \mu)I & \boldsymbol{D}_j(v)^T \\ \boldsymbol{C}_j(v) & \boldsymbol{D}_j(v) & \gamma_j I \end{pmatrix} > 0.$$

If these inequalities are feasible, one can construct a stabilizing controller which bounds the peak-to-peak norm of $z_j = \mathcal{T}_j z_j$ by γ_j. We would like to stress that the converse of this statement is not true since the analysis result involves conservatism.

Note that the synthesis inequalities are formulated in terms of the variables v, λ, and μ; hence they are *non-linear* since $\lambda \boldsymbol{X}(v)$ depends quadratically on λ and v. This problem can be overcome as follows: For fixed $\lambda > 0$, test whether the resulting *linear* matrix inequalities are feasible; if yes, one can stop since the bound γ_j on the peak-to-peak norm has been assured; if the LMI's are infeasible, one has to pick another $\lambda > 0$ and repeat the test.

In practice, it might be advantageous to find the best possible upper bound on the peak-to-peak norm that can be assured with the present analysis result. This would lead to the problem of minimizing γ_j under the synthesis inequality constraints as follows: Perform a line-search over $\lambda > 0$ to minimize $\gamma_j^*(\lambda)$, the minimal value of γ_j if $\lambda > 0$ is held fixed; note that the calculation of $\gamma_j^*(\lambda)$ indeed amounts to solving a genuine LMI problem. The line-search leads to the best achievable upper bound

$$\gamma_j^u = \inf_{\lambda > 0} \gamma_j^*(\lambda).$$

To estimate the conservatism, let us recall that $\|\mathcal{T}_j\|_\infty$ is a lower bound on the peak-to-peak norm of $\mathcal{T}_j$. (Why?) If we calculate the minimal achievable H_∞-norm, say γ_j^l, of $\mathcal{T}_j$, we know that the actual optimal peak-to-peak gain must be contained in the interval

$$[\gamma_j^l, \gamma_j^u].$$

If the length of this interval is small, we have a good estimate of the actual optimal peak-to-peak gain that is achievable by control, and if the interval is large, this estimate is poor.

4.4   Multi-Objective and Mixed Controller Design

In a realistic design problem one is usually not just confronted with a single-objective problem but one has to render various objectives satisfied. As a typical example, one might wish to keep the H_∞ norm of $z_1 = \mathcal{T}_1 w_1$ below a bound γ_1 to ensure robust stability against uncertainties entering as $w_1 = \Delta z_1$ where the stable mapping Δ has L_2-gain smaller than $1/\gamma_1$, and render, at the same time, the H_2-norm of $z_2 = \mathcal{T}_2 w_2$ as small as possible to ensure good performance measured in the H_2-norm (such as guaranteeing small asymptotic variance of z_j against white noise inputs w_j or small energy of the output z_j against pulses as inputs w_j.)

Such a problem would lead to minimizing γ_2 over all controllers which render

$$\sigma(\mathcal{A}) \subset \mathbb{C}^-, \quad \|\mathcal{T}_1\|_\infty < \gamma_1, \quad \|\mathcal{T}_2\|_2 < \gamma_2 \tag{4.4.1}$$

satisfied. This is a **multi-objective** H_2/H_∞ control problem with two performance specifications.

Note that it is often interesting to investigate the trade-off between the H_∞-norm and the H_2-norm constraint. For that purpose one plots the curve of optimal values if varying γ_1 in some

interval $[\gamma_1^l, \gamma_1^u]$ where the lower bound γ_1^l could be taken close to the smallest achievable H_∞-norm of $\mathcal{T}_1$. Note that the optimal value will be non-increasing if increasing γ_1. The actual curve will provide insight in how far one can improve performance by giving up robustness. In practice, it might be numerically advantageous to give up the hard constraints and proceed, alternatively, as follows: For fixed real weights α_1 and α_2, minimize

$$\alpha_1\gamma_1 + \alpha_2\gamma_2$$

over all controllers that satisfy (4.4.1). The larger α_j, the more weight is put on penalizing large values of γ_j, the more the optimization procedure is expected to reduce the corresponding bound γ_j.

Multi-objective control problems as formulated here are hard to solve. Let us briefly sketch one line of approach. The Youla parameterization [17] reveals that the set of all $\mathcal{T}_j$ that can be obtained by internally stabilizing controllers can be parameterized as

$$T_1^j + T_2^j Q T_3^j \quad \text{with } Q \text{ varying freely in } RH_\infty^{p\times q}.$$

Here T_1^j, T_2^j, T_3^j are real-rational proper and stable transfer matrices which can be easily computed in terms of the system description (4.1.1) and an arbitrary stabilizing controller. Recall also that $RH_\infty^{p\times q}$ denotes the algebra of real-rational proper and stable transfer matrices of dimension $p \times q$. With this re-parameterization, the multi-objective control problem then amounts to finding a $Q \in RH_\infty^{p\times q}$ that minimizes γ_2 under the constraints

$$\|T_1^1 + T_2^1 Q T_3^1\|_\infty < \gamma_1, \quad \|T_1^2 + T_2^2 Q T_3^2\|_2 < \gamma_2. \tag{4.4.2}$$

After this re-formulation, we are hence faced with a convex optimization problem in the parameter Q which varies in the infinite-dimensional space RH_∞. A pretty standard Ritz-Galerkin approximation scheme leads to finite-dimensional problems. In fact, consider for a fixed real parameter $a > 0$ the sequence of finite-dimensional subspaces

$$\mathcal{S}_\nu := \left\{ Q_0 + Q_1 \frac{s-a}{s+a} + Q_2 \frac{(s-a)^2}{(s+a)^2} + \cdots + Q_\nu \frac{(s-a)^\nu}{(s+a)^\nu} : Q_0, \ldots, Q_\nu \in \mathbb{R}^{p\times q} \right\}$$

of the space $RH_\infty^{p\times q}$. Let us now denote the infimum of all γ_2 satisfying the constraint (4.4.2) for $Q \in RH_\infty^{p\times q}$ by γ_2^*, and that for $Q \in \mathcal{S}_\nu$ by $\gamma_2(\nu)$. Since $\mathcal{S}_\nu \subset RH_\infty^{p\times q}$, we clearly have

$$\gamma_2^* \leqslant \gamma_2(\nu+1) \leqslant \gamma_2(\nu) \quad \text{for all} \quad \nu = 0, 1, 2 \ldots.$$

Hence solving the optimization problems for increasing ν leads to a non-increasing sequence of values $\gamma(\nu)$ that are all upper bounds on the actual optimum γ_2^*. If we now note that any element of Q can be approximated in the H_∞-norm with arbitrary accuracy by an element in $\mathcal{S}_\nu$ if ν is chosen sufficiently large, it is not surprising that $\gamma_2(\nu)$ actually converges to γ_2^* for $\nu \to \infty$. To be more precise, we need to assume that the strict constraint $\|T_1^1 + T_2^1 Q T_3^1\|_\infty < \gamma_1$ is feasible for $Q \in \mathcal{S}_\nu$ and some ν, and that T_1^1 and T_2^2 or T_2^3 are strictly proper such that $\|T_1^2 + T_2^2 Q T_3^2\|_2$ is finite for all $Q \in RH_\infty^{p\times q}$. Then it is not difficult to show that

$$\lim_{\nu \to \infty} \gamma_2(\nu) = \gamma_2^*.$$

Finally, we observe that computing $\gamma_2(\nu)$ is in fact an LMI problem. (Why?) For more information on this and related problems the reader is referred to [9, 31, 41].

We observe that the approach that is sketched above suffers from two severe disadvantages: First, if improving the approximation accuracy by letting ν grow, the size of the LMI's and the number of variables that are involved grow drastically what renders the corresponding computations slow. Second, increasing ν amounts to a potential increase of the McMillan degree of $Q \in \mathcal{S}_\nu$ what leads to controllers whose McMillan degree cannot be bounded a priori.

In view of these difficulties, it has been proposed to replace the multi-objective control problem by the mixed control problem. To prepare its definition, recall that the conditions (4.4.1) are guaranteed by the existence of symmetric matrices $\mathcal{X}_1$, $\mathcal{X}_2$, Z_2 satisfying

$$\mathcal{X}_1 > 0, \quad \begin{pmatrix} \mathcal{A}^T\mathcal{X}_1 + \mathcal{X}_1\mathcal{A} & \mathcal{X}_1\mathcal{B}_1 & \mathcal{C}_1^T \\ \mathcal{X}_1\mathcal{B}_1 & -\gamma_1 I & \mathcal{D}_1^T \\ \mathcal{C}_1 & \mathcal{D}_1 & -\gamma_1 I \end{pmatrix} < 0$$

$$\mathcal{D}_2 = 0, \quad \begin{pmatrix} \mathcal{A}^T\mathcal{X}_2 + \mathcal{X}_2\mathcal{A} & \mathcal{X}_2\mathcal{B}_2 \\ \mathcal{B}_2^T\mathcal{X}_2 & -\gamma_2 I \end{pmatrix} < 0, \quad \begin{pmatrix} \mathcal{X}_2 & \mathcal{C}_2^T \\ \mathcal{C}_2 & Z_2 \end{pmatrix} > 0, \quad \operatorname{trace}(Z_2) < \gamma_2.$$

If trying to apply the general procedure to derive the synthesis inequalities, there is some trouble since the controller parameter transformation depends on the closed-loop Lyapunov matrix; here two such matrices $\mathcal{X}_1$, $\mathcal{X}_2$ do appear such that the technique breaks down. This observation itself motivates a remedy: Just force the two Lyapunov matrices to be equal. This certainly introduces conservatism that is, in general, hard to quantify. On the positive side, if one can find a common matrix

$$\mathcal{X} = \mathcal{X}_1 = \mathcal{X}_2$$

that satisfies the analysis relations, we can still guarantee (4.4.1) to hold. However, the converse is not true, since (4.4.1) does not imply the existence of common Lyapunov matrix to satisfy the above inequalities.

This discussion leads to the definition of the **mixed** H_2/H_∞ control problem: Minimize γ_2 subject to the existence of $\mathcal{X}$, Z_2 satisfying

$$\begin{pmatrix} \mathcal{A}^T\mathcal{X} + \mathcal{X}\mathcal{A} & \mathcal{X}\mathcal{B}_1 & \mathcal{C}_1^T \\ \mathcal{B}_1^T\mathcal{X} & -\gamma_1 I & \mathcal{D}_1^T \\ \mathcal{C}_1 & \mathcal{D}_1 & -\gamma_1 I \end{pmatrix} < 0$$

$$\mathcal{D}_2 = 0, \quad \begin{pmatrix} \mathcal{A}^T\mathcal{X} + \mathcal{X}\mathcal{A} & \mathcal{X}\mathcal{B}_2 \\ \mathcal{B}_2^T\mathcal{X} & -\gamma_2 I \end{pmatrix} < 0, \quad \begin{pmatrix} \mathcal{X} & \mathcal{C}_2^T \\ \mathcal{C}_2 & Z_2 \end{pmatrix} > 0, \quad \operatorname{trace}(Z_2) < \gamma_2.$$

This problem is amenable to our general procedure. One proves as before that the corre-

sponding synthesis LMI's are

$$
\begin{pmatrix}
\boldsymbol{A}(v)^T + \boldsymbol{A}(v) & \boldsymbol{B}_1(v) & \boldsymbol{C}_1(v)^T \\
\boldsymbol{B}_1(v)^T & -\gamma_1 I & \boldsymbol{D}_1(v)^T \\
\boldsymbol{C}_1(v) & \boldsymbol{D}_1(v) & -\gamma_1 I
\end{pmatrix} < 0
$$

$$
\boldsymbol{D}_2(v) = 0, \quad
\begin{pmatrix}
\boldsymbol{A}(v)^T + \boldsymbol{A}(v) & \boldsymbol{B}_2(v) \\
\boldsymbol{B}_2(v)^T & -\gamma_2 I
\end{pmatrix} < 0, \quad
\begin{pmatrix}
\boldsymbol{X}(v) & \boldsymbol{C}_2(v)^T \\
\boldsymbol{C}_2(v) & Z_2
\end{pmatrix} > 0, \quad \operatorname{trace}(Z_2) < \gamma_2,
$$

and the controller construction remains unchanged.

Let us conclude this section with some important remarks.

- After having solved the synthesis inequalities corresponding to the mixed problem for v and Z_2, one can construct a controller which satisfies (4.4.1) and which has a McMillan degree (size of A_c) that is not larger than (equal to) the size of A.

- For the controller resulting from mixed synthesis one can perform an analysis with different Lyapunov matrices $\mathcal{X}_1$ and $\mathcal{X}_2$ without any conservatism. In general, the actual H_∞-norm of $\mathcal{T}_1$ will be strictly smaller than γ_1, and the H_2-norm will be strictly smaller than the optimal value obtained from solving the mixed problem. Judging a mixed controller should, hence, rather be based on an additional non-conservative and direct analysis.

- Performing synthesis by searching for a common Lyapunov matrix introduces conservatism. Little is known about how to estimate this conservatism a priori. However, the optimal value of the mixed problem is always an upper bound of the optimal value of the actual multi-objective problem.

- Starting from a mixed controller, it has been suggested in [35, 36] how to compute sequences of upper and lower bounds, on the basis of solving LMI problems, that approach the actual optimal value. This allows to provide an a posteriori estimate of the conservatism that is introduced by setting $\mathcal{X}_1$ equal to $\mathcal{X}_2$.

- If starting from different versions of the analysis inequalities (e.g. through scaling the Lyapunov matrix), the artificial constraint $\mathcal{X}_1 = \mathcal{X}_2$ might lead to a different mixed control problem. Therefore, it is recommended to choose those analysis tests that are expected to lead to Lyapunov matrices which are close to each other. However, there is no general rule how to guarantee this property.

- In view of the previous remark, let us sketch one possibility to reduce the conservatism in mixed design. If we multiply the analysis inequalities for stability of $\mathcal{A}$ and for $\|\mathcal{T}_1\|_\infty < \gamma_1$ by an arbitrary real parameter $\alpha > 0$, we obtain

$$
\alpha \mathcal{X}_1 > 0, \quad
\begin{pmatrix}
\mathcal{A}^T(\alpha \mathcal{X}_1) + (\alpha \mathcal{X}_1)\mathcal{A} & (\alpha \mathcal{X}_1)\mathcal{B}_1 & \alpha \mathcal{C}_1^T \\
\mathcal{B}_1^T(\alpha \mathcal{X}_1) & -\alpha \gamma_1 I & \alpha \mathcal{D}_1^T \\
\alpha \mathcal{C}_1 & \alpha \mathcal{D}_1 & -\alpha \gamma_1 I
\end{pmatrix} < 0.
$$

If we multiply the last row and the last column of the second inequality with $\frac{1}{\alpha}$ (what is a congruence transformation) and if we introduce $\mathcal{Y}_1 := \alpha \mathcal{X}_1$, we arrive at the following equivalent version of the analysis inequality for the H_∞-norm constraint:

$$
\mathcal{Y}_1 > 0, \quad
\begin{pmatrix}
\mathcal{A}^T \mathcal{Y}_1 + \mathcal{Y}_1 \mathcal{A} & \mathcal{Y}_1 \mathcal{B}_1 & \mathcal{C}_1^T \\
\mathcal{B}_1^T \mathcal{Y}_1 & -\gamma_1 \alpha I & \mathcal{D}_1^T \\
\mathcal{C}_1 & \mathcal{D}_1 & -\gamma_1/\alpha I
\end{pmatrix} < 0.
$$

Performing mixed synthesis with this analysis inequality leads to optimal values of the mixed H_2/H_∞ problem that depend on α. Each of these values form an upper bound on the actual optimal value of the multi-objective problem such that the best bound is found by performing a line-search over $\alpha > 0$.

- Contrary to previous approaches to the mixed problem, the one presented here does not require identical input- or output-signals of the H_∞ or H_2 channel. In view of their interpretation (uncertainty for H_∞ and performance for H_2), such a restriction is, in general, very unnatural. However, due to this flexibility, it is even more crucial to suitably scale the Lyapunov matrices.

- We can incorporate with ease various other performance or robustness specifications (formulated in terms of linear matrix inequalities) on other channels. Under the constraint of using for all desired specifications the same Lyapunov matrix, the design of a mixed controller is straightforward. Hence, one could conceivably consider a mixture of H_∞, H_2, generalized H_2, and peak-to-peak upper bound requirements on more than one channel. In its flexibility and generality, this approach is unique; however, one should never forget the conservatism that is involved.

- Using the same Lyapunov function might appear less restrictive if viewing the resulting procedure as a *Lyapunov shaping technique*. Indeed, one can start with the most important specification to be imposed on the controller. This amounts to solving a single-objective problem without conservatism. Then one keeps the already achieved property as a constraint and systematically imposes other specifications on other channels of the system to exploit possible additional freedom that is left in designing the controller. Hence, the Lyapunov function is shaped to realize additional specifications.

- Finally, constraints that are not necessarily related to input- output-specifications can be incorporated as well. As a nice example we mention the possibility to place the eigenvalues of $\mathcal{A}$ into an arbitrary LMI region $\{z : \ Q + Pz + P^T \bar{z} < 0\}$. For that purpose one just has to include

$$
\begin{pmatrix}
p_{11} \boldsymbol{X}(v) + q_{11} \boldsymbol{A}(v) + q_{11} \boldsymbol{A}(v)^T & \cdots & p_{1k} \boldsymbol{X}(v) + q_{1k} \boldsymbol{A}(v) + q_{k1} \boldsymbol{A}(v)^T \\
\vdots & \ddots & \vdots \\
p_{k1} \boldsymbol{X}(v) + q_{k1} \boldsymbol{A}(v) + q_{1k} \boldsymbol{A}(v)^T & \cdots & p_{kk} \boldsymbol{X}(v) + q_{kk} \boldsymbol{A}(v) + q_{kk} \boldsymbol{A}(v)^T
\end{pmatrix} < 0
$$

in the set of synthesis LMI (see Chapter 2).

4.5 Elimination of Parameters

The general procedure to derive described in Section 4.2 leads to synthesis inequalities in the variables K, L, M, N and X, Y as well as some auxiliary variables. For specific problems it is often possible to eliminate some of these variables in order to reduce the computation time. For example, since K has the same size as A, eliminating K for a system with McMillan degree 20 would save 400 variables. In view of the fact that, in our experience, present-day solvers are practical for solving problems up to about one thousand variables, parameter elimination might be of paramount importance to be able to solve realistic design problems.

In general, one cannot eliminate any variable that appears in at least two synthesis inequalities. Hence, in mixed design problems, parameter elimination is typically only possible under specific circumstances. In single-objective design problems one has to distinguish various information structures. In output-feedback design problems, it is in general not possible to eliminate X, Y but it might be possible to eliminate some of the variables K, L, M, N if they only appear in one inequality. For example, in quadratic performance problems one can eliminate all the variables K, L, M, N. In state-feedback design, one can typically eliminate in addition X, and for estimation problems one can eliminate Y.

To understand which variables can be eliminated and how this is performed, we turn to a discussion of two topics that will be of relevance, namely the dualization of matrix inequalities and explicit solvability tests for specifically structured LMI's [10, 32].

4.5.1 Dualization

The synthesis inequalities for quadratic performance can be written in the form (4.2.9). The second inequality has the structure

$$\begin{pmatrix} I \\ M \end{pmatrix}^T \begin{pmatrix} Q & S \\ S' & R \end{pmatrix} \begin{pmatrix} I \\ M \end{pmatrix} < 0 \;\; \text{and} \;\; R \geqslant 0. \tag{4.5.1}$$

Let us re-formulate these conditions in geometric terms. For that purpose we abbreviate

$$P = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} \in \mathbb{R}^{(k+l) \times (k+l)}$$

and observe that (4.5.1) is nothing but

$$P < 0 \;\; \text{on} \;\; \text{im} \begin{pmatrix} I \\ M \end{pmatrix} \quad \text{and} \quad P \geqslant 0 \;\; \text{on} \;\; \text{im} \begin{pmatrix} 0 \\ I \end{pmatrix}.$$

Since the direct sum of $\text{im} \begin{pmatrix} I \\ M \end{pmatrix}$ and $\text{im} \begin{pmatrix} 0 \\ I \end{pmatrix}$ spans the whole $\mathbb{R}^{(k+l) \times (k+l)}$, we can apply the following dualization lemma if P is non-singular.

Lemma 4.4 (Dualization Lemma) *Let P be a non-singular symmetric matrix in $\mathbb{R}^{n \times n}$, and let $\mathcal{U}$, $\mathcal{V}$ be two complementary subspaces whose sum equals $\mathbb{R}^n$. Then*

$$x^T P x < 0 \;\; \text{for all} \;\; x \in \mathcal{U} \setminus \{0\} \quad \text{and} \quad x^T P x \geqslant 0 \;\; \text{for all} \;\; x \in \mathcal{V} \tag{4.5.2}$$

is equivalent to

$$x^T P^{-1} x > 0 \;\; \text{for all} \;\; x \in \mathcal{U}^\perp \setminus \{0\} \quad \text{and} \quad x^T P^{-1} x \leqslant 0 \;\; \text{for all} \;\; x \in \mathcal{V}^\perp. \tag{4.5.3}$$

Proof. Since $\mathcal{U} \oplus \mathcal{V} = \mathbb{R}^n$ is equivalent to $\mathcal{U}^\perp \oplus \mathcal{V}^\perp = \mathbb{R}^n$, it suffices to prove that (4.5.2) implies (4.5.3); the converse implication follows by symmetry. Let us assume that (4.5.2) is true. Moreover, let us assume that $\mathcal{U}$ and $\mathcal{V}$ have dimension k and l respectively. We infer from (4.5.2) that P has at least k negative eigenvalues and at least l non-negative eigenvalues. Since $k + l = n$ and since P is non-singular, we infer that P has exactly k negative and l positive eigenvalues. We first prove that P^{-1} is positive definite on $\mathcal{U}^\perp$. We assume, to the contrary, that there exists a vector $y \in \mathcal{U}^\perp \setminus \{0\}$ with $y^T P^{-1} y \geqslant 0$. Define the non-zero vector $z = P^{-1} y$. Then z is not contained in $\mathcal{U}$ since, otherwise, we would conclude from (4.5.2) on the one hand $z^T P z < 0$, and on the other hand $z \perp y = P z$ what implies $z^T P z = 0$. Therefore, the space $\mathcal{U}_e := \text{span}(z) + \mathcal{U}$ has dimension $k + 1$. Moreover, P is positive semi-definite on this space: for any $x \in \mathcal{U}$ we have

$$(z + x)^T P(z + x) = y^T P^{-1} y + y^T x + x^T y + x^T P x = y^T P^{-1} y + x^T P x \geqslant 0.$$

This implies that P has at least $k+1$ non-negative eigenvalues, a contradiction to the already established fact that P has exactly k positive eigenvalues and that 0 is not an eigenvalue of P.

Let us now prove that P^{-1} is negative semi-definite on $\mathcal{V}^\perp$. For that purpose we just observe that $P + \epsilon I$ satisfies

$$x^T (P + \epsilon I) x < 0 \;\; \text{for all} \;\; x \in \mathcal{U} \setminus \{0\} \quad \text{and} \quad x^T (P + \epsilon I) x > 0 \;\; \text{for all} \;\; x \in \mathcal{V} \setminus \{0\}$$

for all small $\epsilon > 0$. Due to what has been already proved, this implies

$$x^T (P + \epsilon I)^{-1} x > 0 \;\text{for all}\; x \in \mathcal{U}^\perp \setminus \{0\} \quad \text{and} \quad x^T (P + \epsilon I)^{-1} x < 0 \;\text{for all}\; x \in \mathcal{V}^\perp \setminus \{0\}$$

for all small ϵ. Since P is non-singular, $(P + \epsilon I)^{-1}$ converges to P^{-1} for $\epsilon \to 0$. After taking the limit, we end up with

$$x^T P^{-1} x \geqslant 0 \;\; \text{for all} \;\; x \in \mathcal{U}^\perp \setminus \{0\} \quad \text{and} \quad x^T P^{-1} x \leqslant 0 \;\; \text{for all} \;\; x \in \mathcal{V}^\perp \setminus \{0\}.$$

Since we already know that the first inequality must be strict, the proof is finished. ∎

Let us hence introduce

$$P^{-1} = \begin{pmatrix} \tilde{Q} & \tilde{S} \\ \tilde{S}^T & \tilde{R} \end{pmatrix} \in \mathbb{R}^{(k+l) \times (k+l)}$$

and observe that

$$\operatorname{im}\begin{pmatrix} I \\ M \end{pmatrix}^{\perp} = \ker \begin{pmatrix} I & M^T \end{pmatrix} = \operatorname{im}\begin{pmatrix} -M^T \\ I \end{pmatrix} \quad \text{as well as} \quad \operatorname{im}\begin{pmatrix} 0 \\ I \end{pmatrix}^{\perp} = \operatorname{im}\begin{pmatrix} I \\ 0 \end{pmatrix}.$$

Hence Lemma 4.4 implies that (4.5.1) is equivalent to

$$\begin{pmatrix} -M^T \\ I \end{pmatrix}^T \begin{pmatrix} \tilde{Q} & \tilde{S} \\ \tilde{S}^T & \tilde{R} \end{pmatrix} \begin{pmatrix} -M^T \\ I \end{pmatrix} > 0 \quad \text{and} \quad \tilde{Q} \leqslant 0. \tag{4.5.4}$$

As an immediate consequence, we arrive at the following dual version of the quadratic performance synthesis inequalities.

Corollary 4.5 *Let* $P_j := \begin{pmatrix} Q_j & S_j \\ S_j^T & R_j \end{pmatrix}$ *be non-singular, and abbreviate* $P_j^{-1} := \begin{pmatrix} \tilde{Q}_j & \tilde{S}_j \\ \tilde{S}_j^T & \tilde{R}_j \end{pmatrix}$.
Then

$$\begin{pmatrix} I & 0 \\ \mathbf{A}(v) & \mathbf{B}_j(v) \\ \hline 0 & I \\ \mathbf{C}_j(v) & \mathbf{D}_j(v) \end{pmatrix}^T \begin{pmatrix} 0 & I & 0 & 0 \\ I & 0 & 0 & 0 \\ \hline 0 & 0 & Q_j & S_j \\ 0 & 0 & S_j^T & R_j \end{pmatrix} \begin{pmatrix} I & 0 \\ \mathbf{A}(v) & \mathbf{B}_j(v) \\ \hline 0 & I \\ \mathbf{C}_j(v) & \mathbf{D}_j(v) \end{pmatrix} < 0, \quad R_j \geqslant 0$$

is equivalent to

$$\begin{pmatrix} -\mathbf{A}(v)^T & -\mathbf{C}(v)^T \\ I & 0 \\ \hline -\mathbf{B}(v)^T & -\mathbf{D}(v)^T \\ 0 & I \end{pmatrix}^T \begin{pmatrix} 0 & I & 0 & 0 \\ I & 0 & 0 & 0 \\ \hline 0 & 0 & \tilde{Q}_j & \tilde{S}_j \\ 0 & 0 & \tilde{S}_j^T & \tilde{R}_j \end{pmatrix} \begin{pmatrix} -\mathbf{A}(v)^T & -\mathbf{C}(v)^T \\ I & 0 \\ \hline -\mathbf{B}(v)^T & -\mathbf{D}(v)^T \\ 0 & I \end{pmatrix} > 0, \quad \tilde{Q}_j \leqslant 0.$$

Remark. Any non-singular performance index $P_j = \begin{pmatrix} Q_j & S_j \\ S_j^T & R_j \end{pmatrix}$ can be inverted to $P_j^{-1} = \begin{pmatrix} \tilde{Q}_j & \tilde{S}_j \\ \tilde{S}_j^T & \tilde{R}_j \end{pmatrix}$. Recall that we required P_j to satisfy $R_j \geqslant 0$ since, otherwise, the synthesis inequalities may not be convex. The above discussion reveals that any non-singular performance index has to satisfy as well $\tilde{Q}_j \leqslant 0$ since, otherwise, we are sure that the synthesis inequalities are not feasible. We stress this point since, in general, $R_j \geqslant 0$ does not imply $\tilde{Q}_j \leqslant 0$. (Take e.g. $P_j > 0$ such that $P_j^{-1} > 0$.)

Similarly, we can dualize the H_2-type synthesis inequalities as formulated in (4.3.4)-(4.3.5).

Corollary 4.6 *For* $\gamma_j > 0$,

$$\begin{pmatrix} I & 0 \\ \mathbf{A}(v) & \mathbf{B}_j(v) \\ \hline 0 & I \end{pmatrix}^T \begin{pmatrix} 0 & I & 0 \\ I & 0 & 0 \\ \hline 0 & 0 & -\gamma_j I \end{pmatrix} \begin{pmatrix} I & 0 \\ \mathbf{A}(v) & \mathbf{B}_j(v) \\ \hline 0 & I \end{pmatrix} < 0$$

if and only if

$$\left(\frac{\begin{matrix} -\boldsymbol{A}(v)^T \\ -\boldsymbol{B}_j(v)^T \end{matrix}}{I} \right)^T \left(\begin{array}{cc|c} 0 & I & 0 \\ I & 0 & 0 \\ \hline 0 & 0 & -\frac{1}{\gamma_j}I \end{array} \right) \left(\frac{\begin{matrix} -\boldsymbol{A}(v)^T \\ -\boldsymbol{B}_j(v)^T \end{matrix}}{I} \right) > 0.$$

For $\boldsymbol{X}(v) > 0$ and $Z_j > 0$,

$$\left(\begin{matrix} I \\ \boldsymbol{C}_j(v) \end{matrix} \right)^T \left(\begin{matrix} -\boldsymbol{X}(v) & 0 \\ 0 & Z_j^{-1} \end{matrix} \right) \left(\begin{matrix} I \\ \boldsymbol{C}_j(v) \end{matrix} \right) < 0$$

if and only if

$$\left(\begin{matrix} -\boldsymbol{C}_j(v)^T \\ I \end{matrix} \right)^T \left(\begin{matrix} -\boldsymbol{X}(v)^{-1} & 0 \\ 0 & Z_j \end{matrix} \right) \left(\begin{matrix} -\boldsymbol{C}_j(v)^T \\ I \end{matrix} \right) > 0.$$

Again, Lemma 4.1 allows to render the first and the second dual inequalities affine in γ_j and $\boldsymbol{X}(v)$ respectively.

4.5.2   Special Linear Matrix Inequalities

Let us now turn to specific linear matrix inequalities for which one can easily derive explicit solvability tests.

We start by a trivial example that is cited for later reference.

Lemma 4.7 *The inequality*

$$\left(\begin{matrix} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} + X & P_{23} \\ P_{31} & P_{32} & P_{33} \end{matrix} \right) < 0$$

in the symmetric unknown X has a solution if and only if

$$\left(\begin{matrix} P_{11} & P_{13} \\ P_{31} & P_{33} \end{matrix} \right) < 0.$$

Proof. The direction 'only if' is obvious by cancelling the second row/column. To prove the converse implication, we just need to observe that any X with

$$X < -P_{22} + \left(\begin{matrix} P_{11} & P_{13} \end{matrix} \right) \left(\begin{matrix} P_{11} & P_{13} \\ P_{31} & P_{33} \end{matrix} \right)^{-1} \left(\begin{matrix} P_{12} \\ P_{32} \end{matrix} \right) < 0$$

(such as $X = -\alpha I$ for sufficiently large $\alpha > 0$) is a solution (Schur).   ■

Remark. This result extends to finding a common solution to a whole system of LMI's, due to the following simple fact: For finitely matrices $Q_1, \ldots, Q_m$, there exists an X with $X < Q_j$, $j = 1, \ldots, m$.

The first of three more advanced results in this vain is just a simple consequence of a Schur complement argument and it can be viewed as a powerful variant of what is often called the technique of 'completing the squares'.

Lemma 4.8 (Projection Lemma) *Let P be a symmetric matrix partitioned into three rows/columns and consider the LMI*

$$\begin{pmatrix} P_{11} & P_{12} + X^T & P_{13} \\ P_{21} + X & P_{22} & P_{23} \\ P_{31} & P_{32} & P_{33} \end{pmatrix} < 0 \qquad (4.5.5)$$

in the unstructured matrix X. There exists a solution X of this LMI iff

$$\begin{pmatrix} P_{11} & P_{13} \\ P_{31} & P_{33} \end{pmatrix} < 0 \quad and \quad \begin{pmatrix} P_{22} & P_{23} \\ P_{32} & P_{33} \end{pmatrix} < 0. \qquad (4.5.6)$$

If (4.5.6) hold, one particular solution is given by

$$X = P_{32}^T P_{33}^{-1} P_{31} - P_{21}. \qquad (4.5.7)$$

Proof. If (4.5.5) has a solution then (4.5.6) just follow from (4.5.5) by canceling the first or second block row/column.

Now suppose that (4.5.6) holds what implies $P_{33} < 0$. We observe that (4.5.5) is equivalent to (Schur complement)

$$\begin{pmatrix} P_{11} & P_{12} + X^T \\ P_{21} + X & P_{22} \end{pmatrix} - \begin{pmatrix} P_{13} \\ P_{23} \end{pmatrix} P_{33}^{-1} \begin{pmatrix} P_{31} & P_{32} \end{pmatrix} < 0.$$

Due to (4.5.6), the diagonal blocks are negative definite. X defined in (4.5.7) just renders the off-diagonal block zero such that it is a solution of the latter matrix inequality. ∎

An even more powerful generalization is the so-called projection lemma.

Lemma 4.9 (Projection Lemma) *For arbitrary A, B and a symmetric P, the LMI*

$$A^T X B + B^T X^T A + P < 0 \qquad (4.5.8)$$

in the unstructured X has a solution if and only if

$$Ax = 0 \ or \ Bx = 0 \quad imply \quad x^T P x < 0 \ or \ x = 0. \qquad (4.5.9)$$

If $A_\perp$ and $B_\perp$ denote arbitrary matrices whose columns form a basis of $\ker(A)$ and $\ker(B)$ respectively, (4.5.9) is equivalent to

$$A_\perp^T P A_\perp < 0 \quad and \quad B_\perp^T P B_\perp < 0. \qquad (4.5.10)$$

We give a full proof of the Projection Lemma since it provides a scheme for constructing a solution X if it exists. It also reveals that, in suitable coordinates, Lemma 4.9 reduces to Lemma 4.8 if the kernels of A and B together span the whole space.

Proof. The proof of 'only if' is trivial. Indeed, let us assume that there exists some X with $A^T X B + B^T X^T A + P < 0$. Then $Ax = 0$ or $Bx = 0$ with $x \neq 0$ imply the desired inequality $0 > x^T(A^T X B + B^T X^T A + P)x = x^T P x$.

For proving 'if', let $S = (S_1\ S_2\ S_3\ S_4)$ be a nonsingular matrix such that the columns of S_3 span $\ker(A) \cap \ker(B)$, those of $(S_1\ S_3)$ span $\ker(A)$, and those of $(S_2\ S_3)$ span $\ker(B)$. Instead of (4.5.8), we consider the equivalent inequality $S^T(4.5.8)S < 0$ which reads as

$$(AS)^T X(BS) + (BS)^T X^T(AS) + S^T P S < 0. \tag{4.5.11}$$

Now note that AS and BS have the structure $(0\ A_2\ 0\ A_4)$ and $(B_1\ 0\ 0\ B_4)$ where $(A_2\ A_4)$ and $(B_1\ B_4)$ have full column rank respectively. The rank properties imply that the equation

$$(AS)^T X(BS) = \begin{pmatrix} 0 \\ A_2^T \\ 0 \\ A_4^T \end{pmatrix} X \begin{pmatrix} B_1 & 0 & 0 & B_4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ Z_{21} & 0 & 0 & Z_{24} \\ 0 & 0 & 0 & 0 \\ Z_{41} & 0 & 0 & Z_{44} \end{pmatrix}$$

has a solution X for arbitrary $Z_{21}, Z_{24}, Z_{41}, Z_{44}$. With $Q := S^T P S$ partitioned accordingly, (4.5.11) hence reads as

$$\left(\begin{array}{ccc|c} Q_{11} & Q_{12} + Z_{21}^T & Q_{13} & Q_{14} + Z_{41}^T \\ Q_{21} + Z_{21} & Q_{22} & Q_{23} & Q_{24} + Z_{24} \\ Q_{31} & Q_{32} & Q_{33} & Q_{34} \\ \hline Q_{41} + Z_{41} & Q_{42} + Z_{24}^T & Q_{43} & Q_{44} + Z_{44} + Z_{44}^T \end{array} \right) < 0 \tag{4.5.12}$$

with *free* blocks $Z_{21}, Z_{24}, Z_{41}, Z_{44}$. Since

$$\ker(AS) = \mathrm{im}\begin{pmatrix} I & 0 \\ 0 & 0 \\ 0 & I \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad \ker(BS) = \mathrm{im}\begin{pmatrix} 0 & 0 \\ I & 0 \\ 0 & I \\ 0 & 0 \end{pmatrix},$$

the hypothesis (4.5.9) just amounts to the conditions

$$\begin{pmatrix} Q_{11} & Q_{13} \\ Q_{31} & Q_{33} \end{pmatrix} < 0 \quad \text{and} \quad \begin{pmatrix} Q_{22} & Q_{23} \\ Q_{32} & Q_{33} \end{pmatrix} < 0.$$

By Lemma 4.8, we can hence find a matrix Z_{21} which renders the marked 3×3 block in (4.5.12) negative definite. The blocks Z_{41} and Z_{24} can be taken arbitrary. After having fixed Z_{21}, Z_{41}, Z_{24}, we can choose Z_{44} according to Lemma 4.7 such that the whole matrix on the left-hand side of (4.5.12) is negative definite. ∎

Remark. We can, of course, replace $<$ everywhere by $>$. It is important to recall that the unknown X is unstructured. If one requires X to have a certain structure (such as being

symmetric), the tests, if existing at all, are much more complicated. There is, however, a generally valid extension of the Projection Lemma to block-triangular unknowns X [29]. Note that the results do not hold true as formulated if just replacing the strict inequalities by non-strict inequalities (as it is sometimes erroneously claimed in the literature)! Again, it is possible to provide a full generalization of the Projection Lemma to non-strict inequalities.

Let

$$P = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} \text{ with } R \geqslant 0 \quad \text{have the inverse} \quad P^{-1} = \begin{pmatrix} \tilde{Q} & \tilde{S} \\ \tilde{S}^T & \tilde{R} \end{pmatrix} \text{ with } \tilde{Q} \leqslant 0 \tag{4.5.13}$$

and let us finally consider the quadratic inequality

$$\begin{pmatrix} I \\ A^T X B + C \end{pmatrix}^T P \begin{pmatrix} I \\ A^T X B + C \end{pmatrix} < 0 \tag{4.5.14}$$

in the unstructured unknown X. According to Lemma 4.4, we can dualize this inequality to

$$\begin{pmatrix} -B^T X^T A - C' \\ I \end{pmatrix}^T P^{-1} \begin{pmatrix} -B^T X^T A - C' \\ I \end{pmatrix} > 0. \tag{4.5.15}$$

It is pretty straightforward to derive necessary conditions for the solvability of (4.5.14). Indeed, let us assume that (4.5.14) holds for some X. If $A_\perp$ and $B_\perp$ denote basis matrices of $\ker(A)$ and $\ker(B)$ respectively, we infer

$$\begin{pmatrix} I \\ A^T X B + C \end{pmatrix} B_\perp = \begin{pmatrix} I \\ C \end{pmatrix} B_\perp \quad \text{and} \quad \begin{pmatrix} -B^T X^T A - C' \\ I \end{pmatrix} A_\perp = \begin{pmatrix} -C' \\ I \end{pmatrix} A_\perp.$$

Since $B_\perp^T (4.5.14) B_\perp < 0$ and $A_\perp^T (4.5.15) A_\perp > 0$, we arrive at the two easily verifiable inequalities

$$B_\perp^T \begin{pmatrix} I \\ C \end{pmatrix}^T P \begin{pmatrix} I \\ C \end{pmatrix} B_\perp < 0 \quad \text{and} \quad A_\perp^T \begin{pmatrix} -C^T \\ I \end{pmatrix}^T P^{-1} \begin{pmatrix} -C^T \\ I \end{pmatrix} A_\perp > 0 \tag{4.5.16}$$

which are necessary for a solution of (4.5.14) to exist. One can constructively prove that they are sufficient [37].

Lemma 4.10 (Elimination Lemma) *Under the hypotheses (4.5.13) on P, the inequality (4.5.14) has a solution if and only if (4.5.16) hold true.*

Proof. It remains to prove that (4.5.16) implies the existence of a solution of (4.5.14).

Let us first reveal that one can assume without loss of generality that $R > 0$ and $\tilde{Q} < 0$. For that purpose we need to have information about the inertia of P. Due to $R \geqslant 0$, P and P^{-1} have $\text{size}(R)$ positive eigenvalues (since none of the eigenvalues can vanish). Similarly,

$\tilde{Q} \leqslant 0$ implies that P^{-1} and P have size($\tilde{Q}$) = size(Q) negative eigenvalues. Let us now consider (4.5.14) with the perturbed data

$$P_\epsilon := \begin{pmatrix} Q & S \\ S^T & R + \epsilon I \end{pmatrix} \quad \text{where} \quad \epsilon > 0$$

is fixed sufficiently small such that (4.5.16) persist to hold for P_ϵ, and such that P_ϵ and P have the same number of positive and negative eigenvalues. Trivially, the right-lower block of P_ϵ is positive definite. The Schur complement $Q - S(R + \epsilon I)^{-1}S^T$ of this right-lower block must be negative definite since P_ϵ has size(Q) negative and size(R) positive eigenvalues. Hence the left-upper block of P_ϵ^{-1} which equals $[Q - S(R + \epsilon I)^{-1}S^T]^{-1}$ is negative definite as well. If the result is proved with $R > 0$ and $\tilde{Q} < 0$, we can conclude that (4.5.14) has a solution X for the perturbed data P_ϵ. Due to $P_0 \leqslant P_\epsilon$, the very same X also satisfies the original inequality for P_0.

Let us hence assume from now on $R > 0$ and $\tilde{Q} < 0$. We observe that the left-hand side of (4.5.14) equals

$$\begin{pmatrix} I \\ C \end{pmatrix}^T P \begin{pmatrix} I \\ C \end{pmatrix} + (A^T X B)^T (S^T + RC) + (S^T + RC)^T (A^T X B) + (A^T X B)^T R (A^T X B).$$

Hence (4.5.14) is equivalent to (Schur)

$$\left(\begin{pmatrix} I \\ C \end{pmatrix}^T P \begin{pmatrix} I \\ C \end{pmatrix} + (A^T X B)^T (S^T + RC) + (S^T + RC)^T (A^T X B) \quad (A^T X B)^T \\ (A^T X B) \quad -R^{-1} \right) < 0$$

or

$$\left(\begin{pmatrix} I \\ C \end{pmatrix}^T P \begin{pmatrix} I \\ C \end{pmatrix} \quad 0 \\ 0 \quad -R^{-1} \right) +$$

$$+ \begin{pmatrix} A(S^T + RC)^T \\ A \end{pmatrix}^T X \begin{pmatrix} B & 0 \end{pmatrix} + \begin{pmatrix} B^T \\ 0 \end{pmatrix} X^T \begin{pmatrix} A(S^T + RC) & A \end{pmatrix} < 0. \quad (4.5.17)$$

The inequality (4.5.17) has the structure as required in the Projection Lemma. We need to show that

$$\begin{pmatrix} B & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0, \quad \begin{pmatrix} x \\ y \end{pmatrix} \neq 0 \qquad (4.5.18)$$

or

$$\begin{pmatrix} A(S^T + RC) & A \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = 0, \quad \begin{pmatrix} x \\ y \end{pmatrix} \neq 0 \qquad (4.5.19)$$

imply

$$\begin{pmatrix} x \\ y \end{pmatrix}^T \left(\begin{pmatrix} I \\ C \end{pmatrix}^T P \begin{pmatrix} I \\ C \end{pmatrix} \quad 0 \\ 0 \quad -I \right) \begin{pmatrix} x \\ y \end{pmatrix} = x^T \begin{pmatrix} I \\ C \end{pmatrix}^T P \begin{pmatrix} I \\ C \end{pmatrix} x - y^T y < 0. \quad (4.5.20)$$

In a first step we show that (4.5.17) and hence (4.5.14) have a solution if $A = I$. Let us assume (4.5.18). Then (4.5.20) is trivial if $x = 0$. For $x \neq 0$ we infer $Bx = 0$ and the first inequality in (4.5.16) implies

$$ x^T \begin{pmatrix} I \\ C \end{pmatrix}^T P \begin{pmatrix} I \\ C \end{pmatrix} x < 0 $$

what shows that (4.5.20) is true. Let us now assume (4.5.19) with $A = I$. We infer $x \neq 0$ and $y = -(S^T + RC)x$. The left-hand side of (4.5.20) is nothing but

$$ x^T \begin{pmatrix} I \\ C \end{pmatrix}^T P \begin{pmatrix} I \\ C \end{pmatrix} x - x^T (S^T + RC)^T R^{-1} (S^T + RC) x = $$

$$ = x^T \begin{pmatrix} I \\ C \end{pmatrix}^T P \begin{pmatrix} I \\ C \end{pmatrix} x - x^T \begin{pmatrix} I \\ C \end{pmatrix}^T \begin{pmatrix} S \\ R \end{pmatrix} R^{-1} \begin{pmatrix} S^T & R \end{pmatrix} \begin{pmatrix} I \\ C \end{pmatrix} x = $$

$$ = x^T \begin{pmatrix} I \\ C \end{pmatrix}^T \left[P - \begin{pmatrix} SR^{-1}S^T & S \\ S^T & R \end{pmatrix} \right] \begin{pmatrix} I \\ C \end{pmatrix} x = x^T (Q - SR^{-1}S^T) x $$

what is indeed negative since $\tilde{Q}^{-1} = Q - SR^{-1}S^T < 0$ and $x \neq 0$. We conclude that, for $A = I$, (4.5.17) and hence

$$ \begin{pmatrix} I \\ XB + C \end{pmatrix}^T P \begin{pmatrix} I \\ XB + C \end{pmatrix} < 0 $$

have a solution.

By symmetry - since one can apply the arguments provided above to the dual inequality (4.5.15) - we can infer that

$$ \begin{pmatrix} I \\ A^T X + C \end{pmatrix}^T P \begin{pmatrix} I \\ A^T X + C \end{pmatrix} < 0 $$

has a solution X. This implies that (4.5.17) has a solution for $B = I$. Therefore, with the Projection Lemma, (4.5.19) implies (4.5.20) for a general A.

In summary, we have proved for general A and B that (4.5.18) or (4.5.19) imply (4.5.20). We can infer the solvability of (4.5.17) or that of (4.5.14). ∎

4.5.3 The Quadratic Performance Problem

For the performance index

$$ P_j = \begin{pmatrix} Q_j & S_j \\ S_j' & R_j \end{pmatrix}, \quad R_j \geqslant 0 \quad \text{with inverse} \quad P_j^{-1} = \begin{pmatrix} \tilde{Q}_j & \tilde{S}_j \\ \tilde{S}_j' & \tilde{R}_j \end{pmatrix}, \quad \tilde{Q}_j \leqslant 0, \qquad (4.5.21) $$

we have derived the following synthesis inequalities:

$$
\boldsymbol{X}(v) > 0, \quad
\begin{pmatrix} I & 0 \\ \boldsymbol{A}(v) & \boldsymbol{B}_j(v) \\ 0 & I \\ \boldsymbol{C}_j(v) & \boldsymbol{D}_j(v) \end{pmatrix}^T
\left(\begin{array}{cc|cc} 0 & I & 0 & 0 \\ I & 0 & 0 & 0 \\ \hline 0 & 0 & Q_j & S_j \\ 0 & 0 & S_j^T & R_j \end{array} \right)
\begin{pmatrix} I & 0 \\ \boldsymbol{A}(v) & \boldsymbol{B}_j(v) \\ 0 & I \\ \boldsymbol{C}_j(v) & \boldsymbol{D}_j(v) \end{pmatrix}
< 0. \qquad (4.5.22)
$$

Due to the specific structure

$$
\left(\begin{array}{c|c} \boldsymbol{A}(v) & \boldsymbol{B}_j(v) \\ \hline \boldsymbol{C}_j(v) & \boldsymbol{D}_j(v) \end{array} \right) =
\left(\begin{array}{cc|c} AY & A & B_j \\ 0 & XA & XB_j \\ \hline C_jY & C_j & D_j \end{array} \right) +
\begin{pmatrix} 0 & B \\ I & 0 \\ 0 & E_j \end{pmatrix}
\begin{pmatrix} K & L \\ M & N \end{pmatrix}
\left(\begin{array}{cc|c} I & 0 & 0 \\ 0 & C & F_j \end{array} \right),
$$

$$
(4.5.23)
$$

it is straightforward to apply Lemma 4.10 to eliminate all the variables $\begin{pmatrix} K & L \\ M & N \end{pmatrix}$. For

that purpose it suffices to compute basis matrices

$$
\Phi_j = \begin{pmatrix} \Phi_j^1 \\ \Phi_j^2 \end{pmatrix} \quad \text{of} \ \ \ker \begin{pmatrix} B^T & E_j^T \end{pmatrix}
\quad \text{and} \quad \Psi_j = \begin{pmatrix} \Psi_j^1 \\ \Psi_j^2 \end{pmatrix} \quad \text{of} \ \ \ker \begin{pmatrix} C & F_j \end{pmatrix}.
$$

Corollary 4.11 *For a performance index with (4.5.21), there exists a solution v of (4.5.22) if and only if there exist symmetric X and Y that satisfy*

$$
\begin{pmatrix} Y & I \\ I & X \end{pmatrix} > 0, \qquad (4.5.24)
$$

$$
\Psi^T
\begin{pmatrix} I & 0 \\ A & B_j \\ 0 & I \\ C_j & D_j \end{pmatrix}^T
\left(\begin{array}{cc|cc} 0 & X & 0 & 0 \\ X & 0 & 0 & 0 \\ \hline 0 & 0 & Q_j & S_j \\ 0 & 0 & S_j^T & R_j \end{array} \right)
\begin{pmatrix} I & 0 \\ A & B_j \\ 0 & I \\ C_j & D_j \end{pmatrix}
\Psi < 0, \qquad (4.5.25)
$$

$$
\Phi^T
\begin{pmatrix} -A^T & -C_j^T \\ I & 0 \\ -B_j^T & -D_j^T \\ 0 & I \end{pmatrix}^T
\left(\begin{array}{cc|cc} 0 & Y & 0 & 0 \\ Y & 0 & 0 & 0 \\ \hline 0 & 0 & \tilde{Q}_j & \tilde{S}_j \\ 0 & 0 & \tilde{S}_j^T & \tilde{R}_j \end{array} \right)
\begin{pmatrix} -A^T & -C_j^T \\ I & 0 \\ -B_j^T & -D_j^T \\ 0 & I \end{pmatrix}
\Phi > 0. \qquad (4.5.26)
$$

Remark. Note that the columns of $\begin{pmatrix} B \\ E_j \end{pmatrix}$ indicate in how far the right-hand side of
(4.1.1) can be modified by control, and the rows of $\begin{pmatrix} C & F_j \end{pmatrix}$ determine those functionals
that provide information about the system state and the disturbance that is available for
control. Roughly speaking, the columns of Φ_j or of Ψ_j indicate what *cannot* be influenced
by control or which information *cannot* be extracted from the measured output. Let us
hence compare (4.5.24)-(4.5.26) with the synthesis inequalities that would be obtained for

$$
\begin{pmatrix} \dot{x} \\ z_1 \\ \vdots \\ z_q \end{pmatrix} =
\left(\begin{array}{c|ccc} A & B_1 & \cdots & B_q \\ \hline C_1 & D_1 & \cdots & D_{1q} \\ \vdots & \vdots & \ddots & \vdots \\ C_q & D_{q1} & \cdots & D_q \end{array} \right)
\begin{pmatrix} x \\ w_1 \\ \vdots \\ w_q \end{pmatrix}
\qquad (4.5.27)
$$

without control input and measurement output. For this system we could choose $\Phi = I$ and $\Psi = I$ to arrive at the synthesis inequalities

$$\begin{pmatrix} Y & I \\ I & X \end{pmatrix} > 0, \tag{4.5.28}$$

$$\begin{pmatrix} I & 0 \\ A & B_j \\ \hline 0 & I \\ C_j & D_j \end{pmatrix}^T \begin{pmatrix} 0 & X & 0 & 0 \\ X & 0 & 0 & 0 \\ \hline 0 & 0 & Q_j & S_j \\ 0 & 0 & S_j^T & R_j \end{pmatrix} \begin{pmatrix} I & 0 \\ A & B_j \\ \hline 0 & I \\ C_j & D_j \end{pmatrix} < 0, \tag{4.5.29}$$

$$\begin{pmatrix} -A^T & -C_j^T \\ I & 0 \\ \hline -B_j^T & -D_j^T \\ 0 & I \end{pmatrix}^T \begin{pmatrix} 0 & Y & 0 & 0 \\ Y & 0 & 0 & 0 \\ \hline 0 & 0 & \tilde{Q}_j & \tilde{S}_j \\ 0 & 0 & \tilde{S}_j^T & \tilde{R}_j \end{pmatrix} \begin{pmatrix} -A^T & -C_j^T \\ I & 0 \\ \hline -B_j^T & -D_j^T \\ 0 & I \end{pmatrix} > 0. \tag{4.5.30}$$

Since there is no control and no measured output, these could be viewed as analysis inequalities for (4.5.27). Hence we have very nicely displayed in how far controls or measurements do influence the synthesis inequalities through Φ_j and Ψ_j. Finally, we note that (4.5.28)-(4.5.30) are equivalent to $X > 0$, (4.5.29) or to $Y > 0$, (4.5.30). Moreover, if dualizing $X > 0$, (4.5.29), we arrive at $Y > 0$, (4.5.30) for $Y := X^{-1}$.

Proof of Corollary 4.11. The first inequality (4.5.24) is just $\boldsymbol{X}(v) > 0$. The inequalities (4.5.25)-(4.5.26) are obtained by simply applying Lemma 4.10 to the second inequality of (4.5.22), viewed as a quadratic matrix inequality in the unknowns $\begin{pmatrix} K & L \\ M & N \end{pmatrix}$. For that purpose we first observe that

$$\ker \begin{pmatrix} 0 & I & 0 \\ B^T & 0 & E_j^T \end{pmatrix}, \quad \ker \begin{pmatrix} I & 0 & 0 \\ 0 & C & F_j \end{pmatrix} \quad \text{have the basis matrices} \quad \begin{pmatrix} \Phi_j^1 \\ 0 \\ \Phi_j^2 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ \Psi_j^1 \\ \Psi_j^2 \end{pmatrix}$$

respectively. Due to

$$\begin{pmatrix} I & 0 \\ \boldsymbol{A}(v) & \boldsymbol{B}_j(v) \\ \hline 0 & I \\ \boldsymbol{C}_j(v) & \boldsymbol{D}_j(v) \end{pmatrix} \begin{pmatrix} 0 \\ \Psi_j^1 \\ \Psi_j^2 \end{pmatrix} = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ AY & A & B_j \\ 0 & XA & XB_j \\ \hline 0 & 0 & I \\ C_jY & C_j & D_j \end{pmatrix} \begin{pmatrix} 0 \\ \Psi_j^1 \\ \Psi_j^2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ I & 0 \\ A & B_j \\ XA & XB_j \\ \hline 0 & I \\ C_j & D_j \end{pmatrix} \Psi,$$

the solvability condition that corresponds to the first inequality in (4.5.16) reads as

$$\Psi^T \begin{pmatrix} 0 & 0 \\ I & 0 \\ A & B_j \\ XA & XB_j \\ \hline 0 & I \\ C_j & D_j \end{pmatrix}^T \begin{pmatrix} 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & Q_j & S_j \\ 0 & 0 & 0 & 0 & S_j^T & R_j \end{pmatrix} \begin{pmatrix} 0 & 0 \\ I & 0 \\ A & B_j \\ XA & XB_j \\ \hline 0 & I \\ C_j & D_j \end{pmatrix} \Psi < 0$$

what simplifies to

$$\Psi^T \begin{pmatrix} I & 0 \\ XA & XB_j \\ \hline 0 & I \\ C_j & D_j \end{pmatrix}^T \left(\begin{array}{cc|cc} 0 & I & 0 & 0 \\ I & 0 & 0 & 0 \\ \hline 0 & 0 & Q_j & S_j \\ 0 & 0 & S_j^T & R_j \end{array} \right) \begin{pmatrix} I & 0 \\ XA & XB_j \\ \hline 0 & I \\ C_j & D_j \end{pmatrix} \Psi < 0.$$

This is clearly nothing but (4.5.25). The very same simple steps lead to (4.5.26). Indeed, we have

$$\begin{pmatrix} \boldsymbol{A}(v)^T & \boldsymbol{C}_j(v)^T \\ I & 0 \\ \hline \boldsymbol{B}_j(v)^T & \boldsymbol{D}_j(v)^T \\ 0 & I \end{pmatrix} \begin{pmatrix} \Phi_j^1 \\ 0 \\ \hline \Phi_j^2 \end{pmatrix} =$$

$$= \left(\begin{array}{cc|c} -YA^T & 0 & -YC_j^T \\ -A^T & -A^T X & -C_j^T \\ I & 0 & 0 \\ 0 & I & 0 \\ \hline -B_j^T & -XB_j^T & -D_j^T \\ 0 & 0 & I \end{array} \right) \begin{pmatrix} \Phi_j^1 \\ 0 \\ \hline \Phi_j^2 \end{pmatrix} = \left(\begin{array}{cc} -YA^T & -YC_1^T \\ -A^T & -C_1^T \\ I & 0 \\ 0 & 0 \\ \hline -B_1^T & -D_1^T \\ 0 & I \end{array} \right) \Phi$$

such that the solvability condition that corresponds to the second inequality in (4.5.16) is

$$\Phi^T \left(\begin{array}{cc} -YA^T & -YC_1^T \\ -A^T & -C_1^T \\ I & 0 \\ 0 & 0 \\ \hline -B_1^T & -D_1^T \\ 0 & I \end{array} \right)^T \left(\begin{array}{cccc|cc} 0 & 0 & I & 0 & 0 & 0 \\ 0 & 0 & 0 & I & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & \tilde{Q}_j & \tilde{S}_j \\ 0 & 0 & 0 & 0 & \tilde{S}_j^T & \tilde{R}_j \end{array} \right) \left(\begin{array}{cc} -YA^T & -YC_1^T \\ -A^T & -C_1^T \\ I & 0 \\ 0 & 0 \\ \hline -B_1^T & -D_1^T \\ 0 & I \end{array} \right) \Phi < 0$$

what simplifies to

$$\Phi^T \left(\begin{array}{cc} -YA^T & -YC_1^T \\ I & 0 \\ \hline -B_1^T & -D_1^T \\ 0 & I \end{array} \right)^T \left(\begin{array}{cc|cc} 0 & I & 0 & 0 \\ I & 0 & 0 & 0 \\ \hline 0 & 0 & \tilde{Q}_j & \tilde{S}_j \\ 0 & 0 & \tilde{S}_j^T & \tilde{R}_j \end{array} \right) \left(\begin{array}{cc} -YA^T & -YC_1^T \\ I & 0 \\ \hline -B_1^T & -D_1^T \\ 0 & I \end{array} \right) \Phi < 0$$

and we arrive at (4.5.26). ∎

Starting from the synthesis inequalities (4.5.22) in the variables X, Y, $\begin{pmatrix} K & L \\ M & N \end{pmatrix}$, we have derived the equivalent inequalities (4.5.24)-(4.5.26) in the variables X, Y only. Testing feasibility of these latter inequalities can hence be accomplished much faster. This is particularly advantageous if optimizing an additional parameter, such as minimizing the sup-optimality level γ in the H_∞ problem.

To conclude this section, let us comment on how to compute the controller after having found solutions X, Y of (4.5.24)-(4.5.26). One possibility is to explicitly solve the quadratic inequality (4.5.22) in $\begin{pmatrix} K & L \\ M & N \end{pmatrix}$ along the lines of the proof of Lemma 4.10, and reconstruct the controller parameters as earlier. One could as well proceed directly: Starting from X and Y, we can compute non-singular U and V with $UV^T = I - XY$, and determine $\mathcal{X} > 0$ by solving the first equation in (4.2.10). Due to (4.1.5), we can apply Lemma 4.10 directly to the analysis inequality

$$\begin{pmatrix} I & 0 \\ \mathcal{A} & \mathcal{B}_j \\ 0 & I \\ \mathcal{C}_j & \mathcal{D}_j \end{pmatrix}^T \left(\begin{array}{cc|cc} 0 & \mathcal{X} & 0 & 0 \\ \mathcal{X} & 0 & 0 & 0 \\ \hline 0 & 0 & Q_j & S_j \\ 0 & 0 & S_j^T & R_j \end{array} \right) \begin{pmatrix} I & 0 \\ \mathcal{A} & \mathcal{B}_j \\ 0 & I \\ \mathcal{C}_j & \mathcal{D}_j \end{pmatrix} < 0$$

if viewing $\begin{pmatrix} A_c & B_c \\ C_c & D_c \end{pmatrix}$ as variables. It is not difficult (and you should provide the details!) to verify the solvability conditions for this quadratic inequality, and to construct an explicit solution along the lines of the proof of Lemma 4.10. Alternatively, one can transform the quadratic inequality to a linear matrix inequality with Lemma 4.1, and apply the Projection Lemma to reconstruct the controller parameters. For the latter step the LMI-Lab offers a standard routine. We conclude that there are many basically equivalent alternative ways to compute a controller once one has determined X and Y.

4.5.4 H_2-Problems

If recalling (4.2.3), we observe that both inequalities in the H_2-synthesis conditions (4.3.3) involve the variables M and N, but only the first one

$$\begin{pmatrix} \boldsymbol{A}(v)^T + \boldsymbol{A}(v) & \boldsymbol{B}_j(v) \\ \boldsymbol{B}_j(v)^T & -\gamma_j I \end{pmatrix} < 0 \tag{4.5.31}$$

is affected by K and L. This might suggest that the latter two variables can be eliminated in the synthesis conditions. Since (4.5.31) is affine in $\begin{pmatrix} K & L \end{pmatrix}$, we can indeed apply the Projection Lemma to eliminate these variables. It is not difficult to arrive at the following alternative synthesis conditions for H_2-type criteria.

Corollary 4.12 *There exists a controller that renders (4.3.2) for some $\mathcal{X}$, Z_j satisfied iff there exist X, Y, M, N, Z_j with $f_j(Z_j) < \gamma_j$, $D_j + E_j N F_j = 0$ and*

$$\begin{pmatrix} Y & I & (C_j Y + E_j M)^T \\ I & X & (C_j + E_j N C)^T \\ C_j Y + E_j M & C_j + E_j N C & Z_j \end{pmatrix} > 0,$$

$$\Psi^T \begin{pmatrix} A^T X + XA & XB_j \\ B_j^T X & -\gamma_j I \end{pmatrix} \Psi < 0, \quad \begin{pmatrix} (AY + BM) + (AY + BM)^T & B_j + BNF_j \\ (B_j + BNF_j)^T & -\gamma_j I \end{pmatrix} < 0. \tag{4.5.32}$$

Proof. We only need to show that the elimination of K and L in (4.5.31) leads to the two inequalities (4.5.32). Let us recall

$$\left(\ \boldsymbol{A}(v)\ \ \boldsymbol{B}_j(v)\ \right) = \left(\begin{array}{cc|c} AY & A & B_j \\ 0 & XA & XB_j \end{array}\right) + \left(\begin{array}{cc} 0 & B \\ I & 0 \end{array}\right)\left(\begin{array}{cc} K & L \\ M & N \end{array}\right)\left(\begin{array}{cc|c} I & 0 & 0 \\ 0 & C & F_j \end{array}\right) =$$

$$= \left(\begin{array}{cc|c} AY+BM & A+BNC & B_j+BNF_j \\ 0 & XA & XB_j \end{array}\right) + \left(\begin{array}{c} 0 \\ I \end{array}\right)\left(\ K\ \ L\ \right)\left(\begin{array}{cc|c} I & 0 & 0 \\ 0 & C & F_j \end{array}\right).$$

Therefore, (4.5.31) is equivalent to

$$\left(\begin{array}{cc|c} AY+YA^T & A & B_j \\ A^T & A^TX+XA & XB_j \\ \hline B_j^T & B_j^TX & -\gamma_j I \end{array}\right) + \mathrm{sym}\left(\left(\begin{array}{c} B \\ 0 \\ 0 \end{array}\right)\left(\ M\ \ N\ \right)\left(\begin{array}{cc|c} I & 0 & 0 \\ 0 & C & F_j \end{array}\right)\right) +$$

$$+ \mathrm{sym}\left(\left(\begin{array}{c} 0 \\ I \\ 0 \end{array}\right)\left(\ K\ \ L\ \right)\left(\begin{array}{cc|c} I & 0 & 0 \\ 0 & C & F_j \end{array}\right)\right) < 0$$

where $\mathrm{sym}\,(M) := M + M^T$ is just an abbreviation to shorten the formulas. Now note that

$$\ker\left(\ 0\ \ I\ |\ 0\ \right), \quad \ker\left(\begin{array}{cc|c} I & 0 & 0 \\ 0 & C & F_j \end{array}\right) \quad \text{have the basis matrices} \quad \left(\begin{array}{cc} I & 0 \\ 0 & 0 \\ 0 & I \end{array}\right), \quad \left(\begin{array}{c} 0 \\ \Psi_j^1 \\ \Psi_j^2 \end{array}\right)$$

respectively. Therefore, the Projection Lemma leads to the two inequalities

$$\left(\begin{array}{c} 0 \\ \Psi_j^1 \\ \Psi_j^2 \end{array}\right)^T \left(\begin{array}{cc|c} AY+YA^T & A & B_j \\ A^T & A^TX+XA & XB_j \\ \hline B_j^T & B_j^TX & -\gamma_j I \end{array}\right)\left(\begin{array}{c} 0 \\ \Psi_j^1 \\ \Psi_j^2 \end{array}\right) < 0$$

and

$$\left(\begin{array}{cc} AY+YA^T & B_j \\ B_j^T & -\gamma_j I \end{array}\right) + \mathrm{sym}\left(\left(\begin{array}{c} B \\ 0 \end{array}\right)\left(\ M\ \ N\ \right)\left(\begin{array}{cc} I & 0 \\ 0 & F_j \end{array}\right)\right) < 0$$

that are easily rewritten to (4.5.32). ∎

If it happens that E_j vanishes, we can also eliminate all variables $\left(\begin{array}{cc} K & L \\ M & N \end{array}\right)$ from the synthesis inequalities. The corresponding results are obtained in a straightforward fashion and their proof is left as an exercise.

Corollary 4.13 *Suppose that $E_j = 0$. Then there exists a controller that renders (4.3.2) for some $\mathcal{X}$, Z_j satisfied iff $D_j = 0$ and there exist X, Y Z_j with $f_j(Z_j) < \gamma_j$ and*

$$\left(\begin{array}{ccc} Y & I & (C_jY)^T \\ I & X & C_j^T \\ C_jY & C_j & Z_j \end{array}\right) > 0,$$

$$\Psi^T\left(\begin{array}{cc} A^TX+XA & XB_j \\ B_j^TX & -\gamma_j I \end{array}\right)\Psi < 0, \quad \left(\begin{array}{cc} \widehat{\Phi} & 0 \\ 0 & I \end{array}\right)^T\left(\begin{array}{cc} AY+YA^T & B_j \\ B_j^T & -\gamma_j I \end{array}\right)\left(\begin{array}{cc} \widehat{\Phi} & 0 \\ 0 & I \end{array}\right) < 0$$

where $\widehat{\Phi}$ *is a basis matrix of* $\ker(B)$.

Remarks.

- Once the synthesis inequalities have been solved, the computation of $\begin{pmatrix} K & L \end{pmatrix}$ or of $\begin{pmatrix} K & L \\ M & N \end{pmatrix}$ can be performed along the lines of the proof of the Projection Lemma.

- It was our main concern to perform the variable elimination with as little computations as possible. They should be read as examples how one can proceed in specific circumstances, and they can be easily extended to various other performance specifications. As an exercise, the reader should eliminate variables in the peak-to-peak upper bound synthesis LMI's.

4.6 State-Feedback Problems

The state-feedback problem is characterized by

$$y = x \quad \text{or} \quad \begin{pmatrix} C & F_1 & \cdots & F_q \end{pmatrix} = \begin{pmatrix} I & 0 & \cdots & 0 \end{pmatrix}.$$

Then the formulas (4.2.3) read as

$$\begin{pmatrix} \boldsymbol{A}(v) & \boldsymbol{B}_j(v) \\ \boldsymbol{C}_j(v) & \boldsymbol{D}_j(v) \end{pmatrix} = \left(\begin{array}{cc|c} AY + BM & A + BN & B_j \\ K & AX + L & XB_j \\ \hline C_j Y + E_j M & C_j + E_j N & D_j \end{array} \right).$$

Note that the variable L only appears in the $(2,2)$-block, and that we can assign an arbitrary matrix in this position by suitably choosing L. Therefore, by varying L, the $(2,2)$ block of

$$\begin{pmatrix} \boldsymbol{A}(v)^T + \boldsymbol{A}(v) & \boldsymbol{B}_j(v) \\ \boldsymbol{B}_j(v) & 0 \end{pmatrix} =$$

$$= \left(\begin{array}{cc|c} (AY + BM) + (AY + BM)^T & (A + BN) + K^T & B_j \\ K + (A + BN)^T & (AX + L) + (AX + L)^T & XB_j \\ \hline B_j^T & B_j^T X & 0 \end{array} \right)$$

varies in the set of all symmetric matrices. This allows to apply Lemma 4.7 in order to eliminate L in synthesis inequalities what leads to a drastic simplification.

Let us illustrate all this for the quadratic performance problem. The corresponding synthesis

inequalities (4.2.7) read as

$$
\begin{pmatrix} Y & I \\ I & X \end{pmatrix} > 0, \quad
\left(\begin{array}{cc|c}
(AY + BM) + (AY + BM)^T & (A + BN) + K^T & B_j \\
K + (A + BN)^T & (AX + L) + (AX + L)^T & XB_j \\ \hline
B_j^T & B_j^T X & 0
\end{array} \right) +
$$

$$
+ \left(\begin{array}{cc|c} 0 & 0 & I \\ C_j Y + E_j M & C_j + E_j N & D_j \end{array} \right)^T P_j
\left(\begin{array}{cc|c} 0 & 0 & I \\ C_j Y + E_j M & C_j + E_j N & D_j \end{array} \right) < 0.
$$

These imply, just by cancelling the second block row/column,

$$
Y > 0, \quad \begin{pmatrix} (AY + BM) + (AY + BM)^T & B_j \\ B_j^T & 0 \end{pmatrix} +
$$

$$
+ \left(\begin{array}{cc} 0 & I \\ C_j Y + E_j M & D_j \end{array} \right)^T P_j
\left(\begin{array}{cc} 0 & I \\ C_j Y + E_j M & D_j \end{array} \right) < 0
$$

or

$$
Y > 0, \quad
\begin{pmatrix} I & 0 \\ AY + BM & B_j \\ 0 & I \\ C_j Y + E_j M & D_j \end{pmatrix}^T
\left(\begin{array}{cc|cc} 0 & I & 0 & 0 \\ I & 0 & 0 & 0 \\ \hline 0 & 0 & Q_j & S_j \\ 0 & 0 & S_j^T & R_j \end{array} \right)
\begin{pmatrix} I & 0 \\ AY + BM & B_j \\ 0 & I \\ C_j Y + E_j M & D_j \end{pmatrix} < 0. \quad (4.6.1)
$$

This is a drastic simplification since only the variables Y and M do appear in the resulting inequalities. It is no problem to reverse the arguments in order to show that the reduced inequalities are equivalent to the full synthesis inequalities.

However, proceeding in a different fashion leads to another fundamental insight: With solutions Y and M of (4.6.1), one can in fact design a *static* controller which solves the quadratic performance problem. Indeed, we just choose

$$
D_c := MY^{-1}
$$

to infer that the static controller $y = D_c u$ leads to a controlled system with the describing matrices

$$
\begin{pmatrix} \mathcal{A} & \mathcal{B}_j \\ \mathcal{C}_j & \mathcal{D}_j \end{pmatrix} =
\begin{pmatrix} A + BD_c & B_j \\ C_j + E_j D_c & D_j \end{pmatrix} =
\begin{pmatrix} (AY + BM)Y^{-1} & B_j \\ (C_j Y + E_j M)Y^{-1} & D_j \end{pmatrix}.
$$

We infer that (4.6.1) is identical to

$$
Y > 0, \quad
\begin{pmatrix} I & 0 \\ \mathcal{A}Y & \mathcal{B}_j \\ 0 & I \\ \mathcal{C}_j Y & \mathcal{D}_j \end{pmatrix}^T
\left(\begin{array}{cc|cc} 0 & I & 0 & 0 \\ I & 0 & 0 & 0 \\ \hline 0 & 0 & Q_j & S_j \\ 0 & 0 & S_j^T & R_j \end{array} \right)
\begin{pmatrix} I & 0 \\ \mathcal{A}Y & \mathcal{B}_j \\ 0 & I \\ \mathcal{C}_j Y & \mathcal{D}_j \end{pmatrix} < 0.
$$

If we perform congruence transformations with Y^{-1} and $\begin{pmatrix} Y^{-1} & 0 \\ 0 & I \end{pmatrix}$, we arrive with $\mathcal{X} :=$ Y^{-1} at

$$\mathcal{X} > 0, \quad \begin{pmatrix} I & 0 \\ \mathcal{X}\mathcal{A} & \mathcal{X}\mathcal{B}_j \\ 0 & I \\ \mathcal{C}_j & \mathcal{D}_j \end{pmatrix}^T \left(\begin{array}{cc|cc} 0 & I & 0 & 0 \\ I & 0 & 0 & 0 \\ \hline 0 & 0 & Q_j & S_j \\ 0 & 0 & S_j^T & R_j \end{array} \right) \begin{pmatrix} I & 0 \\ \mathcal{X}\mathcal{A} & \mathcal{X}\mathcal{B}_j \\ 0 & I \\ \mathcal{C}_j & \mathcal{D}_j \end{pmatrix} < 0.$$

Hence the static gain $\mathcal{D}$ indeed defines a controller which solves the quadratic performance problem.

Corollary 4.14 *Under the state-feedback information structure, there exists a dynamic controller* $\begin{pmatrix} A_c & B_c \\ C_c & D_c \end{pmatrix}$ *and some $\mathcal{X}$ which satisfy (4.2.1) iff there exist solutions Y and M of the inequalities (4.6.1). If Y and M solve (4.6.1), the static state-feedback controller gain*

$$D_c = MY^{-1}$$

and the Lyapunov matrix $\mathcal{X} := Y^{-1}$ render (4.2.1) satisfied.

In literally the same fashion as for output-feedback control, we arrive at the following general procedure to proceed from analysis inequalities to synthesis inequalities, and to construct a static state-feedback controller:

- Rewrite the analysis inequalities in the blocks $\mathcal{X}$, $\mathcal{X}\mathcal{A}$, $\mathcal{X}\mathcal{B}_j$, $\mathcal{C}_j$, $\mathcal{D}_j$ in order to be able to find a (formal) congruence transformation involving $\mathcal{Y}$ which leads to inequalities in the blocks $\mathcal{Y}^T\mathcal{X}\mathcal{Y}$, $\mathcal{Y}^T\mathcal{X}\mathcal{A}\mathcal{Y}$, $\mathcal{Y}^T\mathcal{X}\mathcal{B}_j$, $\mathcal{C}_j\mathcal{Y}$, $\mathcal{D}_j$.

- Perform the substitutions

$$\mathcal{Y}^T\mathcal{X}\mathcal{Y} \to Y \quad \text{and} \quad \begin{pmatrix} \mathcal{Y}^T\mathcal{X}\mathcal{A}\mathcal{Y} & \mathcal{Y}^T\mathcal{X}\mathcal{B}_j \\ \mathcal{C}_j\mathcal{Y} & \mathcal{D}_j \end{pmatrix} \to \begin{pmatrix} AY + BM & B_j \\ C_jY + E_jM & D_j \end{pmatrix}$$

  to arrive at matrix inequalities in the variables Y and M.

- After having solved the synthesis inequalities for Y and M, the static controller gain and the Lyapunov matrix

$$\mathcal{D} = MY^{-1} \quad \text{and} \quad \mathcal{X} = Y^{-1}$$

  render the analysis inequalities satisfied.

As an illustration, starting form the analysis inequalities (4.3.2) for H_2-type problems, the corresponding state-feedback synthesis conditions read as

$$\begin{pmatrix} (AY + BM)^T + (AY + BM) & B_j \\ B_j^T & -\gamma_j I \end{pmatrix} < 0,$$

$$\begin{pmatrix} Y & (C_jY + E_jM)^T \\ C_jY + E_jM & Z_j \end{pmatrix} > 0, \quad f_j(Z_j) < \gamma_j, \quad D_j = 0.$$

All our previous remarks pertaining to the (more complicated) procedure for the output-feedback information structure apply without modification.

In general we can conclude that dynamics in the controller do not offer any advantage over static controllers for state-feedback problems. This is also true for mixed control problems. This statements requires extra attention since our derivation was based on eliminating the variable L which might occur in several matrix inequalities. At this point the remark after Lemma 4.7 comes into play: This particular elimination result also applies to systems of matrix inequalities such that, indeed, the occurrence of L is various inequalities will not harm the arguments.

As earlier, in the single-objective quadratic performance problem by state-feedback, it is possible to eliminate the variable M in (4.6.1). Alternatively, one could as well exploit the particular structure of the system description to simplify the conditions in Theorem 4.11. Both approaches lead to the following result.

Corollary 4.15 *For the state-feedback quadratic performance problem with index satisfying (4.5.21), there exists dynamic controller and some $\mathcal{X}$ with (4.2.1) iff there exists a symmetric Y which solves*

$$Y > 0, \quad \Phi^T \begin{pmatrix} -A^T & -C_j^T \\ I & 0 \\ -B_j^T & -D_j^T \\ 0 & I \end{pmatrix}^T \left(\begin{array}{cc|cc} 0 & Y & 0 & 0 \\ Y & 0 & 0 & 0 \\ \hline 0 & 0 & \tilde{Q}_j & \tilde{S}_j \\ 0 & 0 & \tilde{S}_j^T & \tilde{R}_j \end{array} \right) \begin{pmatrix} -A^T & -C_j^T \\ I & 0 \\ -B_j^T & -D_j^T \\ 0 & I \end{pmatrix} \Phi > 0. \qquad (4.6.2)$$

Remarks. All these results should be viewed as illustrations how to proceed for specific system descriptions. Indeed, another popular choice is the so-called *full information* structure in which both the state and the disturbance are measurable:

$$y = \begin{pmatrix} x \\ w \end{pmatrix}.$$

Similarly, one could consider the corresponding dual versions that are typically related to *estimation problems*, such as e.g.

$$\begin{pmatrix} B \\ E_1 \\ \vdots \\ E_q \end{pmatrix} = \begin{pmatrix} I \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

We have collected all auxiliary results that allow to handle these specific problems without any complications.

4.7 Robust Controller Design

So far we have presented techniques to design controllers for nominal stability and nominal performance. Chapters 2 and 3 have been devoted to a thorough discussion of how to analyze, for a fixed stabilizing controller, robust stability or robust performance. For time-invariant or time-varying parametric uncertainties, we have seen direct tests formulated as searching for constant or parameter-dependent quadratic Lyapunov functions. For much larger classes of uncertainties, we have derived tests in terms of integral quadratic constraints that involve additional variables which have been called scalings or multipliers.

Typically, only those IQC tests with a class of multipliers that admit a state-space description as discussed in the Sections 3.6-3.10 are amenable to a systematic output-feedback controller design procedure which is a reminiscent of the D/K-iteration in μ-theory. This will be the subject of the first section.

In a second section we consider as a particular information structure the robust state-feedback design problem. We will reveal that the search for static state-feedback gains which achieve robust performance can be transformed into a convex optimization problem.

The discussion is confined to the quadratic performance problem since most results can be extended in a pretty straightforward fashion to the other specifications considered in these notes.

4.7.1 Robust Output-Feedback Controller Design

If characterizing robust performance by an IQC, the goal in robust design is to find a controller *and* a multiplier such that, for the closed-loop system, the corresponding IQC test is satisfied. Hence, the multiplier appears as an extra unknown what makes the problem hard if not impossible to solve.

However, if the multiplier is held fixed, searching for a controller amounts to a nominal design problem that can be approached with the techniques described earlier. If the controller is held fixed, the analysis techniques presented in Chapter 3 can be used to find a suitable multiplier. Hence, instead of trying to search for a controller and a multiplier commonly, one iterates between the search for a controller with fixed multiplier and the search for a multiplier with fixed controller. This procedure is known from μ-theory as scalings/controller iteration or D/K iteration.

To be more concrete, we consider the specific *example* of achieving robust quadratic performance against time-varying parametric uncertainties as discussed in Section 3.10.

The uncontrolled unperturbed system is described by (4.1.1). We assume that $w_1 \rightarrow z_1$ is

the uncertainty channel and the uncontrolled uncertain system is described by including

$$w_1(t) = \Delta(t)z_1(t)$$

where $\Delta(.)$ varies in the set of continuous curves satisfying

$$\Delta(t) \in \mathbf{\Delta}_c := \mathrm{co}\{\Delta_1, ..., \Delta_N\} \quad \text{for all} \ \ t \geqslant 0.$$

We assume (w.l.o.g.) that

$$0 \in \mathrm{co}\{\Delta_1, ..., \Delta_N\}.$$

The performance channel is assumed to be given by $w_2 \to z_2$, and the performance index

$$P_p = \begin{pmatrix} Q_p & S_p \\ S_p^T & R_p \end{pmatrix}, \ \ R_p \geqslant 0 \quad \text{with the inverse} \quad \tilde{P}_p^{-1} = \begin{pmatrix} \tilde{Q}_p & \tilde{S}_p \\ \tilde{S}_p^T & \tilde{R}_p \end{pmatrix}, \ \ \tilde{Q}_p \leqslant 0$$

is used to define the quadratic performance specification

$$\int_0^\infty \begin{pmatrix} w_2(t) \\ z_2(t) \end{pmatrix}^T P_p \begin{pmatrix} w_2(t) \\ z_2(t) \end{pmatrix} dt \leqslant -\epsilon \|w_2\|_2^2.$$

The goal is to design a controller that achieves robust stability and robust quadratic performance. We can guarantee both properties by finding a controller, a Lyapunov matrix $\mathcal{X}$, and a multiplier

$$P = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix}, \ \ Q < 0, \ \ \begin{pmatrix} \Delta_j \\ I \end{pmatrix}^T \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} \begin{pmatrix} \Delta_j \\ I \end{pmatrix} > 0 \ \ \text{for all} \ \ j = 1, \ldots, N \quad (4.7.1)$$

that satisfy the inequalities

$$\mathcal{X} > 0, \quad \begin{pmatrix} I & 0 & 0 \\ \mathcal{X}\mathcal{A} & \mathcal{X}\mathcal{B}_1 & \mathcal{X}\mathcal{B}_2 \\ \hline 0 & I & 0 \\ \mathcal{C}_1 & \mathcal{D}_1 & \mathcal{D}_{12} \\ \hline 0 & 0 & I \\ \mathcal{C}_2 & \mathcal{D}_{21} & \mathcal{D}_2 \end{pmatrix}^T \begin{pmatrix} 0 & I & 0 & 0 & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & Q & S & 0 & 0 \\ 0 & 0 & S^T & R & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & Q_p & S_p \\ 0 & 0 & 0 & 0 & S_p^T & R_p \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ \mathcal{X}\mathcal{A} & \mathcal{X}\mathcal{B}_1 & \mathcal{X}\mathcal{B}_2 \\ \hline 0 & I & 0 \\ \mathcal{C}_1 & \mathcal{D}_1 & \mathcal{D}_{12} \\ \hline 0 & 0 & I \\ \mathcal{C}_2 & \mathcal{D}_{21} & \mathcal{D}_2 \end{pmatrix} < 0.$$

(Recall that the condition on the left-upper block of P can be relaxed in particular cases what could reduce the conservatism of the test.)

If we apply the controller parameter transformation of Chapter 4, we arrive that the synthesis matrix inequalities

$$\mathbf{X}(v) > 0, \quad \begin{pmatrix} * \\ * \\ * \\ * \\ * \\ * \end{pmatrix}^T \begin{pmatrix} 0 & I & 0 & 0 & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & Q & S & 0 & 0 \\ 0 & 0 & S^T & R & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & Q_p & S_p \\ 0 & 0 & 0 & 0 & S_p^T & R_p \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ \mathbf{A}(v) & \mathbf{B}_1(v) & \mathbf{B}_2(v) \\ \hline 0 & I & 0 \\ \mathbf{C}_1(v) & \mathbf{D}_1(v) & \mathbf{D}_{12}(v) \\ \hline 0 & 0 & I \\ \mathbf{C}_2(v) & \mathbf{D}_{21}(v) & \mathbf{D}_2(v) \end{pmatrix} < 0.$$

Unfortunately, there is no obvious way how to render these synthesis inequalities convex in *all* variables v, Q, S, R.

This is the reason why we consider, instead, the problem with a scaled uncertainty

$$w_1(t) = [r\Delta(t)]z_1(t), \quad \Delta(t) \in \mathbf{\Delta}_c \tag{4.7.2}$$

where the scaling factor is contained in the interval $[0, 1]$. Due to

$$\begin{pmatrix} r\Delta \\ I \end{pmatrix}^T \begin{pmatrix} Q & rS \\ rS^T & r^2R \end{pmatrix} \begin{pmatrix} r\Delta \\ I \end{pmatrix} = r^2 \begin{pmatrix} \Delta \\ I \end{pmatrix}^T \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} \begin{pmatrix} \Delta \\ I \end{pmatrix},$$

we conclude that the corresponding analysis or synthesis are given by (4.7.1) and

$$\mathcal{X} > 0, \quad \begin{pmatrix} I & 0 & 0 \\ \mathcal{X}\mathcal{A} & \mathcal{X}\mathcal{B}_1 & \mathcal{X}\mathcal{B}_2 \\ 0 & I & 0 \\ \mathcal{C}_1 & \mathcal{D}_1 & \mathcal{D}_{12} \\ 0 & 0 & I \\ \mathcal{C}_2 & \mathcal{D}_{21} & \mathcal{D}_2 \end{pmatrix}^T \begin{pmatrix} 0 & I & 0 & 0 & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & Q & rS & 0 & 0 \\ 0 & 0 & rS^T & r^2R & 0 & 0 \\ 0 & 0 & 0 & 0 & Q_p & S_p \\ 0 & 0 & 0 & 0 & S_p^T & R_p \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ \mathcal{X}\mathcal{A} & \mathcal{X}\mathcal{B}_1 & \mathcal{X}\mathcal{B}_2 \\ 0 & I & 0 \\ \mathcal{C}_1 & \mathcal{D}_1 & \mathcal{D}_{12} \\ 0 & 0 & I \\ \mathcal{C}_2 & \mathcal{D}_{21} & \mathcal{D}_2 \end{pmatrix} < 0 \tag{4.7.3}$$

or

$$\boldsymbol{X}(v) > 0, \quad \begin{pmatrix} * \\ * \\ * \\ * \\ * \\ * \end{pmatrix}^T \begin{pmatrix} 0 & I & 0 & 0 & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & Q & rS & 0 & 0 \\ 0 & 0 & rS^T & r^2R & 0 & 0 \\ 0 & 0 & 0 & 0 & Q_p & S_p \\ 0 & 0 & 0 & 0 & S_p^T & R_p \end{pmatrix} \begin{pmatrix} I & 0 & 0 \\ \boldsymbol{A}(v) & \boldsymbol{B}_1(v) & \boldsymbol{B}_2(v) \\ 0 & I & 0 \\ \boldsymbol{C}_1(v) & \boldsymbol{D}_1(v) & \boldsymbol{D}_{12}(v) \\ 0 & 0 & I \\ \boldsymbol{C}_2(v) & \boldsymbol{D}_{21}(v) & \boldsymbol{D}_2(v) \end{pmatrix} < 0. \tag{4.7.4}$$

For $r = 0$, we hence have to solve the nominal quadratic performance synthesis inequalities. If they are not solvable, the robust quadratic performance synthesis problem is not solvable either and we can stop. If they are solvable, the idea is to try to increase, keeping the synthesis inequalities feasible, the parameter r from zero to one. Increasing r is achieved by alternatingly maximizing r over v satisfying (4.7.4) (for fixed P) and by varying $\mathcal{X}$ and P in (4.7.3) (for a fixed controller).

The maximization of r proceeds along the following steps:

Initialization. Perform a nominal quadratic performance design by solving (4.7.4) for $r = 0$. Proceed if these inequalities are feasible and compute a corresponding controller.

After this initial phase, the iteration is started. The $j - 1$-st step of the iteration leads to a controller, a Lyapunov matrix $\mathcal{X}$, and a multiplier P that satisfy the inequalities (4.7.1) and (4.7.3) for the parameter $r = r_{j-1}$. Then it proceeds as follows:

First step: Fix the controller and maximize r by varying the Lyapunov matrix $\mathcal{X}$ and the scaling such that such that (4.7.1) and (4.7.3) hold. The maximal radius is denoted as $\widehat{r}_j$ and it satisfies $r_{j-1} \leqslant \widehat{r}_j$.

Second step: Fix the resulting scaling P and find the largest r by varying the variables v in (4.7.4). The obtained maximum r_j clearly satisfies $\widehat{r}_j \leqslant r_j$.

The iteration defines a sequence of radii

$$r_1 \leqslant r_2 \leqslant r_3 \leqslant \cdots$$

and a corresponding controller that guarantee robust stability and robust quadratic performance for all uncertainties (4.7.2) with radius $r = r_j$.

If we are in the lucky situation that there is an index for which $r_j \geqslant 1$, the corresponding controller is robustly performing for all uncertainties with values in $\boldsymbol{\Delta}_c$ as desired, and we are done. However, if $r_j < 1$ for all indices, we cannot guarantee robust performance for $r = 1$, but we still have a guarantee of robust performance for $r = r_j$!

Before entering a brief discussion of this procedure, let us include the following remarks on the start-up and on the computations. If the nominal performance synthesis problem has a solution, the LMI's (4.7.1)-(4.7.3) do have a solution $\mathcal{X}$ and P for the resulting controller and for some - possibly small - $r > 0$; this just follows by continuity. Hence the iteration does not get stuck after the first step. Secondly, for a fixed r, the first step of the iteration amounts to solving an analysis problem, and finding a solution v of (4.7.4) can be converted to an LMI problem. Therefore, the maximization of r can be performed by bisection.

Even if the inequalities (4.7.1)-(4.7.4) are solvable for $r = 1$, it can happen the the limit of r_j is smaller than one. As a remedy, one could consider another parameter to maximize, or one could modify the iteration scheme that has been sketched above. For example, it is possible to take the fine structure of the involved functions into account and to suggest other variable combinations that render the resulting iteration steps convex. Unfortunately, one cannot give general recommendations for modifications which guarantee success.

Remark. It should be noted that the controller/multiplier iteration can be extended to all robust performance tests that are based on families of dynamic IQC's which are described by real rational multipliers. Technically, one just requires a parametrization of the multipliers such that the corresponding analysis test (for a fixed controller) and the controller synthesis (for a fixed multiplier) both reduce to solving standard LMI problems.

4.7.2   Robust State-Feedback Controller Design

For the same set-up as in the previous section we consider the corresponding synthesis problem if the state of the underlying system is measurable. According to our discussion in

Section 4.6, the resulting synthesis inequalities read as

$$Q < 0, \quad \begin{pmatrix} \Delta_j \\ I \end{pmatrix}^T \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} \begin{pmatrix} \Delta_j \\ I \end{pmatrix} > 0 \quad \text{for all} \quad j = 1, \dots, N$$

and

$$Y > 0, \quad \begin{pmatrix} * \\ * \\ * \\ * \\ * \\ * \end{pmatrix}^T \left(\begin{array}{cc|cc|cc} 0 & I & 0 & 0 & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & Q & S & 0 & 0 \\ 0 & 0 & S^T & R & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & Q_p & S_p \\ 0 & 0 & 0 & 0 & S_p^T & R_p \end{array} \right) \begin{pmatrix} I & 0 & 0 \\ AY + BM & B_1 & B_2 \\ \hline 0 & I & 0 \\ C_1Y + E_1M & D_1 & D_{12} \\ \hline 0 & 0 & I \\ C_2Y + E_1M & D_{21} & D_2 \end{pmatrix} < 0$$

in the variables Y, M, Q, S, R.

In this form these inequalities are *not convex*. However, we can can apply the Dualization Lemma (Section 4.5.1) to arrive at the equivalent inequalities

$$\tilde{R} > 0, \quad \begin{pmatrix} I \\ -\Delta_j^T \end{pmatrix}^T \begin{pmatrix} \tilde{Q} & \tilde{S} \\ \tilde{S}^T & \tilde{R} \end{pmatrix} \begin{pmatrix} I \\ -\Delta_j^T \end{pmatrix} < 0 \quad \text{for all} \quad j = 1, \dots, N$$

and $Y > 0$,

$$* \left(\begin{array}{cc|cc|cc} 0 & I & 0 & 0 & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & \tilde{Q} & \tilde{S} & 0 & 0 \\ 0 & 0 & \tilde{S}^T & \tilde{R} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & \tilde{Q}_p & \tilde{S}_p \\ 0 & 0 & 0 & 0 & \tilde{S}_p^T & \tilde{R}_p \end{array} \right) \begin{pmatrix} -(AY+BM)^T & -(C_1Y+E_1M)^T & -(C_2Y+E_2M)^T \\ I & 0 & 0 \\ -B_1^T & -D_1^T & -D_{21}^T \\ \hline 0 & I & 0 \\ -B_2^T & -D_{12}^T & -D_2^T \\ \hline 0 & 0 & I \end{pmatrix} > 0$$

in the variables Y, M, $\tilde{Q}$, $\tilde{S}$, $\tilde{R}$. It turns out that these dual inequalities are all *affine* in the unknowns. Testing feasibility hence amounts to solving a standard LMI problem. If the LMI's are feasible, a robust static state-feedback gain is given by $\mathcal{D} = MY^{-1}$. This is one of the very few lucky instances in the world of designing robust controllers!

4.7.3 Affine Parameter Dependence

Let us finally consider the system

$$\begin{pmatrix} \dot{x} \\ z \\ y \end{pmatrix} = \left(\begin{array}{c|cc} A(\Delta(t)) & B_1(\Delta(t)) & B(\Delta(t)) \\ \hline C_1(\Delta(t)) & D(\Delta(t)) & E(\Delta(t)) \\ C(\Delta(t)) & F(\Delta(t)) & 0 \end{array} \right) \begin{pmatrix} x \\ w \\ u \end{pmatrix}, \quad \Delta(t) \in \mathrm{co}\{\Delta_1, \dots, \Delta_N\}$$

where the describing matrices depend *affinely* on the time-varying parameters. If designing output-feedback controllers, there is no systematic alternative to pulling out the uncertainties and applying the scalings techniques as in Section 4.7.1.

For robust state-feedback design there is an alternative without scalings. One just needs to directly solve the system of LMI's

$$
Y > 0, \quad
\begin{pmatrix} * \\ * \\ * \\ * \end{pmatrix}^T
\left(\begin{array}{cc|cc} 0 & I & 0 & 0 \\ I & 0 & 0 & 0 \\ \hline 0 & 0 & Q_p & S_p \\ 0 & 0 & S_p^T & R_p \end{array}\right)
\left(\begin{array}{cc} I & 0 \\ A(\Delta_j)Y + B(\Delta_j)M & B_1(\Delta_j) \\ \hline 0 & I \\ C_1(\Delta_j)Y + E(\Delta_j)M & D(\Delta_j) \end{array}\right) < 0, \; j = 1, \ldots, N
$$

(4.7.5)

in the variables Y and M.

For the controller gain $D_c = MY^{-1}$ we obtain

$$
Y > 0, \quad
\begin{pmatrix} * \\ * \\ * \\ * \end{pmatrix}^T
\left(\begin{array}{cc|cc} 0 & I & 0 & 0 \\ I & 0 & 0 & 0 \\ \hline 0 & 0 & Q_p & S_p \\ 0 & 0 & S_p^T & R_p \end{array}\right)
\left(\begin{array}{cc} I & 0 \\ (A(\Delta_j) + B(\Delta_j)D_c)Y & B_1(\Delta_j) \\ \hline 0 & I \\ (C_1(\Delta_j) + E(\Delta_j)D_c)Y & D(\Delta_j) \end{array}\right) < 0, \; j = 1, \ldots, N
$$

A convexity argument leads to

$$
Y > 0, \quad
\begin{pmatrix} * \\ * \\ * \\ * \end{pmatrix}^T
\left(\begin{array}{cc|cc} 0 & I & 0 & 0 \\ I & 0 & 0 & 0 \\ \hline 0 & 0 & Q_p & S_p \\ 0 & 0 & S_p^T & R_p \end{array}\right)
\left(\begin{array}{cc} I & 0 \\ (A(\Delta(t)) + B(\Delta(t))D_c)Y & B_1(\Delta(t)) \\ \hline 0 & I \\ (C_1(\Delta(t)) + E(\Delta(t))D_c)Y & D(\Delta(t)) \end{array}\right) < 0
$$

for all parameter curves $\Delta(t) \in \mathrm{co}\{\Delta_1, ..., \Delta_N\}$, and we can perform a congruence transformation as in Section 4.6 to get

$$
\mathcal{X} > 0, \quad
\begin{pmatrix} * \\ * \\ * \\ * \end{pmatrix}^T
\left(\begin{array}{cc|cc} 0 & I & 0 & 0 \\ I & 0 & 0 & 0 \\ \hline 0 & 0 & Q_p & S_p \\ 0 & 0 & S_p^T & R_p \end{array}\right)
\left(\begin{array}{cc} I & 0 \\ \mathcal{X}(A(\Delta(t)) + B(\Delta(t))D_c) & \mathcal{X}B_1(\Delta(t)) \\ \hline 0 & I \\ (C_1(\Delta(t)) + E(\Delta(t))D_c) & D(\Delta(t)) \end{array}\right) < 0.
$$

These two inequalities imply, in turn, robust exponential stability and robust quadratic performance for the controlled system as seen in Section 3.10.2.

We have proved that it suffices to directly solve the LMI's (4.7.5) to compute a robust static state-feedback controller. Hence, if the system's parameter dependence is affine, we have found two equivalent sets of synthesis inequalities that differ in the number of the involved variables and in the sizes of the LMI's that are involved. In practice, the correct choice is dictated by whatever system can be solved faster, more efficiently, or numerically more reliably.

Remark. Here is the reason why it is possible to directly solve the robust performance problem by state-feedback without scalings, and why this technique does, unfortunately, not extend to output-feedback control: The linearizing controller parameter transformation for state-feedback problems *does not involve the matrices that describe the open-loop system*, whereas that for that for ouptut-feedback problems indeed depends on the matrices A, B, C of the open-loop system description. (See also Exercise 2.)

Let us conclude this section by stressing, again, that these techniques find straightforward extensions to other performance specifications. As an exercise, the reader is asked to work out the details of the corresponding results for the robust H_2-synthesis problem by state- or output-feedback.

4.8   Discrete-Time Systems

Everything what has been said so far can be easily extended to discrete time-design problems. This is particularly surprising since, in the literature, discrete-time problem solutions often seem much more involved and harder to master than their continuous-time counterparts.

Our general procedure to step from analysis to synthesis as well as the technique to recover the controller need no change at all; in particular, the concrete formulas for the block substitutions do not change. The elimination of transformed controller parameters proceeds in the same fashion on the basis of the Projection Lemma or the Elimination Lemma and the specialized version thereof.

Only as a example we consider the problem discussed in [13]: the mixed H_2/H_∞ problem with different disturbance inputs and controlled outputs in discrete-time.

It is well-known [13] that $\mathcal{A}$ has all its eigenvalues in the unit disk, that the discrete time H_2-norm of

$$\mathcal{C}_1(zI - \mathcal{A})^{-1}\mathcal{B}_1 + \mathcal{D}_1$$

is smaller than γ_1, and that the discrete time H_∞-norm of

$$\mathcal{C}_2(zI - \mathcal{A})^{-1}\mathcal{B}_2 + \mathcal{D}_2$$

is smaller than γ_2 iff there exist symmetric matrices $\mathcal{X}_1$, $\mathcal{X}_2$, and Z with $\text{trace}(Z) < \gamma_1$ and

$$\begin{pmatrix} \mathcal{X}_1 & \mathcal{X}_1\mathcal{A} & \mathcal{X}_1\mathcal{B}_1 \\ \mathcal{A}^T\mathcal{X}_1 & \mathcal{X}_1 & 0 \\ \mathcal{B}_1^T\mathcal{X}_1 & 0 & \gamma_1 I \end{pmatrix} > 0, \begin{pmatrix} \mathcal{X}_1 & 0 & \mathcal{C}_1^T \\ 0 & I & \mathcal{D}_1^T \\ \mathcal{C}_1 & \mathcal{D}_1 & Z \end{pmatrix} > 0, \begin{pmatrix} \mathcal{X}_2 & 0 & \mathcal{A}^T\mathcal{X}_2 & \mathcal{C}_2^T \\ 0 & \gamma_2 I & \mathcal{B}_2^T\mathcal{X}_2 & \mathcal{D}_2^T \\ \mathcal{X}_2\mathcal{A} & \mathcal{X}_2\mathcal{B}_2 & \mathcal{X}_2 & 0 \\ \mathcal{C}_2 & \mathcal{D}_2 & 0 & \gamma_2 I \end{pmatrix} > 0.$$

$$(4.8.1)$$

Note that we have transformed these analysis LMI's such that they are affine in the blocks that will be transformed for synthesis.

The mixed problem consists of searching for a controller that renders these inequalities satisfied with a common Lyapunov function $\mathcal{X} = \mathcal{X}_1 = \mathcal{X}_2$. The solution is immediate: Perform congruence transformations of (4.8.1) with

$$\mathrm{diag}(\mathcal{Y}, \mathcal{Y}, I), \quad \mathrm{diag}(\mathcal{Y}, I, I), \quad \mathrm{diag}(\mathcal{Y}, I, \mathcal{Y}, I)$$

and read off the synthesis LMI's using (4.2.3). After solving the synthesis LMI's, we stress again that the controller construction proceeds along *the same steps* as in Theorem 4.2. The inclusion of pole constraints for arbitrary LMI regions (related, of course, to discrete time stability) and other criteria poses no extra problems.

4.9 Exercises

Exercise 1

Derive an LMI solution of the H_∞-problem for the system

$$\begin{pmatrix} \dot{x} \\ z_1 \\ y \end{pmatrix} = \begin{pmatrix} A & B_1 & B \\ C_1 & D_1 & E \\ C & F & 0 \end{pmatrix} \begin{pmatrix} x \\ w_1 \\ u \end{pmatrix}$$

with

$$C = \begin{pmatrix} I \\ 0 \end{pmatrix}, \quad F = \begin{pmatrix} 0 \\ I \end{pmatrix} \quad \text{such that} \quad y = \begin{pmatrix} x \\ w_1 \end{pmatrix}.$$

(This is the so-called full information problem.)

Exercise 2 (Nominal and Robust Estimation)

Consider the system

$$\begin{pmatrix} \dot{x} \\ z \\ y \end{pmatrix} = \begin{pmatrix} A & B_1 \\ C_1 & D_1 \\ C & F \end{pmatrix} \begin{pmatrix} x \\ w \end{pmatrix}$$

and inter-connect it with the estimator

$$\begin{pmatrix} \dot{x}_c \\ \hat{z} \end{pmatrix} = \begin{pmatrix} A_c & B_c \\ C_c & D_c \end{pmatrix} \begin{pmatrix} x_c \\ y \end{pmatrix} \tag{4.9.1}$$

where both A and A_c are Hurwitz. The goal in optimal estimation is to design an estimator which keeps $z - \hat{z}$ as small as possible for all disturbances w in a certain class. Out of the multitude of possibilities, we choose the L_2-gain of $w \to z - \hat{z}$ (for zero initial condition of both the system and the estimator) as the measure of the estimation quality.

This leads to the following problem formulation: Given $\gamma > 0$, test whether there exists an estimator which renders

$$\sup_{w \in L_2, \ w \neq 0} \frac{\|z - \hat{z}\|_2}{\|w\|_2} < \gamma \tag{4.9.2}$$

satisfied. If yes, reveal how to design an estimator that leads to this property.

1. Show that the estimation problem is a specialization of the general output-feedback H_∞-design problem.

2. Due to the specific structure of the open-loop system, show that there exists a linearizing transformation of the estimator parameters which does not involve any matrices that describe the open-loop system.

Hint: To find the transformation, proceed as in the proof of Theorem 4.2 with the factorization

$$\mathcal{Y}^T \mathcal{X} = \mathcal{Z} \ \text{ where } \ \mathcal{Y}^T = \begin{pmatrix} I & Y^{-1}V \\ I & 0 \end{pmatrix}, \ \ \mathcal{Z} = \begin{pmatrix} Y^{-1} & 0 \\ X & U \end{pmatrix},$$

and consider as before the blocks $\mathcal{Y}^T \mathcal{X} \mathcal{A} \mathcal{Y}$, $\mathcal{Y}^T \mathcal{X} \mathcal{B}$, $\mathcal{C} \mathcal{Y}$.

3. Now assume that the system is affected by time-varying uncertain parameters as

$$\begin{pmatrix} \dot{x} \\ z \\ y \end{pmatrix} = \begin{pmatrix} A(\Delta(t)) & B_1(\Delta(t)) \\ C_1(\Delta(t)) & D_1(\Delta(t)) \\ C(\Delta(t)) & F(\Delta(t)) \end{pmatrix} \begin{pmatrix} x \\ w \end{pmatrix}$$

where

$$\begin{pmatrix} A(\Delta) & B_1(\Delta) \\ C_1(\Delta) & D_1(\Delta) \\ C(\Delta) & F(\Delta) \end{pmatrix} \ \text{ is affine in } \ \Delta \ \text{ and } \ \Delta(t) \in \text{co}\{\Delta_1, ..., \Delta_N\}.$$

Derive LMI conditions for the existence of an estimator that guarantees (4.9.2) for all uncertainties, and show how to actually compute such an estimator if the LMI's are feasible.

Hint: Recall what we have discussed for the state-feedback problem in Section 4.7.3.

4. Suppose that the uncertainty enters rationally, and that it has been pulled out to arrive at the LFT representation

$$\begin{pmatrix} \dot{x} \\ z_1 \\ z \\ y \end{pmatrix} = \begin{pmatrix} A & B_1 & B_2 \\ C_1 & D_1 & D_{12} \\ C_2 & D_{21} & D_2 \\ C & F_1 & F_2 \end{pmatrix} \begin{pmatrix} x \\ w_1 \\ w \end{pmatrix}, \ \ w_1(t) = \Delta(t) z_1(t), \ \ \Delta(t) \in \text{co}\{\Delta_1, ..., \Delta_N\}$$

of the uncertain system. Derive synthesis inequalities with full-block scalings that guarantee the existence of an estimator that guarantees (4.9.2) for all uncertainties and reveal how to actually compute such an estimator if the LMI's are feasible. What happens if $D_1 = 0$ such that the uncertainty enters affinely?

Hint: The results should be formulated analogously to what we have done in Section 4.7.2. There are two possibilities to proceed: You can either just use the transformation (4.2.10) to obtain synthesis inequalities that can be rendered convex by an additional congruence transformation, or you can employ the alternative parameter transformation as derived in part 2 of this exercise to directly obtain a convex test.

Exercise 3

This is an exercise on robust control. To reduce the complexity of programming, we consider a non-dynamic system only.

Suppose you have given the algebraic uncertain system

$$
\begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ \hline z \\ y_1 \\ y_2 \end{pmatrix} = \left(\begin{array}{cccc|c|cc} 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 0.5 & 0 & 0.5 & 0 & 1 & 0 & 1 \\ 2a & 0 & a & 0 & 1 & 0 & 0 \\ 0 & -2a & 0 & -a & 1 & 0 & 0 \\ \hline 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{array} \right) \begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \\ \hline w \\ u_1 \\ u_2 \end{pmatrix},
$$

with a time-varying uncertainty

$$
\begin{pmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{pmatrix} = \begin{pmatrix} \delta_1(t) & & & 0 \\ & \delta_1(t) & & \\ & & \delta_2(t) & \\ 0 & & & \delta_2(t) \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix}, \quad |\delta_1(t)| \leqslant 0.7, \quad |\delta_2(t)| \leqslant 0.7.
$$

As the performance measure we choose the L_2-gain of the channel $w \to z$.

1. For the uncontrolled system and for each $a \in [0,1]$, find the minimal robust L_2-gain level of the channel $w \to z$ by applying the robust performance analysis test in Chapter 3 with the following class of scalings $P = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix}$:

   - P is as in μ-theory: Q, S, R are block-diagonal, $Q < 0$, R is related to Q (how?), and S is skew-symmetric.
   - P is general with $Q < 0$.
   - P is general with $Q_1 < 0$, $Q_2 < 0$, where Q_j denote the blocks $Q(1:2, 1:2)$ and $Q(3:4, 3:4)$ in Matlab notation.

   Draw plots of the corresponding optimal values versus the parameter a and comment!

2. For $a = 0.9$, apply the controller

$$
\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & k \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}
$$

   and perform the analysis test with the largest class of scalings for $k \in [-1,1]$. Plot the resulting optimal value over k and comment.

3. Perform a controller/scaling iteration to minimize the optimal values for the controller structures

$$
\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & k_2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} k_1 & k_{12} \\ k_{21} & k_2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.
$$

   Start from gain zero and plot the optimal values that can are reached in each step of the iteration to reveal how they decrease. Comment on the convergence.

4. With the last full controller from the previous exercise for a performance level that is close to the limit, redo the analysis of the first part. Plot the curves and comment.

Exercise 4

This is a simulation exercise that involves the synthesis of an active controller for the suspension system in Exercise 7 of Chapter 1. We consider the rear wheel of a tractor-trailer combination as is depicted in Figure 4.2. Here m_1 represents tire, wheel and rear axle mass,
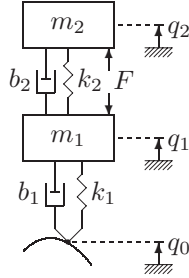

Figure 4.2: Active suspension system

m_2 denotes a fraction of the semitrailer mass. The deflection variables q_i are properly scaled so that $q_2 - q_1 = 0$ and $q_1 - q_0 = 0$ in steady state. The system is modeled by the state space equations

$$\dot{x} = Ax + B \begin{pmatrix} q_0 \\ F \end{pmatrix}$$

$$z = Cx + D \begin{pmatrix} q_0 \\ F \end{pmatrix}$$

where

$$A = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -\frac{k_1+k_2}{m_1} & \frac{k_2}{m_1} & -\frac{b_1+b_2}{m_1} & \frac{b_2}{m_1} \\ \frac{k_2}{m_2} & -\frac{k_2}{m_2} & \frac{b_2}{m_2} & -\frac{b_2}{m_2} \end{pmatrix}; \qquad B = \begin{pmatrix} \frac{b_1}{m_1} & 0 \\ 0 & 0 \\ \frac{k_1}{m_1} - \frac{b_1}{m_1}\frac{b_1+b_2}{m_1} & -\frac{1}{m_1} \\ \frac{b_1 b_2}{m_1 m_2} & \frac{1}{m_2} \end{pmatrix}$$

$$C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{k_2}{m_2} & -\frac{k_2}{m_2} & \frac{b_2}{m_2} & -\frac{b_2}{m_2} \\ -1 & 1 & 0 & 0 \end{pmatrix}; \qquad D = \begin{pmatrix} -1 & 0 \\ 0 & 1 \\ \frac{b_1 b_2}{m_1 m_2} & \frac{1}{m_2} \\ 0 & 0 \end{pmatrix}.$$

Here, $x = \begin{pmatrix} q_1 & q_2 & \dot{q}_1 - b_1 q_0/m_1 & \dot{q}_2 \end{pmatrix}^\top$ and $z = \begin{pmatrix} q_1 - q_0 & F & \ddot{q}_2 & q_2 - q_1 \end{pmatrix}^\top$ define the state and the to-be-controlled output, respectively. The control input is the force F, the exogenous input is the road profile q_0.

Suppose that the masses $m_1 = 1.5 \times 10^3$ and $m_2 = 10 \times 10^3$, the tire damping $b_1 = 1.7 \times 10^3$ and the tire stiffness $k_1 = 5 \times 10^6$. The suspension damping is a time-varying uncertain quantity with

$$b_2(t) \in [50 \times 10^3 - \bar{b}, 50 \times 10^3 + \bar{b}], \quad t \geqslant 0 \qquad (4.9.3)$$

and the suspension stiffness is a time-varying uncertainty parameter with

$$k_2(t) \in [500 \times 10^3 - \bar{k}, 500 \times 10^3 + \bar{k}], \quad t \geqslant 0. \qquad (4.9.4)$$

The aim of the exercise is to design an active suspension control system that generates the force F as a (causal) function of the variable $y = \begin{pmatrix} \ddot{q}_2 & q_2 - q_1 \end{pmatrix}^{\top}$.

The main objective of the controller design is to achieve low levels of acceleration throughout the vehicle ($\ddot{q}_2$), bounded suspension deflection ($q_2 - q_1$ and $q_1 - q_0$) and bounded dynamic tire force (F).

1. Suppose first that $\bar{k} = \bar{b} = 0$ (i.e., no uncertainty n the parameters) and let the road profile be represented by $q_0 = W_{q_0} \tilde{q}_0$ where $\tilde{q}_0 \in \mathcal{L}_2$ is equalized in frequency and W_{q_0} is the transfer function

$$W_{q_0}(s) = \frac{0.01}{0.4s + 1}$$

reflecting the quality of the road when the vehicle drives at constant speed. Define the to-be-controlled output $\tilde{z} = W_z z$ where W_z is a weighting matrix with transfer function

$$W_z(s) = \begin{pmatrix} 200 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 \\ 0 & 0 & \frac{0.0318s + 0.4}{3.16 \times 10^{-4} s^2 + 0.0314s + 1} & 0 \\ 0 & 0 & 0 & 100 \end{pmatrix}.$$

The weight on the chassis acceleration reflects the human sensitivity to vertical accelerations. Use the routines `ltisys`, `smult` and `sdiag` to implement the generalized plant

$$P : \begin{pmatrix} q_0 \\ F \end{pmatrix} \mapsto \begin{pmatrix} \tilde{z} \\ y \end{pmatrix}$$

and synthesize with the routine `hinflmi` a controller which minimizes the H_∞ norm of the closed-loop transfer function $\mathcal{T} : \tilde{q}_0 \mapsto \tilde{z}$.

2. Construct with this controller the closed-loop system which maps q_0 to z (not $\tilde{q}_0$ to $\tilde{z}$!) and validate the controlled system by plotting the four frequency responses of the closed-loop system and the four responses to a step with amplitude 0.3 (meter). (See the routines `slft`, `ssub` and `splot`). What are your conclusions about the behavior of this active suspension system?

3. Partition the output z of the system into

$$z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix}; \quad z_1 = \begin{pmatrix} q_1 - q_0 \\ F \end{pmatrix}; \quad z_2 = \begin{pmatrix} \ddot{q}_2 \\ q_2 - q_1 \end{pmatrix}.$$

and let the weights on the signal components be as in the first part of this exercise. Suppose again that $\bar{k} = \bar{b} = 0$. Let $\mathcal{T}_i$, $i = 1, 2$ be the transfer function mapping $\tilde{q}_0 \mapsto \tilde{z}_i$. We wish to obtain insight in the achievable trade-offs between upper bounds of $\|\mathcal{T}_1\|_\infty$ and $\|\mathcal{T}_2\|_2$. To do this,

(a) Calculate the minimal achievable H_∞ norm of $\mathcal{T}_1$.

(b) Calculate the minimal achievable H_2 norm of $\mathcal{T}_2$.

(c) Calculate the minimal achievable H_2 norm of $\mathcal{T}_2$ subject to the bound $\|\mathcal{T}_1\|_\infty < \gamma_1$ where γ_1 takes some[1] values in the interval $[0.15, 0.30]$.

Make a plot of the Pareto optimal performances, i.e, plot the minimal achievable H_2 norm of $\mathcal{T}_2$ as function of γ_1. (See the routine `hinfmix` for details).

4. (*This part is optional*). We now incorporate the uncertainty in the parameters k_2 and b_2. Let $\bar{k} = 50000$ and $\bar{b} = 5000$. Synthesize a gain scheduled H_∞ controller for the uncertain system such that

- the closed-loop system is stable for all parameter trajectories $k_2(t)$ and $b_2(t)$ satisfying (4.9.3) and eqrefk2unc.

- the worst-case H_∞ performance from $\tilde{q}_0$ to $\tilde{z}$ does not exceed the level $\gamma > 0$.

To do this, use the same weighting functions as in part (a) of this exercise. (See the routine `hinfgs`). Construct the parameter dependent closed-loop system using `slft` and create an .m-file which generates a realization of the parameters $b_2(t)$ and $k_2(t)$ according to their specification. Make a time simulation on the time interval $[0, 5]$ (in seconds) of the step-response of this system with time-varying parameters. (Use `pdsimul` for the latter).

---

[1]Slightly depending on your patience and the length of your coffee breaks, I suggest about 5.

Chapter 5

Linear Parameterically Varying Systems

Linear parameterically varying (LPV) systems are linear systems whose describing matrices depend on a time-varying parameter such that both the parameter itself and its rate of variation are known to be contained in pre-specified sets.

In robust control, the goal is to find *one fixed* controller that achieves robust stability and robust performance for all possible parameter variations, irrespective of which specific parameter curve does indeed perturb the system.

Instead, in LPV control, it is assumed that the parameter (and, possibly, its rate of variation), although not known a priori, is (are) on-line measurable. Hence the actual parameter value (and its derivative) can be used as extra information to control the system - the controller will turn out to depend on the parameter as well. We will actually choose also an LPV structure for the controller to be designed.

We would like to stress the decisive distinction to the control of time-varying systems: In the standard techniques to controlling time-varying systems, the model description is assumed to be known a priori over the whole time interval $[0, \infty)$. In LPV control, the model is assumed to be known, at time instant t, only over the interval $[0, t]$.

The techniques we would like to develop closely resemble those for robust control we have investigated earlier. It is possible to apply them

- to control certain classes of nonlinear systems

- to provide a systematic procedure for gain-scheduling

177

with guarantees for stability and performance.

Before we explore these applications in more detail we would like to start presenting the available problem setups and solution techniques to LPV control.

5.1 General Parameter Dependence

Suppose that $\boldsymbol{\delta}_c, \dot{\boldsymbol{\delta}}_c \subset \mathbb{R}^m$ are two parameter sets such that

$$\boldsymbol{\delta}_c \times \dot{\boldsymbol{\delta}}_c \ \text{ is compact,}$$

and that the matrix valued function

$$\left(\begin{array}{c|cc} A(p) & B_p(p) & B(p) \\ \hline C_p(p) & D_p(p) & E(p) \\ C(p) & F(p) & 0 \end{array} \right) \quad \text{is continuous in } \ p \in \boldsymbol{\delta}_c. \qquad (5.1.1)$$

Consider the *Linear Parameterically Varying (LPV) system* that is described as

$$\left(\begin{array}{c} \dot{x} \\ \hline z_p \\ y \end{array} \right) = \left(\begin{array}{c|cc} A(\delta(t)) & B_p(\delta(t)) & B(\delta(t)) \\ \hline C_p(\delta(t)) & D_p(\delta(t)) & E(\delta(t)) \\ C(\delta(t)) & F(\delta(t)) & 0 \end{array} \right) \left(\begin{array}{c} x \\ \hline w_p \\ u \end{array} \right), \quad \delta(t) \in \boldsymbol{\delta}_c, \ \ \dot{\delta}(t) \in \dot{\boldsymbol{\delta}}_c. \qquad (5.1.2)$$

We actually mean the family of systems that is obtained if letting $\delta(.)$ vary in the set of continuously differentiable parameter curves

$$\delta : [0, \infty) \to \mathbb{R}^m \ \text{ with } \ \delta(t) \in \boldsymbol{\delta}_c, \ \ \dot{\delta}(t) \in \dot{\boldsymbol{\delta}}_c \ \text{ for all } \ t \geqslant 0.$$

The signals admit the same interpretations as in Chapter 4: u is the control input, y is the measured output available for control, and $w_p \to z_p$ denotes the performance channel.

In *LPV control*, it is assumed that the parameter $\delta(t)$ is on-line measurable. Hence the actual value of $\delta(t)$ can be taken as extra information for the controller to achieve the desired design goal.

In view of the specific structure of the system description, we assume that the controller admits a similar structure. In fact, an *LPV controller* is defined by functions

$$\left(\begin{array}{cc} A_c(p) & B_c(p) \\ C_c(p) & D_c(p) \end{array} \right) \quad \text{that are continuous in } \ p \in \boldsymbol{\delta}_c \qquad (5.1.3)$$

as

$$\left(\begin{array}{c} \dot{x}_c \\ u \end{array} \right) = \left(\begin{array}{cc} A_c(\delta(t)) & B_c(\delta(t)) \\ C_c(\delta(t)) & D_c(\delta(t)) \end{array} \right) \left(\begin{array}{c} x_c \\ y \end{array} \right)$$

with the following interpretation: It evolves according to linear dynamics that are defined at time-instant t via the actually measured value of $\delta(t)$.

Note that a robust controller would be simply defined with a constant matrix

$$\left(\begin{array}{cc} A_c & B_c \\ C_c & D_c \end{array} \right)$$

that does not depend on δ what clarifies the difference between robust controllers and LPV controllers.

The controlled system admits the description

$$\left(\begin{array}{c} \dot{\xi} \\ z_p \end{array} \right) = \left(\begin{array}{cc} \mathcal{A}(\delta(t)) & \mathcal{B}(\delta(t)) \\ \mathcal{C}(\delta(t)) & \mathcal{D}(\delta(t)) \end{array} \right) \left(\begin{array}{c} \xi \\ w_p \end{array} \right), \quad \delta(t) \in \boldsymbol{\delta}_c, \quad \dot{\delta}(t) \in \dot{\boldsymbol{\delta}}_c \qquad (5.1.4)$$

where the function

$$\left(\begin{array}{cc} \mathcal{A}(p) & \mathcal{B}(p) \\ \mathcal{C}(p) & \mathcal{D}(p) \end{array} \right) \quad \text{is continuous in} \quad p \in \boldsymbol{\delta}_c$$

and given as

$$\left(\begin{array}{cc|c} A(p) + B(p)D_c(p)C(p) & B(p)C_c(p) & B_p(p) + B(p)D_c(p)F(p) \\ B_c(p)C(p) & A_c(p) & B_c(p)F(p) \\ \hline C_p(p) + E(p)D_c(p)C(p) & E(p)C_c(p) & D_p(p) + E(p)D_c(p)F(p) \end{array} \right)$$

or

$$\left(\begin{array}{cc|c} A(p) & 0 & B_p(p) \\ 0 & 0 & 0 \\ \hline C_p(p) & 0 & D_p(p) \end{array} \right) + \left(\begin{array}{cc} 0 & B(p) \\ I & 0 \\ 0 & E(p) \end{array} \right) \left(\begin{array}{cc} A_c(p) & B_c(p) \\ C_c(p) & D_c(p) \end{array} \right) \left(\begin{array}{cc|c} 0 & I & 0 \\ C(p) & 0 & F(p) \end{array} \right).$$

To evaluate performance, we concentrate again on the quadratic specification

$$\int_0^\infty \left(\begin{array}{c} w(t) \\ z(t) \end{array} \right)^T P_p \left(\begin{array}{c} w(t) \\ z(t) \end{array} \right) dt \leqslant -\epsilon \|w\|_2^2 \qquad (5.1.5)$$

with an index

$$P_p = \left(\begin{array}{cc} Q_p & S_p \\ S_p^T & R_p \end{array} \right), \quad R_p \geqslant 0 \quad \text{that has the inverse} \quad \tilde{P}_p^{-1} = \left(\begin{array}{cc} \tilde{Q}_p & \tilde{S}_p \\ \tilde{S}_p^T & \tilde{R}_p \end{array} \right), \quad \tilde{Q}_p \leqslant 0.$$

In order to abbreviate the formulation of the analysis result we introduce the following differential operator.

Definition 5.1 If $X : \boldsymbol{\delta}_c \ni p \to X(p) \in \mathbb{R}^{n \times n}$ is continuously differentiable, the continuous mapping

$$\partial X : \boldsymbol{\delta}_c \times \dot{\boldsymbol{\delta}}_c \to \mathbb{R}^{n \times n} \quad \text{is defined as} \quad \partial X(p, q) := \sum_{j=1}^m \frac{\partial X}{\partial p_j}(p) q_j.$$

Note that this definition is simply motivated by the fact that, along any continuously differentiable parameter curve $\delta(.)$, we have

$$\frac{d}{dt}X(\delta(t)) = \sum_{j=1}^{m} \frac{\partial X}{\partial p_j}(\delta(t))\dot{\delta}_j(t) = \partial X(\delta(t), \dot{\delta}(t)). \tag{5.1.6}$$

(We carefully wrote down the definitions and relations, and one should read all this correctly. X and ∂X are functions of the parameters $p \in \boldsymbol{\delta}_c$ and $q \in \dot{\boldsymbol{\delta}}_c$ respectively. In the definition of ∂X, no time-trajectories are involved. The definition of ∂X is just tailored to obtain the property (5.1.6) if plugging in a function of time.)

In view of the former discussion, the following analysis result comes as no surprise.

Theorem 5.2 *Suppose there exists a continuously differentiable $\mathcal{X}(p)$ defined for $p \in \boldsymbol{\delta}_c$ such that for all $p \in \boldsymbol{\delta}_c$ and $q \in \dot{\boldsymbol{\delta}}_c$ one has*

$$\mathcal{X}(p) > 0, \quad \begin{pmatrix} \partial\mathcal{X}(p,q) + \mathcal{A}(p)^T\mathcal{X}(p) + \mathcal{X}(p)\mathcal{A}(p) & \mathcal{X}(p)\mathcal{B}(p) \\ \mathcal{B}(p)^T\mathcal{X}(p) & 0 \end{pmatrix} +$$

$$+ \begin{pmatrix} 0 & I \\ \mathcal{C}(p) & \mathcal{D}(p) \end{pmatrix}^T P_p \begin{pmatrix} 0 & I \\ \mathcal{C}(p) & \mathcal{D}(p) \end{pmatrix} < 0. \tag{5.1.7}$$

Then there exists an $\epsilon > 0$ such that, for each parameter curve with $\delta(t) \in \boldsymbol{\delta}_c$ and $\dot{\delta}(t) \in \dot{\boldsymbol{\delta}}_c$, the system (5.1.4) is exponentially stable and satisfies (5.1.5) if the initial condition is zero and if $w_p \in L_2$.

In view of our preparations the proof is a simple exercise that is left to the reader.

We can now use the same procedure as for LTI systems to arrive at the corresponding synthesis result. It is just required to obey that all the matrices are actually functions of $p \in \boldsymbol{\delta}_c$ or of $(p,q) \in \boldsymbol{\delta}_c \times \dot{\boldsymbol{\delta}}_c$. If partitioning

$$\mathcal{X} = \begin{pmatrix} X & U \\ U^T & * \end{pmatrix}, \quad \mathcal{X}^{-1} = \begin{pmatrix} Y & V \\ V^T & * \end{pmatrix},$$

we can again assume w.l.o.g. that U, V have full row rank. (Note that this requires the compactness hypothesis on $\boldsymbol{\delta}_c$ and $\dot{\boldsymbol{\delta}}_c$. Why?) With

$$\mathcal{Y} = \begin{pmatrix} Y & I \\ V^T & 0 \end{pmatrix} \quad \text{and} \quad \mathcal{Z} = \begin{pmatrix} I & 0 \\ X & U \end{pmatrix}$$

we obtain the identities

$$\mathcal{Y}^T\mathcal{X} = \mathcal{Z} \quad \text{and} \quad I - XY = UV^T.$$

If we apply the differential operator ∂ to the first functional identity, we arrive at $(\partial\mathcal{Y})^T\mathcal{X} + \mathcal{Y}^T(\partial\mathcal{X}) = \partial\mathcal{Z}$. (Do the simple calculations. Note that ∂ is not the usual differentiation such that you cannot apply the standard product rule.) If we right-multiply $\mathcal{Y}$, this leads to

$$\mathcal{Y}^T(\partial\mathcal{X})\mathcal{Y} = (\partial\mathcal{Z})\mathcal{Y} - (\partial\mathcal{Y})^T\mathcal{Z}^T = \begin{pmatrix} 0 & 0 \\ \partial X & \partial U \end{pmatrix}\begin{pmatrix} Y & I \\ V^T & 0 \end{pmatrix} - \begin{pmatrix} \partial Y & \partial V \\ 0 & 0 \end{pmatrix}\begin{pmatrix} I & X \\ 0 & U^T \end{pmatrix}$$

and hence to

$$\mathcal{Y}^T(\partial\mathcal{X})\mathcal{Y} = \begin{pmatrix} -\partial Y & -(\partial Y)X - (\partial V)U^T \\ (\partial X)Y + (\partial U)V^T & \partial X \end{pmatrix}.$$

If we introduce the transformed controller parameters

$$\begin{pmatrix} K & L \\ M & N \end{pmatrix} = \begin{pmatrix} U & XB \\ 0 & I \end{pmatrix}\begin{pmatrix} A_c & B_c \\ C_c & D_c \end{pmatrix}\begin{pmatrix} V^T & 0 \\ CY & I \end{pmatrix} + \begin{pmatrix} XAY & 0 \\ 0 & 0 \end{pmatrix} +$$
$$+ \begin{pmatrix} (\partial X)Y + (\partial U)V^T & 0 \\ 0 & 0 \end{pmatrix},$$

a brief calculation reveals that

$$\mathcal{Y}^T(\partial\mathcal{X} + \mathcal{A}^T\mathcal{X} + \mathcal{X}\mathcal{A})\mathcal{Y} = \begin{pmatrix} -\partial Y + \operatorname{sym}(AY + BM) & (A + BNC) + K^T \\ (A + BNC)^T + K & \partial X + \operatorname{sym}(AX + LC) \end{pmatrix}$$
$$\mathcal{Y}^T\mathcal{X}\mathcal{B} = \begin{pmatrix} B_p + BNF \\ XB_p + LF \end{pmatrix}, \quad C\mathcal{Y} = \begin{pmatrix} C_pY + EM & C_p + ENC \end{pmatrix}, \quad \mathcal{D} = D_p + ENF$$

where we used again the abbreviation $\operatorname{sym}(M) = M + M^T$. If compared to a parameter inde-
pendent Lyapunov function, we have modified the transformation to K by $(\partial X)Y + (\partial U)V^T$
in order to eliminate this extra term that appears from the congruence transformation of
$\partial\mathcal{X}$. If $\mathcal{X}$ is does not depend on p, $\partial\mathcal{X}$ vanishes identically and the original transformation
is recovered.

We observe that L, M, N are functions of $p \in \boldsymbol{\delta}_c$ only, whereas K also depends on $q \in \dot{\boldsymbol{\delta}}_c$.
In fact, this function has the structure

$$K(p, q) = K_0(p) + \sum_{i=1}^{m} K_i(p)q_i \tag{5.1.8}$$

(why?) and, hence, it is fully described by specifying

$$K_i(p), \quad i = 0, 1, \ldots, m$$

that depend, as well, on $p \in \boldsymbol{\delta}_c$ only.

Literally as in Theorem 4.2 one can now prove the following synthesis result for LPV systems.

Theorem 5.3 *If there exists an LPV controller defined by (5.1.3) and a continuously dif-
ferentiable $\mathcal{X}(.)$ defined for $p \in \boldsymbol{\delta}_c$ that satisfy (5.1.7), then there exist continuously differen-
tiable functions X, Y and continuous functions K_i, L, M, N defined on $\boldsymbol{\delta}_c$ such that, with
K given by (5.1.8), the inequalities*

$$\begin{pmatrix} Y & I \\ I & X \end{pmatrix} > 0 \tag{5.1.9}$$

and

$$\left(\begin{array}{cc|c} -\partial Y + \mathrm{sym}\,(AY + BM) & (A + BNC) + K^T & B_p + BNF \\ (A + BNC)^T + K & \partial X + \mathrm{sym}\,(AX + LC) & XB_p + LF \\ \hline (B_p + BNF)^T & (XB_p + LF)^T & 0 \end{array} \right) + $$

$$+ \left(\begin{array}{c} * \\ * \end{array} \right)^T P_p \left(\begin{array}{cc|c} 0 & 0 & I \\ C_pY + EM & C_p + ENC & D_p + ENF \end{array} \right) < 0 \quad (5.1.10)$$

hold on $\boldsymbol{\delta}_c \times \dot{\boldsymbol{\delta}}_c$. Conversely, suppose the continuously differentiable X, Y and the continuous K_i, defining K as in (5.1.8), L, M, N satisfy these synthesis inequalities. Then one can factorize $I - XY = UV^T$ with continuously differentiable square and nonsingular U, V, and

$$\mathcal{X} = \left(\begin{array}{cc} Y & V \\ I & 0 \end{array} \right)^{-1} \left(\begin{array}{cc} I & 0 \\ X & U \end{array} \right) \quad (5.1.11)$$

$$\left(\begin{array}{cc} A_c & B_c \\ C_c & D_c \end{array} \right) = \left(\begin{array}{cc} U & XB \\ 0 & I \end{array} \right)^{-1} \left(\begin{array}{cc} K - XAY - [(\partial X)Y + (\partial U)V^T] & L \\ M & N \end{array} \right) \left(\begin{array}{cc} V^T & 0 \\ CY & I \end{array} \right)^{-1} $$

$$(5.1.12)$$

render the analysis inequalities (5.1.7) satisfied.

Remark. Note that the formula (5.1.12) just emerges from the modified controller parameter transformation. We observe that the matrices B_c, C_c, D_c are functions of $p \in \boldsymbol{\delta}_c$ only. Due to the dependence of K on q and due to the extra term $U^{-1}[(\partial X)Y + (\partial U)V^T]V^{-T}$ in the formula for A_c, this latter matrix is a function that depends both on $p \in \boldsymbol{\delta}_c$ and $q \in \dot{\boldsymbol{\delta}}_c$. It has the same structure as K and can be written as

$$A_c(p, q) = A_0(p) + \sum_{i=1}^{m} A_i(p)q_i.$$

A straightforward calculation reveals that

$$A_i = U^{-1}[K_i V^{-T} - \frac{\partial X}{\partial p_i} Y V^{-T} - \frac{\partial U}{\partial p_i}], \quad i = 1, \ldots, m.$$

Hence, to implement this controller, one indeed requires not only to measure $\delta(t)$ but also its rate of variation $\dot{\delta}(t)$. However, one could possibly exploit the freedom in choosing U and V to render $A_i = 0$ such that A_c does not depend on q any more. Recall that U and V need to be related by $I - XY = UV^T$; hence let us choose

$$V^T := U^{-1}(I - XY).$$

This leads to

$$A_i = U^{-1}[(K_i - \frac{\partial X}{\partial p_i} Y)(I - XY)^{-1}U - \frac{\partial U}{\partial p_i}], \quad i = 1, \ldots, m.$$

Therefore, U should be chosen as a nonsingular solution of the system of first order partial differential equations

$$\frac{\partial U}{\partial p_i}(p) = [K_i(p) - \frac{\partial X}{\partial p_i}(p)Y(p)](I - X(p)Y(p))^{-1}U(p), \quad j = 1, \ldots, m.$$

This leads to $A_i = 0$ such that the implementation of the LPV controller does not require any on-line measurements of the rate of the parameter variations. First order partial differential equations can be solved by the method of characteristics [11]. We cannot go into further details at this point.

In order to construct a controller that solves the LPV problem, one has to verify the solvability of the synthesis inequalities in the unknown *functions* X, Y, K_i, L, M, N, and for designing a controller, one has to find functions that solve them.

However, standard algorithms do not allow to solve functional inequalities directly. Hence we need to include a discussion of how to reduce these functional inequalities to finitely many LMI's in real variables.

First step. Since $q \in \dot{\boldsymbol{\delta}}_c$ enters the inequality (5.1.10) affinely, we can replace the set $\dot{\boldsymbol{\delta}}_c$, if convex, by its extreme points. Let us make the, in practice non-restrictive, assumption that this set has finitely many generators:

$$\dot{\boldsymbol{\delta}}_c = \mathrm{co}\{\dot{\delta}^1, \ldots, \dot{\delta}^k\}.$$

Solving (5.1.9)-(5.1.10) over $(p,q) \in \boldsymbol{\delta}_c \times \dot{\boldsymbol{\delta}}_c$ is equivalent to solving (5.1.9)-(5.1.10) for

$$p \in \boldsymbol{\delta}_c, \quad q \in \{\dot{\delta}^1, \ldots, \dot{\delta}^k\}. \tag{5.1.13}$$

Second step. Instead of searching over the set of all continuous functions, we restrict the search to a finite dimensional subspace thereof, as is standard in Ritz-Galerkin techniques. Let us hence choose basis functions

$$f_1(.), \ldots, f_l(.) \quad \text{that are continuously differentiable on} \quad \boldsymbol{\delta}_c.$$

Then all the functions to be found are assumed to belong to the subspace spanned by the functions f_j. This leads to the Ansatz

$$X(p) = \sum_{j=1}^{l} X_j f_j(p), \quad Y(p) = \sum_{j=1}^{l} Y_j f_j(p)$$

$$K_i(p) = \sum_{j=1}^{l} K_j^i f_j(p), \quad i = 0, 1, \ldots, m,$$

$$L(p) = \sum_{j=1}^{l} L_j f_j(p), \quad M(p) = \sum_{j=1}^{l} M_j f_j(p), \quad N(p) = \sum_{j=1}^{l} N_j f_j(p).$$

We observe

$$\partial X(p,q) = \sum_{j=1}^{l} X_j \, \partial f_j(p,q), \quad \partial Y(p,q) = \sum_{j=1}^{l} Y_j \, \partial f_j(p,q).$$

If we plug these formulas into the inequalities (5.1.9)-(5.1.10), we observe that all the co-efficient matrices enter affinely. After this substitution, (5.1.9)-(5.1.10) turns out to be a family of linear matrix inequalities in the

$$\text{matrix variables} \quad X_j, Y_j, K_j^i, L_j, M_j, N_j$$

that is parameterized by (5.1.13). The variables of this system of LMI's are now real numbers; however, since the parameter p still varies in the infinite set $\boldsymbol{\delta}_c$, we have to solve infinitely many LMI's. This is, in fact, a so-called semi-infinite (not infinite dimensional as often claimed) convex optimization problem.

Third step. To reduce the semi-infinite system of LMI's to finitely many LMI's, the presently chosen route is to just fix a *finite subset*

$$\boldsymbol{\delta}_{\text{finite}} \subset \boldsymbol{\delta}_c$$

and solve the LMI system in those points only. Hence the resulting family of LMI's is parameterized by

$$p \in \boldsymbol{\delta}_{\text{finite}} \quad \text{and} \quad q \in \{\dot{\delta}^1, \dots, \dot{\delta}^k\}.$$

We end up with a finite family of linear matrix inequalities in real valued unknowns that can be solved by standard algorithms. Since a systematic choice of points $\boldsymbol{\delta}_{\text{finite}}$ is obtained by gridding the parameter set, this last step is often called the gridding phase, and the whole procedure is said to be a gridding technique.

Remark on the second step. Due to Weierstraß' approximation theorem, one can choose a sequence of functions $f_1, f_2, \dots$ on $\boldsymbol{\delta}_c$ such that the union of the subspaces

$$\mathcal{S}_\nu = \text{span}\{f_1, \dots, f_\nu\}$$

is *dense* in the set of all continuously differentiable mappings on $\boldsymbol{\delta}_c$ with respect to the norm

$$\|f\| = \max\{|f(p)| \mid p \in \boldsymbol{\delta}_c\} + \sum_{j=1}^{m} \max\{|\frac{\partial f}{\partial p_j}(p)| \mid p \in \boldsymbol{\delta}_c\}.$$

This implies that, given any continuously differentiable g on $\boldsymbol{\delta}_c$ and any accuracy level $\epsilon > 0$, one can find an index ν_0 such that there exists an $f \in \mathcal{S}_{\nu_0}$ for which

$$\forall p \in \boldsymbol{\delta}_c, \ q \in \dot{\boldsymbol{\delta}}_c : \ |g(p) - f(p)| \leqslant \epsilon, \ |\partial g(p,q) - \partial f(p,q)| \leqslant \epsilon.$$

(Provide the details.) Functions in the subspace $\mathcal{S}_\nu$ hence approximate any function g and its image ∂g under the differential operator ∂ up to arbitrary accuracy, if the index ν is chosen sufficiently large.

Therefore, if (5.1.9)-(5.1.10) viewed as functional inequalities do have a solution, then they have a solution if restricting the search over the finite dimensional subspace $\mathcal{S}_\nu$ for sufficiently large ν, i.e., if incorporating sufficiently many basis functions. However, the number of basis functions determines the number of variables in the resulting LMI problem. To keep the number of unknowns small requires an efficient choice of the basis functions what is, in theory and practice, a difficult problem for which one can hardly give any general recipes.

Remark on the third step. By compactness of $\boldsymbol{\delta}_c$ and continuity of all functions, solving the LMI's for $p \in \boldsymbol{\delta}_c$ or for $p \in \boldsymbol{\delta}_{\text{finite}}$ is equivalent if only the points are chosen sufficiently dense. A measure of density is the infimal ϵ such that the balls of radius ϵ around each of the finitely many points in $\boldsymbol{\delta}_{\text{finite}}$ already cover $\boldsymbol{\delta}_c$:

$$\boldsymbol{\delta}_c \subset \bigcup_{p_0 \in \boldsymbol{\delta}_{\text{finite}}} \{u \mid \|p - p_0\| \leqslant \epsilon\}.$$

If the data functions describing the system are also differentiable in δ, one can apply the mean value theorem to provide explicit estimates of the accuracy of the required approximation. Again, however, it is important to observe that the number of LMI's to solve depends on the number of grid-points; hence one has to keep this number small in order to avoid large LMI's.

Remark on extensions. Only slight adaptations are required to treat all the other performance specifications (such as bounds on the L_2-gain and on the analogue of the H_2-norm or generalized H_2-norm for time-varying systems) as well as the corresponding mixed problems as discussed in Chapter 4 in full generality. Note also that, for single-objective problems, the techniques to eliminate parameters literally apply; there is no need go into the details. In particular for solving gain-scheduling problems, it is important to observe that one can as well let the performance index depend on the measured parameter without any additional difficulty. As a designer, one can hence ask *different performance properties* in *different parameter ranges* what has considerable relevance in practical controller design.

Remark on robust LPV control. As another important extension we mention *robust LPV design*. It might happen that some parameters are indeed on-line measurable, whereas others have to be considered as unknown perturbations with which the controller cannot be scheduled. Again, it is straightforward to extend the robustness design techniques that have been presented in Chapter 4 from LTI systems and controllers to LPV systems and controllers. This even allows to include dynamic uncertainties if using IQC's to capture their properties. Note that the scalings that appear in such techniques constitute extra problem variables. In many circumstances it causes no extra technical difficulties to let these scalings also depend on the scheduling parameter what reduces the conservatism.

5.2   Affine Parameter Dependence

Suppose that the matrices (5.1.1) describing the system are *affine* functions on the set

$$\boldsymbol{\delta}_c = \mathrm{co}\{\delta^1, \ldots, \delta^k\}.$$

In that case we intend to search, as well, for an LPV controller that is defined with *affine* functions (5.1.3). Note that the describing matrices for the cosed-loop system are also *affine* in the parameter if

$$\begin{pmatrix} B \\ E \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} C & F \end{pmatrix} \quad \text{are parameter independent}$$

what is assumed from now on. Finally, we let $\mathcal{X}$ in Theorem 5.2 be *constant*.

Since $R_p \geqslant 0$, we infer that (5.1.7) is satisfied if and only if it holds for the generators $p = \delta^j$ of the set $\boldsymbol{\delta}_c$. Therefore, the analysis inequalities reduce to the finite set of LMI's

$$\mathcal{X} > 0, \quad \begin{pmatrix} \mathcal{A}(\delta^j)^T \mathcal{X} + \mathcal{X} \mathcal{A}(\delta^j) & \mathcal{X} \mathcal{B}(\delta^j) \\ \mathcal{B}(\delta^j)^T \mathcal{X} & 0 \end{pmatrix} +$$

$$+ \begin{pmatrix} 0 & I \\ \mathcal{C}(\delta^j) & \mathcal{D}(\delta^j) \end{pmatrix}^T P_p \begin{pmatrix} 0 & I \\ \mathcal{C}(\delta^j) & \mathcal{D}(\delta^j) \end{pmatrix} < 0 \quad \text{for all} \quad j = 1, \ldots, k.$$

Under the present structural assumptions, the affine functions $\begin{pmatrix} A_c & B_c \\ C_c & D_c \end{pmatrix}$ are transformed into affine functions $\begin{pmatrix} K & L \\ M & N \end{pmatrix}$ under the controller parameter transformation as considered in the previous section.

Then the synthesis inequalities (5.1.9)-(5.1.10) whose variables are the constant X and Y and the affine functions $\begin{pmatrix} K & L \\ M & N \end{pmatrix}$ turn out to be *affine* in the parameter p. This implies for the synthesis inequalities that we can replace the search over $\boldsymbol{\delta}_c$ without loss of generality by the search over the generators δ^j of this set. Therefore, solving the design problem amounts to testing whether the LMI's

$$\begin{pmatrix} Y & I \\ I & X \end{pmatrix} > 0$$

and

$$\begin{pmatrix} \mathrm{sym}\left(A(\delta^j)Y + BM(\delta^j)\right) & * & * \\ (A(\delta^j) + BN(\delta^j)C)^T + K(\delta^j) & \mathrm{sym}\left(A(\delta^j)X + L(\delta^j)C\right) & * \\ (B_p(\delta^j) + BN(\delta^j)F)^T & (XB_p(\delta^j) + L(\delta^j)F)^T & 0 \end{pmatrix} +$$

$$+ \begin{pmatrix} * \\ * \end{pmatrix}^T P_p \begin{pmatrix} 0 & 0 & I \\ C_p(\delta^j)Y + EM(\delta^j) & C_p(\delta^j) + EN(\delta^j)C & D_p(\delta^j) + EN(\delta^j)F \end{pmatrix} < 0$$

for $j = 1, \ldots, k$ admit a solution.

Since affine, the functions K, L, M, N are parameterized as

$$\left(\begin{array}{cc} K(p) & L(p) \\ M(p) & N(p) \end{array} \right) = \left(\begin{array}{cc} K_0 & L_0 \\ M_0 & N_0 \end{array} \right) + \sum_{i=1}^{m} \left(\begin{array}{cc} K_i & L_i \\ M_i & N_i \end{array} \right) p_i$$

with real matrices K_i, L_i, M_i, N_i. Hence, the synthesis inequalities form genuine linear matrix inequalities that can be solved by standard algorithms.

5.3 LFT System Description

Similarly as for our discussion of robust controller design, let us assume in this section that the LPV system is described as and LTI system

$$\left(\begin{array}{c} \dot{x} \\ z_u \\ z_p \\ y \end{array} \right) = \left(\begin{array}{c|ccc} A & B_u & B_p & B \\ \hline C_u & D_{uu} & D_{up} & E_u \\ C_p & D_{pu} & D_{pp} & E_p \\ \hline C & F_u & F_p & 0 \end{array} \right) \left(\begin{array}{c} x \\ w_u \\ w_p \\ u \end{array} \right)$$

(5.3.1)

in wich the parameter enters via the uncertainty channel $w_u \to z_u$ as

$$w_u(t) = \Delta(t) z_u(t), \quad \Delta(t) \in \mathbf{\Delta}_c.$$

(5.3.2)

The size and the structure of the possible parameter values $\Delta(t)$ is captured by the convex set

$$\mathbf{\Delta}_c := \mathrm{co}\{\Delta_1, ..., \Delta_N\}$$

whose generators Δ_j are given explicitly. We assume w.l.o.g. that $0 \in \mathbf{\Delta}_c$. As before, we concentrate on the quadratic performance specification with index P_p imposed on the performance channel $w_p \to z_p$.

Adjusted to the structure of (5.3.1)-(5.3.2), we assume that the measured parameter curve enters the controller also in a linear fractional fashion. Therefore, we assume that the to-be-designed LPV controller is defined by scheduling the LTI system

$$\dot{x}_c = A_c x_c + B_c \left(\begin{array}{c} y \\ w_c \end{array} \right), \quad \left(\begin{array}{c} u \\ z_c \end{array} \right) = C_c x_c + D_c \left(\begin{array}{c} y \\ w_c \end{array} \right)$$

(5.3.3)

with the actual parameter curve entering as

$$w_c(t) = \Delta_c(\Delta(t)) z_c(t).$$

(5.3.4)

The LPV controller is hence parameterized through the matrices A_c, B_c, C_c, D_c, and through a possibly non-linear matrix-valued scheduling function

$$\Delta_c(\Delta) \in \mathbb{R}^{n_r \times n_c} \quad \text{defined on} \quad \mathbf{\Delta}_c.$$
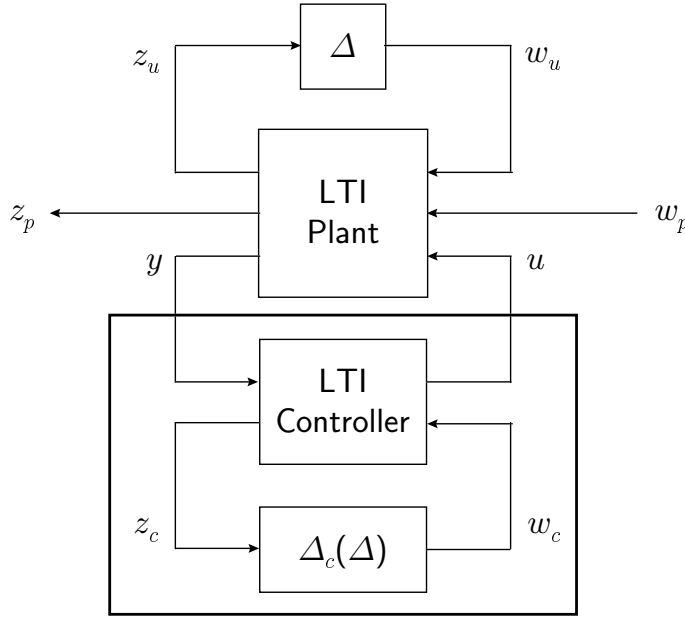
Figure 5.1: LPV system and LPV controller with LFT description

Figure 5.1 illustrates this configuration.

The goal is to construct an LPV controller such that, for all admissible parameter curves, the controlled system is exponentially stable and, the quadratic performance specification with index P_p for the channel $w_p \to z_p$ is satisfied.

The solution of this problem is approached with a simple trick. In fact, the controlled system can, alternatively, be obtained by scheduling the LTI system

$$
\begin{pmatrix} \dot{x} \\ z_u \\ z_c \\ z_p \\ y \\ w_c \end{pmatrix} = \left(\begin{array}{cc|cc|cc} A & B_u & 0 & B_p & B & 0 \\ C_u & D_{uu} & 0 & D_{up} & E_u & 0 \\ 0 & 0 & 0 & 0 & 0 & I_{n_c} \\ \hline C_p & D_{pu} & 0 & D_{uu} & E_p & 0 \\ C & F_u & 0 & F_p & 0 & 0 \\ 0 & 0 & I_{n_r} & 0 & 0 & 0 \end{array} \right) \begin{pmatrix} x \\ w_u \\ w_c \\ w_p \\ u \\ z_c \end{pmatrix}
\tag{5.3.5}
$$

with the parameter as

$$
\begin{pmatrix} w_1 \\ w_c \end{pmatrix} = \begin{pmatrix} \Delta(t) & 0 \\ 0 & \Delta_c(\Delta(t)) \end{pmatrix} \begin{pmatrix} z_1 \\ z_c \end{pmatrix},
\tag{5.3.6}
$$

and then controlling this parameter dependent system with the LTI controller (5.3.3). Alternatively, we can interconnect the LTI system (5.3.5) with the LTI controller (5.3.3) to
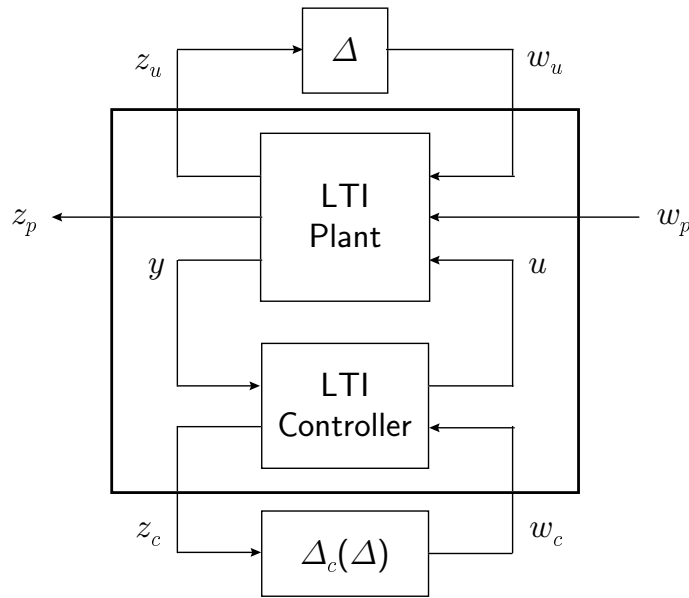
Figure 5.2: LPV system and LPV controller: Alternative Interpretation

arrive at the LTI system

$$
\begin{pmatrix} \dot{x} \\ z_u \\ z_c \\ z_p \end{pmatrix} = \left(\begin{array}{c|ccc} \mathcal{A} & \mathcal{B}_u & \mathcal{B}_c & \mathcal{B}_p \\ \hline \mathcal{C}_u & \mathcal{D}_{uu} & \mathcal{D}_{uc} & \mathcal{D}_{up} \\ \mathcal{C}_c & \mathcal{D}_{cu} & \mathcal{D}_{cc} & \mathcal{D}_{cp} \\ \mathcal{C}_p & \mathcal{D}_{pu} & \mathcal{D}_{pc} & \mathcal{D}_{pp} \end{array} \right) \begin{pmatrix} x \\ w_u \\ w_c \\ w_p \end{pmatrix}, \tag{5.3.7}
$$

and then re-connect the parameter as (5.3.6). This latter interconnection order is illustrated in Figure 5.2.

Note that (5.3.5) is an extension of the original system (5.3.1) with an additional uncertainty channel $w_c \to z_c$ and with an additional control channel $z_c \to w_c$; the number n_r and n_c of the components of w_c and z_c dictate the size of the identity matrices I_{n_r} and I_{n_c} that are indicated by their respective indices.

Once the scheduling function $\Delta_c(\Delta)$ has been fixed, it turns out that (5.3.3) is a *robust controller* for the system (5.3.5) with uncertainty (5.3.6). The genuine robust control problem in which the parameter is not measured on-line would relate to the situation that $n_r = 0$ and $n_c = 0$ such that (5.3.5) and (5.3.1) are identical. In LPV control we have the extra freedom of being able to first extend the system as in (5.3.5) and design for this extended system a robust controller. It will turn out that this extra freedom will render the corresponding synthesis inequalities convex.

Before we embark on a solution of the LPV problem, let us include some further comments on the corresponding genuine robust control problem. We have seen in section 4.7.1 that the search for a robust controller leads to the problem of having to solve the matrix inequalities

$$
\boldsymbol{X}(v) > 0, \quad
\begin{pmatrix} * \\ * \\ * \\ * \\ * \\ * \end{pmatrix}^T
\left(\begin{array}{cc|cc|cc} 0 & I & 0 & 0 & 0 & 0 \\ I & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & Q & S & 0 & 0 \\ 0 & 0 & S^T & R & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & Q_p & S_p \\ 0 & 0 & 0 & 0 & S_p^T & R_p \end{array} \right)
\left(\begin{array}{ccc} I & 0 & 0 \\ \boldsymbol{A}(v) & \boldsymbol{B}_u(v) & \boldsymbol{B}_p(v) \\ \hline 0 & I & 0 \\ \boldsymbol{C}_u(v) & \boldsymbol{D}_{uu}(v) & \boldsymbol{D}_{up}(v) \\ \hline 0 & 0 & I \\ \boldsymbol{C}_p(v) & \boldsymbol{D}_{pu}(v) & \boldsymbol{D}_{pp}(v) \end{array} \right) < 0
$$

$$
\begin{pmatrix} \Delta \\ I \end{pmatrix}^T \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} \begin{pmatrix} \Delta \\ I \end{pmatrix} > 0 \quad \text{for all} \ \ \Delta \in \boldsymbol{\Delta}_c
$$

in the parameter v and in the multiplier $P = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix}$.

Recall from our earlier discussion that one of the difficulties is a numerical tractable parameterization of the set of multipliers. This was the reason to introduce, at the expense of conservatism, the following subset of multipliers that admits a description in terms of finitely many LMI's:

$$
\boldsymbol{P} := \left\{ P = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix} \mid Q < 0, \ \begin{pmatrix} \Delta_j \\ I \end{pmatrix}^T P \begin{pmatrix} \Delta_j \\ I \end{pmatrix} > 0 \ \ \text{for} \ \ j = 1, \ldots, N \right\}. \quad (5.3.8)
$$

Even after confining the search to v and $P \in \boldsymbol{P}$, no technique is known how to solve the resulting still non-convex synthesis inequalities by standard algorithms.

In contrast to what we have seen for state-feedback design, the same is true of the dual inequalities that read as

$$
\boldsymbol{X}(v) > 0, \quad
\begin{pmatrix} * \\ * \\ * \\ * \\ * \\ * \end{pmatrix}^T
\left(\begin{array}{cc|cc|cc} 0 & X & 0 & 0 & 0 & 0 \\ X & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & \tilde{Q} & \tilde{S} & 0 & 0 \\ X & 0 & \tilde{S}^T & \tilde{R} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & \tilde{Q}_p & \tilde{S}_p \\ 0 & 0 & 0 & 0 & \tilde{S}_p^T & \tilde{R}_p \end{array} \right)
\left(\begin{array}{ccc} -\boldsymbol{A}(v)^T & -\boldsymbol{C}_u(v)^T & -\boldsymbol{C}_p(v)^T \\ I & 0 & 0 \\ \hline -\boldsymbol{B}_u(v)^T & -\boldsymbol{D}_{uu}(v)^T & -\boldsymbol{D}_{pu}(v)^T \\ 0 & I & 0 \\ \hline -\boldsymbol{B}_p(v)^T & -\boldsymbol{D}_{up}(v)^T & -\boldsymbol{D}_{pp}(v)^T \\ 0 & 0 & I \end{array} \right) > 0
$$

$$
\begin{pmatrix} I \\ -\Delta^T \end{pmatrix}^T \begin{pmatrix} \tilde{Q} & \tilde{S} \\ \tilde{S}^T & \tilde{R} \end{pmatrix} \begin{pmatrix} I \\ -\Delta^T \end{pmatrix} < 0 \quad \text{for all} \ \ \Delta \in \boldsymbol{\Delta}_c.
$$

Again, even confining the search to the set of multipliers

$$
\tilde{\boldsymbol{P}} := \left\{ \tilde{P} = \begin{pmatrix} \tilde{Q} & \tilde{S} \\ \tilde{S}^T & \tilde{R} \end{pmatrix} \mid \tilde{R} > 0, \ \begin{pmatrix} I \\ -\Delta_j^T \end{pmatrix}^T \tilde{P} \begin{pmatrix} I \\ -\Delta_j^T \end{pmatrix} < 0 \ \ \text{for} \ \ j = 1, \ldots, N \right\} \quad (5.3.9)
$$

does not lead to a convex feasibility problem.

Since non-convexity is caused by the multiplication of functions that depend on v with the multipliers, one could be lead to the idea that it might help to eliminate as many of the variables that are involved in v as possible. We can indeed apply the technique exposed in Section 4.5.3 and eliminate K, L, M, N.

For that purpose one needs to compute basis matrices

$$\Phi = \begin{pmatrix} \Phi^1 \\ \Phi^2 \\ \Phi^3 \end{pmatrix} \quad \text{of} \quad \ker \begin{pmatrix} B^T & E_u^T & E_p^T \end{pmatrix} \quad \text{and} \quad \Psi = \begin{pmatrix} \Psi^1 \\ \Psi^2 \\ \Psi^3 \end{pmatrix} \quad \text{of} \quad \ker \begin{pmatrix} C & F_u & F_p \end{pmatrix}$$

respectively. After elimination, the synthesis inequalities read as

$$\begin{pmatrix} Y & I \\ I & X \end{pmatrix} > 0, \tag{5.3.10}$$

$$\Psi^T \begin{pmatrix} I & 0 & 0 \\ A & B_u & B_p \\ 0 & I & 0 \\ C_u & D_{uu} & D_{up} \\ 0 & 0 & I \\ C_p & D_{pu} & D_{pp} \end{pmatrix}^T \left(\begin{array}{cc|cc|cc} 0 & X & 0 & 0 & 0 & 0 \\ X & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & Q & S & 0 & 0 \\ 0 & 0 & S^T & R & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & Q_p & S_p \\ 0 & 0 & 0 & 0 & S_p^T & R_p \end{array} \right) \begin{pmatrix} I & 0 & 0 \\ A & B_u & B_p \\ 0 & I & 0 \\ C_u & D_{uu} & D_{up} \\ 0 & 0 & I \\ C_p & D_{pu} & D_{pp} \end{pmatrix} \Psi < 0, \tag{5.3.11}$$

$$\Phi^T \begin{pmatrix} -A^T & -C_u^T & -C_p^T \\ I & 0 & 0 \\ -B_u^T & -D_{uu}^T & -D_{pu}^T \\ 0 & I & 0 \\ -B_p^T & -D_{pu}^T & -D_{pp}^T \\ 0 & 0 & I \end{pmatrix}^T \left(\begin{array}{cc|cc|cc} 0 & X & 0 & 0 & 0 & 0 \\ X & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & \tilde{Q} & \tilde{S} & 0 & 0 \\ 0 & 0 & \tilde{S}^T & \tilde{R} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & \tilde{Q}_p & \tilde{S}_p \\ 0 & 0 & 0 & 0 & \tilde{S}_p^T & \tilde{R}_p \end{array} \right) \begin{pmatrix} -A^T & -C_u^T & -C_p^T \\ I & 0 & 0 \\ -B_u^T & -D_{uu}^T & -D_{pu}^T \\ 0 & I & 0 \\ -B_p^T & -D_{pu}^T & -D_{pp}^T \\ 0 & 0 & I \end{pmatrix} \Phi > 0$$

$$\tag{5.3.12}$$

in the variables X, Y, and in the multiplier P and $\tilde{P}$ that are coupled as

$$\tilde{P} = \begin{pmatrix} \tilde{Q} & \tilde{S} \\ \tilde{S}^T & \tilde{R} \end{pmatrix} = \begin{pmatrix} Q & S \\ S^T & R \end{pmatrix}^{-1} = P^{-1}. \tag{5.3.13}$$

Hence, after elimination, it turns out that the inequalities (5.3.10)-(5.3.12) are indeed affine in the unknowns X, Y, P and $\tilde{P}$. Unfortunately, non-convexity re-appears through the coupling (5.3.13) of the multipliers P and $\tilde{P}$.

Let us now turn back to the LPV problem where we allow, via the scheduling function $\Delta_c(\Delta)$ in the controller, extra freedom in the design process.

For guaranteeing stability and performance of the controlled system, we employ *extended*

multipliers adjusted to the extended uncertainty structure (5.3.6) that are given as

$$P_e = \left(\begin{array}{c|c} Q_e & S_e \\ \hline S_e^T & R_e \end{array} \right) = \left(\begin{array}{cc|cc} Q & Q_{12} & S & S_{12} \\ Q_{21} & Q_{22} & S_{21} & S_{22} \\ \hline * & * & R & R_{12} \\ * & * & R_{21} & R_{22} \end{array} \right) \quad \text{with} \ \ Q_e < 0, \ \ R_e > 0 \qquad (5.3.14)$$

and that satisfy

$$\left(\begin{array}{cc} \Delta & 0 \\ 0 & \Delta_c(\Delta) \\ I & 0 \\ 0 & I \end{array} \right) P_e \left(\begin{array}{cc} \Delta & 0 \\ 0 & \Delta_c(\Delta) \\ I & 0 \\ 0 & I \end{array} \right) > 0 \ \ \text{for all} \ \ \Delta \in \boldsymbol{\Delta}. \qquad (5.3.15)$$

The corresponding dual multipliers $\tilde{P}_e = P_e^{-1}$ are partitioned similarly as

$$\tilde{P}_e = \left(\begin{array}{c|c} \tilde{Q}_e & \tilde{S}_e \\ \hline \tilde{S}_e^T & \tilde{R}_e \end{array} \right) = \left(\begin{array}{cc|cc} \tilde{Q} & \tilde{Q}_{12} & \tilde{S} & \tilde{S}_{12} \\ \tilde{Q}_{21} & \tilde{Q}_{22} & \tilde{S}_{21} & \tilde{S}_{22} \\ \hline * & * & \tilde{R} & \tilde{R}_{12} \\ * & * & \tilde{R}_{21} & \tilde{R}_{12} \end{array} \right) \quad \text{with} \ \ \tilde{Q}_e < 0, \ \ \tilde{R}_e > 0 \qquad (5.3.16)$$

and they satisfy

$$\left(\begin{array}{cc} I & 0 \\ 0 & I \\ -\Delta^T & 0 \\ 0 & -\Delta_c(\Delta)^T \end{array} \right)^T P_e \left(\begin{array}{cc} I & 0 \\ 0 & I \\ -\Delta^T & 0 \\ 0 & -\Delta_c(\Delta)^T \end{array} \right) > 0 \ \ \text{for all} \ \ \Delta \in \boldsymbol{\Delta}.$$

As indicated by our notation, we observe that

$$\left(\begin{array}{cc} Q & S \\ S^T & R \end{array} \right) \in \boldsymbol{P} \ \ \text{and} \ \ \left(\begin{array}{cc} \tilde{Q} & \tilde{S} \\ \tilde{S}^T & \tilde{R} \end{array} \right) \in \tilde{\boldsymbol{P}}$$

for the corresponding sub-matrices of P_e and $\tilde{P}_e$ respectively.

If we recall the description (5.3.6)-(5.3.7) of the controlled LPV system, the desired exponential stability and quadratic performance property is satisfied if we can find a Lyapunov matrix $\mathcal{X}$ and an extended scaling P_e with (5.3.14)-(5.3.15) such that

$$\mathcal{X} > 0, \quad \left(\begin{array}{c} * \\ * \\ * \\ * \\ * \\ * \\ * \\ * \end{array} \right)^T \left(\begin{array}{cc|cccc|cc} 0 & \mathcal{X} & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathcal{X} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & Q & Q_{12} & S & S_{12} & 0 & 0 \\ 0 & 0 & Q_{21} & Q_{22} & S_{21} & S_{22} & 0 & 0 \\ 0 & 0 & * & * & R & R_{12} & 0 & 0 \\ 0 & 0 & * & * & R_{21} & R_{22} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & Q_p & S_p \\ 0 & 0 & 0 & 0 & 0 & 0 & S_p^T & R_p \end{array} \right) \left(\begin{array}{cccc} I & 0 & 0 & 0 \\ \mathcal{A} & \mathcal{B}_u & \mathcal{B}_c & \mathcal{B}_p \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \\ \mathcal{C}_u & \mathcal{D}_{uu} & \mathcal{D}_{uc} & \mathcal{D}_{up} \\ \mathcal{C}_c & \mathcal{D}_{cu} & \mathcal{D}_{cc} & \mathcal{D}_{cp} \\ 0 & 0 & 0 & I \\ \mathcal{C}_p & \mathcal{D}_{pu} & \mathcal{D}_{pc} & \mathcal{D}_{pp} \end{array} \right) < 0.$$

$$(5.3.17)$$

We are now ready to formulate an LMI test for the existence of an LPV controller such that the controlled LPV system fulfills this latter analysis test.

Theorem 5.4 *The following statements are equivalent:*

1. *There exists a controller (5.3.3) and a scheduling function $\Delta_c(\Delta)$ such that the controlled system as described by (5.3.4)-(5.3.7) admits a Lyapunov matrix $\mathcal{X}$ and a multiplier (5.3.14)-(5.3.15) that satisfy (5.3.17).*

2. *There exist X, Y and multipliers $P \in \boldsymbol{P}$, $\tilde{P} \in \tilde{\boldsymbol{P}}$ that satisfy the linear matrix inequalities (5.3.10)-(5.3.12).*

Proof.   Let us first prove $1 \Rightarrow 2$. We can apply the technique as described in Section 4.5.3 to eliminate the controller parameters in the inequality (5.3.17). According to Corollary 4.11, this leads to the coupling condition (4.5.24) and to the two synthesis inequalities (4.5.25)-(4.5.26). The whole point is to show that the latter two inequalities can indeed be simplified to (5.3.11)-(5.3.12). Let us illustrate this simplification for the first inequality only since a duality argument leads to the same conclusions for the second one.

With

$$\Psi_e = \begin{pmatrix} \Psi^1 \\ \Psi^2 \\ 0 \\ \Psi^3 \end{pmatrix} \quad \text{as a basis matrix of } \ker \begin{pmatrix} C & F_u & 0 & F_p \\ 0 & 0 & I_{n_r} & 0 \end{pmatrix},$$

the inequality that corresponds to (4.5.24) reads as

$$\Psi_e^T \begin{pmatrix} * \\ * \\ * \\ * \\ * \\ * \\ * \\ * \end{pmatrix}^T \left(\begin{array}{cc|cccc|cc} 0 & X & 0 & 0 & 0 & 0 & 0 & 0 \\ X & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & Q & Q_{12} & S & S_{12} & 0 & 0 \\ 0 & 0 & Q_{21} & Q_{22} & S_{21} & S_{22} & 0 & 0 \\ 0 & 0 & * & * & R & R_{12} & 0 & 0 \\ 0 & 0 & * & * & R_{21} & R_{22} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & Q_p & S_p \\ 0 & 0 & 0 & 0 & 0 & 0 & S_p^T & R_p \end{array} \right) \left(\begin{array}{cccc} I & 0 & 0 & 0 \\ A & B_u & 0 & B_p \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \\ C_u & D_{uu} & 0 & D_{up} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & I \\ C_p & D_{pu} & D_{pc} & D_{pp} \end{array} \right) \Psi_e < 0.$$

Due to the zero block in Ψ_e, it is obvious that this is the same as

$$\Psi^T \begin{pmatrix} * \\ * \\ * \\ * \\ * \\ * \\ * \\ * \end{pmatrix}^T \left(\begin{array}{cc|cccc|cc} 0 & X & 0 & 0 & 0 & 0 & 0 & 0 \\ X & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & Q & Q_{12} & S & S_{12} & 0 & 0 \\ 0 & 0 & Q_{21} & Q_{22} & S_{21} & S_{22} & 0 & 0 \\ 0 & 0 & * & * & R & R_{12} & 0 & 0 \\ 0 & 0 & * & * & R_{21} & R_{22} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & Q_p & S_p \\ 0 & 0 & 0 & 0 & 0 & 0 & S_p^T & R_p \end{array} \right) \left(\begin{array}{ccc} I & 0 & 0 \\ A & B_u & B_p \\ 0 & I & 0 \\ 0 & 0 & 0 \\ C_u & D_{uu} & D_{up} \\ 0 & 0 & 0 \\ 0 & 0 & I \\ C_p & D_{pu} & D_{pp} \end{array} \right) \Psi < 0.$$

The two zero block rows in the outer factors allow to simplify this latter inequality to (5.3.11), what finishes the proof of $1 \Rightarrow 2$.

The constructive proof of $2 \Rightarrow 1$ is more involved and proceeds in three steps. Let us assume that we have computed solutions X, Y and $P \in \boldsymbol{P}$, $\tilde{P} \in \tilde{\boldsymbol{P}}$ with (5.3.10)-(5.3.12).

First step: Extension of Scalings. Since $P \in \boldsymbol{P}$ and $\tilde{P} \in \tilde{\boldsymbol{P}}$, let us recall that we have

$$\begin{pmatrix} \Delta \\ I \end{pmatrix}^T P \begin{pmatrix} \Delta \\ I \end{pmatrix} > 0 \ \text{ and } \ \begin{pmatrix} I \\ -\Delta^T \end{pmatrix}^T \tilde{P} \begin{pmatrix} I \\ -\Delta^T \end{pmatrix} < 0 \ \text{ for all } \ \Delta \in \boldsymbol{\Delta}. \qquad (5.3.18)$$

Due to $0 \in \boldsymbol{\Delta}_c$, we get $R > 0$ and $\tilde{Q} < 0$. Hence we conclude for the diagonal blocks of P that $Q < 0$ and $R > 0$, and for the diagonal blocks of $\tilde{P}$ that $\tilde{Q} > 0$ and $\tilde{R} < 0$. If we introduce

$$Z = \begin{pmatrix} I \\ 0 \end{pmatrix} \ \text{ and } \ \tilde{Z} = \begin{pmatrix} 0 \\ I \end{pmatrix}$$

with the same row partition as P, these properties can be expressed as

$$Z^T P Z < 0, \ \tilde{Z}^T P \tilde{Z} > 0 \ \text{ and } \ Z^T \tilde{P} Z < 0, \ \tilde{Z}^T \tilde{P} \tilde{Z} > 0. \qquad (5.3.19)$$

If we observe that $\mathrm{im}(\tilde{Z})$ is the orthogonal complement of $\mathrm{im}(Z)$, we can apply the Dualization Lemma to infer

$$\tilde{Z}^T P^{-1} \tilde{Z} > 0, \ Z^T P^{-1} Z < 0 \ \text{ and } \ \tilde{Z}^T \tilde{P}^{-1} \tilde{Z} > 0, \ Z^T \tilde{P}^{-1} Z < 0. \qquad (5.3.20)$$

For the given P and $\tilde{P}$, we try to find an extension P_e with (5.3.14) such that the dual multiplier $\tilde{P}_e = P_e^{-1}$ is related to the given $\tilde{P}$ as in (5.3.16). After a suitable permutation, this amounts to finding an extension

$$\begin{pmatrix} P & T \\ T^T & T^T N T \end{pmatrix} \ \text{ with } \ \begin{pmatrix} \tilde{P} & * \\ * & * \end{pmatrix} = \begin{pmatrix} P & T \\ T^T & T^T N T \end{pmatrix}^{-1}, \qquad (5.3.21)$$

where the specific parameterization of the new blocks in terms of a non-singular matrix T and some symmetric N will turn out convenient. Such an extension is very simple to obtain. However, we also need to obey the positivity/negativity constraints in (5.3.14) that amount to

$$\begin{pmatrix} Z & 0 \\ 0 & Z \end{pmatrix}^T \begin{pmatrix} P & T \\ T^T & T^T N T \end{pmatrix} \begin{pmatrix} Z & 0 \\ 0 & Z \end{pmatrix} < 0 \qquad (5.3.22)$$

and

$$\begin{pmatrix} \tilde{Z} & 0 \\ 0 & \tilde{Z} \end{pmatrix}^T \begin{pmatrix} P & T \\ T^T & T^T N T \end{pmatrix} \begin{pmatrix} \tilde{Z} & 0 \\ 0 & \tilde{Z} \end{pmatrix} > 0. \qquad (5.3.23)$$

We can assume w.l.o.g. (perturb, if necessary) that $P - \tilde{P}^{-1}$ is non-singular. Then we set

$$N = (P - \tilde{P}^{-1})^{-1}$$

and observe that (5.3.21) holds for any non-singular T.

The main goal is to adjust T to render (5.3.22)-(5.3.23) satisfied. We will in fact construct the sub-blocks $T_1 = TZ$ and $T_2 = T\tilde{Z}$ of $T = (T_1 \ T_2)$. Due to (5.3.19), the conditions (5.3.22)-(5.3.23) read in terms of these blocks as (Schur)

$$T_1^T \left[N - Z(Z^T P Z)^{-1} Z^T \right] T_1 < 0 \ \text{ and } \ T_2^T \left[N - \tilde{Z}(\tilde{Z}^T P \tilde{Z})^{-1} \tilde{Z}^T \right] T_2 > 0. \qquad (5.3.24)$$

If we denote by $n_+(S)$, $n_-(S)$ the number of positive, negative eigenvalues of the symmetric matrix S, we hence have to calculate $n_-(N - Z(Z^T P Z)^{-1} Z^T)$ and $n_+(N - \tilde{Z}(\tilde{Z}^T P \tilde{Z})^{-1} \tilde{Z}^T)$. Simple Schur complement arguments reveal that

$$n_- \begin{pmatrix} Z^T P Z & Z^T \\ Z & N \end{pmatrix} = n_-(Z^T P Z) + n_-(N - Z(Z^T P Z)^{-1} Z^T) =$$

$$= n_-(N) + n_-(Z^T P Z - Z^T N^{-1} Z) = n_-(N) + n_-(Z^T \tilde{P}^{-1} Z).$$

Since $Z^T P Z$ and $Z^T \tilde{P}^{-1} Z$ have the same size and are both negative definite by (5.3.19) and (5.3.20), we conclude $n_-(Z^T P Z) = n_-(Z^T \tilde{P}^{-1} Z)$. This leads to

$$n_-(N - Z(Z^T P Z)^{-1} Z^T) = n_-(N).$$

Literally the same arguments will reveal

$$n_+(N - \tilde{Z}(\tilde{Z}^T P \tilde{Z})^{-1} \tilde{Z}^T) = n_+(N).$$

These two relations imply that there exist T_1, T_2 with $n_-(N)$, $n_+(N)$ columns that satisfy (5.3.24). Hence the matrix $T = (T_1 \ T_2)$ has $n_+(N) + n_-(N)$ columns. Since the number of rows of T_1, T_2, Z, $\tilde{Z}$, N are all identical, T is actually a *square* matrix. We can assume w.l.o.g. - by perturbing T_1 or T_2 if necessary - that the square matrix T is non-singular.

This finishes the construction of the extended multiplier (5.3.14). Let us observe that the dimensions of Q_{22}/R_{22} equal the number of columns of T_1/T_2 which are, in turn, given by the integers $n_-(N)/n_+(N)$.

Second Step: Construction of the scheduling function. Let us fix Δ and let us apply the Elimination Lemma to (5.3.15) with $\Delta_c(\Delta)$ viewed as the unknown. We observe that the solvability conditions of the Elimination Lemma just amount to the two inequalities (5.3.18). We conclude that for any $\Delta \in \mathbf{\Delta}$ one can indeed compute a $\Delta_c(\Delta)$ which satisfies (5.3.15).

Due to the structural simplicity, we can even provide an explicit formula which shows that $\Delta_c(\Delta)$ can be selected to depend smoothly on Δ. Indeed, by a straightforward Schur-complement argument, (5.3.15) is equivalent to

$$\left(\begin{array}{cc|cc} U_{11} & U_{12} & (W_{11} + \Delta)^T & W_{21}^T \\ U_{21} & U_{22} & W_{12}^T & (W_{22} + \Delta_c(\Delta))^T \\ \hline W_{11} + \Delta & W_{12} & V_{11} & V_{12} \\ W_{21} & W_{22} + \Delta_c(\Delta) & V_{21} & V_{22} \end{array} \right) > 0$$

for $U = R_e - S_e^T Q_e^{-1} S_e > 0$, $V = -Q_e^{-1} > 0$, $W = Q_e^{-1} S_e$. Obviously this can be rewritten to

$$\begin{pmatrix} U_{22} & * \\ W_{22} + \Delta_c(\Delta) & V_{22} \end{pmatrix} - \begin{pmatrix} U_{21} & W_{12}^T \\ W_{21} & V_{21} \end{pmatrix} \begin{pmatrix} U_{11} & (W_{11} + \Delta)^T \\ W_{11} + \Delta & V_{11} \end{pmatrix}^{-1} \begin{pmatrix} U_{12} & W_{21}^T \\ W_{12} & V_{12} \end{pmatrix} > 0$$

in which $\Delta_c(\Delta)$ only appears in the off-diagonal position. Since we are sure that there does indeed exist a $\Delta_c(\Delta)$ that renders the inequality satisfied, the diagonal blocks must be positive definite. If we then choose $\Delta_c(\Delta)$ such that the off-diagonal block vanishes, we have found a solution of the inequality; this leads to the following explicit formula

$$\Delta_c(\Delta) = -W_{22} + \begin{pmatrix} W_{21} & V_{21} \end{pmatrix} \begin{pmatrix} U_{11} & * \\ W_{11} + \Delta & V_{11} \end{pmatrix}^{-1} \begin{pmatrix} U_{12} \\ W_{12} \end{pmatrix}$$

for the scheduling function. We note that $\Delta_c(\Delta)$ has the dimension $n_-(N) \times n_+(N)$.

Third Step: LTI controller construction. After having constructed the scalings, the last step is to construct an LTI controller and Lyapunov matrix that render the inequality (5.3.17) satisfied. We are confronted with a standard nominal quadratic design problem of which we are sure that it admits a solution, and for which the controller construction proceed along the steps that have been intensively discussed in Chapter 4. ∎

We have shown that the LMI's that needed to be solved for designing an LPV controller are *identical* to those for designing a robust controller, with the only exception that the coupling condition (5.3.13) drops out. Therefore, the search for X and Y and for the multipliers $P \in \boldsymbol{P}$ and $\tilde{P} \in \tilde{\boldsymbol{P}}$ to satisfy (5.3.10)-(5.3.12) amounts to testing the feasibility of standard LMI's. Moreover, the controller construction in the proof of Theorem 5.4 is constructive. Hence we conclude that we have found a full solution to the quadratic performance LPV control problem (including L_2-gain and dissipativity specifications) for full block scalings P_e that satisfy $Q_e < 0$. The more interesting general case without this still restrictive negativity hypotheses is dealt with in future work.

Remarks.

- The proof reveals that the scheduling function $\Delta_c(\Delta)$ has a many rows/colums as there are negative/positive eigenvalues of $P - \tilde{P}^{-1}$ (if assuming w.l.o.g. that the latter is non-singular.) If it happens that $P - \tilde{P}^{-1}$ is positive or negative definite, there is no need to schedule the controller at all; we obtain a controller that solves the robust quadratic performance problem.

- Previous approaches to the LPV problem [1, 8, 21, 39] were based on $\Delta_c(\Delta) = \Delta$ such that the controller is scheduled with an identical copy of the parameters. These results were based on block-diagonal parameter matrices and multipliers that were as well assumed block-diagonal. The use of full block scalings [33] require the extension to a more general scheduling function that is - as seen a posteriori - a quadratic function of the parameter Δ.

- It is possible to extend the procedure to H_2-control and to the other performance specifications in these notes. However, this requires restrictive hypotheses on the system description. The extension to general mixed problems seems nontrivial and is open in its full generality.

5.4 A Sketch of Possible Applications

It is obvious how to apply robust or LPV control techniques in linear design: If the underlying system is affected, possibly in a nonlinear fashion, by some possibly time-varying parameter (such as varying resonance poles and alike), one could strive

- either for designing a robust controller if the actual parameter changes are not available as on-line information

- or for constructing an LPV controller if the parameter (and its rate of variation) can be measured on-line.

As such the presented techniques can be a useful extension to the nominal design specifications that have been considered previously.

In a brief final and informal discussion we would like to point out possible applications of robust and LPV control techniques to the control of nonlinear systems:

- They clearly apply if one can systematically embed a nonlinear system in a class of linear systems that admit an LPV parameterization.

- Even if it is required to perform a heuristic linearization step, they can improve classical gain-scheduling design schemes for nonlinear systems since they lead to a one-shot construction of a family of linear controllers.

5.4.1 From Nonlinear Systems to LPV Systems

In order to apply the techniques discussed in these notes to nonlinear systems, one uses variations of what is often called *global linearization*.

Consider a nonlinear system described by

$$\dot{x} = f(x) \tag{5.4.1}$$

where we assume that $f : \mathbb{R}^n \to \mathbb{R}^n$ is a smooth vector field.

If $f(0) = 0$, it is often possible to rewrite $f(x) = A(x)x$ with a smooth matrix valued mapping $A(.)$. If one can guarantee that the LPV system

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

is exponentially stable, we can conclude that the nonlinear system

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

has 0 as a globally exponentially stable equilibrium. Note that one can and should impose a priori bounds on the state-trajectories such as $x(t) \in M$ for some set M such that the stability of the LPV system only has to be assured for $\delta(t) \in M$; of course, one can then only conclude stability for trajectories of the nonlinear system that remain in M.

A slightly more general procedure allows to consider arbitrary system trajectories instead of equilibrium points (or constant trajectories) only. In fact, suppose $x_1(.)$ and $x_2(.)$ are two trajectories of (5.4.1). By the mean-value theorem, there exist

$$\eta_j(t) \in \mathrm{co}\{x_1(t), x_2(t)\}$$

such that

$$\dot{x}_1(t) - \dot{x}_2(t) = f(x_1(t)) - f(x_2(t)) = \begin{pmatrix} \frac{\partial f_1}{\partial x}(\eta_1(t)) \\ \vdots \\ \frac{\partial f_n}{\partial x}(\eta_n(t)) \end{pmatrix} (x_1(t) - x_2(t)).$$

Therefore, the increment $\xi(t) = x_1(t) - x_2(t)$ satisfies the LPV system

$$\dot{\xi}(t) = A(\eta_1(t), \ldots, \eta_n(t))\xi(t)$$

with parameters $\eta_1, \ldots, \eta_n$. Once this LPV system is shown to be exponentially stable, one can conclude that $\xi(t) = x_1(t) - x_2(t)$ converges exponentially to zero for $t \to \infty$. If $x_2(.)$ is a nominal system trajectory (such as an equilibrium point or a given trajectory to be investigated), we can conclude that $x_1(t)$ approaches this nominal trajectory exponentially.

Finally, the following procedure is often referred to as *global linearization*. Let

$$\mathcal{F} \quad \text{be the closure of} \quad \mathrm{co}\{f_x(x) \mid x \in \mathbb{R}^n\}.$$

Clearly, $\mathcal{F}$ is a closed and convex subset of $\mathbb{R}^{n \times n}$. It is not difficult to see that any pair of trajectories $x_1(.)$, $x_2(.)$ of (5.4.1) satisfies the linear differential inclusion

$$\dot{x}_1(t) - \dot{x}_2(t) \in \mathcal{F}(x_1(t) - x_2(t)). \tag{5.4.2}$$

Proof. Fix any t and consider the closed convex set

$$\mathcal{F}[x_1(t) - x_2(t)] \subset \mathbb{R}^n.$$

Suppose this set is contained in the negative half-space defined by the vector $y \in \mathbb{R}^n$:

$$y^T \mathcal{F}[x_1(t) - x_2(t)] \leqslant 0.$$

Due to the mean-value theorem, there exists a $\xi \in \mathrm{co}\{x_1(t), x_2(t)\}$ with

$$y^T[\dot{x}_1(t) - \dot{x}_2(t)] = y^T[f(x_1(t)) - f(x_2(t))] = y^T f_x(\xi)[x_1(t) - x_2(t)].$$

Since $f_x(\xi) \in \mathcal{F}$, we infer

$$y^T[\dot{x}_1(t) - \dot{x}_2(t)] \leqslant 0.$$

Hence $\dot{x}_1(t) - \dot{x}_2(t)$ is contained, as well, in the negative half-space defined by y. Since $\mathcal{F}$ is closed and convex, we can indeed infer (5.4.2) as desired. ∎

To analyze the stability of the differential inclusion, one can cover the set $\mathcal{F}$ by the convex hull of finitely many matrices A_j and apply the techniques that have been presented in these notes.

Remarks. Of course, there are many other possibilities to embed nonlinear systems in a family of linear systems that depend on a time-varying parameter. Since there is no general recipe to transform a given problem to the LPV scenario, we have only sketched a few ideas. Although we concentrated on stability analysis, these ideas straightforwardly extend to various nominal or robust performance design problems what is a considerable advantage over other techniques for nonlinear systems. This is particularly important since, in practical problems, non-linearities are often highly structured and not all states enter non-linearly. For example, in a stabilization problem, one might arrive at a system

$$\dot{x} = A(y)x + B(y)u, \quad y = Cx$$

where u is the control input and y is the measured output that captures, as well, those states that enter the system non-linearly. We can use the LPV techniques to design a stabilizing LPV controller for this system. Since y is the scheduling variable, this controller will depend, in general, non-linearly on y; hence LPV control amounts to a systematic technique to design nonlinear controllers for nonlinear systems 'whose non-linearities can be measured'.

5.4.2 Gain-Scheduling

A typical engineering technique to attack design problems for nonlinear systems proceeds as follows: Linearize the system around a couple of operating points, design good linear controllers for each of these points, and then glue these linear controllers together to control the nonlinear system.

Although this scheme seems to work reasonably well in many practical circumstances, there are considerable drawbacks:

- There is no general recipe how to glue controllers together. It is hard to discriminate between several conceivable controller interpolation techniques.

- It is not clear how to design the linear controllers such that, after interpolation, the overall controlled system shows the desired performance.

- There are no guarantees whatsoever that the overall system is even stabilized, not to speak of guarantees for performance. Only through nonlinear simulations one can roughly assess that the chosen design scenario has been successful.

Based on LPV techniques, one can provide a recipe to systematically design a family of linear controllers that is scheduled on the operating point without the need for ad-hoc interpolation strategies. Moreover, one can provide, at least for the linearized family of systems, guarantees for stability and performance, even if the system undergoes rapid changes of the operating condition.

Again, we just look at the stabilization problem and observe that the extensions to include as well performance specifications are straightforward.

Suppose a nonlinear system

$$\dot{x} = a(x, u), \quad y = c(x, u) - r \tag{5.4.3}$$

has x as its state, u as its control, r as a reference input, and y as a tracking error output that is also the measured output. We assume that, for each reference input r, the system admits a unique equilibrium (operating condition)

$$0 = a(x_0(r), u_0(r)), \quad 0 = c(x_0(r), u_0(r)) - r$$

such that $x_0(.)$, $u_0(.)$ are smooth in r. (In general, one applies the implicit function theorem to guarantee the existence of such a parameterized family of equilibria under certain conditions. In practice, the calculation of these operating points is the first step to be done.)

The next step is to linearize the the system around each operating point to obtain

$$\dot{x} = f_x(x_0(r), u_0(r))x + f_u(x_0(r), u_0(r))u, \quad y = c_x(x_0(r), u_0(r))x + c_u(x_0(r), u_0(r))u - r.$$

This is indeed a family of linear systems that is parameterized by r.

In standard gain-scheduling, linear techniques are used to find, for each r, a good tracking controller for each of these systems, and the resulting controllers are then somehow interpolated.

At this point we can exploit the LPV techniques to systematically design an LPV controller that achieves good tracking for all reference trajectories in a certain class, even if these references vary quickly with time. This systematic approach directly leads to a family of linear systems, where the interpolation step is taken care of by the algorithm. Still, however, one has to confirm by nonlinear simulations that the resulting LPV controller works well for the original nonlinear system. Note that the Taylor linearization can sometimes be replaced by global linearization (as discussed in the previous section) what leads to a priori guarantees for the controlled nonlinear system.

Again, this was only a very brief sketch of ideas to apply LPV control in gain-scheduling, and we refer to [14] for a broader exposition of gain-scheduling in nonlinear control.

Bibliography

[1] P. Apkarian and P. Gahinet. A convex characterization of gain-scheduled $\mathcal{H}_\infty$ controllers. 40:853–864, 1995.

[2] Stephen Boyd, Laurent El Ghaoui, Eric Feron, and Venkataramanan Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*. SIAM, 1994.

[3] M. Chilali and P. Gahinet. H_∞ design with pole placement constraints: an lmi approach. *IEEE Transactions on Automatic Control*, 41(3):358–367, 1996.

[4] R. D'Andrea. Generalized L_2 synthesis: a new framework for control design. In *Proc. 35th IEEE Conference on Decision and Control*, pages 4655–4660, 1996.

[5] Pascal Gahinet, Pierre Apkarian, and Mahmoud Chilali. Affine parameter-dependent Lyapunov functions for real parametric uncertainty. In *Proc. 33rd Conference on Decision and Control*, pages 2026–2031, 1994.

[6] Pascal Gahinet, Arkadi Nemirovski, Alan J. Laub, and Mahmoud Chilali. *LMI Control Toolbox*. The Mathworks Inc., 1995.

[7] Wolfgang Hahn. *Stability of Motion*. Springer-Verlag, New York, 1967.

[8] A. Helmersson. *Methods for Robust Gain-Scheduling*. PhD thesis, Linköping University, Sweden, 1995.

[9] H.A. Hindi, B. Hassibi, and S.P. Boyd. Multiobjective $\mathcal{H}_2/\mathcal{H}_\infty$-optimal control via finite dimensional Q-parametrization and linear matrix inequalities. pages 3244–3248, 1998.

[10] T. Iwasaki and S. Hara. Well-posedness of feedback systems: insights into exact robustness analysis and approximate computations. *IEEE Trans. Aut. Control*, 43:619–630, 1998.

[11] F. John. *Partial Differential Equations*. Springer, New York, 4th edition, 1982.

[12] N Kamarkar. A new polynomial-time algorithm for linear programming. *Combinatorica*, 4:373–395, 1984.

[13] I. Kaminer, P.P. Khargonekar, and M.A. Rotea. Mixed H_2/H_∞ control for discrete time systems via convex optimization. *Automatica*, 29:57–70, 1993.

[14] H.K. Khalil. *Nonlinear Systems*. Prentic Hall, second edition, 1996.

[15] David G. Luenberger. *Optimization by Vector Space Methods*. John Wiley and Sons, Inc, New York, 1969.

[16] A.M. Lyapunov. *The General Problem of the Stability of Motion*. Taylor and Francis Ltd, London, 1992. with a bibliography of Lyapunov by V.I. Smirnov and a bibliography of Lyapunov's work by J.F. Barrett.

[17] J.M. Maciejowski. *Multivariable Feedback Design*. Addison-Wesley, Wokingham, England, 1989.

[18] I. Masubuchi, A. Ohara, and N. Suda. Lmi-based controller synthesis: a unified formulation and solution. In *Proceedings of American Control Conference*, 1995.

[19] A Megretsky and S. Treil. Power distribution inequalities in optimization and robustness of uncertain systems. *J. on Math. Systems Estimation and Control*, 3:301–319, 1993.

[20] Yurii Nesterov and Arkadii Nemirovskii. *Interior-Point Polynomial Algorithms in Convex Programming*. SIAM, 1994.

[21] A. Packard. Gain scheduling via linear fractional transformations. *Syst. Contr. Letters*, (2):79–92, 1994.

[22] F. Paganini. Robust H_2 performance: Guaranteeing margins for lqg regulators. Preprint.

[23] K. Poolla and A. Tikku. Robust performance against time-varying structured perturbations. *IEEE Transactions on Automatic Control*, 40:1589–1602, 1995.

[24] A. Rantzer and A. Megretzky. Stability criteria based on integral quadratic constraints. In *35th IEEE Conference on Decision and Control*, pages 215–220, 1996.

[25] A. Wayne Roberts and Dale E. Varberg. *Convex Functions*. Academic Press, New York, 1973.

[26] M.A. Rotea. The generalized H_2 control problem. *Automatica*, 29:373–385, 1993.

[27] N. Rouche, P Habets, and M. Laloy. *Stability Theory by Liapunov's Direct Method*. Springer-Verlag, New York, 1977.

[28] W. Rudin. *Real and Complex Analysis*. McGraw-Hill, 1986.

[29] C.W. Scherer. A complete algebraic solvaility test for the nonstrict lyapunov inequality. *Systems and Control Letters*, 25:327–335, 1995.

[30] C.W. Scherer. *Mixed H_2/H_∞ control*, pages 173–216. Springer-Verlag, Berlin, 1995.

[31] C.W. Scherer. Multiobjective H_2/H_∞ control. *IEEE Trans. Autom. Control*, 40:1054–1062, 1995.

[32] C.W. Scherer. Robust generalized H_2 control for uncertain and LPV systems with general scalings. *Proc. 35nd IEEE CDC 1996, Kobe, Japan*, (0):3970–3975, 1996.

[33] C.W. Scherer. A full block S-procedure with applications. In *36th IEEE Conference on Decision and Control*, 1997.

[34] C.W. Scherer. A full block S-procedure with applications. pages 1510–1515, 1997.

[35] C.W. Scherer. From mixed to multi-objective control. *submitted for the CDC*, 1999.

[36] C.W. Scherer. Lower bounds in multi-objective H_2/H_∞ problems. *submitted for CDC.*, 1999.

[37] C.W. Scherer. *Robust Mixed Control and LPV Control with Full Block Scalings*. SIAM, 1999.

[38] C.W. Scherer, P. Gahinet, and M. Chilali. Multi-objective output-feedback control. *IEEE Transactions on Automatic Control*, 1997. to appear.

[39] G. Scorletti and L. El Ghaoui. Improved linear matrix inequality conditions for gain scheduling. pages 3626–3631, 1995.

[40] J. Shamma. Robust stability with time-varying structured uncertainty. *IEEE Transactions on Automatic Control*, 39:714–724, 1994.

[41] M. Sznaier. An exact solution to general SISO mixed $\mathcal{H}_\infty/\mathcal{H}_2$ problems via convex optimization. 39:2511–2517, 1994.

[42] H.L. Trentelman and J.C. Willems. The dissipation inequality and the algebraic Riccati equation. Technical report, Eindhoven University of Technology, 1990.

[43] Jan Van Tiel. *Convex Analysis: an Introductory Text*. John Wiley and Sons, Chichester, 1984.

[44] Siep Weiland. Dissipative dynamical systems: a behavioral approach. In G Conte, A.M. Perdon, and B. Wyman, editors, *New Trends in Systems Theory*, pages 707–715. Birkhauser, 1991.

[45] Siep Weiland. Dissipative dynamical systems in a behavioral context. *Mathematical Models and Methods in Applied Sciences*, 1, No.1:1–25, 1991.

[46] J.C. Willems. Least squares stationary optimal control and the algebraic Riccati equation. *IEEE Transactions on Automatic Control*, 16 (6):621–634, 1971.

[47] J.C. Willems. Dissipative dynamical systems, part i: General theory. *Arch. Ratinal Mech. Anal.*, 45:321–351, 1972.

[48] J.C. Willems. Dissipative dynamical systems, part ii: Linear systems with quadratic supply rates. *Arch. Rational Mech. Anal.*, 45:352–393, 1972.

[49] T. Yoshizawa. *Stability Theory by Liapunov's Second Method.* The Mathematical Society of Japan, Tokyo, 1966.