Input-Output Analysis of Feedback Loops with Saturation Nonlinearities

B.G. Romanchuk
Magdalene College

Dissertation submitted for the
degree of Doctor of Philosophy

University of Cambridge
Cambridge, England

February 22, 1995

Revised April 24, 1995

Abstract

In this dissertation, we examine the input/output properties of linear systems with saturation in feedback configurations. This is a class of nonlinear systems which consists of linear operators cascaded with memoryless saturation nonlinearities. This form of nonlinearity is usually introduced to simulate the effect of actuator limitations.

As the analysis often relies on state space methods, an initial examination is made of the controllability/observability properties of this class of systems. A key difference to the case of purely linear systems is that one cannot globally stabilise an asymptotically unstable plant with a purely bounded feedback, which justifies the need for semi-global stability results.

Attention is then turned to operator based analysis, which is an extension of the sort used in linear $\mathcal{H}_\infty$ robust control theory. One generalisation of the $\mathcal{H}_\infty$ norm of a linear operator is the induced 2-norm of the nonlinear operator. The computation of this norm by the use of dissipation functions is examined in this dissertation.

By restricting the nonlinearity present to be the piecewise linear standard saturation function, the test conditions on a candidate dissipation function for a single input system are found to be a set of differential inequalities. Dissipation theory is also applied to discrete time systems. The test condition which results for discrete time systems is in a more implicit form, however. A suboptimal existence theorem is developed for a subclass of systems which shows the equivalence of finite gain stability and the convergence of a specific dissipation function.

A second generalisation of the $\mathcal{H}_\infty$ norm of a linear system is the incremental gain of the nonlinear system. A test condition to bound the incremental gain is developed using the solution of linear matrix inequalities. The computational advantage of using the easily computed linear matrix inequality solution is offset by the fact that it appears to be a quite conservative measure. A general construction is performed on discrete time systems which proves that open loop unstable plants cannot be stabilised in an incremental gain sense. Although the analysis has been made on a restricted class of systems, the basic principle of this lack of stabilisability seems readily generalisable.

Finally, analytic examples of the use of the techniques developed in this dissertation are examined. The inter-relationship between stability, the incremental gain and the induced norm is illustrated. Additionally, the importance of the use of bounded input sets in the induced norm computation is made evident.

Dedicated to the memory of my sister, Brenda Anne Romanchuk.

Preface to this revision

This document represents a revision to the dissertation submitted to the University of Cambridge for consideration for the Ph.D. degree. Most of the changes are of the nature of removing typographical errors, and minor corrections. A certain amount of motivational discussions have been added where the material had been found to be more difficult to follow. However, there have been some more significant changes made, which will be noted for those who are concerned about referencing matters.

- Appendix D is a new chapter; it consists of the proofs of a series of Lemmas in Chapter 5. These proofs had been moved to make the presentation more readily followed.
- Sections 5.3 and 5.5 are new; they consist of simpler cases of more general results. Their addition should make the presentation much easier to follow. Theorem 5.4.1 was numbered 5.3.1, and Theorem 5.6.8 was numbered 5.4.10 in the non-revised version which is to be deposited in the University Library.
- The history of Theorem 3.5.8 (open loop null stabilisability) has been corrected to include newly found early discrete time results.
- Section 3.3.2, the open loop configuration definition, has been added. Hence, all the following subsections in Section 3.3 used to have subsection numbers one less than in this revised version.
- Appendix C (the proof of Theorem 4.5.3) has been somewhat expanded, and the numbering of the Lemmas has changed.
- The presentation of Section 8.6 has been changed, and the material split into two subsections, whereas previously it had been in the one section only.
- The comments on uniform observability in Remark 3.6.3 are new, and they provide a better motivation as to why the $\mathcal{L}_p$-detectability definition was introduced.

B.G. Romanchuk
Magdalene College
Cambridge
April 24, 1995

Preface

I would very much like to thank Dr. Malcolm C. Smith for the advice and help he has given to me during my years at Cambridge. He, Professor K. Glover and Dr. J. Maciejowski were kept busy with my questions regarding unusual properties of linear systems. I had visited the Institute of Mathematics and its Applications and the University of Minnesota during 1993, which was kindly arranged by Tryphon Georgiou. I also would like to thank Professors Caines and Zames for making it possible for me to work at McGill University during a visit to Montréal.

I would like to thank my parents and sister Brenda for their support before and during my years at Cambridge. I also would like to thank the following bodies formally for financial aid I received from them during my stay at Cambridge: The Natural Sciences and Engineering Research Council (Canada); Fonds pour la Formation de Chercheurs et l'Aide à la Recherche (Québec); the Cambridge Commonwealth Trust; and the Committee of Vice-Chancellors and Principals of the United Kingdom (ORS Award). I also am grateful for travel grants received from Magdalene College, and the Cambridge Philosophical Society.

I should also thank my fellow Ph.D. students and research associates during my years at the Department of Engineering: Alan, Anita, Frank, Gavin, Geir, George, Giles, Glenn, Haig, Honichi, Jie, John, Markus, Mihai, Nelson, Paul, Phil, Richard L., Rick, Rob, Sanjay, Sharon, Simon, Tung, and Y.C. Chu. Special thanks goes out to Richard Williams for agreeing to proofread draft copies of this dissertation. The 107 Boys provided immeasurable support during my stay in Cambridge. I would also like to thank G.F. Will, J.L. Hooker, P.J. O'Rourke, E. James, G. King, and S. Meier for their efforts in keeping me sane over the years. Additionally, I would like to thank Sue Bennett for her help; especially for preparing the forms by which I communicated with the aforementioned funding bodies.

Finally, as required by University Statute, I declare that this dissertation is not substantially the same as any that I have submitted for a degree at any other university, is the result of my own work and includes nothing which is the result of work done in collaboration.

B.G. Romanchuk
Magdalene College
Cambridge
February 22, 1995

Contents

Notation

$\mathbb{R}$		The real line
$\mathbb{R}_+$		$\{x: x \in \mathbb{R}, x \geq 0\}$
$\mathbb{R}_-$		$\{x: x \in \mathbb{R}, x \leq 0\}$
$\mathbb{Z}$		The integers
$\mathbb{Z}_+$		$\{k: k \in \mathbb{Z}, k \geq 0\}$
$\mathbb{C}$		The complex plane
Ω^C		The set complement of Ω
$\bar{\Omega}$		The closure of Ω
$\bar{x}$		The complex conjugate of x
$\mathrm{Re}\{x\}$		The real part of x
$\mathrm{Im}\{x\}$		The imaginary part of x
x'		The transpose of vector x
$P > 0$		The symmetric matrix P is positive definite
$P \geq 0$		The symmetric matrix P is positive semi-definite
$\rho(A)$		Spectral radius of A
$x[k]$		Square brackets used for purely discrete time indices
$\frac{dV}{dx}$		The derivative of V, expressed as a row vector
$\bar{\sigma}(A)$		The maximal singular value of matrix A
$\|x\|$	6	The Euclidean (2-) Norm of x
$\|x\|_p$	6	The $\mathcal{L}_{p+}$ or ℓ_{p+} norm of x
$\mathcal{W}_{p,q}(N)$	7	$\mathcal{L}_{p+} \bigcap \{x: \|x\|_q \leq N\}$ (resp. $\ell_{p+} \bigcap \{x: \|x\|_q \leq N\}$)
$\|\mathbf{\Upsilon}\|_{i,p}$	7	The induced norm of $\mathbf{\Upsilon}$ over $\mathcal{L}_{p+}$ or ℓ_{p+}
$\|\mathbf{\Upsilon}\|_{\Delta,p}$	7	The incremental gain of $\mathbf{\Upsilon}$ over $\mathcal{L}_{p+}$ or ℓ_{p+}
$\mathbf{P}_T$	7	The truncation operator at time T
$\Delta x(t)$	9	Represents $\dot{x}(t)$ or $x[t+1]$
$\mathcal{L}u$	9	The Laplace Transform of u
$\mathcal{L}^{-1}u$	9	The Inverse Laplace Transform of u
$\mathcal{Z}u$	9	The z-Transform of u
$\mathcal{Z}^{-1}u$	9	The Inverse z-Transform of u
$\left[\begin{array}{c\|c} A & B \\ \hline C & D \end{array}\right]$	10	The transfer function realised by $D + C(sI - A)^{-1}B$ (resp. z)
$\sigma(x)$	14	A saturation function
$\sigma_0(x)$	15	The standard saturation
$x(t;\xi)$	56	The induced (null input) trajectory with initial condition ξ
div	111	The Euclidean divisor
mod	111	The Modulo operation (integer math)
$\partial f(x)$	116	Generalised gradient of f at x

Chapter 1

Introduction

1.1 Overview

One of the main driving forces behind the development of modern control theory was the realisation that controllers designed solely from optimisation concerns exhibited a lack of robustness with respect to modelling uncertainty, both in theory and in practice. Small mismatches from the model used for design to the actual plant could cause a serious loss of performance, or even stability. For linear control theory in the 1980's, a major field of research was $\mathcal{H}_\infty$ control (and related topics), which addressed these robustness issues. The control algorithms developed amount to sophisticated generalisations of classical design methods, and have proven effective in practice, especially for multi-input/ multi-output systems.

Nonlinear control theory lacks many of the mathematical tools which are available to linear systems, and its development consequently lags that of linear theory. The success of linear robust control methods has led to an interest in extending the work to nonlinear systems. The robust control theory for linear systems is usually approached using input/output or operator theoretic methods, and so there has been renewed interest in the study of nonlinear input/output systems.

The difficulty of analysing generic nonlinear systems leads to the idea of restricting the class of systems studied. That is the approach taken in this dissertation, where attention is focussed on the class of linear systems with saturation, which is a subclass of the systems which have piecewise linear state dynamics. As almost all actuators have limits to their control effort, the problem of the analysis of systems with saturation nonlinearities has good practical motivation. The performance degradation that linear controllers experience in the presence of the saturation on the control inputs can be viewed as a generalisation of the "anti-windup" problem. Industrial engineers have had to rely on their intuition when approaching this problem, but problems are faced when the systems are multiple input, or the controller has a fairly complicated dynamic structure.

As will be seen in later chapters of this dissertation, it appears that the key to extending $\mathcal{H}_\infty$ methods to nonlinear systems is the ability to solve various partial differential inequalities over the state space. Unfortunately, the problems seem to be generally quite difficult to solve numerically, and this constrains the practical effectiveness of the analysis. It is conjectured that the piecewise linear systems which are the focus of study herein may have solutions to these partial differential inequalities which are of a form which may be more easily computed, possibly in a purely algebraic way.

The theory developed in this dissertation is almost exclusively aimed at analysis problems, and a good deal of the work examines the mathematical framework used. This was a deliberate choice, as there are a great deal of fundamental issues which need to be addressed before synthesis algorithms can be developed in a methodical fashion.

1.2 Outline

We will now give a general outline of the contents of this dissertation.

Chapter 2 This chapter contains background mathematical material, and develops the notation used throughout.

Chapter 3 This chapter describes the systems studied in this dissertation (linear systems with saturation), and describes various problem configurations encountered. Most of the contents of the chapter are an examination of controllability and observability properties of these systems. Most of the material is a survey of other's work, although more specialised new results are developed.

Chapter 4 Chapter 4 is a survey of stability results, most of which are for continuous time systems (discrete time stability results are dealt with in Chapter 6). The first part of the chapter discusses state space stability, and Lyapunov theory. The latter part of the chapter is of more importance for operator theoretic analysis. The Small Gain Theorem is given, which is based upon operator gains. In order to compute these operator gains, dissipation theory ideas are used, the background of which is surveyed.

Chapter 5 The problem of the computation of the induced norm for linear systems with saturation is examined here. Although there is a survey of other researchers' results, most of this chapter consists of new results on the use of dissipation functions to bound the induced norm. A test condition is given, as well as an algorithm which generates functions which will meet the condition, although in a suboptimal fashion.

Chapter 6 The problem of the induced norm computation for discrete time systems is developed in Chapter 6. The results are not as strong as is the case for the continuous time systems, but it was possible to make the development parallel the continuous time case closely. Additionally, the chapter contains some survey material, such as discrete time Lyapunov theory, and basic dissipation results.

Chapter 7 The incremental gain of linear systems with saturation is examined, for both continuous and discrete time systems. Negative results reported by other researchers are surveyed, and those conclusions are then amplified in the new results which form the remainder of the chapter. It is shown to be possible to bound the incremental gain using linear matrix inequalities; however general test signals have been constructed which show that the incremental gain will not exist for unstable plants with constrained inputs.

This seems to indicate that the incremental gain will be of more limited use in stability analysis.

Chapter 8 In this chapter, examples of the use of the analytical techniques developed elsewhere in this dissertation are given. The examples seem to give a good qualitative feel for the inter-relationship between the bounds for the operator gains given, as well as the usefulness of bounded input sets.

Chapter 9 This is a brief concluding chapter; contributions made in the dissertation are summarised, and recommendations for future work given.

Appendices In Appendices A,B,C, and D, proofs are given for results which had been deferred from the main body of the text.

Chapter 2

Mathematical Preliminaries

2.1 Introduction

This chapter sets forth the basic mathematical tools which are used throughout the remainder of this dissertation. It is not self-contained, as it is assumed that the reader will be familiar with most of this mathematics. General references are cited which contain fuller expositions of these topics.

In Section 2.2, important notation and conventions used are listed. Section 2.3 gives the definitions used for the signal sets, of obvious importance when treating systems as input/output operators. Basic definitions used in classifying systems are given in Section 2.4. As the theory developed in this dissertation relies heavily on linear system properties, some of the basic notation associated with linear systems is given in Section 2.5. Finally, as they are part of a relatively new research area, linear matrix inequalities are introduced in Section 2.6.

2.2 Preliminary Definitions

It is assumed that the reader is familiar the with basic definitions and concepts of real analysis, and so they will not be listed in full here. Most introductory real analysis texts should be sufficient; see (for example) the text by W. Rudin [39]. We will now present some of the notation and results used.

The real line is denoted $\mathbb{R}$; the subset $\{x: x \in \mathbb{R}, x \geq 0\}$ is denoted $\mathbb{R}_+$. The integers are denoted $\mathbb{Z}$, with the subset $\{x: x \in \mathbb{Z}, x \geq 0\}$ denoted $\mathbb{Z}_+$. The complex field is denoted $\mathbb{C}$, with $i \triangleq \sqrt{-1}$. The real and imaginary parts of x are denoted $\mathrm{Re}\{x\}$, $\mathrm{Im}\{x\}$, respectively.

The signum function is used at various points; denoted $\mathrm{sign}(\cdot)$, it is defined here as:

$$\mathrm{sign}(x) = \begin{cases} -1 & x < 0, \\ 1 & x \geq 0. \end{cases}$$

The important point to be noted is that we have defined $\mathrm{sign}(0)$ as 1; some statements given here are reliant upon having $\mathrm{sign}(0) \neq 0$.

The integral used throughout is the Lebesgue Integral; some familiarity with Lebesgue Measure theory would be useful at some points.

Continuity concepts are heavily used throughout. Of importance is the definition of uniform Lipschitz continuity, and the Balzano-Weierstrass Theorem. A function f from a metric space A to a metric space B has a uniform Lipschitz constant κ if for all $x, y \in A$, $d_B(f(x), f(y)) \leq \kappa d_A(x, y)$. An important corollary to the Balzano-Weierstrass theorem states that if a bounded function f is continuous on a compact subset of $\mathbb{R}^n$, then it achieves its infimal and supremal value upon that set (i.e., the maximum and minimum exist).

The derivative of a function $f: \mathbb{R}^n \mapsto \mathbb{R}$ at a point x is denoted $\left.\frac{df}{dx}\right|_x$, and if it exists, it is equal to a vector A in $\mathbb{R}^{1\times n}$ such that:

$$\lim_{\delta\to 0} \frac{f(x+\delta) - f(x) - A\delta}{\|\delta\|} = 0.$$

A function which is continuous on its domain is said to lie in $\mathcal{C}^0$, and if it is continuously differentiable k times, it lies in $\mathcal{C}^k$.

Basic linear algebra is used throughout, especially since we are attempting to inter-relate the results with the equivalent linear system theory. The maximal singular value of a matrix A is denoted $\bar{\sigma}(A)$, and is defined:

$$\bar{\sigma}(A) = \max_{\|x\|=1} \|Ax\|.$$

A square symmetric matrix P is *positive definite*, denoted $P > 0$, if for all nonzero x of appropriate dimension, $x'Px > 0$. Such a matrix is *positive semi-definite* if the preceding strict inequalities are relaxed to non-strict ones. A matrix is said to be *Hurwitz* if all its eigenvalues lie in the open left half plane. A matrix is *neutrally stable* if all its eigenvalues lie in the closed left half plane, with no repeated eigenvalues on the imaginary axis. Finally, a matrix A is *skew-symmetric* if $A = -A'$.

2.3 Signals and Systems

Although state space analysis is heavily used throughout this dissertation, the motivation for the work is based on operator theoretic, or input/output based approaches to control. In such approaches, the actual system dynamics (expressed as differential or difference equations) are not the point of concern, rather the signals which are viewed as inputs and outputs are examined, usually in terms of them lying in signal sets, usually normed spaces. Systems are to be thought of as operators on the signal sets, and not in terms of their state space realisation.

Definition 2.3.1: A *system* is an operator which maps from a signal set $\mathcal{U}$ (the input set) to a signal set $\mathcal{Y}$ (the output set).

Remark 2.3.2: We will distinguish operators by setting them in bold type; they can thus be distinguished from other quantities, such as the transfer function corresponding to a linear operator.

Definition 2.3.3: A system is said to be *input/output stable* if its domain of definition is its entire input set.

Remark 2.3.4: This is a weaker condition than finite gain stability, which takes into account signal size.

A generic definition of signal sets is not sufficient for any non-trivial theory to be developed, and so we will now define the signal spaces and sets used.

Definition 2.3.5: The space norm used throughout this dissertation is the Euclidean (or 2-) norm. For all x in $\mathbb{R}^n$ (resp. $\mathbb{C}^n$), the norm is given by $\|x\| = \sqrt{x'x}$ (resp. $\sqrt{\bar{x}'x}$), with x' denoting the transpose, and $\bar{x}$ the complex conjugate. Henceforth, the Euclidean norm is the only norm used which will not be distinguished by a subscript.

Definition 2.3.6: The $\mathcal{L}_{p+}$ *norm* (with $p \in \mathbb{Z}_+$, and $0 < p < \infty$) is defined as:

$$\|x\|_p = \left(\int_0^\infty \|x(t)\|^p \, dt \right)^{\frac{1}{p}}, \tag{2.1}$$

with $x(t) \in \mathbb{R}^m$, and with the integral in (2.1) being the Lebesgue integral. The space $\mathcal{L}_{p+}$ is the set of all functions $x \colon \mathbb{R}_+ \mapsto \mathbb{R}^m$ such that the $\mathcal{L}_{p+}$ norm of x is finite.

Definition 2.3.7: The $\mathcal{L}_{\infty+}$ *norm* is defined:

$$\|x\|_\infty = \operatorname{ess}\sup_{t \in \mathbb{R}_+} \{\|x(t)\|\}.$$

Remark 2.3.8: We follow the usual practice and identify all signals which lie in an equivalence class whose members have difference of measure zero. Additionally, we suppress the dimension of the signal vector; this is assumed to be clear from the context.

Definition 2.3.9: The ℓ_{p+} *norm* (with $p \in \mathbb{Z}_+$, and $0 < p < \infty$) is defined as:

$$\|x\|_p = \left(\sum_{k=0}^{\infty} \|x[k]\|^p \right)^{\frac{1}{p}},$$

with $x[k] \in \mathbb{R}^m$. The space ℓ_{p+} is the set of all functions $x \colon \mathbb{Z}_+ \mapsto \mathbb{R}^m$ such that their ℓ_{p+} norm is finite.

Definition 2.3.10: The $\ell_{\infty+}$ *norm* is defined:

$$\|x\|_\infty = \sup_{k \in \mathbb{Z}_+} \{\|x[k]\|\},$$

with $x[k] \in \mathbb{R}^m$. The space $\ell_{\infty+}$ is the set of all functions $x \colon \mathbb{Z}_+ \mapsto \mathbb{R}^m$ such that their $\ell_{\infty+}$ norm is finite.

The sets $\mathcal{W}_{p,q}(N)$ will now be introduced, which are bounded subsets of $\mathcal{L}_{p+}$ (resp. ℓ_{p+}). Although they are somewhat non-standard, their use seems to be necessary for the analysis of linear systems with saturation nonlinearities. A motivation for their introduction follows from the analysis which is found in Chapter 8.

Definition 2.3.11: The set $\mathcal{W}_{p,q}(N)$ is the intersection of the sets, $\mathcal{L}_{p+} \bigcap \{w: \|w\|_q \le N\}$. The discrete time counterpart is also denoted $\mathcal{W}_{p,q}(n)$, which is the set $\ell_{p+} \bigcap \{w: \|w\|_q \le N\}$.

It may have been preferable to distinguish continuous time and discrete time input sets; however as this dissertation is concerned solely with purely continuous time or discrete time systems, it should be apparent from the context which sets are referred to.

The last signal sets to be defined are the *extended spaces*. They are not used explicitly very often within this dissertation, rather their use is to be implicitly understood when discussing system outputs which may be unbounded. In order to define the extended spaces, the truncation operator is needed, which is defined first.

Definition 2.3.12: The *truncation operator at time T*, denoted $\mathbf{P}_T$, is defined by:

$$(\mathbf{P}_T w)(t) = \begin{cases} w(t) & 0 \le t < T, \\ 0 & t \ge T. \end{cases}$$

This definition holds for both continuous and discrete time.

Definition 2.3.13: The *extended space* $\mathcal{L}_{p,e}$ (resp. $\ell_{p,e}$) is the set of signals such that for all $T \in \mathbb{R}_+$ (resp. $\mathbb{Z}_+$),

$$\{x: \mathbf{P}_T x \in \mathcal{L}_{p+} (\text{resp. } \ell_{p+})\}.$$

We will now define the extremely important concepts of operator norms and incremental gain (the incremental gain has similar properties to an operator norm). They are a measure of "signal magnitude throughput", and can be used in basic stability analysis of closed loop systems.

Definition 2.3.14: The *induced norm* of a stable system $\mathbf{\Upsilon}$ from an $\mathcal{L}_{p+}$ space (or $\mathcal{W}_{p,q}(N)$ set) to another $\mathcal{L}_{p+}$ space, denoted $\|\mathbf{\Upsilon}\|_{i,p}$, is the quantity (if it exists):

$$\|\mathbf{\Upsilon}\|_{i,p} = \sup_{w \in \mathcal{L}_{p+} (\text{resp. } \mathcal{W}_{p,q}(N)), w \neq 0} \frac{\|\mathbf{\Upsilon} w\|_p}{\|w\|_p}.$$

The notation used is the same for discrete time for the appropriate ℓ_{p+} spaces.

Definition 2.3.15: The *incremental gain* of a stable system $\mathbf{\Upsilon}$ from an $\mathcal{L}_{p+}$ space (resp. $\mathcal{W}_{p,q}(N)$ set) to another $\mathcal{L}_{p+}$ space, denoted $\|\mathbf{\Upsilon}\|_{\Delta,p}$, is the quantity (if it exists):

$$\|\mathbf{\Upsilon}\|_{\Delta,p} = \sup_{w,\tilde{w} \in \mathcal{L}_{p+} (\text{resp. } \mathcal{W}_{p,q}(N)), w \neq \tilde{w}} \frac{\|\mathbf{\Upsilon}\tilde{w} - \mathbf{\Upsilon} w\|_p}{\|\tilde{w} - w\|_p}.$$

The notation used is the same for discrete time systems with the appropriate ℓ_{p+} spaces replacing $\mathcal{L}_{p+}$ spaces.

It is trivial that $\|\mathbf{\Upsilon}\|_{i,p} \le \|\mathbf{\Upsilon}\|_{\Delta,p}$ (provided that $\mathbf{\Upsilon} 0 = 0$); however it will be found that for linear systems with saturation, the difference between the two may be large. This is greatly distinguished from the case for linear systems, where the two are equal.

2.4 Systems Definitions

When developing theory on system operators, it is usually necessary to add restrictions on the class to be studied. Without such added assumptions, it is quite difficult to characterise system properties. We will now give the systems definitions used in this dissertation, defined in terms of the properties of operators.

We will almost exclusively confine our attention to *shift invariant* system operators. An operator is shift invariant if it commutes with the shift operator with time shift T for all $T \in \mathbb{R}_+$ ($T \in \mathbb{Z}_+$ for discrete time systems); the *shift operator by time* T, being defined by:

$$(\mathbf{S}_T w)(t) = \begin{cases} 0 & 0 \leq t < T, \\ w(t-T) & t \geq T. \end{cases}$$

A system operator $\boldsymbol{\Upsilon}$ is said to be *causal* if for all T, $\mathbf{P}_T \boldsymbol{\Upsilon} \mathbf{P}_T = \mathbf{P}_T \boldsymbol{\Upsilon}$. Finally, a system operator $\boldsymbol{\Upsilon}$ is *memoryless* if it is causal, and if for all T, $(\mathbf{I} - \mathbf{P}_T)\boldsymbol{\Upsilon}(\mathbf{I} - \mathbf{P}_T) = (\mathbf{I} - \mathbf{P}_T)\boldsymbol{\Upsilon}$.

The theory of linear systems is quite heavily developed, and the work in this dissertation is allied quite closely to it. Formally, a system operator $\boldsymbol{\Upsilon}$ is *linear* if for all x, y in an input set, and for all $\alpha, \beta \in \mathbb{R}$, we have:

$$\boldsymbol{\Upsilon}(\alpha x + \beta y) = \alpha \boldsymbol{\Upsilon} x + \beta \boldsymbol{\Upsilon} y,$$

if $\alpha x + \beta y$ also lies in the input set.

Most of the operators studied in this dissertation are in fact generated by state space realisations. This is extremely useful for nonlinear systems with memory, as it is then possible to generate the system outputs via simulations on a digital computer. Linear system theory (with shift invariant systems) can be dealt with purely in the frequency domain, and so the operator theory can overlook state space representations if so desired. This is generally not possible for nonlinear systems, however. Formally, a *realisation* of a continuous time system $\boldsymbol{\Upsilon}: w \mapsto z$ (with initial condition $\xi_0 \in \mathbb{R}^n$) is a set of differential equations over $\mathbb{R}_+$ of the form:

$$\begin{aligned} \dot{x}(t) &= f(x(t), w(t), t), && (2.2) \\ x(0) &= \xi_0, && (2.3) \\ z(t) &= h(x(t), w(t), t), && (2.4) \end{aligned}$$

with $x(t) \in \mathbb{R}^n$. It is assumed that $f(0,0,t) = 0$ and $h(0,0,t) = 0$ for all $t \in \mathbb{R}_+$. The operator defined by this initial value problem is said to be *finite dimensional* (or of *dimension* n). Similarly, the discrete time *realisation* of a system $\boldsymbol{\Upsilon} w: \mapsto z$ (with initial condition $\xi_0 \in \mathbb{R}^n$) is a set of difference equations over $\mathbb{Z}_+$ of the form:

$$\begin{aligned} x[k+1] &= f(x[k], w[k], k), && (2.5) \\ x[0] &= \xi_0, && (2.6) \\ z[k] &= h(x[k], z[k], k), && (2.7) \end{aligned}$$

with $x[k] \in \mathbb{R}^n$. Once again, we assume that $f(0,0,k) = 0$ and $h(0,0,k) = 0$ for all $k \in \mathbb{Z}_+$.

In the preceding discussion, we have neglected an important point: does the state space realisation give rise to a solution, and is the solution in fact unique? In Section 3.4, a well known result is quoted to show that for the classes of systems of interest in this dissertation, existence and uniqueness are ensured. Hence, the existence and uniqueness of the solution is assumed herein, although this is not necessarily true for generic nonlinear state space systems.

The state trajectories and corresponding outputs should be considered to reside in $\mathcal{L}_{p,e}$ (resp. $\ell_{p,e}$) spaces, so that they should remain well defined even if they have unbounded p norms (except in cases with finite escape times, for instance). Otherwise, the states and outputs could be considered as not existing, and some manipulations (such as generating destabilising inputs) would be ill posed mathematically. Therefore, state trajectories and corresponding outputs are assumed to lie in extended spaces, and so it will be possible to refer to outputs corresponding to inputs lying outside the domain of the operators. Most results are given in a form for which the use of extended spaces is not necessary, however.

With regards to the initial conditions, it is generally assumed that $\xi_0 = 0$ for the operators studied herein. This, along with the conditions that $f(0,0,t) = 0$, $h(0,0,t) = 0$, ensures that $\boldsymbol{\Upsilon} 0 = 0$, which is essentially necessary for the induced norm to be defined. The definition of induced norm could be relaxed to a formulation that takes into account non-zero initial conditions, but this was not felt to be necessary.

Remark 2.4.1: We will henceforth make use of the following notation in order to express simultaneously continuous and discrete time results:

$$\Delta x(t) = f(x(t), w(t), t),$$

which is equivalent to $\dot{x}(t) = f(x(t), w(t), t)$ (continuous time) or $x[t+1] = f(x[t], w[t], t)$ (discrete time).

The following is a well known statement regarding the systems properties of operators generated by state space realisations.

Proposition 2.4.2: A system operator $\boldsymbol{\Upsilon}$ generated by a finite dimensional realisation will be causal. Additionally, if the functions f, h in equations (2.2) - (2.4) (resp. (2.5) - (2.7)) are independent of t (resp. k), $\boldsymbol{\Upsilon}$ is shift invariant.

2.5 Linear Systems and the Frequency Domain

Frequency domain techniques are not extensively used herein as the present level of analysis is quite heavily tied to state space techniques. However, some use is made of them, and so they will be briefly introduced. Further details can be easily found in most texts on linear control, such as [21],[62].

The Laplace and z-transforms of continuous and discrete time signals are both used herein. If $u(t)$ is a continuous time signal (resp. discrete time), then the transform is denoted $\hat{u}(s) = (\mathcal{L}\{u\})(s)$ (resp. $\hat{u}(z) = (\mathcal{Z}\{u\})(z)$). The inverse transform is denoted $u(t) = \mathcal{L}^{-1}\{\hat{u}(s)\}$ (resp. $u[k] = \mathcal{Z}^{-1}\{\hat{u}(z)\}$).

Although not extensively used herein, it is necessary to define the Hardy Space $\mathcal{H}_2$. It is the set of all (matrix valued) functions F defined on $\mathbb{C}$, analytic on $\text{Re}\{s\} > 0$, and with the indefinite integral

$$\sqrt{\sup_{\alpha>0} \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{Trace}\{\bar{F}'(\alpha + i\omega) F(\alpha + i\omega)\} d\omega} \tag{2.8}$$

finite. This is a closed subspace with regards to the set of matrix valued functions F with the integral in (2.8) finite. The value of the integral in (2.8) is a norm over $\mathcal{H}_2$, and in fact the Laplace Transform yields an isometric isomorphism between the $\mathcal{L}_{2+}$ norm and this norm for signals and their transforms. For discrete time systems, the set $\mathcal{H}_2$ is defined similarly, with the corresponding integral in (2.8) being taken over circles with $r > 1$.

When describing a shift invariant linear operator, it is possible to do so in terms of a time domain convolution, or as a frequency domain transfer function. Convolution representations are not used herein, instead linear systems are specified as transfer functions, or as the corresponding realisations. The following propositions are a summary of well known facts on linear systems, and are included to clarify the notation used herein.

Proposition 2.5.1: A linear, shift invariant continuous time (resp. discrete time) system $\mathbf{G}$ can be represented as:

$$\mathbf{G}u = \mathcal{L}^{-1}(G(s)\mathcal{L}\{u\}) \qquad (\text{resp. } \mathcal{Z}^{-1}(G(z)\mathcal{Z}\{u\})\ ),$$

with $G(s)$ (resp. $G(z)$) termed its *transfer function.*

If a linear, shift invariant continuous time (resp. discrete time) system $\mathbf{G}$ has a finite dimensional state space representation, the state equations can be written:

$$\begin{aligned} \Delta x(t) &= Ax(t) + Bu(t), \\ x(0) &= 0, \\ z(t) &= Cx(t) + Du(t), \end{aligned}$$

with A, B, C, D appropriately sized matrices. Additionally, we have that its transfer function is specified by:

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

(resp.

$$G(z) = D + C(zI - A)^{-1}B,).$$

We can use the notation that $[A, B, C, D]$ is a representation of $\mathbf{G}$, or that

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right].$$

Additionally, the transfer function is *rational* in the complex variable s.

Finally, we will need to introduce the important $\mathcal{H}_\infty$ and $\mathcal{RH}_\infty$ spaces. If F is a matrix valued function on $\mathbb{C}$, then $F \in \mathcal{H}_\infty$ if F is analytic on the open right half plane, and

$$\|F\|_\infty \triangleq \text{ess} \sup_{\text{Re}\{s\}>0} \bar{\sigma}(F(s)), \tag{2.9}$$

is finite. This is a closed subspace with respect to functions which are essentially bounded along the imaginary axis. The set $\mathcal{RH}_\infty \subset \mathcal{H}_\infty$ is the space of functions in $\mathcal{H}_\infty$ which are rational. The sets $\mathcal{H}_\infty$ and $\mathcal{RH}_\infty$ corresponding to discrete time systems are Similarly defined, except that the functions are analytic outside the unit disk, and the supremum in (2.9) is taken over circles with $r > 1$.

In fact if $G(s)$ (resp. $G(z)$) is in $\mathcal{H}_\infty$, then $\|\mathbf{G}\|_{i,2} = \|G(s)\|_\infty$ (resp. $\|G(z)\|_\infty$). If the functions are not rational, however, it is necessary to determine the infinity norm of the transfer function solely in the frequency domain, which is not readily extensible to nonlinear systems. With rational functions, it is possible to use state space methods to determine the norm (see Sections 4.6 and 6.3.1), a technique which has been extended for nonlinear systems.

2.6 Linear Matrix Inequalities

It is beyond the scope of this dissertation to seriously treat the topic of linear matrix inequalities (often abbreviated as LMI's). They are now an area of intense research interest, as they are believed to be an extension of the quite useful algebraic Riccati equation. Analytic solutions to the inequalities generally do not exist, but numerical schemes to find feasible solutions (if any exist) have been developed, and are the subject of current research. A useful recent reference which is an introduction to the theory is that by Boyd, El Ghaoui, Feron and Balakrishnan [4]. The authors in [4] stress that as long as analytic solutions are not demanded, reduction of an analysis problem to a linear matrix inequality is sufficient in some sense to have "solved" the problem. This is precisely the viewpoint taken here in Chapter 7; the problem of generating a bound for the incremental gain of a class of systems is reduced to a linear matrix inequality. However, in Appendix B, a somewhat unusual construction is made which gives an interesting systems insight into the interpretation of the non-existence of the solution to a coupled pair of inequalities.

We will now give some preliminary definitions taken from [4]. A *linear matrix inequality* is of the form:

$$F(x) \triangleq F_0 + \sum_{i=1}^{m} x_i F_i > 0, \tag{2.10}$$

with $F_i \in \mathbb{R}^{n\times n}$, and $x \in \mathbb{R}^m$ the variable. Once again, the inequality in (2.10) is interpreted as requiring that $F(x)$ be positive definite. A *non-strict linear matrix inequality* is not surprisingly the result of the relaxation of the condition on $F(x)$ to be positive semi-definite.

It is possible to express the multiple linear matrix inequalities $F^{(1)}(x) > 0, \ldots, F^{(p)}(x) > 0$ as a single linear matrix inequality, $\mathrm{diag}(F^{(1)}(x), \ldots, F^{(p)}(x)) > 0$. As such, multiple coupled inequalities are treated as a single problem with no difficulties.

An important manipulation upon linear matrix inequalities is the use of *Schur complements*. A simple formulation is that the linear matrix inequality

$$\begin{bmatrix} Q(x) & S(x) \\ S(x)' & R(x) \end{bmatrix} > 0,$$

with $Q(x) = Q(x)'$, $R(x) = R(x)'$, and $S(x)$ dependent in an affine fashion upon x, is equivalent to:

$$R(x) > 0, \qquad Q(x) - S(x)R(x)^{-1}S(x)' > 0.$$

For example, we can express

$$A'P + PA + PBR^{-1}B'P + Q < 0,$$

(with P the symmetric variable matrix, and with $Q = Q'$, $R = R' > 0$) as

$$\begin{bmatrix} -A'P - PA - Q & PB \\ B'P & R \end{bmatrix} > 0.$$

Discussion of the computational aspects of linear matrix inequalities are beyond the scope of this dissertation. However, their solution appears to be computationally tractable. Algorithms exist which have been shown to be of polynomial complexity, and in fact seem to be quite efficient in practice. The simplest in principle is the ellipsoid algorithm, which has some use in proofs. Since the late 1980's, interior point methods have been developed which are much more efficient. Their development has motivated the interest in this field. The reader is once again referred to [4], and references therein.

Chapter 3

Fundamentals of the Theory of Linear Systems With Saturation

3.1 Introduction

In this chapter, we will examine the basic systems properties of the class of systems to be studied in this dissertation, being linear systems with saturation. With regards to these systems properties, linear systems with saturation are characterised in a fashion very similar to purely linear systems.

In Section 3.2, the class of operators which are referred to as saturation operators is characterised. These operators are memoryless, and when cascaded with linear operators, form the class of nonlinear plants which are the objects of study in this dissertation.

Section 3.3 gives a quick summary of the types of problems which can result from the study of the feedback loops with saturation nonlinearities present. It is noted that there are different points in the block diagram to introduce the nonlinearities, and the precise positioning will give rise to quite different behaviour.

The fundamental properties of the state space realisations of these systems are analysed in Sections 3.4 - 3.6. The conditions under which solutions exist and are unique is examined, being a necessity for the analysis to be well posed. The controllability/observability conditions are then examined, and compared to those for linear systems. Some more specialised new results on these topics are developed; otherwise the theory in this chapter represents a survey of that available in the literature.

3.2 Linear Systems with Saturation

The main class of systems studied in this thesis are linear systems with saturation nonlinearities. A (single input) *linear system with saturation* is a system whose describing operator is of the form:

$$\mathbf{G}u = \mathbf{G_1}u + \mathbf{G_2}\sigma u, \tag{3.1}$$

where $\mathbf{G_1}, \mathbf{G_2}$, are shift invariant linear operators, and σ is a memoryless *saturation operator* (Definition 3.2.4). The extension of this definition to multiple inputs is straightforward, although the notation could prove confusing in general - there are are great many possible ways of introducing a nonlinearity into the equations. In order to keep analysis tractable, single input cases will be the primary concern, and extensions to multiple inputs treated secondarily. Additionally, we will now add a well posedness assumption on the operators studied herein. It amounts to assuming that there are no cancellations of unstable poles between the two parts of the system.

Assumption 3.2.1: We will assume that for all the linear systems with saturation, if $\mathbf{G} = \mathbf{G_1} + \mathbf{G_2}\sigma$ (with $\mathbf{G_1}, \mathbf{G_2}$ linear operators), then $\mathbf{G}u \in \mathcal{L}_{p+} \implies \mathbf{G_1}u \in \mathcal{L}_{p+}, \mathbf{G_2}u \in \mathcal{L}_{p+}$. Also, if it specified that the system is finite dimensional, this applies to both $G_1(s)$ and $G_2(s)$, as well as $G(s)$.

Remark 3.2.2: The form of system chosen in (3.1) is admittedly unusual, and it has been chosen to allow the study of some systems of interest. It basically amounts to being a generic single input system with a piecewise linear right hand side to the dynamics equation, and is not claimed to be immediately representative of real world systems in general.

Remark 3.2.3: We should also comment on the decision to restrict our attention to single input systems. The computational techniques, being state space based, are not inherently SISO in character. In fact, matrix notation is intentionally used throughout, admittedly at some cost in terms of the burden upon the reader. The inherent problem is the issue of *multiple nonlinearities*, and the related issue of directionality compensation. Although MIMO anti-windup analysis/synthesis is an important problem, the single input case would need to be understood first. When the usual MIMO anti-windup synthesis is attempted, directionality compensation is quite often added in front of the standard saturations (see the dissertation by P.J. Campo ([7, Section 6.2.4]) for a comprehensive discussion of this topic). However, the author wishes to avoid their use. Other than for for the analysis of existing anti-windup schemes, there is no reason to introduce these nonlinearities into one's analysis scheme, as it seems readily apparent that a well posed nonlinear $\mathcal{H}_\infty$ synthesis method would inherently build the directionality compensation into the nonlinear controller. To put the previous remark into perspective, it must be admitted that the computational base to allow the practical synthesis of such controllers does not exist at present, and is an on-going research problem.

In order to compare our results to those arising from linear theory, we will need to specify what we mean by the local linearisation of the nonlinear operator. It should be noted that in order for this definition to make sense, the saturation operator σ would need to be differentiable at the origin, with the derivative normalised to 1. This in fact will hold for all cases in this dissertation where the local linearisation is referred to, but care would need to be taken in general.

Definition 3.2.4: For a fixed linear system with saturation $\mathbf{G} = \mathbf{G_1} + \mathbf{G_2}\sigma$, the *local linearisation* is the linear operator $\mathbf{G_L} = \mathbf{G_1} + \mathbf{G_2}$.

The (single input) saturation operator σ is a memoryless operator with 1 dimensional input and output spaces, which is defined by:

$$(\sigma u)(t) = \sigma(u(t)), \forall t \in \mathbb{R}_p,$$

with σ being a member of the class of *saturation functions* (defined below), which map $\mathbb{R}$ to a subset of $\mathbb{R}$.

Many of the original results developed herein will be given in terms of one specific saturation function, called the *standard saturation function*, denoted σ_0, which is defined :

$$\sigma_0(x) = \begin{cases} -1 & x < -1, \\ x & |x| \leq 1, \\ 1 & x > 1. \end{cases}$$

The usefulness of this function is that it represents one of the simplest definitions of the functions in its class, and it is possible to exploit the fact that it is piecewise linear. It is of course non-differentiable, and so sometimes approximating functions are used, which have better smoothness properties; for example tanh. The usual defining condition for the class of saturation functions will now be given.

Definition 3.2.5: A function $\sigma\colon \mathbb{R} \mapsto \mathbb{R}$ is a *saturation function* if the following conditions hold:

1. σ is absolutely continuous, bounded, and $(1+|x|)\left|\left.\frac{d\sigma}{dx}\right|_x\right|$ is bounded almost everywhere;
2. if $x \neq 0$, $x\sigma(x) > 0$;
3. $\liminf_{x\to 0} \frac{\sigma(x)}{x} > 0$;
4. the function stays away from zero at infinity, i.e. $\liminf_{|x|\to\infty} |\sigma(x)| > 0$;
5. it is assumed here that the function is normalised at the origin - if $\dot{\sigma}(0)$ exists, then $\dot{\sigma}(0) = 1$.

It is possible to show that these conditions are equivalent to some others, allowing other equivalent definitions of the class. The following characterisation is taken from [34].

Lemma 3.2.6: A function σ satisfies Condition 1 of Definition 3.2.5 if and only if it satisfies the two conditions:

1. σ is globally Lipschitz;
2. there exists a constant $C_1 \geq 0$ such that for all $x, y \in \mathbb{R}$,

$$|x||\sigma(x+y) - \sigma(x)| \leq C_1|y|.$$

Definition 3.2.7: The constant $\kappa > 0$ is an *S-bound* for the saturation function σ if there exists $a, b > 0$ and a measurable function $\tau\colon \mathbb{R} \mapsto [a, b]$ such that:

1. $|\sigma(t) - \sigma(s)| \leq \kappa|t-s|$;
2. $|t||\sigma(t+s) - \sigma(t)| \leq \kappa|s|$;
3. $|\sigma(t) - \tau(t)t| \leq \kappa t\sigma(t)$.

(It is shown in [34] that such a bound always exists).

3.3 Classification of Problem Types

This dissertation is primarily a study of the analysis problems associated with linear systems with saturation in closed loop configurations. This is opposed to the synthesis problem, in which one attempts to systematically construct controllers which achieve some objective when placed in the closed loop with a specified plant. However, the analysis problem needs to be addressed first, as it is first necessary to develop mathematical tools which assess the performance of an arbitrary controller, to allow comparison between designs.

The new results presented here are almost exclusively associated with the gain computation problem for linear systems with saturation. This is associated with the disturbance rejection problem, in which one wishes to design a controller which keeps the plant output near some set point, taken here to be the origin. This is the most natural problem for an input/output based approach to control, as the gain of operators is a fundamental property which allows the development of robust stability theorems. The robustness comes from the fact that one can show that the closed loop may remain stable, even if the plant model is perturbed. The model perturbations are expressed as unknown, norm bounded operators, and stability guarantees given in terms of how large their norm may be. This approach has been quite successful in the linear theory, with $\mathcal{H}_\infty$ theory being that which results if the signal space being used is taken to be $\mathcal{L}_{2+}$.

In order to keep nomenclature clear, and aid the survey of other work, we will now list the main problem types addressed in this thesis. The naming scheme is a reasonable extrapolation of the linear systems terminology, although the presence of the nonlinearity(ies) means that quite distinct system configurations can result from the basic feedback loop. For the cases other than the standard feedback configuration, the definition will be in terms of a single input (or none, in the state stabilisation case) for ease of presentation. The special cases' definitions could be generalised to include higher dimension inputs, but it may occur that with multiple inputs, the resulting system will be a hybrid and less amenable to analysis.

3.3.1 The Standard Feedback Configuration

The most generic configuration addressed is that of the *standard feedback configuration*, with $\mathbf{G}$ being a linear system with saturation. The block diagram representation is given in Figure 3.1. The equations of the closed loop are given in operator form by:

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{G} & \mathbf{I} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

The controller $\mathbf{K}$ is not necessarily restricted to be a linear system with saturation, although this may be assumed in some cases to allow the analysis to be tractable. There are two problems which arise: the *analysis problem*, in which the operator norm is measured from an external disturbance to internal signals for a fixed $\mathbf{G}, \mathbf{K}$, and the *synthesis problem*, in which the plant is fixed, and the controller is chosen so that the closed operator meets some performance criterion, usually expressed as a condition upon the operator norms. For the synthesis problem, there are two sub-problems: the *state feedback* and the *output feedback* problems. The state feedback problem is an easier problem, as it is assumed the controller can

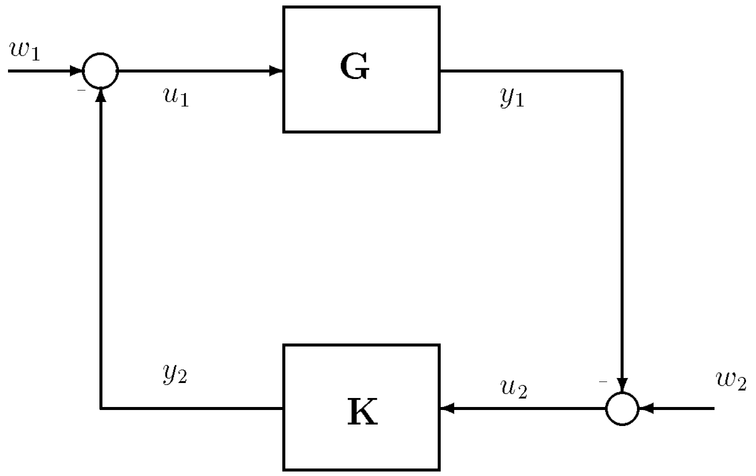

Figure 3.1: Standard Feedback Configuration

be designed with access to the plant state, as opposed to output feedback controllers, which only have access to the plant output, which is some function of the state and feedthrough on the inputs.

3.3.2 Open Loop Configuration

The most trivial case is the *open loop configuration* (Figure 3.2), where the output is given by $\mathcal{L}\{z\} = G(s)\mathcal{L}\{\sigma(w)\}$. It is trivial that the operator norm is unbounded if $G(s)$ is unstable, and if G is in $\mathcal{RH}_\infty$, then $\|P\|_{i,2} \leq \|G\|_\infty \|\sigma\|_{i,2}$ (similarly for other induced norms, and the discrete time case). A state space realisation is given by:

$$\begin{aligned} \Delta x &= Ax + B\sigma(w), \\ z &= Cx, \\ x(0) &= \xi_0. \end{aligned}$$

The importance of this configuration is that if the disturbance is interpreted as a control ouput, with no other disturbances, one studies this configuration to determine controllability (stabilisability), observability (etc.) properties.

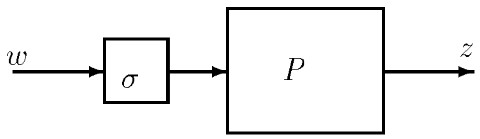

Figure 3.2: Open Loop Configuration

3.3.3 State Stabilisation with Input Saturation Configuration

In the state stabilisation with input saturation configuration, the objective of the control is to steer the plant state to the origin, in the absence of external disturbances. This problem

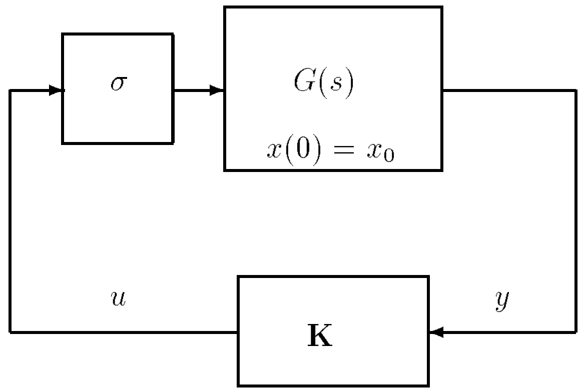

Figure 3.3: State Stabilisation with Input Saturation Configuration

is of considerable interest, however, as it historically has attracted the most analysis. (This is quite often referred to as the Lur'e problem). The defining state equations are:

$$\begin{aligned} \Delta x &= Ax + B\sigma(-\mathbf{K}y), \\ y &= F(x), \\ x(0) &= \xi_0, \end{aligned}$$

with $\mathbf{K}y$ the controller output (the block diagram is given in Figure 3.3). The problem (once again) can be either of the analysis or synthesis form; i.e., either one verifies whether a controller will stabilise the plant from some initial state(s), or one attempts to design a controller which stabilises the plant over some set of initial conditions.

The state stabilisation problem is somewhat unusual in the framework of this thesis in that for the problem to be non-trivial, a non-zero initial state is assumed for the plant, and there is no external disturbance. For the other problem configurations, the stability of the systems is specified as having a finite gain from the external disturbance to the output signals; here the stability is necessarily given in terms of the state space stability of the realisation of the closed loop.

3.3.4 Control Saturation Only Configuration

The control saturation (only) is a basic extension to the state stabilisation problem, as an external disturbance is to allowed to enter in a linear fashion. The defining equations are given by:

$$\begin{aligned} \Delta x &= Ax + Bw + B\sigma(-\mathbf{K}y), \\ y &= F(x, w - \sigma(\mathbf{K}y)), \\ z &= Cx + D(w - \sigma(\mathbf{K}y)), \end{aligned}$$

with the block diagram representation given in Figure 3.4. The analysis problem of concern here is to determine the operator norm of the mapping from w to output signals; the corresponding synthesis problem is to find a controller which meets some closed loop performance condition.

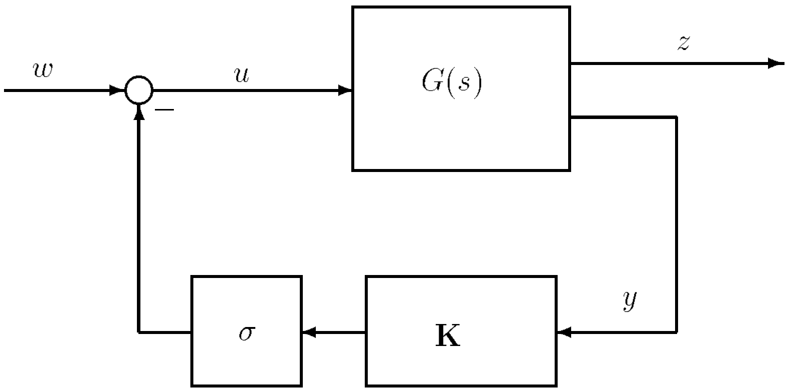

Figure 3.4: Control Saturation Only Configuration

This problem has attracted some recent attention by other researchers; their work is surveyed in Section 5.8. As noted there, the disturbance rejection may be simpler to analyse, but it appears that problems may arise due to the nature of the configuration. If the problem were viewed as a game, it is distinctly "unfair", as the disturbance "player" does not have its signal strength limited, and can override the control action with a bounded input. This will cause difficulties if unstable plants are encountered.

3.3.5 Control and Disturbance Saturation Configuration

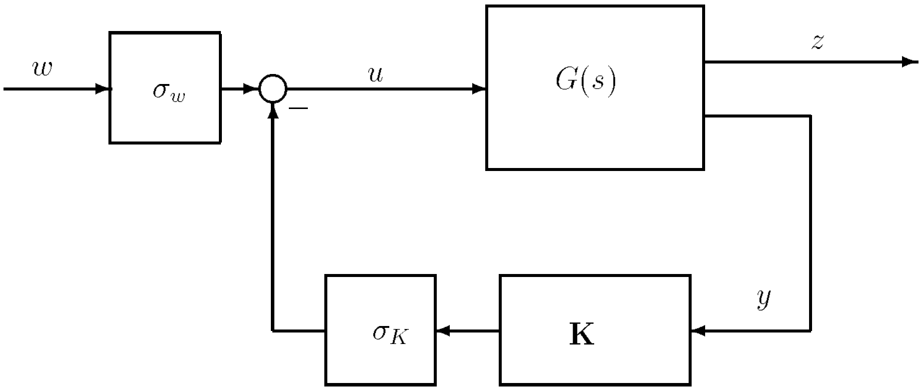

Figure 3.5: Control and Disturbance Saturation Configuration

The control and disturbance saturation configuration is more symmetric than is the case for control saturation (only), as the external disturbance enters the system through a saturation element (as shown in Figure 3.5). The system equations are of the form:

$$\begin{aligned} \Delta x &= Ax + B\sigma_w(w) + B\sigma_K(-\mathbf{K}y), \\ y &= F(x, \sigma(w) - \sigma(\mathbf{K}y)), \\ z &= Cx + D(\sigma(w) - \sigma(\mathbf{K}y)). \end{aligned}$$

Once again, the objective of the analysis is to compute the operator norm from the external disturbance w to internal signals.

The control and disturbance saturation configuration may allow more interesting analysis possibilities than is the case when only the controller output saturates, especially if the two saturation functions σ^w, σ^K have differing S-bounds. The present problem configuration seems somewhat contrived, however, and could be viewed as being the control only saturation case, with the disturbance lying in a $\mathcal{W}_{p,\infty}(N)$ set. If the analysis in the control only saturation configuration is to seriously address unstable plants, the use of bounded signal sets seems necessary, and so this problem may need to be studied despite its unusual appearance.

3.3.6 Plant Input Saturation Configuration

The plant input saturation configuration is the problem configuration which will receive the most attention in this dissertation. In it, the control output and external disturbance are summed together before entering the saturation element (as in Figure 3.6). The defining equations are given by:

$$\begin{aligned} \Delta x &= Ax + B\sigma(w - \mathbf{K}y), && (3.2)\\ y &= F(x, w - \mathbf{K}y)), && (3.3)\\ z &= Cx + D\sigma(w - \mathbf{K}y). && (3.4) \end{aligned}$$

Additionally, some results are given where there is a linear feedthrough term into the state dynamics; i.e.:

$$\begin{aligned} \Delta x &= Ax + B_1(w - \mathbf{K}y) + B_2\sigma(w - \mathbf{K}y), && (3.5)\\ z &= Cx + D_1(w - \mathbf{K}y) + D_2\sigma(w - \mathbf{K}y), && (3.6) \end{aligned}$$

although this does not correspond to the block diagram in Figure 3.6. The effect of the input $u = w - \mathbf{K}y$ is still symmetric in w, $-\mathbf{K}y$.

The control problem to be studied which arises from this configuration is the determination whether the system is input/output stable with respect to the disturbance w. The closed loop induced norm and incremental gain computation with respect to w is the topic of study in Chapters 5 - 7.

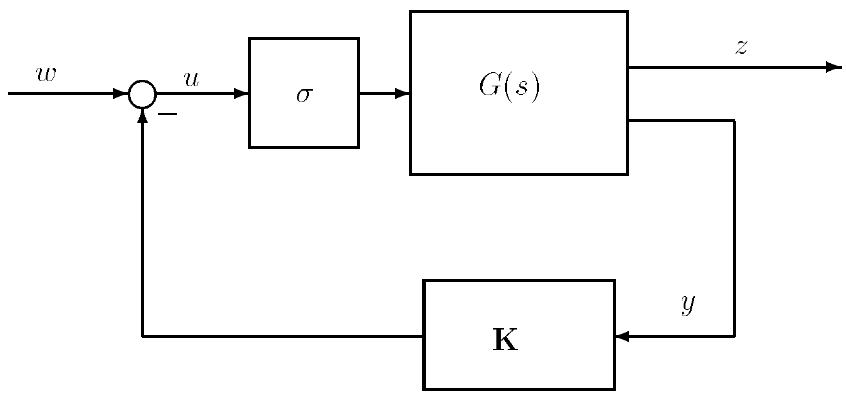

Figure 3.6: Plant Input Saturation Configuration

Remark 3.3.1: It has been commented that this configuration does not appear to be too natural a model for systems with actuator limitations, it would seem that "disturbance rejection" would more naturally be modelled by the controller saturation configuration. This was understood, but its study has been concentrated upon as it provides a more interesting mathematical problem. The mathematical configuration will arise from "disturbances" due to reference signals and output noise (with a linear controller), and may be interpreted in that light. Realistic problems will likely be a hybrid. Additionally, the B_1,D_1 terms have been added to allow the analysis of certain problems of interest (the Parallel Projection Operator (Section 8.3), and the PI anti-windup scheme in [1]).

3.3.7 Plant Output Saturation Configuration

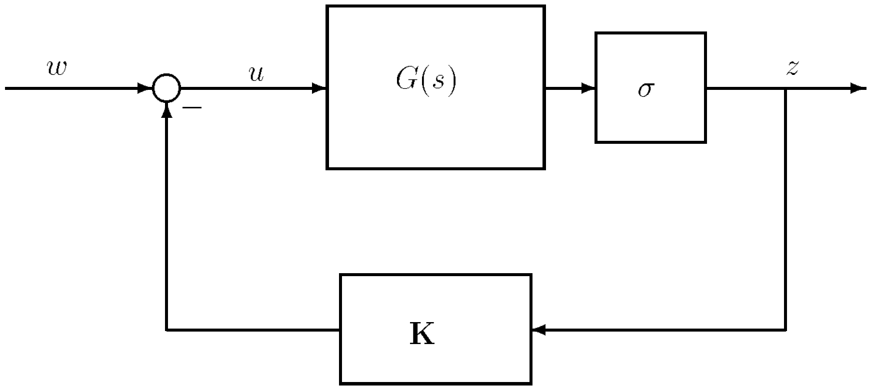

Figure 3.7: Plant Output Saturation Configuration

The plant output saturation configuration is that which arises when the saturation element follows the plant (as shown in Figure 3.7). It should be noted that this problem is not a specialisation of the standard feedback loop (in the framework of this thesis), as the plant is realised by the cascading of two linear systems with saturation. The defining equations are given by:

$$\begin{aligned}\Delta x &= Ax + B(w - \mathbf{K}z),\\ x(0) &= \xi_0,\\ z &= \sigma(Cx),\end{aligned}$$

with $z(t) \in \mathbb{R}$.

The most natural problem which seems to arise from this configuration is that of observer synthesis; i.e., given the output z and known controller output $\mathbf{K}z$, construct the best estimate of the state trajectory x. The performance of the observer could be given as the operator norm from the external disturbance to the state estimate error. As the control enters the state dynamic equation in a linear fashion, the control synthesis problem would appear to be a straightforward linear problem, if the state were accurately estimated.

This configuration seems somewhat more contrived than those arising from the saturation of control outputs, as actuator saturation seems a more natural condition. However, the study of this configuration may be needed as a simplest case for the analysis of feedback loops where it is assumed that the controller is a linear system with saturation, or the output is taken to be a piecewise linear function of the state.

3.4 Existence and Uniqueness of Solutions

A fundamental theoretical problem which needs to be examined when one defines operators in terms of a state space system is to ensure the result is well defined as an operator - i.e., that a solution exists, and is unique. When the system is defined on discrete time, it is obvious that there is a unique solution for linear systems with saturation, as the right hand side is single valued and the verification of a unique solution can be done inductively. When the system is continuous time, the problem is more difficult, but is a fairly standard problem in the mathematical theory of differential equations, and a basic result will now be presented. It is assumed that the controller is restricted to be a linear system with saturation - if a more generic nonlinear controller is to be studied, separate analysis will be required to verify that the problem remains well posed.

Note that if the controller is a linear system with saturation, the defining state space differential equation can be rewritten as:

$$\dot{x}(t) = Ax(t) + B_1\sigma(C_1x(t)) + B_2\sigma(C_2x(t) + w(t)) + B_3w(t), \tag{3.7}$$

where w is an external signal, and with A, B_i, C_i, appropriately sized matrices. As there is no structure on the nature of the external disturbance, we require that it is locally integrable, the definition of which will now be given. The existence and uniqueness will then be stated.

Definition 3.4.1: A signal w is *locally integrable* over an interval $\mathcal{I} \subseteq \mathbb{R}_+$ if for all $s < t$, $s, t \in \mathcal{I}$,

$$\int_s^t \|w(\tau)\| \, d\tau < \infty.$$

Theorem 3.4.2: For a differential equation of the form given by equation (3.7), if the signal w is locally integrable over an interval $\mathcal{I} \subseteq \mathbb{R}$, the solution x exists and is unique over $\mathcal{I}$.
Proof: This theorem is a specialisation of standard results, for example see [45, Proposition C.3.8]. □

Although local integrability implies that the signal is only (locally) in $\mathcal{L}_1$, this is sufficient for our purposes for all signals which are locally in $\mathcal{L}_{p+}$, for $p \geq 1$. This can be seen as although there are signals which are in $\mathcal{L}_{2+}$ (say) which are not in $\mathcal{L}_{1,+}$, any finite truncation will be, and so we have existence and uniqueness over any finite interval. This is because that problems occur over finite intervals only if the signals get large, but the square of a signal greater than 1 is greater than the signal itself.

3.5 Controllability

Once we are satisfied that the state space system will act in a well defined manner in response to external driving signals, we can then examine properties of the realisation which are of more engineering interest. The first properties we will examine are perhaps the most fundamental when it comes to controller (as opposed to observer) design: null controllability and reachability. The notion of null controllability is extremely important for state space approaches based on non-zero initial values: from which states can we return the state trajectory to the origin? Reachability is the converse problem: which states can be reached with the state trajectory starting at the origin? Reachability is of importance for input/output approaches which assume a zero initial condition; which is a typical assumption made in this dissertation.

The controllability/reachability properties of linear systems with saturation are quite distinct depending upon whether the system configuration is open or closed loop. This is unlike the case for linear systems, where one treats the realisation of a closed loop transfer function as if it were an open loop problem.

We will now formally define the controllability properties utilised in this dissertation, being null controllability, reachability, and complete controllability. They refer to the properties of a n-dimensional realisation, $\Delta x = f(x, w, t)$, with w an external disturbance signal.

Definition 3.5.1: A finite dimensional realisation (with its state trajectory denoted x) is termed *null controllable from a state* ξ, if for fixed $x(0) = \xi$, there exists an external signal w such that for some finite time T, $x(T) = 0$.

Definition 3.5.2: A state ξ is *reachable* for finite dimensional realisation (with an external driving signal w) if there exists a signal w and a finite time T such that if $x(0) = 0$, $x(T) = \xi$. (The point ξ may then be referred to as *reachable by time* T).

Definition 3.5.3: A state space realisation is *completely controllable over a set* Ω ($\subseteq \mathbb{R}^n$) if for all $\xi \in \Omega$, the system is null controllable from ξ as well as ξ being a reachable state.

Remark 3.5.4: It is possible to consider asymptotic null controllability, in which the state trajectory goes to 0 in the limit as T goes to ∞. In the linear case, this corresponds to the presence of stable uncontrollable modes.

3.5.1 Controllability of Linear Systems

The criterion for determining whether a linear shift invariant system is completely controllable is well known, and is in the form of a rank condition on the system matrices. The criterion is usually credited to Kalman, and can be found in most basic texts discussing the state space theory of linear systems; for example, see [8], [45]. It is an interesting and useful fact that the criterion for controllability is the same for both the continuous and discrete time case. Additionally, the distinction made here between open and closed loop controllability is not a concern, as if the controller is also a finite dimensional shift invariant linear system, its dynamics may be readily appended to that of the plant.

Theorem 3.5.5: A (finite dimensional, shift invariant) linear system whose state dynamic equation is $\Delta x = Ax + Bu$, is globally completely controllable if and only if

$$\operatorname{rank}\left\{[B, AB, A^2B, \ldots, A^{n-1}B]\right\} = n. \tag{3.8}$$

Proof: See [8], [45]. □

Remark 3.5.6: We can bend our terminology slightly, and say that if the rank condition in equation (3.8) holds, the matrices A, B form a *controllable pair*, or that $[A, B]$ is controllable.

3.5.2 Open Loop Controllability

We will now survey the controllability results for linear systems with saturation in open loop configurations. They have been known for some time, and give fundamental limitations on the ability to stabilise linear plants with constrained inputs. The class of systems for which the controllability results are given will now be fixed. The state dynamics are generated by:

$$\Delta x(t) = Ax(t) + B\sigma(u(t)), \tag{3.9}$$

an open loop equation.

We will now state a useful lemma, and the important result on null controllability for these open loop systems.

Lemma 3.5.7: The reachable set by a time t for the system generated by (3.9), with $x(0) = 0$, is both convex and compact.
Proof: See [29, p. 164]. □

Theorem 3.5.8: A system generated by the equation (3.9) is globally null controllable if and only if $[A, B]$ is controllable, and A has all its eigenvalues lying in the closed left half plane (continuous time) or in the closed unit disk (discrete time).
Proof: See [44] for the combined discrete and continuous time results. □

Remark 3.5.9: It is important to note that polynomially unstable modes are in fact stabilisable using bounded controls; these are modes with eigenvalues on the stability boundary, and corresponding to non-trivial Jordan blocks.

Remark 3.5.10: The author has ignored the issue of asymptotic null controllability, in which one adds asymptotically stable uncontrollable modes. These do not effect input-output behaviour with zero initial conditions.

This theorem has quite a long history, which befits its fundamental importance for the study of saturating systems. Although the continuous time results are better known, the discrete time version of the results appeared first in the papers by C. A. Desoer and J. Wing ([12], [59]). That an open loop asymptotically stable dynamics matrix was a sufficient condition for null controllability was well known earlier, as noted in the discussion in the paper [28] by E. B. Lee. The necessity conditions for the continuous time case are ussually credited as first appearing in the text by E.B. Lee and L. Markus ([29]), although it was only proved in its full generality in the single input case. The papers of Brammer [5], and Schmitendorf and Barmish [42] removed the assumptions, and the paper by Sontag [44] gave an algebraic approach for both continuous time and discrete time systems. An important component of these works is the extension to more general constraint sets; such a level of generality is not of concern to the problems studied in this dissertation, however. In the paper by Brammer [5], which extends the preliminary results of Saperstone and Yorke [41], reachability is characterised by using the convexity of the reachable set.

Theorem 3.5.11: A continuous time system generated by the equation (3.9) with zero initial conditions is globally reachable if and only if $[A, B]$ is controllable and A has no eigenvalues with strictly negative real part.
Proof: See [5]. □

Remark 3.5.12: The result can be seen to follow from Theorem 3.5.8 if we examine the negative time version of the system - i.e., replace A, B with $-A, -B$. This result also indicates that the only linear systems with saturation which are globally completely controllable are those with eigenvalues lying solely on the imaginary axis. Additionally, the paper [5] extends to general constraint sets the readily apparent fact that if the local linearisation is locally controllable, the system is locally completely controllable. The author has not found specifically the extension to discrete time in the literature, and thus the proof for that extension will be given for Theorem 3.5.13, mainly for illustrative purposes. It is not claimed to be original, however.

Theorem 3.5.13: A discrete time linear system generated by (3.9) (with zero initial conditions) is globally reachable if and only if $[A, B]$ is controllable, and all eigenvalues of A have modulus greater than or equal to 1.
Proof: We will first show that if the A matrix is not invertible, then the system is not globally reachable. If A^{-1} does not exist, it is a basic fact from linear algebra that A has an eigenvalue at 0, and there exists an invertible matrix T such that we can write $\hat{x} = Tx$, with $\hat{x}$ defined by the dynamic equation:

$$\begin{aligned} \hat{x}[k+1] &= \begin{bmatrix} \hat{a}_{11} & \cdots & \hat{a}_{1,n} \\ & \cdots & \\ \hat{a}_{(n-1),1} & \cdots & \hat{a}_{(n-1),n} \\ 0 & \cdots & 0 \end{bmatrix} \hat{x}[k] + \hat{B}\sigma(u[k]), \\ \hat{x}[0] &= 0. \end{aligned}$$

If we examine the equation for $\hat{x}_n$, we see that $\hat{x}_n[k+1] = \hat{B}_n\sigma(u[k])$. As there exists some bound K for σ, we have that $\|\hat{x}_n\|_\infty \leq K\hat{B}_n$. The transformed state $\hat{x}$ is thus confined to some subset of $\mathbb{R}^n$. As x is related to $\hat{x}$ by a linear transformation, this implies that x is also confined to a subset of $\mathbb{R}^n$, showing that the lack of invertibility of A ensures that the system is not globally reachable (this of course is a subset of the case where A has eigenvalues lying in the unit disk).

We will then turn to the case where A^{-1} exists. We can write the "negative time" dynamics as

$$x[k] = A^{-1}x[k+1] - A^{-1}B\sigma(u[k]).$$

Take as a new state space system,

$$y[k+1] = A^{-1}y[k] - A^{-1}B\sigma(u_r[k+1]). \tag{3.10}$$

Fix an input u_r to the system (3.10), defined over a finite interval $0, \ldots, T$, with $y[0] = y_0$, $y[T] = y_f$. Then if the input to the forward time system (3.9) u,i is taken to be $0, \ldots, T$ as $u[k] = u_r[T-k]$, it is easy to verify that $x[0] = y_f$, $x[T] = y_0$. As such, we have that the forward time system (3.9) is globally reachable if and only if the reversed time system (3.10) is globally null controllable. By Theorem 3.5.8, this is true if and only if $[A^{-1}, -A^{-1}B]$ is controllable, and all eigenvalues of A^{-1} lie on the closed unit disk. This is then equivalent to $[A, B]$ being controllable and the eigenvalues of A all lying outside the unit disk, as the controllability matrix's rank is invariant under pre-multiplication by $-A^{-n}$, and the eigenvalues of A^{-1} are the multiplicative inverses of the eigenvalues of A.□

The importance of these results is that from an input-output point of view, global state space results are of limited use. If the system is open loop asymptotically stable, the state is confined to some ellipsoid in the state space, and if the system is (purely) open loop unstable, there is some ellipsoid beyond which the system is not stabilisable. In the latter case, some *a priori* method of determining whether or not it is realistic that the state will in fact be able to make such a large excursion so as to make the system non-stabilisable is needed. If so, the control problem as posed is untenable, otherwise the restrictive assumptions needed to confine the state to some compact set is justifiable.

Although the previously stated results have been long known, they have not addressed the problem of determining precisely which subset of $\mathbb{R}^n$ can be reached by the use of a bounded input. The paper by J.E. Gayek [17] surveys the work on this topic, and notes that there are essentially two approaches - the exact determination of the reachable/controllable set by the use of optimal control techniques, or the use of Lyapunov-like functions to overbound the reachable sets. The first approach gives rise to fairly complicated results, especially in discrete time. The second broad approach is much more closely allied to the basic systems analysis tools used in this dissertation, and are of more interest for present purposes.

We will now state a fairly generic result used to bound reachable sets, which is based on the use of a linear matrix inequality. This was taken from [4].

Theorem 3.5.14: If given a (continuous time) linear system with saturation with realisation:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B\sigma(u(t)), \\ x(0) &= 0, \end{aligned}$$

with κ a S-bound for σ, we have that $x(t) \in \Omega$ for all $t \in \mathbb{R}_+$, where $\Omega = \{\xi : \xi^T P \xi \leq 1\}$ if there exists a symmetric, positive definite matrix P and positive scalar α, which solve the linear matrix inequality problem:

$$\begin{bmatrix} A^T P + PA + \alpha P & \kappa PB \\ \kappa B^T P & -\alpha I \end{bmatrix} \leq 0, \tag{3.11}$$

$$P > 0, \alpha \geq 0.$$

Proof: The proof for this statement may be found in [4, Section 6.1.3]. □.

Remark 3.5.15: Characterising the inequality (3.11) as an LMI is not purely correct, as the product of αP appears. Rather, one must determine for which (preferably small) α a solution exists.

By using the technique of reversing the flow of time, if A has all poles in the open right half plane, we see that the set over which A is null controllable is overbounded by an ellipsoid.

3.5.3 Closed Loop Controllability

The study of controllability and reachability for closed loop configurations has not been extensively studied in the literature. Most studies have examined over which set a controller can stabilise the plant (or conversely, to which states the controller can drive the plant); for the problems studied in this dissertation, we are more interested in studying which states an external disturbance can drive the state trajectory to, with a fixed controller in place.

As usual, for a single input problem, we can assume that we can append any controller dynamics to that of the plant, and then then treat the controller as a state feedback. If we are allowing general $\mathcal{L}_{p+}$ (resp. ℓ_{p+}) disturbances, the open loop results would then give a good guide as to the form of the reachable/controllable set, as we could merely override the control action and then treat it as an open loop problem. For example, if the (augmented) plant dynamics matrix is Hurwitz, the reachable set is overbounded by some ellipsoid.

As we are more interested in the effect of using $\mathcal{W}_{p,q}(N)$ sets, the problem then becomes more intricate. In that case, it is most likely that the control action cannot be overridden in the large, and more involved constructions would be necessary to reach states near the reachable boundary.

Once again, the set determination problem could be dealt with using optimal control methods without much extra difficulty; it is most likely that a general analysis would be extremely difficult to do, however. At present, the simplest and most consistent method is to use Lyapunov-like functions to overbound the reachable set. This will be discussed in Sections 5.7 and 6.6, when the necessary Lyapunov theory has been put in place.

3.6 Observability

The dual property to controllability is observability, which is also of fundamental significance in the output feedback problem. The usual concern is to determine whether it is possible to construct an observer - a system which uses measurements to attempts to construct an estimate for the plant state. Observability is of importance in its own right for use in monitoring system status; however it is also plays a role in controller design. A fairly standard interpretation of modern robust controller implementations (as the result of an $\mathcal{H}_\infty$ design, say) is that the controller will consist of a state feedback controller following an observer, both of which are generated by appropriate Riccati equations.

The primary concern in this dissertation is the stabilisation problem, so the observer construction problem is not addressed. For present purposes, a weaker condition is sufficient for analysis, which amounts to a well posedness constraint. This condition appears to be non-standard, and is referred to here as $\mathcal{L}_p$-detectability. For linear systems, $\mathcal{L}_p$-detectability is equivalent to the usual definition of detectability, with an added controllability requirement.

The two definitions of observability conditions studied in this dissertation will now be formally stated. The first, distinguishability, is stated here for completeness and for comparison purposes (as there are many differing definitions for observability put forth for nonlinear systems); the second definition ($\mathcal{L}_p$- detectability) is the one usually referred to in the results developed in this dissertation.

Definition 3.6.1: An n-dimensional state space system is *distinguishable* if for all states $x, y \in$

$\mathbb{R}^n$, $x \neq y$, there exists some input defined over some finite interval $\mathcal{I}$ such that the system output over $\mathcal{I}$ for the system with initial condition x does not equal that when the initial condition is y.

Definition 3.6.2: The state space operator $\mathbf{\Upsilon}_x\colon w \mapsto x$ which corresponds to a realisation of an operator $\mathbf{\Upsilon}$ is *$\mathcal{L}_p$-detectable* (resp. *ℓ_p-detectable*) if for $w \in \mathcal{L}_{p+}$ (resp. ℓ_{p+}), $\mathbf{\Upsilon} w \in \mathcal{L}_{p+}$ (resp. ℓ_{p+}) implies that $\mathbf{\Upsilon}_x w \in \mathcal{L}_{p+}$ (resp. ℓ_{p+}).

The caveat that the driving inputs are bounded is important for the $\mathcal{L}_p$-detectability condition to make sense, as (for example) if there are feedthrough terms, a pathological choice of an unbounded input could be used to cancel out the state read out component.

It appears that this detectability condition is non-standard; as most approaches focus either solely on state space or operator stability. One of the closest definitions used is that by Willems in [58], which he terms *uniform observability*. This amounts to a similar condition to that given here, but with no external input. His definition is aimed at the generation of Lyapunov functions which are used to show state space stability. The concern here is to give some form of state space convergence with inputs of non-compact support. One may also note the development of work similar to Willems in the paper by Hill and Moylan [24]. That paper is a study of the input/output problem when there may be multiple equilibria, a possibility which is not addressed here.

Remark 3.6.3: We will now give an example as to why the $\mathcal{L}_p$-detectability definition is better suited than the *uniform observability* definitions used in [58],[24] for the study of input-output systems. In those papers, the observability criterion is developed in terms of the null input trajectories. However, this is not felt to be as useful in the study of the operators themselves, at least in a direct form. Take as an example, the operator $\mathbf{\Upsilon}\colon w \mapsto z$, generated by:

$$\dot{z}(t) = f(z(t), w(t)), \qquad z(0) = 0.$$

As the readout function is the identity, the system trivially meets any sensible observability criterion. However, we then append the state x_{n+1}, with the dynamics of x_{n+1} generated by:

$$\dot{x}_{n+1}(t) = \begin{cases} -x_{n+1}(t) & w(t) = 0, \\ x_{n+1}(t) + |w(t)| & w(t) \neq 0, \end{cases}$$

with $x_{n+1}(0) = 0$ (the output is still independent of x_{n+1}). It is easy to see that x_{n+1} will go to $+\infty$ for all w with non-compact support, yet will go to the origin with $w \equiv 0$, and hence meets the sort of definitions used in [58],[24]. This is clearly objectionable. The obvious problem is that f is no longer continuous, and so one can argue that this is too pathological an example. But it possible to redefine the dynamics for $w \leq \epsilon$, and give a definition continuous in w. In such a case, it should be possible to prove that uniform stability is sufficient for w in $\mathcal{L}_{p+}$ spaces, with $p < \infty$, but ill-posed behaviour will still occur for $p = \infty$. Hence, even though it is likely that one will be able to show the equivalence for many classes of systems of interest, and such proofs may have their use, $\mathcal{L}_p$-detectability still seems to be of more fundamental interest.

3.6.1 Observability of Linear Systems

When the observability definitions are applied to finite dimensional, shift invariant systems, algebraic conditions on the dynamics and readout matrices result. We will now give the formal defining conditions on these linear systems, and then relate the algebraic conditions to the observability definitions used in this dissertation.

Definition 3.6.4: An n-dimensional linear state space system of the form,

$$\begin{aligned} \Delta x &= Ax + Bu, \\ z &= Cx + Du, \end{aligned}$$

is said to be *observable* if

$$\operatorname{rank} \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} = n.$$

In this case, we may also say that $[C, A]$ form an observable pair.

Definition 3.6.5: For a system as in Definition 3.6.4, it is said to be *detectable* if

$$\operatorname{rank} \begin{bmatrix} \lambda I - A \\ C \end{bmatrix} = n,$$

for all $\lambda \geq 0$ (continuous time) or $|\lambda| \geq 1$ (discrete time). We may say that $[C, A]$ is detectable.

Theorem 3.6.6: For a finite dimensional, linear shift invariant state space system in continuous (resp. discrete) time, the system is distinguishable if and only if $[C, A]$ is an observable pair.

Proof: This is a standard result, for example, see [45, p. 200]. □

Theorem 3.6.7: Fix a linear shift invariant system with a finite dimensional transfer function $G(s)$(resp. $G(z)$) $= \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$, with $[A, B]$ controllable. The realisation is $\mathcal{L}_p$ (resp. ℓ_p)-detectable if and only if $[C, A]$ is detectable.

Proof: (The author has not seen any proof of this result in the literature; and so it will be shown here.) We will first show that $[C, A]$ detectable implies that the system is $\mathcal{L}_p$ (resp. ℓ_p)-detectable. Without any loss of generality, assume that $D = 0$, as for any input $u \in \mathcal{L}_{p+}$ (resp. ℓ_{p+}), $Du \in \mathcal{L}_{p+}$ (resp. ℓ_{p+}) and may be subtracted away without affecting whether z lies in the normed output space. Then, using the fact that the system is controllable and detectable, we can express $G(s) = C(sI - A)^{-1}B$ as the right coprime fraction $G(s) = N(s)M^{-1}(s)$, with N, M realised by:

$$\begin{bmatrix} M(s) \\ N(s) \end{bmatrix} = \left[\begin{array}{c|c} A - BF & B \\ \hline \begin{bmatrix} -F \\ C \end{bmatrix} & \begin{bmatrix} I \\ 0 \end{bmatrix} \end{array}\right],$$

with F a stabilising state feedback ([54, p. 83]). Then, denote the transfer function of the mapping to the state by $G_x(s) = (sI - A)^{-1}B$, with right coprime fractional representation $G_x(s) = N_x(s)M_x^{-1}(s)$. Examining their realisations, we see that we can choose $M_x(s) = M(s)$, and $N(s) = CN_x(s)$. By then using the result of Vidyasagar ([53]) that for all $\mathcal{L}_{p+}$ signal spaces, the domain of the plant is given by $\mathbf{M}\mathcal{L}_{p+}$. As $G(s)$ and $G_x(s)$ share a right coprime fraction, we can then assert that their domains are the same, and so the $\mathcal{L}_p$-detectability condition holds. The proof is the same in discrete time.

The proof to show that $\mathcal{L}_p$ (resp. ℓ_p)-detectability implies that $[C, A]$ is detectable is done by contradiction. Assume that the system is $\mathcal{L}_p$-detectable, but $[C, A]$ is not detectable. Then, we have that there exists some $\xi \in \mathbb{R}^n$ such that $Ce^{At}\xi = 0$ for all $t \geq 0$, and $\lim_{t\to\infty} e^{At}\xi \to \infty$. By the controllability assumption, for an arbitrary $T > 0$, there exists an essentially bounded u defined on $[0, T]$ such that $x(T) = \xi$. Then, taking $u(t) = 0$ for all $t > T$, we have that $u, z \in \mathcal{L}_{p+}$. However, $x \notin \mathcal{L}_{p+}$, contradicting the $\mathcal{L}_p$-detectability assumption. The proof for discrete time is an analogous construction. □

It may appear eccentric to add the controllability assumption to Theorem 3.6.7, however as the definition is given in terms of input/output behaviour, the B matrix therefore needs to be accounted for. As the usual assumption made in this dissertation is that there is a zero initial state, it is not a restriction to assume that $[A, B]$ is controllable, as uncontrollable modes cannot be excited and affect input/output behaviour.

3.6.2 Open Loop Observability

We will now examine the observability properties of linear systems with saturation in an open loop configuration. It is perhaps not surprising that the observability follows from the observability of the local linearisation.

Theorem 3.6.8: The realisation defined by:

$$\begin{aligned} \Delta x &= Ax + B\sigma(u), \\ z &= Cx + D\sigma(u), \end{aligned}$$

is (open loop) distinguishable if $[C, A]$ is observable, and is $\mathcal{L}_p$ (resp. ℓ_p)-detectable if $[C, A]$ is detectable and $[A, B]$ is controllable.

Proof: It is a basic result of linear system theory that the system is distinguishable with the null input if and only if $[C, A]$ is observable; hence the addition of the input nonlinearity does not affect the distinguishability of the system. The fact that $[C, A]$ detectable implies that the system is $\mathcal{L}_p$ (resp. ℓ_p)-detectable is shown by contradiction. Assume there exists some $u \in \mathcal{L}_p$ (resp. ℓ_p) such that $x \notin \mathcal{L}_p$ (resp. ℓ_p), $z \in \mathcal{L}_p$ (resp. ℓ_p). If we then define $v = \sigma(u)$, as σ is a bounded operator, then $v \in \mathcal{L}_p$ (resp. ℓ_p). Take the linear system:

$$\begin{aligned} \Delta x_l &= Ax_l + Bv, \\ z_l &= Cx_l + Dv, \end{aligned}$$

with $x_l(0) = x(0)$, and we have that $x_l = x$, $z_l = z$. By applying Theorem 3.6.7, we arrive at a contradiction. □

3.6.3 Observability of Output Saturated Systems

The case of plant output saturation gives rise to a more interesting open loop problem, as the nonlinearity directly affects the output signal. The paper by Schwarzschild, Sontag and Hautus [43] addresses this problem for continuous and discrete time systems. There seems to be very few other examinations of this problem in the literature, and on its face may be only of mathematical interest as a "dual" problem to that of input saturation. As the multiple output case is more intricate, the results for the single output case only are given after a preliminary definition. It should be noted that it is only a necessary condition for the discrete time case.

Definition 3.6.9: For a matrix triple A, B, C, the *Markov parameter sequence*, denoted $\mathcal{A}$, is given by:

$$\mathcal{A} = \{CB, CAB, CA^2B, \ldots\}.$$

Theorem 3.6.10: The continuous time, single output, output saturated system generated by:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t), \\ z(t) &= \sigma_0(Cx(t)), \end{aligned}$$

is distinguishable if and only if $[C, A]$ is observable, and either

1. $\mathcal{A} \not\equiv 0$, or
2. $\inf_{t \geq 0} |Ce^{At}x| = 0, \; \forall x \in \mathbb{R}^n$.

Proof: See [43]. □

Theorem 3.6.11: If the discrete time, single output, output saturated system generated by:

$$\begin{aligned} x[k+1] &= Ax[k] + Bu[k], \\ z[k] &= \sigma_0(Cx[k]), \end{aligned}$$

is distinguishable, then $[C, A]$ is observable, and either

1. $\mathcal{A} \not\equiv 0$, or
2. for all $x \in \mathbb{R}^n$, there exists some $k \geq 0$ such that $|CA^kx| < 1$.

Proof: See [43]. □

3.6.4 Closed Loop Observability

It appears that the topic of observability for linear systems with saturation in a closed loop configuration has not been previously addressed. What will be presented here is a preliminary result on $\mathcal{L}_p$-detectability which has proved sufficient for present purposes.

Our main concern in this dissertation is the study of single input, plant input saturation closed loop systems. The feedback is assumed to be of the state feedback form, without much loss of generality (since the assumed absence of output noise allows one to append controller dynamics to that of the plant). However, even for purely linear systems, it is possible to choose a state feedback which destroys the observability of an open loop observable plant. This is not possible for an output feedback controller. This property still holds for the nonlinear systems in question, which will now be shown.

Theorem 3.6.12: Fix a single input plant input saturation configuration system (defined by (3.2) - (3.4)), with $D = 0$, and with the saturation function being the standard saturation. Assume that the control feedback $\mathbf{K}$ is memoryless, and is a linear output feedback (i.e., $K(\xi) = K_1 C\xi$, with K_1 a matrix of appropriate dimension). Then the detectability of $[C, A]$, the controllability of $[A, B]$, and the stability of the matrix $A - BK_1C$ (i.e., the closed loop local linearisation) implies that the (nonlinear) closed loop realisation is $\mathcal{L}_p$-detectable (for all p).

Proof: In order to show that the system is $\mathcal{L}_p$-detectable, assume that the input w and output z lie in $\mathcal{L}_{p+}$. We now need to verify that x lies in $\mathcal{L}_{p+}$.

Note that $u = w - Kx = w - K_1 z$, and hence $u \in \mathcal{L}_{p+}$. It then follows that $\sigma_0(u) \in \mathcal{L}_{p+}$, and $u - \sigma_0(u) \in \mathcal{L}_{p+}$. Let $m = \sigma_0(u) - u$. It is then interesting to note that the signals x, z are the same as would arise if the signal $\tilde{w} = w + m$ were input to the equivalent linear system. As $\tilde{w} \in \mathcal{L}_{p+}$, the assumed stability of the equivalent closed loop linear system, and its observability, imply that $x, z \in \mathcal{L}_{p+}$. $\square$

Remark 3.6.13: It is a trivial re-statement to give the discrete time proof. Additionally, although the proof seems straightforward and the details are omitted, care should be taken when using the device employed here: the replacement of the nonlinearity with a cancelling external signal, and then using the stability of the resulting linear system to prove some stability result. There is normally no way of ensuring *a priori* that the signal m lies in a bounded signal space; however, that is easily achieved here.

Chapter 4

Survey Of Stabilisation Results For Nonlinear Systems

4.1 Introduction

This chapter is a survey of results taken from the literature on the stability analysis of nonlinear systems. Some of the results will be for generic nonlinear systems, some more specialised for the class of linear systems with saturation. The analysis tools developed are basic ones used throughout nonlinear system analysis, and to a large extent they underlie much of the linear system stability theory. In the case of linear systems, it is usually possible to reduce the stability results to algebraic tests, which then become the focus of concern.

In Section 4.2, the well known Lyapunov stability test, relying on Lyapunov functions, is motivated. The generic form of potential functions are of extreme importance throughout the analysis methods used for the remainder of this dissertation. For this reason, the framework is quite detailed, and care has been taken to develop the theory in terms of compact subsets of the state space, for reasons which will be readily apparent.

When state space stability of linear systems with saturation is examined, they seem to exhibit unusual global properties. In Section 4.3 results on the stabilisation problem for these plants are briefly touched upon. They are viewed to be of more importance to the synthesis problem, which is beyond the scope of this dissertation.

From the operator theoretic point of view, the theory in Section 4.4 is quite important, being the Small Gain Theorem, and its corollary, the Circle Criterion. The Small Gain Theorem is similar to the Contraction Mapping Theorem, and it allows one to guarantee input/output stability for closed loops in the presence of uncertainties over which one has little knowledge, other than norm bounds.

In order to apply the Contraction Mapping based theory, one needs to be able to bound operator gains. A typical method to compute the gains is via the use of dissipation results, the basic theory on which was developed in the early 1970's. The theory of dissipative systems has become quite mathematically sophisticated in recent approaches, due to technical difficulties in the computation of storage functions. An introduction to dissipative system concepts is given in Sections 4.5- 4.8.

Finally, in Section 4.8, some comments are made on the use of linear matrix inequalities to find dissipation functions. This area has been attracting some recent research interest, and gives rise to efficient computational schemes.

4.2 Lyapunov Stability

4.2.1 Lyapunov's Direct Method

One of the oldest and most generic methods of determining the state space stability of a nonlinear system is via the use of *Lyapunov functions*, which is variously referred to as Lyapunov's Direct Method, or his Second Method. This method is generally applicable to what are often referred to in the control literature as autonomous systems, or systems with no external inputs (in the usual usage in mathematics, an autonomous differential equation refers to the corresponding system as being shift invariant), although the result can be readily adapted to apply to a fixed input by treating the system as time varying, and using the time varying version of this theory. However, as we shall treat almost exclusively shift invariant systems, the result will presented as such, which allows some advantages in presentation. Additionally, by weakening "definiteness" assumptions, one can characterise weaker forms of stability. However, when one is concerned with input/output stability for $\mathcal{L}_p$ norms ($p < \infty$) asymptotic state space stability is essential. The main results of interest for this weaker concept of stability are usually referred to as the *LaSalle Invariance Principle*. For a fuller discussion of Lyapunov and LaSalle stability theory, see the texts [45], [52]. It should perhaps also be noted that the text by Lyapunov is now available in an English translation [35].

Assumption 4.2.1: The set $\Omega \subseteq \mathbb{R}^n$ used in the following results is assumed to either be bounded and simply connected, or else if it is unbounded, $\Omega = \mathbb{R}^n$.

Definition 4.2.2: The function $V: \Omega \mapsto \mathbb{R}_+$ is a *candidate Lyapunov function* for a subset $\Omega \subseteq \mathbb{R}^n$ if V is continuously differentiable, and there exists continuous, nondecreasing, functions $\alpha, \beta: \mathbb{R}_+ \mapsto \mathbb{R}$ such that:

1. $\alpha(0) = \beta(0) = 0$, $\alpha(y), \beta(y) > 0$ when $y \neq 0$;
2. $V(0) = 0$;
3. $\alpha(\|\xi\|) \leq V(\xi) \leq \beta(\|\xi\|)$ for all $\xi \in \Omega$;
4. if Ω is unbounded, it is required that $\alpha(y) \to \infty$ uniformly as $y \to \infty$;
5. or else if Ω is bounded, for all $\xi \in \partial V$, $V(\xi)$ equals some constant κ.

Remark 4.2.3: The rationale for the last requirement is to ensure that the set Ω will remain invariant if the candidate function is actually a valid Lyapunov function.

Definition 4.2.4: A state space system with unique trajectories generated by $\Delta x(t) = f(x)$ ($x(t) \in \mathbb{R}^n$), $x(0) = \xi_0$ with $f(0) = 0$, is said to be *uniformly asymptotically stable* over the set $\Omega \subseteq \mathbb{R}^n$ if:

1. for all $\epsilon, t_1 > 0$, there exists some $\delta(\epsilon)$ such that $\|x(t_1;\xi_0)\| < \delta(\epsilon) \implies \|x(t;\xi_0)\| < \epsilon$ for all $t > t_1$;
2. for all $\xi_0 \in \Omega$, $\lim_{t\to\infty} \|x(t;\xi_0)\| \to 0$.

Remark 4.2.5: The Lyapunov stability theorem will now be given for continuous time systems; the discrete time case is addressed in Section 6.2.

Theorem 4.2.6 (Lyapunov): Fix an initial value state space system $\dot{x}(t) = f(x)$, $x(0) = \xi_0 \in \Omega$, with unique trajectories. Then if there exists a candidate Lyapunov function V such that for all $\xi \in \Omega$,

$$\left.\frac{dV}{dx}\right|_{\xi} f(\xi) \le M(\xi) < 0, \forall \xi \in \Omega \setminus 0, \tag{4.1}$$

with $M\colon \Omega \mapsto \mathbb{R}_-$ continuous and negative definite, then the system is uniformly asymptotically stable over Ω.

Remark 4.2.7: The term $\left.\frac{dV}{dx}\right|_{\xi} f(\xi)$ is the *Lie Derivative* of V with respect to f, and will often be denoted $\dot{V}(x)$ - i.e., the time derivative of $V(x(t))$.

It is usual to note at this point that this is only a sufficient condition for stability - if a given candidate Lyapunov function fails the test, it tells us nothing. However, converse Lyapunov theorems have been developed, which tell us that such a function necessarily exists if the system is stable (in a sufficiently strong sense) - the difficulty is that the actual computation of the function can be quite numerically intensive, and the results are seen to be more of theoretical interest than being practical. Linear systems are the exception to this rule, and the computation is well adapted to solution on digital computers - as will be discussed in Section 4.2.2. From an input/output based approach, the characterisation of the converse theorems found in the papers by Willems [55], and Hill and Moylan [23] are of interest to the present work. The results in these papers are extended to provide an existence theorem when the problem of induced norm computation is addressed in Chapter 5.

4.2.2 Lyapunov Functions and Linear Systems

Finding a Lyapunov function for linear systems is made relatively easy as we can assume that it is quadratic in form. This, coupled with the fact that the corresponding test is turned into an algebraic test on the dynamics matrix means that it is possible to develop stability theory for linear systems in a straightforward and computationally efficient fashion. The motivation as to why the test over the state space is turned into a single algebraic test is that a purely quadratic function is characterised completely by its second derivative, which is a fixed matrix.

Proposition 4.2.8: Fix a candidate Lyapunov function of the form

$$V(x) = x'Px,$$

with $P = P' \in \mathbb{R}^{n \times n}$. Then, if $\dot{x}(t) = Ax$, the Lie derivative of V with respect to Ax is given by:

$$\dot{V}(x(t)) = x'A'Px + x'PAx.$$

Theorem 4.2.9 (Lyapunov): Fix a linear system $\dot{x} = Ax$. The following statements are then equivalent:

1. A is Hurwitz;

2. for any symmetric $Q > 0$, there exists a symmetric matrix P which solves the *algebraic Lyapunov equation*:

$$A'P + PA = -Q;$$

3. the system is uniformly asymptotically stable.

Proof: This is a standard result of linear theory; for a proof, see [8, Theorem 8-20]. □

4.3 State Space Stabilisation of Linear Systems with Saturation

Having addressed the main method of determining the stability of nonlinear systems, we will now briefly examine the state space stabilisation problem for linear systems with saturation. It is an obvious consequence of the controllability results in Section 3.5 that unstable linear plants are not globally stabilisable with purely constrained inputs, and stable plants can be trivially "stabilised" with null controllers. Hence, if one is concerned with global stabilisation of such systems, the only nontrivial results are for systems whose poles only lie upon the axis. And in fact, one can specialise further and examine n-th order integrators, as the problems arise from the order of the poles, rather than their location upon the imaginary axis.

In his 1969 paper [16], A.T. Fuller examined the stabilisation problem for integrator chains of various orders. He demonstrated that for first and second order integrators, there are linear state feedback controllers which globally stabilise the plant; however, third order and higher integrators are not globally stabilisable by purely linear state feedbacks. This result was independently arrived at in the paper by Sussman and Yang [48], although in the latter case the proof had been extended to more general saturation functions. This led to the paper by Teel [50], in which it was shown by the use of a memoryless state feedback controller consisting of cascaded linear mappings and standard saturation functions, that it is possible to globally stabilise integrator chains of arbitrary length. The proof is made in an inductive fashion, and the strategy consists of stabilising the state block by block by the means of a linear subcontroller, with each block of consisting of one or two states. Later work by Sussman, Sontag and Yang in [46, 47] extended this construction to include generic saturation functions. At present, we are primarily concerned with the analysis problem, and so these results will be summarised in a simplified theorem which will now be given. The important point at present is that it is possible to globally stabilise all systems whose poles lie in the closed left half plane (only).

Theorem 4.3.1: Fix a system generated by: $\dot{x} = Ax + \sigma(u)$, $x(0) = \xi_0$. Then if the system is null controllable (i.e., all the eigenvalues of A are non-positive), there exists a state feedback of the form:

$$u(x) = a_1 x + \sigma(a_2 x + \sigma(a_3 x + \ldots)),$$

(with the function description terminating after a finite number of terms) which globally stabilises the plant.
Proof: See the papers [50], [46], [47] for details. □

For large systems with many saturation and similar nonlinearities (i.e., relays, hysteresis elements) present, the determination of whether a system is state space stable is quite a difficult problem. In [37] (and related publications) N.B.O.L. Pettit and P.E. Wellstead describe a graphical method of describing the state trajectories for such systems. The state space is divided into regions over which the system dynamics are piecewise linear. The behaviour of trajectories over the regions, and the inter-relation between regions is then characterised using graph theory, which is extremely useful for algorithm development on digital computers.

4.4 Contraction Mapping Based Stability Techniques

The difficulty with Lyapunov function based techniques is that they are inextricably bound to a specific state space realisation of the plant. Some relaxation of this condition can be made, and there are a variety of techniques to allow slow time variation, parameter drift, adaptive controls, etc., the limitation largely being the ingenuity of the analyst in bounding away uncertainties. However, it is nearly impossible to cope with uncertainties which change the dimension of the state space - such as the effects of unmodelled dynamics and time delays. Such model perturbations will destroy the topological structure of the system, the support upon which Lyapunov/LaSalle stability theory ultimately rests.

This defect in purely state space based analysis is a driving motivation to study systems from an input/output based perspective. By using a variant of the Contraction Mapping principle, one can guarantee the closed loop stability of a system which has a general perturbation, the condition being expressed in terms of operator norms. In the control community, the main result is usually referred to as the Small Gain Theorem, which is attributed independently to Zames [60, 61] and Sandberg [40]. The Contraction Mapping Theorem, which the Small Gain Theorem was inspired by, was due to Banach. It is quite often used in the theory of differential equations for proving existence/uniqueness of solutions. It will be stated next.

Theorem 4.4.1 (Banach): Let $f\colon M \mapsto M$ be a mapping from a Banach Space M with metric d to itself. Then, if there exists a $\kappa < 1$ such that

$$d(f(x), f(y)) \leq \kappa d(x, y),$$

for all $x, y \in M$, then there exists a unique *fixed point* p such that $f(p) = p$.
Proof: For a proof, see (amongst others) [49, Theorem 9.4.3]. □

4.4.1 An Invertibility Condition

Before presenting the usual version of the Small Gain Theorem, we shall state an invertibility condition on nonlinear operators, which is a prototypical "small gain" argument; other problem configurations are cast into this basic form, and the result is applied. The theorem itself comes in two variants - depending upon whether the incremental gain or the induced norm of the operator is used. The incremental gain version of the result is stronger, as it is then possible to guarantee that the inverse operator exists. The induced norm version of the theorem is weaker in that it cannot be guaranteed that the inverse exists or is unique, and nothing is said about continuity/causality properties. Linear systems are once again in a favourable situation as the induced norm and incremental gain are equal, and so the induced norm version of the theorem need not be used. The two versions will now be stated, along with their proof, due to their importance.

Theorem 4.4.2 (Incremental Gain Version): Fix a system operator $\boldsymbol{\Upsilon}: M \mapsto M$ over some complete normed space M, with $\boldsymbol{\Upsilon}0 = 0$. Then, if $\|\boldsymbol{\Upsilon}\|_\Delta < 1$, $(\mathbf{I} - \boldsymbol{\Upsilon})^{-1}$ is a well defined system operator, and in fact

$$\left\|(\mathbf{I} - \boldsymbol{\Upsilon})^{-1}\right\|_\Delta \leq \frac{1}{1 - \|\boldsymbol{\Upsilon}\|_\Delta}.$$

Additionally, if $\boldsymbol{\Upsilon}$ is causal, $(\mathbf{I} - \boldsymbol{\Upsilon})^{-1}$ is also causal.

Proof: Originally in [60] the full version of this statement was noted but the proof disregarded (presumably for presentation/space reasons), the following proof is given in the text by Willems [55] (which gives a fairly exhaustive treatment of the operator theoretic definition of stability). To prove the statement, we need to show that we can solve $x = \boldsymbol{\Upsilon}x + u$ for all $u \in M$. Take the sequence $x_{n+1} = \boldsymbol{\Upsilon}x_n + u$ for $n \geq 0$, $x_0 = 0$. As $\|x_{n+1} - x_n\| < \|x_n - x_{n-1}\|$, x_n is a Cauchy Sequence, and hence converges to a solution x in M. The solution is unique, as if there were two distinct solutions x_1, x_2, with:

$$\begin{aligned} x_1 &= \boldsymbol{\Upsilon}x_1 + u, \\ x_2 &= \boldsymbol{\Upsilon}x_2 + u, \end{aligned}$$

it then results that $\|x_1 - x_2\| = \|\boldsymbol{\Upsilon}x_1 - \boldsymbol{\Upsilon}x_2\| < \|x_1 - x_2\|$, a contradiction. The incremental gain property follows as if we take the distinct u_1, u_2, by examining the construction of the solution, we can inductively show that $\|x_1 - x_2\| \leq \|\boldsymbol{\Upsilon}\|_\Delta \|u_1 - u_2\| + \|\boldsymbol{\Upsilon}\|_\Delta^2 \|u_1 - u_2\| + \ldots = \frac{1}{1 - \|\boldsymbol{\Upsilon}\|_\Delta} \|u_1 - u_2\|$. The causality condition is a straightforward consequence of the inductive construction. □

Theorem 4.4.3 (Induced Norm Version): Fix a system operator $\boldsymbol{\Upsilon}: M \mapsto M$ over some complete normed space M, with $\boldsymbol{\Upsilon}0 = 0$. Then, if $\|\boldsymbol{\Upsilon}\|_i < 1$, and if a solution to the equation $x = \boldsymbol{\Upsilon}x + u$ exists, then,

$$\|x\| \leq \frac{1}{1 - \|\boldsymbol{\Upsilon}\|_i} \|u\|.$$

Hence, if we can independently ensure that a unique solution exists for all $u \in M$, we can assert that $(\mathbf{I} - \boldsymbol{\Upsilon})^{-1}$ exists (trivial by the previous assumption), and

$$\left\|(\mathbf{I} - \boldsymbol{\Upsilon})^{-1}\right\|_i \leq \frac{1}{1 - \|\boldsymbol{\Upsilon}\|_i}.$$

Proof: This follows in straightforward fashion from the Triangle Inequality; we have from the defining equation that

$$\|u\| = \|(\mathbf{I} - \boldsymbol{\Upsilon})x\|_i \geq (1 - \|\boldsymbol{\Upsilon}\|_i) \|x\|_i ,$$

verifying the claim. For more details, see [55]. $\square$

It will be seen in Chapter 7 that having a finite incremental gain is a strong property; in order for the analysis to be somewhat less conservative, the induced norm version of the Theorem needs to be used, and the existence/uniqueness/causality conditions established separately.

4.4.2 Closed Loop Stability and Small Gain

We will now apply the invertibility conditions in Section 4.4.1 to give closed loop stability results. We will then present an important specialisation of the result - the Circle Criterion, which has been a fairly standard analysis technique in the analysis of linear systems in a feedback configuration with a nonlinearity.

Theorem 4.4.4: Take the system given in Figure 4.1, with all signals residing on appropriately sized complete normed spaces. The operator

$$\begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{G} & \mathbf{I} \end{bmatrix}^{-1} : \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \mapsto \begin{bmatrix} u_1 \\ u_2 \end{bmatrix},$$

exists and has a finite incremental gain if $\|\mathbf{G}\|_\Delta \|\mathbf{K}\|_\Delta < 1$. Otherwise, if it can be guaranteed that $\begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{G} & \mathbf{I} \end{bmatrix}^{-1}$ exists as an operator, then it has a finite induced norm if $\|\mathbf{G}\|_i \|\mathbf{K}\|_i < 1$.

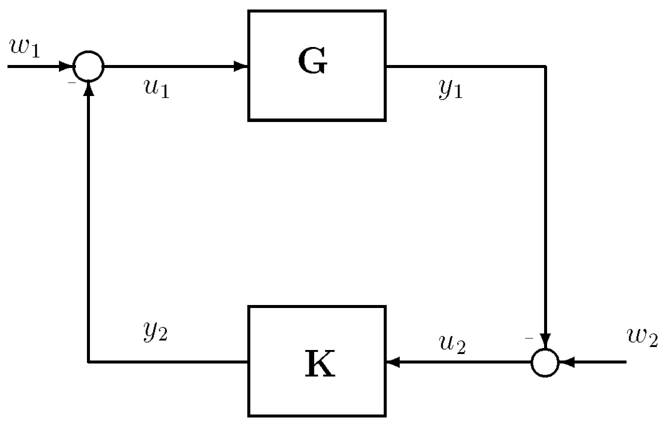

Figure 4.1: Closed Loop Configuration

Proof: This follows from Theorems 4.4.2 and 4.4.3, by noting that we can write

$$\begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{G} & \mathbf{I} \end{bmatrix} = \mathbf{I} + \begin{bmatrix} \mathbf{0} & \mathbf{K} \\ \mathbf{G} & \mathbf{0} \end{bmatrix}.$$

If either $\|\mathbf{G}\|_\Delta$ or $\|\mathbf{K}\|_\Delta$ (resp. $\|\mathbf{G}\|_i$ or $\|\mathbf{K}\|_i$) are greater than 1, we need to apply an appropriate scaling to the signals first. □

4.4.3 The Circle and Popov Criteria

It was noted in the paper by Zames [61] that the stability analysis techniques which can be gathered under the titles of the Circle and Popov Criteria can be derived from the Small Gain Theorem, and it is probably most profitable to view them as corollaries. As such, the results will be presented here in a summary fashion, which may be somewhat unfair as they have been amongst the most useful non-approximative results available for the analysis of linear systems with sector nonlinearities. The results will only be presented for continuous time in order to keep the presentation compact.

The usefulness of these results partially derives from the fact that they can be viewed as a generalisation of the Nyquist Criterion, a major result of "classical control", where design techniques were often derived from graphical stability criteria. The Nyquist technique examines the image of the imaginary axis under the mapping of the transfer function, and the stability condition is expressed in terms of avoidance/encirclements of the -1 point in the complex plane. When a sector bounded nonlinearity is introduced, the added uncertainty makes the condition an appropriate encirclement/avoidance criterion of a disk in the complex plane (with the disk possibly degenerate, becoming either a point or a half plane at the extremes of the possible sector size).

Definition 4.4.5: The continuous function $\phi: \mathbb{R}_+ \times \mathbb{R} \mapsto \mathbb{R}$ lies in the sector $[a, b]$ if:

$$ax^2 \leq x\phi(t, x) \leq bx^2,$$

for all $x \in \mathbb{R}, t \in \mathbb{R}_+$. The memoryless operator ϕ lies in the sector $[a, b]$ it is defined by $(\phi u)(t) = \phi(t, u(t))$, with $\phi(\cdot, \cdot)$ a continuous function within the sector $[a, b]$.

Definition 4.4.6: Fix a rational function $G: \mathbb{C} \mapsto \mathbb{C}$. Fix a positive constant R, and the corresponding *s-plane contour* is the (oriented) contour in the complex plane which runs along the semi-circle of radius R centred on the origin, along with the connecting segment $[-iR, iR]$ along the imaginary axis, with the orientation clockwise. Additionally, if $G(s)$ has any poles along the imaginary axis (all poles must be isolated, as $G(s)$ is rational), the curve makes a semi-circular detour of radius ϵ into the right half plane to avoid the pole. The *Nyquist Locus* is the limiting set of the image under G of the s-plane contour as $R \to \infty$, $\epsilon \to 0$. In cases where an encirclement criterion is to be examined, the number of encirclements of the locus is the limiting value as $\epsilon \to 0$.

Theorem 4.4.7 (Circle Criterion): Take the feedback configuration given in Figure 4.2, with ϕ a memoryless operator in the sector $[a, b]$, and $G(s)$ realised by a strictly proper finite dimensional shift invariant linear system (all signals are scalar). The system has a finite induced $\mathcal{L}_{2+}$ norm if the appropriate following condition holds.

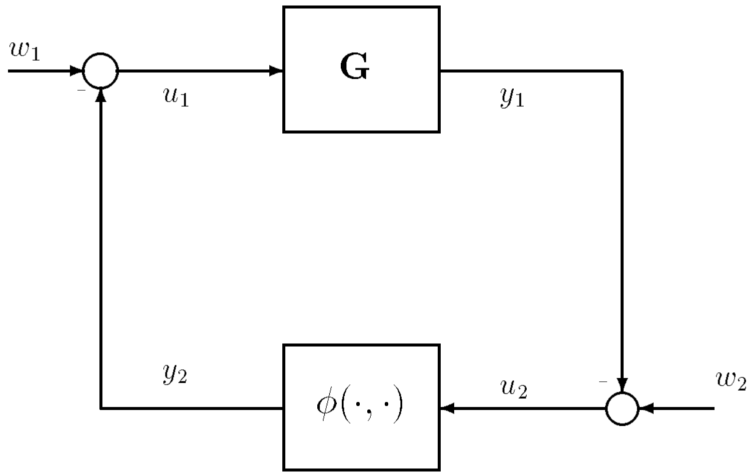

Figure 4.2: Circle/Popov Criterion System Configuration

1. If $0 = a < b$: $G(s)$ must have all poles lying in the open left half plane, as well the Nyquist Locus of $G(s)$ must lie in the half plane: $\{x: x \in \mathbb{C}, \mathrm{Re}\{x\} < \frac{1}{b}\}$.

2. If $0 < a < b$: The system has finite norm if the Nyquist Locus of $G(s)$ is bounded away from the disk D in the complex plane which is centred on the real axis, which passes through the points $-\frac{1}{a}$, $-\frac{1}{b}$. Additionally if ν is the number of poles of $G(s)$ in the closed right half plane, the locus encircles the disk D (counterclockwise sense) ν times.

3. If $a < 0 < b$: $G(s)$ must have all its poles lying in the open left half plane, as well as having the Nyquist Locus lying within the open disk in the complex plane which is centred on the real axis, and whose boundary passes through the points $-\frac{1}{a}, -\frac{1}{b}$.

4. If $a < b \leq 0$: Here, make the appropriate inversions of the signals and sector definitions to put the transformed system in the form of cases 1- 3.

Proof: This characterisation was largely taken from [52, Section 6.7]; the characterisation used here of the Nyquist Locus was taken from Frederick and Carlson [15, Section 11.2]. See also [55] and [61]. □

In Section 8.2.2, a bound for the incremental gain of a specific system is derived from the use of the Small Gain Theorem. In order to derive the bound, the technique of loop shifting is used. This is the same technique used to prove the Circle Criterion, and so the stability test is equivalent to that given in Theorem 4.4.7, although in Theorem 4.4.7, no mention is made of the predicted closed loop norm bound.

We will now state the Popov Criterion, in which the nonlinearity is assumed to be shift invariant. The test condition to determine stability is weaker, as the proof takes advantage of the added knowledge on the nonlinearity. The theorem is stated here for completeness, as it will not be otherwise referred to herein.

Theorem 4.4.8 (Popov): Take the system as in Theorem 4.4.7, with the added assumptions that $G(s)$, and $\frac{dG(s)}{ds}$ are in $\mathcal{RH}_\infty$, and ϕ is shift invariant and lies in the sector $[0, b]$. The system is then input/output stable (w.r.t. $\mathcal{L}_{2+}$) if there exist positive constants δ, κ, such that:

$$\mathrm{Re}\{(1 + i\omega\kappa G(i\omega)\} + \frac{1}{b} \geq \delta > 0,$$

for all $\omega \in \mathbb{R}$.

Proof: This characterisation was taken from [52]. □

4.5 Dissipative System Theory and Induced Norm Computation

4.5.1 Introduction

In order to advance beyond the estimate provided by the Small Gain Theorem, we need to use more direct analysis techniques. The method to be examined at length herein is the use of dissipation functions, which amounts to a generalisation of Lyapunov functions.

J.C. Willems is generally credited with the introduction of dissipation functions in the control literature in [56, 57]. Additionally, he addressed the related topic of the generation of Lyapunov functions for input/output stable systems in [58]. Additional background work in this area appeared in the papers [36], [23], [24], by D.J. Hill and P.J. Moylan.

There has been a good deal of recent work on the problem of the computation of dissipation functions for the class of nonlinear affine systems (systems of the form $\dot{x}(t) = f(x) + g(x)u$). The paper by A.J. van der Schaft [51] has provided the inspiration for the dissipation technique used in this work, although obviously it requires some adaptation as the main effect of the nonlinearity here is that the input is no longer affine. The computational problem of determining dissipation functions for these systems gives rise to a Hamilton-Jacobi inequality, which is solved via dynamic programming methods. It turns out that the previously used smoothness assumptions on the candidate functions were too strong for the numerical solutions which result. The lack of differentiability is acceptable as long as the function is differentiable almost everywhere. However, there is a lack of uniqueness to the solution to the Hamilton-Jacobi equations unless added conditions are added on how the function behaves where it is non-differentiable. This problem was dealt with by the introduction of the notion of *viscosity solutions* by M.G. Crandall and P.-J. Lions in [11]. In this work, and that which has flowed from it, non-smooth analysis is used, using a generalisation of the derivative. Such lack of differentiability is a feature of the solutions to the computational method given in Chapter 5; this is due to the non-differentiability of the standard saturation function. Other recent examples of work in this area are the papers by Ball, Helton and Walker [3],[2], M.R. James [27], [26], and the text by W.H. Fleming and H.M. Soner [14].

4.5.2 Results on Dissipation Theory

We will now give a fairly basic introduction to dissipative system theory. If we were interested in developing the ideas of viscosity solutions, more effort would be required, as we would need to introduce set-valued derivatives, etc. For the theory used in this dissertation, such detail is not necessary. We will now define the generalisation of Lyapunov functions, termed dissipation functions. If a function meets the condition in Theorem 4.5.3 for a given choice of γ, it is then termed a valid dissipation function, and implies that γ is an upper bound for the induced $\mathcal{L}_{2+}$ norm.

Definition 4.5.1: A function $V\colon \mathbb{R}^n \mapsto \mathbb{R}_+$ is a *candidate function* if V is uniformly Lipschitz over any compact set, and $V(0) = 0$.

The following result is an important characterisation of the differentiability of Lipschitz continuous functions. Although it is expected that candidate functions will not be differentiable over sets of measure zero, we are actually able to neglect any problems over such sets, and the generic lack of differentiability is not a problem.

Theorem 4.5.2: A Lipschitz continuous function is differentiable almost everywhere.
Proof: This theorem is due to Rademacher, see [63, Theorem 2.2.1] for details. □

Theorem 4.5.3: Fix the shift-invariant input/output system $\boldsymbol{\Upsilon}\colon w \mapsto z$ ($w(t) \in \mathbb{R}^m$, $z(t) \in \mathbb{R}^p$) which is generated by the differential equations:

$$\begin{aligned} \dot{x}(t) &= f(x(t), w(t)), \\ x(0) &= 0, \\ z(t) &= g(x(t), w(t)), \end{aligned}$$

with $f(\xi, \omega)$ uniformly Lipschitz with respect to ξ, ω, for ξ lying in a compact set. Fix the input set $\mathcal{W}_{2,q}(N)$, for which it is assumed that its reachable set Ω has compact closure. Additionally, we assume that this realisation of $\boldsymbol{\Upsilon}$ is $\mathcal{L}_2$-detectable. Fix a candidate function V. Let $\Omega_d \subseteq \Omega$ be the set where $\frac{dV}{dx}$ exists. Then if $\gamma > \|D\|$ is fixed, the Hamiltonian H^* is defined on Ω_d:

$$H^*(\xi) := \sup_{\omega \in \mathbb{R}} \left\{ H(\xi, \omega) := 2 \left.\frac{dV}{dx}\right|_{\xi} f(\xi, \omega) - \gamma^2 \|\omega\|^2 + \|g(\xi, \omega)\|^2 \right\}. \tag{4.2}$$

Then $\|\boldsymbol{\Upsilon}\|_{i,2} \leq \gamma$ (with respect to $\mathcal{W}_{2,q}(N)$) is implied by $H^*(\xi) \leq 0$ for all $\xi \in \Omega_d$.
Proof: See Appendix C. □

Remark 4.5.4: It is typical to indicate the dependence upon the derivative (or co-state) in the Hamiltonian (4.2), however that has been suppressed herein, as the candidate functions are normally treated as pre-specified data.

The basis for this theorem is originally due to Willems ([58]), however later work (as in Lions [32], or James [26]) has studied the consequences of relaxing the condition that the function be differentiable everywhere. Functions which are differentiable only almost everywhere are termed *generalised solutions*. Most recent work is concerned with viscosity solutions, which are specific generalised solutions which meet added conditions so that uniqueness is ensured.

The basic dissipation theorem itself is not immediately useful for our purposes, as the need to take the supremum pointwise over $\mathbb{R}$ is undesirable. By using the technique of *the completion of squares*, it is possible to convert the pointwise optimisation in Ω into a pointwise algebraic test. This is a long established technique in linear system theory, and gives rise to the algebraic Riccati equation. For nonlinear affine systems, a recent usage of the technique appears in [51]. However, the result will be slightly recast in terms of its notation to make the result more useful for our purposes.

Theorem 4.5.5: Fix a a single input system operator $\boldsymbol{\Upsilon}: w \mapsto z$ which is generated by the equations:

$$\begin{aligned} \dot{x} &= f(x) + B(w - K(x)), && (4.3)\\ x(0) &= 0, && (4.4)\\ z &= h(x) + D(w - K(x)), && (4.5) \end{aligned}$$

i.e., it is affine in $u := w - K(x)$. Given a candidate function V which is differentiable at ξ, and a fixed choice for the quantity γ. We can then define the real-valued functions θ,Φ,and H:

$$\theta(\xi) = -K(\xi) + \frac{1}{\gamma^2 - \|D\|^2}\left(B' \left.\frac{d^T V}{dx}\right|_{\xi} + D'h(\xi) - \|D\|^2 K(\xi)\right),$$

$$\begin{aligned} \Phi(\xi) &= 2 \left.\frac{dV}{dx}\right|_{\xi} (f(\xi) + B\theta(\xi)) - \gamma^2 \|K(\xi) + \theta(\xi)\|^2 + \|h(\xi) + D\theta(\xi)\|^2, \\ H(\xi, \omega) &= 2 \left.\frac{dV}{dx}\right|_{\xi} (f(\xi) + B(\omega - K(\xi))) - \gamma^2 \|\omega\|^2 + \|h(\xi) + D(\omega - K(\xi))\|^2 \end{aligned}$$

(recall (4.2)). Then,

$$H(\xi, \omega) = -(\gamma^2 - \|D\|^2) \|\omega - K(\xi) - \theta(\xi)\|^2 + \Phi(\xi),$$

and hence $H(\xi, \omega) \leq 0$ if and only if $\Phi(\xi) \leq 0$.

Proof: (Although done in [51], the algebra is reproduced here for convenience)

$$\begin{aligned} H(\xi, \omega) - \Phi(\xi) &= 2 \left.\frac{dV}{dx}\right|_{\xi} B(\omega - K(\xi) - \theta(\xi)) - \gamma^2 \|\omega\|^2 + \gamma^2 \|K(\xi) + \theta(\xi)\|^2 \\ &\quad + \|h(\xi) + D(\omega - K(\xi))\|^2 - \|h(\xi) + D\theta(\xi)\|^2 \\ &= -(\gamma^2 - \|D\|^2) \|\omega\|^2 + 2 \left(\left.\frac{dV}{dx}\right|_{\xi} B + D'h(\xi) - \|D\|^2 K(\xi)\right) \omega \\ &\quad - \frac{1}{\gamma^2 - \|D\|^2} \left\| B' \left.\frac{d^T V}{dx}\right|_{\xi} + D'h(\xi) - \|D\|^2 K(\xi)\right\|^2 \\ &= -(\gamma^2 - \|D\|^2) \|\omega - K(\xi) - \theta(\xi)\|^2 . \end{aligned}$$

The remainder of the theorem statement trivially follows. $\square$

Examination of the proof gives the motivation as to what the function $\theta(\xi)$ represents: it is the maximising input with respect to a given $V(\xi)$ for the Hamiltonian at ξ. Of course, one cannot say that it is an "optimal" choice, as such terminology would only seem to make sense with respect to the generation of maximising $\mathcal{L}_{2+}$ signals.

4.6 Linear Systems and the Computation of the $\mathcal{H}_\infty$ Norm

If the system studied is linear and shift invariant, it is possible to convert the test in Theorem 4.5.5 into a matrix algebraic condition. This makes the norm computation problem for linear systems straightforward. In fact, the induced $\mathcal{L}_{2+}$ norm of stable linear systems is the $\mathcal{H}_\infty$ norm of the transfer function matrix, and can be determined from frequency domain techniques, allowing its determination for infinite dimensional systems. We will examine the finite dimensional case only, and the associated algebraic state space technique.

Theorem 4.6.1: Take the (finite dimensional, shift invariant) linear system $\boldsymbol{\Upsilon}_L$ which has the realisation $\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$, with A Hurwitz. Then $\|\boldsymbol{\Upsilon}_L\|_{i,2} < \gamma$ if and only if $\bar{\sigma}(D) < \gamma$ and there exists a solution $P \geq 0$ to the algebraic Riccati equation:

$$(A + BR^{-1}D'C)'P + P(A + BR^{-1}D'C) + PBR^{-1}B'P + C'(I + DR^{-1}D')C = 0. \quad (4.6)$$

with $R = \gamma^2 I - D'D$, and such that $A + BR^{-1}(D'C + B'P)$ is Hurwitz. If $[C, A]$ is observable, we may strengthen the condition to $P > 0$.

Proof: This is a standard result, see [21, Theorem 3.7.1]. The strengthening of the condition on P due to observability is noted in [62, Chapter 11]. □

Note that as a solution is guaranteed for all suboptimal γ, it is possible to iteratively improve the norm estimate via a bisection algorithm. This makes it possible to determine the norm within an arbitrary tolerance, neglecting the usual computational error effects.

4.6.1 Lower Bounds for the Induced $\mathcal{L}_{2+}$ Norm

We will now state a result giving a lower bound for the induced $\mathcal{L}_{2+}$ norm, using the properties of the linearisation at the origin. The result is itself straightforward, and can be presumed to be essentially known. However, the method of proof given in this dissertation requires some care, and the author is unaware of any specific reference using this technique. As such, the proof is given in Appendix A for completeness.

Proposition 4.6.2: Fix a feedback loop in the standard feedback configuration (Section 3.3.1), with $\mathbf{G}$,$\mathbf{K}$ linear systems with the standard saturation (multiple inputs allowed), and let $\mathbf{G_L}, \mathbf{K_L}$ be their respective local linearisations. It is assumed that the linear operators have rational transfer functions. Then, the induced $\mathcal{L}_{2+}$ norm can be characterised as:

1. if $\begin{bmatrix} I & K_L(s) \\ G_L(s) & I \end{bmatrix}^{-1} \in \mathcal{RH}_\infty$ and $\left\| \begin{bmatrix} I & K_L(s) \\ G_L(s) & I \end{bmatrix}^{-1} \right\|_\infty = \gamma_l^*$, then

   $$\left\| \begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{G} & \mathbf{I} \end{bmatrix}^{-1} \right\|_{i,2} \geq \gamma_l^*;$$

2. if $\begin{bmatrix} I & K_L(s) \\ G_L(s) & I \end{bmatrix}^{-1} \notin \mathcal{RH}_\infty$, then $\begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{G} & \mathbf{I} \end{bmatrix}^{-1}$ does not have a finite induced $\mathcal{L}_{2+}$ norm.

Proof: See Appendix A. $\square$

4.7 Norm Computation via Dynamic Programming

There has been considerable interest in the norm computation problem using dynamic programming methods. It had been known that one could in principle find particular dissipation functions using a dynamic program, the functions being the storage functions of the system. Unfortunately, a lack of uniqueness is encountered, causing concern with numerical implementation of the technique. The use of set-valued generalisations of the derivative, leading to the notion of viscosity solutions, removes the problems associated with this lack of uniqueness. Additionally, increasing power of digital computers has allowed the computations to be attempted in what is viewed as reasonable time.

Fix a system operator $\mathbf{\Upsilon}: w \mapsto z$ generated by a finite dimensional shift invariant state space realisation with zero initial conditions and with an induced $\mathcal{L}_{2+}$ norm less than or equal γ. Then, we can define *available storage* and *required supply* functions V_a and V_r respectively:

$$V_a(\xi) = -\frac{1}{2} \inf_{w \in \mathcal{L}_{2+}, T \geq 0, x(0)=\xi} \int_0^T (\gamma^2 \|w(t)\|^2 - \|z(t)\|^2) dt,$$

$$V_r(\xi) = \frac{1}{2} \inf_{w \in \mathcal{L}_{2-}, T_- \leq 0, x(T_-)=\xi} \int_{T_-}^0 (-\gamma^2 \|w(t)\|^2 + \|z(t)\|^2) dt.$$

It is an important property of V_a and V_r that

$$0 \leq V_a(\xi) \leq V(\xi) \leq V_r(\xi),$$

for all valid dissipation functions for the same value of γ. These functions, and their theoretical importance was noted in [58], and [23]. The use of viscosity solutions has allowed numerical computation of these functions by the use of finite difference methods. Recent publications on this topic include the paper by M.R. James [27], and the text by W.H. Fleming and H.M. Soner [14, Chapter IX].

4.8 Linear Matrix Inequalities and Algebraic Norm Bound Computations

The controller saturation (only) configuration has received a great deal of attention from researchers over the years. Usually referred to as the *Lur'e problem*, much of the effort has gone into state space stability methods. This problem seems to have an intrinsic quadratic nature (see Section 5.8 for further comments on the induced $\mathcal{L}_{2+}$ norm problem for these systems, and the quadratic nature of the problem), and so appears amenable to algebraic solution via linear matrix inequalities. As our focus is on the operator gain computation problem, and we are more interested in the plant input saturation configuration, we will only cite an example of recent linear matrix inequality based methods. The following results are quoted from the text [4]; an interested reader can consult further references therein for more detail.

The operator $\boldsymbol{\Upsilon}: w \mapsto z$ is generated by:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B_1 w(t) + B_2 \phi(Kx(t)), \\ x(0) &= 0, \\ z(t) &= Cx(t), \end{aligned}$$

with $A \in \mathbb{R}^{n\times n}$, $B_1, B_2 \in \mathbb{R}^{n\times 1}$, $K \in \mathbb{R}^{1\times n}$, $C \in \mathbb{R}^{p\times n}$, and with $\phi(\cdot)$ each lying in the sector $[0,1]$, i.e.,for all $x \in \mathbb{R}$,

$$0 \le x\phi(x) \le x^2.$$

We use dissipation functions of the form:

$$V(\xi) = \xi' P\xi + 2\lambda \int_0^{K\xi} \phi(\alpha) d\alpha, \tag{4.7}$$

(a form of function proposed by Lur'e and Postnikov). By ensuring that $P > 0$, and $\lambda \ge 0$, $V(\xi) \ge \xi' P \xi > 0$ for $\xi \ne 0$.

Theorem 4.8.1: The function V defined in (4.7) is a valid Lyapunov function over R^n if the linear matrix inequality in the variables $P > 0$, $\lambda \ge 0$, $\tau \ge 0$,

$$\begin{bmatrix} A'P + PA & PB_2 + A'K'\lambda + K'\tau \\ B_2'P + \lambda KA + \tau K & \lambda K B_2 + B_2' K' \lambda - 2\tau \end{bmatrix} < 0$$

has a solution.
Proof: This is taken from [4, Section 8.1.1]. □

Remark 4.8.2: The variable τ is a scaling factor.

Theorem 4.8.3: The function V defined in (4.7) is a valid dissipation function determining an induced $\mathcal{L}_{2+}$ norm bound γ if the linear matrix inequality in the variables $\gamma > 0$, $P > 0$, $\lambda \ge 0$, $\tau \ge 0$

$$\begin{bmatrix} A'P + PA + C'C & PB_2 + A'K'\lambda + K\tau & PB_1 \\ B_2'P + \lambda KA + \tau K & \lambda K B_2 + B_2' K' \lambda - 2\tau & \lambda K B_1 \\ B_1' P & B_1' K' \lambda & \gamma^2 I \end{bmatrix} \le 0$$

has a solution.
Proof: This is taken from [4, Section 8.1.4]. □

Chapter 5

Induced Norm Computation for Continuous Time Systems

5.1 Introduction

In this chapter, we will examine the induced norm computation problem for the plant input saturation configuration in continuous time. The corresponding problem in discrete time will be considered in Chapter 6. The analysis for this class of systems contrasts with the more generic results in Chapter 4.

The main justification for the study of the induced norm computation problem is that it is hoped that it will be possible to use the induced norm version of the Small Gain Theorem to develop a robust control theory for this class of nonlinear systems. At present, the basic computational issues are a subject of research, and the issues raised by controller synthesis cannot yet be adequately addressed with a high level of confidence. There are many differences from the corresponding linear computation problem, especially when unstable plants are allowed.

In Section 5.2, results for neutrally stable plants will be reviewed, taken from the paper by Liu, Chitour and Sontag [33]. The results surveyed appear to be the first which directly addresses the norm computation problem for the plant input saturation configuration. The analysis takes in all $\mathcal{L}_{p+}$ norms, and with generic saturation nonlinearities, but at the cost of restriction of the plant and controller class.

The results in Sections 5.4 and 5.6 are new results on the induced $\mathcal{L}_{2+}$ norm computation problem for the plant input saturation configuration (the results were first announced in [38]). A test condition is developed to determine whether a given candidate function in fact is a dissipation function for a given norm bound. The condition appears quite complicated, and is the result of the completion of squares along the three piecewise linear segments of the standard saturation function. A heuristic method for generating a (suboptimal) valid candidate function is given. The result is presented as an existence result, which inter-relates state space and input/output stability. Simplified versions of the results in these sections are first given in Sections 5.3 and 5.5. Examples of its application are found in Chapter 8.

An important problem to be considered when addressing these computational problems is the determination of the reachable set. A preliminary means of using a dissipation function to overbound the reachable set is discussed in Section 5.7.

Finally, in Section 5.8, some brief comments are made regarding the norm computation problem for the controller only saturation configuration. This problem has been the subject of study of other researchers, and their (preliminary) results are noted.

5.2 Induced Norm Computation For Neutrally Stable Plants

The recent paper by Liu, Chitour and Sontag [33] is one of the first to treat the norm computation problem of linear systems with saturation without using a small gain argument or relying on sector estimates. It was the first set of results that allowed definite conclusions to be drawn with regards to the input/output stability of linear systems with poles on the imaginary axis which are stabilised by state feedback, the case where the Circle Criterion makes no prediction.

In [33], induced norm bounds are given for all $\mathcal{L}_{p+}$ norms (continuous time only is addressed). These bounds hold for all saturation functions. Also, the results hold for multiple input plants, a case which is not addressed in the new results detailed here. However, in order to allow the level of generality in [33], the analysis has been given in terms of a quite specific class of plants and controls, the plants being neutrally stable and a specific feedback controller fixed (a standard controller construction used in passivity arguments). Also, in terms of the actual bounds computed, the results are of most interest only for plants with skew symmetric dynamics matrices. Such a dynamics matrix has all its eigenvalues on the imaginary axis with simple Jordan blocks; if the plant has any open left half plane poles, those modes are not attempted to be controlled, and the induced norm of such a subsystem is the open loop norm.

The main result of [33] will now be stated, with some simplifications (such as we assume here that the realisation is controllable as the preference here is to treat systems as arising from transfer functions).

Theorem 5.2.1: Fix a controllable pair $[A, B]$, with A neutrally stable and lying in $\mathbb{R}^{n\times n}$, and $B \in \mathbb{R}^{n\times m}$. Denote as $\sigma\colon \mathbb{R}^m \mapsto \mathbb{R}^m$ a function which consists of m componentwise saturation functions, and $\phi\colon \mathbb{R}^k \mapsto \mathbb{R}^n$ a globally Lipschitz continuous function with $\phi(0) = 0$. There then exists an $F \in \mathbb{R}^{m\times n}$ and an $\epsilon > 0$ such that the operator mapping (u, v) to x generated by:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B\sigma(Fx + w) + \epsilon\phi(v(t)), \\ x(0) &= 0, \end{aligned}$$

is $\mathcal{L}_{p+}$ stable for all $p \in [1, \infty]$.

Proof: See [33]. □

Additionally, for $p = 1, 2$, [33] gives comparison results which bound the induced norm in terms of the closed loop linear system. These bounds are applicable directly only to skew symmetric plants.

Theorem 5.2.2: Let $\mathbf{\Upsilon}: w \mapsto x$ be generated by the equations:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B\sigma(-B'x(t) + w(t)), && (5.1) \\ x(0) &= 0, && (5.2) \end{aligned}$$

and $\mathbf{\Upsilon}_L$ the linear operator generated by replacing σ in equation (5.1) by the identity function. Additionally, we assume that $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, with $A = -A'$, and $A - BB'$ Hurwitz. Then, if κ is an S-Bound for σ, then we have that:

1. $\|\mathbf{\Upsilon}\|_{i,1} \leq \kappa(1 + \sqrt{m}\kappa)\,\|\mathbf{\Upsilon}_L\|_{i,1}$, and
2. $\|\mathbf{\Upsilon}\|_{i,2} \leq 2\sqrt{n}\kappa(1 + \sqrt{m}\kappa)\,\|\mathbf{\Upsilon}_L\|_{i,2}$.

Proof: See [33, Corollary 1]. □

5.3 Induced $\mathcal{L}_{2+}$ Norm Bounding: Simple Case

In order for the theory of linear systems with saturation to be fully developed, there will be a need to be able to analyse fairly generic problem configurations. In comparison to the results in Section 5.2, the new results presented in the next two sections will be given for general linear plants. In order to be able to do the analysis cleanly, however, the nonlinearity has been restricted to be the standard saturation, and only the (single input) induced $\mathcal{L}_{2+}$ norm is considered. The extension to multiple inputs appears relatively straightforward, however the notational aspect could get quite involved, although the computational aspects when implemented may be of a similar order of complexity as the single input case.

We will now state, without proof, the test condition on candidate functions for a simple problem configuration. Theorem 5.3.1 is in fact a consequence of the more general result in Theorem 5.4.1. This simpler configuration gives rise to a much simpler theorem statement, and a more obvious interpretation. The general case studied in Section 5.4 arose from the study of problem configurations of interest, and the more general proof is necessary.

The main results in these next two sections are only test condition for dissipation functions; the main problem of actually computing such function is examined later. The key point to be noted is that the general pointwise test on the Hamiltonian expression (4.2) is converted pointwise into algebraic tests, each corresponding to a piecewise linear segment of the nonlinearity. Hence, in the case of the standard saturation, three tests result. For the simpler class studied in this section, one of the tests is redundant, and there are two algebraic tests to be satisfied, which correspond to the algebraic Riccati and Lyapunov equations.

In Sections 5.5, 5.6, a heuristic technique is given to generate continuous candidate functions. It is hoped that it will be possible to compute candidate functions using algebraic techniques, avoiding the use of imposing a grid over the state space. The use of discretisations over the state space to compute functions is undesirable, as the number of grid points grows at the power of the state dimension.

The operator $\mathbf{\Upsilon}: w \mapsto z$ studied in this section is assumed to be generated by the state space realisation:

$$\dot{x}(t) = Ax(t) + B\sigma_0(w(t) - Kx(t)), \qquad (5.3)$$

$$
\begin{aligned}
x(0) &= 0, & (5.4)\\
z(t) &= Cx(t), & (5.5)
\end{aligned}
$$

with $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times 1}$, $C \in \mathbb{R}^{p\times n}$, $K \in \mathbb{R}^{1\times n}$, and $[A, B, C]$ minimal. Without loss of generality, we also assume that $A - BK$ is Hurwitz (otherwise, by Proposition 4.6.2, the induced $\mathcal{L}_{2+}$ norm is infinite).

Theorem 5.3.1: Fix an operator $\Upsilon: w \mapsto z$ generated by (5.3) - (5.5). We assume that the realisation is $\mathcal{L}_2$-detectable, and has a reachable set Ω for inputs w lying in $\mathcal{W}_{2,q}(N)$. Then $\Upsilon: w \mapsto z$ will have an upper bound γ to its induced $\mathcal{L}_{2+}$ norm over $\mathcal{W}_{2,q}(N)$, if there exists a candidate function V, whose derivative exists on the set Ω_d (with $\Omega \setminus \Omega_d$ having zero measure), and

$$\sup_{\xi\in\Omega_d} \left\{ H^*(\xi) = \max\{\Phi_A(\xi), \tilde{\Phi}_B(\xi)\} \right\} \leq 0$$

(recall (4.2) for the definition of H^*), with $\Phi_\bullet$ defined next. Define $\Phi_A(\xi)$ for $\xi \in \Omega_d$ as:

$$
\begin{aligned}
\Phi_A(\xi) &:= 2\left.\frac{dV}{dx}\right|_\xi (A\xi + B\sigma_0(\theta_a(\xi))) - \gamma^2(K\xi + \sigma_0(\theta_a(\xi)))^2 + \xi' C' C\xi \\
&= H(\xi, K\xi + \sigma_0(\theta_a(\xi))),
\end{aligned}
$$

with $\theta_a(\xi) := -K\xi + \frac{1}{\gamma^2} B^T \left.\frac{d^T V}{dx}\right|_\xi$. $\tilde{\Phi}_B(\xi)$ (with $\xi \in \Omega_d$) is defined as:

$$
\begin{aligned}
\tilde{\Phi}_B(\xi) &= 2\left.\frac{dV}{dx}\right|_\xi (A\xi + B\sigma_0(-K\xi)) + \xi' C' C\xi \\
&= H(\xi, 0).
\end{aligned}
$$

Proof: This is a consequence of Theorem 5.4.1, and Remark 5.4.2. □

An important point with regards to this theorem is that the determination whether the realisation is $\mathcal{L}_2$-detectable, and the exact determination of the reachable set, can be still considered open problems, although there are preliminary results on these issues herein.

Remark 5.3.2: The reader should note that the definition given for $\tilde{\Phi}_B$ does differ from Φ_B defined in Remark 5.4.2. It is a trivial consequence of the observations in Remark 5.4.2 that $\Phi_A(\xi), \Phi_B(\xi) \leq 0$ is equivalent to $\Phi_A(\xi), \tilde{\Phi}_B(\xi) \leq 0$. As a concrete example, note that the plot of $\Phi_B(\xi)$ in Figure 8.6 is negative definite near 0, not negative semi-definite.

There is a fairly simple interpretation of this result. The $\tilde{\Phi}_B$ expression corresponds to the Hamiltonian with a 0 external input, and gives rise to a Lyapunov style condition, reducing exactly to the algebraic Lyapunov equation in the linear case. When the $\mathcal{L}_2$-detectability assumption is also considered, it is possible to provide a straightforward proof that the existence of a valid candidate dissipation function ensures state space asymptotic stability. This in fact will be brought out in Section 5.5.

The Φ_A has an obvious affinity to the algebraic Riccati equation, and is the maximum value of $H(\xi, \omega)$, with $\omega \in [K\xi - 1, K\xi + 1]$. This is the range of external inputs which give rise to a non-saturating input to the saturation element. The "worst case" disturbance at the point ξ (with a loose interpretation of "worst case") is $\omega^* = \frac{1}{\gamma^2} B^T \left.\frac{dV}{dx}\right|_\xi$ for a non-constrained input. However, if the sum of the "worst case" disturbance and the control input lies outside the linear region, the best choice is to take the appropriate end point of the interval $[-1, 1]$. One can easily derive this from noting that the optimisation with respect to ω is a quadratic function. Additionally, one can then see that if $|K\xi| \leq 1$, $\Phi_A(\xi) \leq 0 \implies \tilde{\Phi}_B(\xi) \leq 0$. Hence, for small ξ, the quadratic determined by the algebraic Riccati equation will meet these conditions.

For this class of systems, one could derive these conditions from first principles, and not use the completion of squares used in the proof of Theorem 5.4.1. However, such a technique is not readily generalisable for non-zero D matrices, and gives rise to a faulty intuition as to how to interpret more general cases.

Examples of the use of this analysis is found in Chapter 8.

5.4 Induced $\mathcal{L}_{2+}$ Norm Bounding: General Case

We will now prove the test condition for a more general class of operators than those considered in Section 5.3. The operator $\mathbf{\Upsilon}: w \mapsto z$ is assumed to be generated by the state space realisation:

$$
\begin{aligned}
\dot{x}(t) &= Ax(t) + B_1 u(t) + B_2 \sigma_0(u(t)), && (5.6)\\
x(0) &= 0, && (5.7)\\
u(t) &\triangleq w(t) - K(x(t)), && (5.8)\\
z(t) &= h(x(t)) + D_1 u(t) + D_2 \sigma_0(u(t)) =: h(x(t)) + d(u(t)) =: g(x(t), u(t)), && (5.9)
\end{aligned}
$$

with $A \in \mathbb{R}^{n\times n}$, $B_i \in \mathbb{R}^{n\times 1}$, $D_i \in \mathbb{R}^{p\times 1}$, and $h: \mathbb{R}^n \mapsto \mathbb{R}^p$, $K: \mathbb{R}^n \mapsto \mathbb{R}$ continuous mappings with $h(0) = 0$, $K(0) = 0$. Additionally, we define $B \triangleq B_1 + B_2$, $D \triangleq D_1 + D_2$.

We will first note a condition on the minimum choice of γ needed for the algebraic manipulations made here to be valid. For linear systems, it is necessary to specify that γ is greater than the norm of D, and this needs to be extended for the nonlinear feedthrough term. It is assumed that $\gamma \geq \max\{\|D\|, \|D_1\|\}$. The assumption that $\gamma > \|D\|$ is not a restriction, as this is a consequence of Proposition 4.6.2. The assumption that $\gamma \geq \|D_1\|$ may be perhaps unnecessary for some systems with some choices of $\mathcal{W}_{2,q}(N)$, however it is apparent that it will be a lower bound in cases where a large enough disturbance signal is allowed to extract the saturated feedthrough behaviour of the system.

Theorem 5.4.1: Fix an operator $\mathbf{\Upsilon}: w \mapsto z$ generated by (5.6) - (5.9). We assume that the realisation is $\mathcal{L}_2$-detectable, and has a reachable set Ω for inputs w lying in $\mathcal{W}_{2,q}(N)$. Then $\mathbf{\Upsilon}: w \mapsto z$ will have an upper bound γ to its induced $\mathcal{L}_{2+}$ norm over $\mathcal{W}_{2,q}(N)$, with $\gamma > \max\{\|D\|, \|D_1\|\}$, if there exists a candidate function V, whose derivative exists on the set Ω_d (with $\Omega \setminus \Omega_d$ having zero measure), and

$$\sup_{\xi\in\Omega_d} \left\{ H^*(\xi) = \max\{\Phi_A(\xi), \Phi_B^+(\xi), \Phi_B^-(\xi)\} \right\} \leq 0 \qquad (5.10)$$

(recall the definition of $H^*(\xi)$ in (4.2)). The functions $\Phi_A, \Phi_B^+, \Phi_B^-$, (and functions upon which their definitions depend), will now be defined upon the set Ω_d. Let

$$\Phi_A(\xi) = \begin{cases} \Phi_{A.1}(\xi) & \text{if } |\theta_a(\xi)| \leq 1, \\ \Phi_{A.2}(\xi) & \text{if } |\theta_a(\xi)| > 1, \end{cases}$$

where $\theta_a, \Phi_{A.1}, \Phi_{A.2}$ are defined as:

$$\theta_a(\xi) = -K(\xi) + \frac{1}{\gamma^2 - \|D\|^2} \left(B^T \left. \frac{d^T V}{dx} \right|_\xi + D^T h(\xi) - \|D\|^2 K(\xi) \right),$$

$$\begin{aligned} \Phi_{A.1}(\xi) &= 2 \left. \frac{dV}{dx} \right|_\xi (A\xi + B\theta_a(\xi)) - \gamma^2 \|K(\xi) + \theta_a(\xi)\|^2 + \|g(\xi, \theta_a(\xi))\|^2, \\ \Phi_{A.2}(\xi) &= 2 \left. \frac{dV}{dx} \right|_\xi (A\xi + B\text{sign}(\theta_a(\xi))) - \gamma^2 \|K(\xi) + \text{sign}(\theta_a(\xi))\|^2 \\ &\quad + \|g(\xi, \text{sign}(\theta_a(\xi)))\|^2. \end{aligned}$$

The functions Φ_B^+, Φ_B^- are defined as:

$$\begin{aligned} \Phi_B^+(\xi) &= \begin{cases} \Phi_{B.1}^+(\xi) & \text{if } \theta_b^+(\xi) \geq 1, \\ \Phi_{B.2}^+(\xi) & \text{if } \theta_b^+(\xi) < 1, \end{cases} \\ \Phi_B^-(\xi) &= \begin{cases} \Phi_{B.1}^-(\xi) & \text{if } \theta_b^-(\xi) \leq -1, \\ \Phi_{B.2}^-(\xi) & \text{if } \theta_b^-(\xi) > -1. \end{cases} \end{aligned}$$

The functions $\theta_b^+, \theta_b^-, \Phi_{B.1}^+, \Phi_{B.2}^-, \Phi_{B.2}^+, \Phi_{B.1}^-$ are defined:

$$\theta_b^\pm(\xi) = -K(\xi) + \frac{1}{\gamma^2 - \|D_1\|^2} \left(B_1' \left. \frac{d^T V}{dx} \right|_\xi + D_1'(h(\xi) \pm D_2) - \|D_1\|^2 K(\xi) \right),$$

$$\begin{aligned} \Phi_{B.1}^\pm(\xi) &= 2 \left. \frac{dV}{dx} \right|_\xi \left(A\xi + B_1\theta_b^\pm(\xi) \pm B_2 \right) - \gamma^2 \left\| K(\xi) + \theta_b^\pm(\xi) \right\|^2 \\ &\quad + \left\| h(\xi) + D_1\theta_b^\pm(\xi) \pm D_2 \right\|^2, \\ \Phi_{B.2}^\pm(\xi) &= 2 \left. \frac{dV}{dx} \right|_\xi (A\xi \pm B) - \gamma^2 \|K(\xi) \pm 1\|^2 + \|h(\xi) \pm D\|^2. \end{aligned}$$

Proof:

It will be shown that $H(\xi, \omega) \leq 0$ is implied by (5.10) holding. Then by Theorem 4.5.3, the norm condition follows.

Fix any point $\xi \in \Omega_d$, and a choice of $\omega \in \mathbb{R}$. One of three cases is true:

(1) $\omega - K(\xi) > 1$,

(2) $-1 \leq \omega - K(\xi) \leq 1$,

(3) $\omega - K(\xi) < -1$.

Examine case (1), and note that case (3) is symmetric. Then, there are two possibilities; either $\theta_b^+(\xi) \geq 1$, or $\theta_b^+(\xi) < 1$.

In either possible case, we have

$$\begin{aligned} H(\xi,\omega) &= 2\left.\frac{dV}{dx}\right|_{\xi}(A\xi + B_1(\omega - K(\xi)) + B_2) - \gamma^2\|\omega\|^2 \\ &\quad + \|h(\xi) + D_1(\omega - K(\xi)) + D_2\|^2. \end{aligned}$$

If $\theta_b^+(\xi) \geq 1$, by the condition that $\Phi_B^+(\xi) \leq 0$, we have that $\Phi_{B.1}^+(\xi) \leq 0$. It then is a straightforward application of Theorem 4.5.5 to see that this implies $H(\xi,\omega) \leq 0$. If $\theta_b^+(\xi) < 1$, we have by assumption that $\Phi_{B.2}^+(\xi) \leq 0$. Then,

$$\begin{aligned} H(\xi,\omega) &\leq H(\xi,\omega) - \Phi_{B.2}^+(\xi) \\ &= (H(\xi,\omega) - \Phi_{B.1}^+(\xi)) - (\Phi_{B.2}^+(\xi) - \Phi_{B.1}^+(\xi)). \end{aligned}$$

By applying Theorem 4.5.5, we arrive at

$$\begin{aligned} H(\xi,\omega) &\leq -(\gamma^2 - \|D_1\|^2)\left(\left\|\omega - K(\xi) - \theta_b^+(\xi)\right\|^2 - \left\|1 - \theta_b^+(\xi)\right\|^2\right) \\ &= -\Big\{(\gamma^2 - \|D_1\|^2)\left(\left\|\omega - K(\xi) - \theta_b^+(\xi)\right\| - \left\|1 - \theta_b^+(\xi)\right\|\right) \\ &\quad \left(\left\|\omega - K(\xi) - \theta_b^+(\xi)\right\| + \left\|1 - \theta_b^+(\xi)\right\|\right)\Big\} \\ &\leq 0, \end{aligned}$$

as we have that $\omega - K(\xi) > 1$, and $\theta_b^+(\xi) < 1$.

For case (2) (i.e. $|\omega - K(\xi)| \leq 1$), there are (once again) two possibilities: either $|\theta_a(\xi)| \leq 1$ or $|\theta_a(\xi)| > 1$. If $|\theta_a(\xi)| \leq 1$, it is a trivial application of Theorem 4.5.5 to see that $H(\xi,\omega) \leq 0$ follows from the assumption that $\Phi_A(\xi) = \Phi_{A.1}(\xi) \leq 0$.

If $|\theta_a(\xi)| > 1$, by the stated condition, $\Phi_{A.2}(\xi)$ is less than or equal 0. Then,

$$\begin{aligned} H(\xi,\omega) &\leq H(\xi,\omega) - \Phi_{A.2}(\xi) \\ &= (H(\xi,\omega) - \Phi_{A.1}(\xi)) - (\Phi_{A.2}(\xi) - \Phi_{A.1}(\xi)). \end{aligned}$$

By the use of Theorem 4.5.5,

$$\begin{aligned} H(\xi,\omega) &\leq -(\gamma^2 - \|D\|^2)\left(\|\omega - K(\xi) - \theta_a(\xi)\|^2 - \|\mathrm{sign}(\theta_a(\xi)) - \theta_a(\xi)\|^2\right) \\ &= -\Big\{(\gamma^2 - \|D\|^2)(\|\omega - K(\xi) - \theta_a(\xi)\| - \|\theta_a(\xi)\| + 1) \\ &\quad (\|\omega - K(\xi) - \theta_a(\xi)\| + \|\theta_a(\xi)\| - 1)\} \\ &\leq 0, \end{aligned}$$

by the Triangle Inequality. We have thus shown that $H(\xi,\omega) \leq 0$, and hence the norm condition follows. □

Remark 5.4.2: Some of the test conditions given in Theorem 5.4.1 are under various conditions redundant, and the tests collapse down to a lesser number, in some cases one. The following observations describe this.

1. We first note that $\Phi_{A.1}(\xi) \leq 0 \implies \Phi_A(\xi), \Phi_{A.2}(\xi) \leq 0$. This follows as $\Phi_{A.2}(\xi) - \Phi_{A.1}(\xi) = -(\gamma^2 - \|D\|^2) \|\text{sign}(\theta_a(\xi)) - \theta_a(\xi)\|^2$.

2. $\Phi^{\pm}_{B.2} \leq \Phi_{A.2}$, hence $\Phi_A(\xi) \leq 0 \implies \Phi^{\pm}_{B.2}(\xi) \leq 0$.

3. If $|\frac{D_1' D_2}{\gamma^2 - \|D_1\|^2}| \leq 2$, it is not possible for $\theta_b^+(\xi) > 1$ and $\theta_b^-(\xi) < -1$ simultaneously, and one (or both) of the conditions that $\Phi^{\pm}_B(\xi) \leq 0$ will be implied by $\Phi_A(\xi) \leq 0$.

4. If $D_1' D_2 = 0$, $\theta_b^+ = \theta_b^-$. It is then possible to define $\Phi_B(\xi), \theta_b(\xi)$ as follows:

$$\theta_b(\xi) = -K(\xi) + \frac{1}{\gamma^2 - \|D_1\|^2} \left(B_1' \left. \frac{d^T V}{dx} \right|_\xi + D_1' h(\xi) - \|D_1\|^2 K(\xi) \right),$$

$$\Phi_B(\xi) = \begin{cases} \Phi_{B.1}(\xi) & \text{if } |\theta_b(\xi)| \geq 1, \\ \Phi_{B.2}(\xi) & \text{otherwise,} \end{cases}$$

   with

$$\begin{aligned} \Phi_{B.1}(\xi) &= 2 \left. \frac{dV}{dx} \right|_\xi (A\xi + B_1 \theta_b(\xi) + B_2 \text{sign}(\theta_b(\xi))) - \gamma^2 \|K(\xi) + \theta_b(\xi)\|^2 \\ &\quad + \|g(\xi, \theta_b(\xi))\|^2, \\ \Phi_{B.2}(\xi) &= 2 \left. \frac{dV}{dx} \right|_\xi (A\xi + B \text{sign}(\theta_b(\xi))) - \gamma^2 \|K(\xi) + \text{sign}(\theta_b(\xi))\|^2 \\ &\quad + \|g(\xi, \text{sign}(\theta_b(\xi)))\|^2. \end{aligned}$$

   The norm condition then amounts to requiring that $\Phi_A(\xi), \Phi_B(\xi) \leq 0$ holds for all ξ in Ω_d.

The difference in the interpretation between Φ_B (Remark 5.4.2) and $\tilde{\Phi}_B$ (Section 5.3) is that $\tilde{\Phi}_B$ corresponds to a null external input, whereas Φ_B corresponds to the worst case external input which gives rise to u lying on the two flat segements of the input nonlinearity. Hence, Φ_B is negative definite near the origin in Figure 8.6.a The interpretation of the functions $\Phi_A, \Phi_B^{\pm}$ is that $\Phi_A(\xi)$ is the maximum value of $H(\xi, \omega)$ for $|\omega - K(\xi)| \leq 1$, $\Phi_B^+(\xi)$ is the maximum value of $H(\xi, \omega)$ for $\omega - K(\xi) \geq 1$, and $\Phi_B^-(\xi)$ is the maximum value of $H(\xi, \omega)$ for $\omega - K(\xi) \leq 1$. Hence the maximum of these functions equals H^*.

5.5 A Sub-Optimal Solution: Simple Case

In this section, a heuristic approach to generating candidate functions explicitly is given. The computation is reminiscent of the inverse Lyapunov theorems developed in the papers by Willems [57], and by Hill and Moylan [24]. In fact, the present treatment captures the Lyapunov results by letting γ tend to infinity (or setting $\psi(\xi) \equiv 0$), just as the Riccati equation tends to the Lyapunov equation in linear system theory. The result is presented as an existence result, which in fact shows the relationship between state space (null input) and stability in the input/output sense.

It is not claimed that this technique will compute the optimal bound for the norm, as it is not presented as an iterative technique. It may be possible to iteratively improve the results, as there are free parameters which may be tuned; there are no guarantees on such a procedure, however. The contribution of the present work is to set out the framework of the problem at hand, and to show that the computation appears to be tractable.

It is believed that the heuristics behind the present technique are reasonable, and in fact the technique is able to determine the optimal value in basic examples in Chapter 8.

The results stated in this section are a simplified version of those in Section 5.6, and so will be stated without proof. In fact, the results in Section 5.6 are for a simplified class of systems from the general case, but with the addition of D matrices, which cause difficulty in the algebra.

The motivation for the algorithm is found in Section 5.6.1, as the means of proof offer clues as to why the functions of this form are valid dissipation functions. Additionally, the comments in Remark 5.6.10 should be noted: the result has been presented as an existence result, which is stylistically pleasing, but it is not believed to be the sole merit of this work. There are a good number of corollaries and properties which follow from it, and the manipulations used in its proof are felt to be useful as general tools.

We are once again examining the class of operators $\mathbf{\Upsilon}\colon w \mapsto z$ with realisations determined by (5.3) - (5.5).

In order to define the candidate function, we first need to define an auxiliary function ψ. Fix $\epsilon > 0$. Then for some $\gamma_l > \left\| \left[\begin{array}{c|c} A - BK & B \\ \hline C & 0 \end{array} \right] \right\|_\infty$ (remember that we necessarily assume that $A - BK$ is Hurwitz), let $P > 0$ be the solution to the algebraic Riccati equation:

$$(A - BK)'P + P(A - BK) + \frac{1}{\gamma_l^2} PBB'P + C'C + \epsilon I_n = 0.$$

Then $\psi(\xi) := \min\{0, 2\xi' P(A\xi - B\sigma_0(K\xi)) + \xi' C' C\xi\}$.

In order to state the theorem, we will need the following two definitions.

Definition 5.5.1: Let the *induced trajectory* from the point $\xi \in \bar{\Omega}$ (written as $x(t;\xi)$), be the trajectory defined by:

$$\begin{aligned} \dot{x}(t) &= Ax(t) - B\sigma_0(Kx(t)), \\ x(0) &= \xi. \end{aligned}$$

Definition 5.5.2: A $\Omega \subset \mathbb{R}^n$ is said to have a *stable boundary* if for all points $\xi \in \partial\Omega$, there exists a finite $t(\xi) > 0$ such that $x(t;\xi) \in \Omega$.

Theorem 5.5.3: Fix an operator $\mathbf{\Upsilon}\colon w \mapsto z$ which is of the form given by (5.3) - (5.5). We also assume the realisation of $\mathbf{\Upsilon}$ is $\mathcal{L}_2$-detectable, has a reachable set Ω with compact closure for $w \in \mathcal{W}_{2,q}(N)$, and that the set Ω has a stable boundary. The following statements are then equivalent:

1. for all $w \in \mathcal{W}_{2,q}(N)$, $\mathbf{\Upsilon} w$ is an element of $\mathcal{L}_{2+}$;

2. for all $\xi \in \Omega$, $x(t;\xi)$ is an element of $\mathcal{L}_{2+}$;

3. the function $V\colon \bar{\Omega} \mapsto \mathbb{R}_+$ defined:

$$V(\xi) = \frac{1}{2}\int_0^\infty \left(\|Cx(t;\xi)\|^2 - \psi(x(t;\xi))\right) dt,$$

is uniformly Lipschitz on $\bar{\Omega}$, and satisfies the conditions of Theorem 5.3.1 over the set Ω for some finite γ. Hence, γ is an upper bound for the induced $\mathcal{L}_{2+}$ norm over $\mathcal{W}_{2,g}(N)$.

Remark 5.5.4: The assumption of the stable boundary in Theorem 5.6.8 is somewhat unsatisfactory, but essentially amounts to requiring that there be no equilibria nor limit cycles along the boundary of the reachable set. It is most likely that if such was the case, the induced norm would be infinite, as one could steer the state arbitrarily close to the equilibrium/limit cycle, resulting in a state trajectory with unbounded $\mathcal{L}_{2+}$ norm. If this can be shown (as expected), the assumption could then be dropped, as it is incompatible with statements 1. - 3. in the theorem.

5.6 A Sub-Optimal Solution: General Case

In order to cleanly present the results, the general construction has been made on a subclass of the systems in Section 5.4. It is believed that these results could be extended for the entire class of systems studied in Section 5.4, but at the cost of making the presentation quite difficult. In addition, some auxiliary results are presented, some of which deal with the linear system at the origin, amounting to local results.

The class of operators $\boldsymbol{\Upsilon}\colon w \mapsto z$ from a subset of the operators studied in Section 5.4, and the state space realisation has been restricted to be of the form:

$$\dot{x}(t) = Ax(t) + B\sigma_0(u(t)), \tag{5.11}$$

$$x(0) = 0, \tag{5.12}$$

$$u(t) \triangleq w(t) - Kx(t), \tag{5.13}$$

$$z(t) = Cx(t) + D_1u(t) + D_2\sigma_0(u(t)) =: h(x(t)) + d(u(t)) =: g(x(t), u(t)), \tag{5.14}$$

with $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times 1}$, $D_i \in \mathbb{R}^{p\times 1}$, and $C \in \mathbb{R}^{p\times n}$, $K \in \mathbb{R}^{1\times n}$. We (necessarily) assume that the local linearisation is closed loop stable (i.e., $A - BK$ is Hurwitz), and that $C'D = 0$, $D_1'D_2 = 0$.

There will be a variety of consequences of relaxing these assumptions. The assumption that $A - BK$ is Hurwitz is not a restriction, as this is necessary if $\boldsymbol{\Upsilon}$ is to have finite induced $\mathcal{L}_{2+}$ norm, by the use of Proposition 4.6.2. Relaxing the conditions that the readout function and controller are linear would be fairly basic; if they are linear over a neighbourhood around the origin, the proofs would go through with no real technical changes. If they are not, there would have to be some verification of the local properties of the origin in a manner reminiscent of the work on nonlinear affine systems. The assumption that $B_1 = 0$, and the orthogonality on the output matrices are solely introduced to reduce the effort involved in the algebra.

In order to define the candidate function, use is made of the solution for the local linearisation. Let $\tilde{\gamma}_l > \max\{\|D\|, \|D_1\|\}$ be fixed such that the solution $P > 0$ to the algebraic Riccati equation exists:

$$(A+B\Theta_L)'P+P(A+B\Theta_L)-\tilde{\gamma}_l^2(K+\Theta_L)'(K+\Theta_L)+(C+D\Theta_L)'(C+D\Theta_L)+Q = 0, \quad (5.15)$$

with $\Theta_L := -K + \frac{1}{\tilde{\gamma}^2 - \|D\|^2}(B'P - \|D\|^2 K)$, Q an arbitrary positive definite matrix. Choose $\gamma_l \geq \tilde{\gamma}_l$ such that

$$\frac{1}{\gamma_l^2 - \|D\|^2}(B'P - \|D\|^2 K)'(B'P - \|D\|^2 K) + Q > \frac{\|D_1\|^4}{\gamma_l^2 - \|D_1\|^2} K'K \quad (5.16)$$

(the requirement in (5.16) is made so that some of the algebraic manipulations used here are in a more convenient form, and could be relaxed). Note that equation (5.15) is the same as (4.6); we have just rewritten it so that it is closer to the form of the $\Phi_{A.1}$ expression. We can then define the corresponding linear solution $V_l: \mathbb{R}^n \mapsto \mathbb{R}_+$ as $V_l(\xi) := \frac{1}{2}\xi' P \xi$. Finally, let $\psi: \mathbb{R}^n \mapsto \mathbb{R}_-$ be defined:

$$\begin{aligned}\psi(\xi) \quad := \min \quad & \left\{ \frac{-1}{\gamma_l^2 - \|D_1\|^2} \left\| \|D_1\|^2 K\xi \right\|^2, \right. \\ & \left. 2\xi' P(A\xi - B\sigma_0(K\xi)) + \|C\xi - d(K\xi)\|^2 \right\}. \end{aligned} \quad (5.17)$$

Remark 5.6.1: The Q matrix is an arbitrary matrix needed for ensuring boundedness of expressions, it may be thought of as ϵI ($\epsilon > 0$). The arbitrary form is useful, as there may exist useful heuristics to its choice - such as aligning it mainly in the direction of a matrix perpendicular to C.

Lemma 5.6.2: For all $Q > 0$, there exists a finite $\gamma_l \geq \tilde{\gamma}_l$ such that there is a solution P to the equation (5.15) and with the additional inequality (5.16) also holding.
Proof: Deferred to Appendix D. □

Remark 5.6.3: The function ψ has some useful properties which should be noted. Firstly, $\psi(\xi)$ is uniformly Lipschitz over a compact set, as it is formed by the sums and products of functions which are individually uniformly Lipschitz over compact sets. Secondly, by applying the completion of squares to the linear system, and using the assumed relation in equation (5.16), we have that $\psi(\xi) = -\frac{1}{\gamma_l^2 - \|D\|^2} \left\| B'P\xi - \|D\|^2 K\xi \right\|^2 - \xi' Q \xi$ if $|K\xi| \leq 1$.

Let $P_\infty = P_\infty' > 0$ be the solution to the Lyapunov equation of the linear dynamic system, i. e. :

$$(A - BK)'P_\infty + P_\infty(A - BK) = -I_n.$$

Then, let $\Omega_l \subset \mathbb{R}^n$ be the set

$$\Omega_l = \{\xi : \xi' P_\infty \xi \leq b\},$$

with $b = \min_{\xi : |K\xi| = 1} \xi' P_\infty \xi$.

At this point, it may be useful to refer back to the definition of the induced trajectory denoted $x(t; \xi)$ (Definition 5.5.1).

Lemma 5.6.4: For all $\xi \in \Omega_l$, $x(t;\xi) \in \Omega_l$, and $|Kx(t;\xi)| \leq 1$ for all $t \in \mathbb{R}_+$.
Proof: Deferred to Appendix D. □

Lemma 5.6.5: Fix an autonomous state space realisation $\dot{x}(t) = f(x(t))$, with f Lipschitz continuous with constant κ. Then if x, y are trajectories with initial conditions $x(0) = x_0$, $y(0) = y_0$, then

$$\|x(t) - y(t)\| \leq e^{\kappa t} \|x_0 - y_0\|.$$

Proof: This estimation result follows from the Bellman-Gronwall Inequality; see [45, Lemma C.3.1, Theorem 37]. □

Lemma 5.6.6: If $x(t;\xi) \in \mathcal{L}_{2+}$ for all $x \in \bar{\Omega}$, there then exists a finite constant T_Ω such that $T_\Omega = \max_{\xi \in \bar{\Omega}} T(\xi)$, with $T: \bar{\Omega} \mapsto \mathbb{R}_+$ defined as

$$T(\xi) = \min_{t \in R_+} \{t : x(t;\xi) \in \Omega_l\}.$$

Proof: Deferred to Appendix D. □

Lemma 5.6.7: Let $U: \bar{\Omega} \mapsto \mathbb{R}$ be defined:

$$U(\xi) = \int_0^\infty \alpha(x(t;\xi))dt,$$

with $\alpha: \bar{\Omega} \mapsto \mathbb{R}$ having a (uniform) Lipshitz constant k, and with $\alpha(0) = 0$. Then if $x(t;\xi) \in \mathcal{L}_{2+}$ for all ξ in $\bar{\Omega}$, U is uniformly Lipschitz.
Proof: Deferred to Appendix D. □

Theorem 5.6.8: Fix an operator $\mathbf{\Upsilon}: w \mapsto z$ which is of the form given by (5.11) - (5.14), with the noted conditions on the matrices. We also assume the realisation of $\mathbf{\Upsilon}$ is $\mathcal{L}_2$-detectable, has a reachable set Ω with compact closure for $w \in \mathcal{W}_{2,q}(N)$, and that the set Ω has a stable boundary (Definition 5.5.2). The following statements are then equivalent:

1. for all $w \in \mathcal{W}_{2,q}(N)$, $\mathbf{\Upsilon} w$ is an element of $\mathcal{L}_{2+}$;
2. for all $\xi \in \Omega$, $x(t;\xi)$ is an element of $\mathcal{L}_{2+}$;
3. the function $V: \bar{\Omega} \mapsto \mathbb{R}_+$ defined (note the definition of ψ in (5.17)):

$$V(\xi) = \frac{1}{2} \int_0^\infty \left(\|g(x(t;\xi), -Kx(t;\xi))\|^2 - \psi(x(t;\xi)) \right) dt, \tag{5.18}$$

is uniformly Lipschitz on $\bar{\Omega}$, and satisfies the conditions of Theorem 5.4.1 over the set Ω for some finite γ. Hence, γ is an upper bound for the induced $\mathcal{L}_{2+}$ norm over $\mathcal{W}_{2,q}(N)$.

Proof: (1) $\Longrightarrow$ (2): Pick any $\xi \in \Omega$. By the assumption of reachability, there exists some $w \in \mathcal{W}_{2,q}(N)$, such that for some $t \in \mathbb{R}_+$ we have that $(\mathbf{\Upsilon}_x w)(t) = \xi$. Since the system is causal, we also have that $(\mathbf{\Upsilon}_x \mathbf{P}_t w)(t) = \xi$. Since $\|\mathbf{P}_t w\|_\infty \leq \|w\|_\infty$, $\mathbf{P}_t w$ is also an element of $\mathcal{W}_{2,q}(N)$. As (1) implies that $\mathbf{\Upsilon}\mathbf{P}_t w$ lies in $\mathcal{L}_{2+}$, by the observability assumption, $\mathbf{\Upsilon}_x \mathbf{P}_t w \in \mathcal{L}_{2+}$. By then noting that $\mathbf{\Upsilon}_x \mathbf{P}_t w = \mathbf{P}_t \mathbf{\Upsilon}_x \mathbf{P}_t w + (\mathbf{I} - \mathbf{P}_t)\mathbf{\Upsilon}_x \mathbf{P}_t w$, and that $x(\tau;\xi) = (\mathbf{\Upsilon}_x \mathbf{P}_t w)(t+\tau)$, we have that

$$\|x(t;\xi)\|_2 = \|(\mathbf{I} - \mathbf{P}_t)\mathbf{\Upsilon}_x \mathbf{P}_t w\|_2 \leq \|\mathbf{\Upsilon}_x \mathbf{P}_t w\|_2 . \tag{5.19}$$

Since $\mathbf{\Upsilon}_x \mathbf{P}_t w$ is in $\mathcal{L}_{2+}$, equation (5.19) implies that $x(t;\xi) \in \mathcal{L}_{2+}$.

(2) $\Longrightarrow$ (3): By the use of (2), and the assumption that Ω has a stable boundary, we have that $x(t;\xi) \in \mathcal{L}_{2+}$ for all $\xi \in \bar{\Omega}$. By applying Lemma 5.6.7, we see that V defined in (5.18) is uniformly Lipschitz. It only remains to be shown that there exists a finite choice of $\gamma \geq \gamma_l$ such that the conditions of Theorem 5.4.1 hold.

The claim will be shown by verifying that for any fixed point in Ω_d there exists a choice of γ which forces Φ_A, Φ_B (see Remark 5.4.2) to be less than or equal to zero, expressed in terms of functions which remain bounded over $\bar{\Omega}$.

To prove the above claim, fix any point $\xi \in \Omega_d$. It first should be noted that by its construction, V satisfies the partial differential equation:

$$2\left.\frac{dV}{dx}\right|_{\xi}(A\xi - B\sigma_0(K\xi)) + \|g(\xi, -K\xi)\|^2 = \psi(\xi). \tag{5.20}$$

There are then three possible cases to examine, which are considered in turn.

1. The first possibility is that ξ lies in the set $\Omega_l \bigcap \{\xi : (1 + \frac{\|D\|^2}{\gamma_l^2 - \|D\|^2})|K\xi| < 1\}$. Over the set Ω_l, $V(\xi)$ is forced to equal $V_l(\xi)$ by the construction of $\psi(\xi)$. It is easy to verify that if γ is chosen to be γ_l,$\Phi_A(\xi) = -\xi^T Q \xi$. And over this subset, we have that $|\theta_b(\xi)| < 1$ for all choices of $\gamma \geq \gamma_l$, and so $\Phi_B(\xi) = \Phi_{B.2}(\xi)$. By noting Remark 5.4.2, this implies that $\Phi_B(\xi) \leq 0$ for all γ greater than or equal γ_l.

2. The second case is where $|K\xi| \leq 1$, but $\xi \notin \Omega_l \bigcap \{\xi : (1 + \frac{\|D\|^2}{\gamma_l^2 - \|D\|^2})|K\xi| < 1\}$. Examine $\Phi_{A.1}$ (if $|\theta_a(\xi)| > 1$, the condition will be weaker than that indicated here): by noting the constitutive equation for V (equation (5.20)), the relation on ψ given in Remark 5.6.3, and performing the manipulations analogous to the completion of squares, it results that

$$\begin{aligned}\Phi_{A.1}(\xi) \quad &= \quad \frac{1}{\gamma^2 - \|D\|^2}\left\|\left.\frac{dV}{dx}\right|_{\xi} B - \|D\|^2 K\xi\right\|^2 \\ &\quad - \frac{1}{\gamma_l^2 - \|D\|^2}\left\|\xi^T P B - \|D\|^2 K\xi\right\|^2 - \xi^T Q \xi.\end{aligned}$$

Therefore, if γ is chosen such that

$$\gamma \geq (\gamma_l^2 - \|D\|^2)\frac{\left\|\left.\frac{dV}{dx}\right|_{\xi} B - \|D\|^2 K\xi\right\|^2}{\left\|\xi^T P B - \|D\|^2 K\xi\right\|^2 + \xi^T Q \xi} + \|D\|^2,$$

then $\Phi_{A.1}(\xi) \leq 0$. Since V is uniformly Lipschitz, there is a bound on $\left\| \frac{dV}{dx} \right\|$ over Ω_d, hence the choice function will be bounded. When $\Phi_B(\xi)$ is examined, there are two possibilities, depending whether $\left(1 + \frac{\|D_1\|^2}{\gamma^2 - \|D_1\|^2}\right) |K\xi| \leq 1$. If so, $\Phi_B(\xi) = \Phi_{B.2}(\xi)$, and then $\Phi_B(\xi) \leq 0 \iff \Phi_A(\xi) \leq 0$. If not, by examining (5.20), the condition is:

$$\begin{aligned}\Phi_B(\xi) &= 2 \left.\frac{dV}{dx}\right|_\xi B(K\xi - \operatorname{sign}(K\xi)) - \gamma^2 \left\| \frac{\|D_1\|^2}{\gamma^2 - \|D_1\|^2} K\xi \right\|^2 \\ &\quad + \left\| C\xi - d\left((1 + \frac{\|D_1\|^2}{\gamma^2 - \|D_1\|^2})K\xi\right) \right\|^2 - \|C\xi - DK\xi\|^2 + \psi(\xi).\end{aligned}$$

As $\left|\left(1 + \frac{\|D_1\|^2}{\gamma^2 - \|D_1\|^2}\right) K\xi\right| > 1$, $1 \geq |K\xi| > 1 - \frac{\|D_1\|^2}{\gamma^2}$. If the set Z is defined $Z = \bar{\Omega} \bigcap \{\xi : 1 - \frac{\|D_1\|^2}{\gamma_l^2} \leq |K\xi| \leq 1\}$, then a conservative choice of γ to force $\Phi_B(\xi) < 0$ is

$$\gamma^2 > \frac{\max_{\xi \in Z} 2 \|D_1\|^2 \left\{ \left\| \left.\frac{dV}{dx}\right|_\xi B \right\| + \|D_2\|^2 \right\}}{\min_{\xi \in Z} \left\{ \frac{\gamma_l^2}{\gamma_l^2 - \|D_1\|^2} \left\| \|D_1\| K\xi \right\|^2 - \psi(\xi) \right\}}.$$

This expression is bounded, as $\left\| \frac{dV}{dx} \right\|$ is bounded.

3. The final possibility is that $|K\xi| > 1$. Here, $|\theta_b(\xi) > 1$. Then, examining Φ_B, we have that

$$\Phi_B(\xi) = -\gamma^2 \|K\xi + \theta_b(\xi)\|^2 + \|C\xi + d(\theta_b(\xi))\|^2 - \|C\xi + d(K\xi)\|^2 - \psi(\xi).$$

Since $\psi(\xi) \leq \frac{1}{\gamma_l^2 - \|D_1\|^2} \left\| \|D_1\|^2 K\xi \right\|^2$,

$$\Phi_B(\xi) \leq \frac{1}{\gamma^2 - \|D_1\|^2} \left\| \|D_1\|^2 K\xi \right\|^2 - \frac{1}{\gamma_l^2 - \|D_1\|^2} \left\| \|D_1\|^2 K\xi \right\|^2 \leq 0.$$

By using arguments parallel to that in the previous case, $\Phi_A(\xi)$ can be bounded above by zero over some set $\Omega_d \bigcap \{\xi : 1 < |K\xi| \leq 1 + \Delta\}$, for some $\Delta > 0$. Then, over the remainder of the set, Φ_A can be conservatively bounded as

$$\Phi_A(\xi) \leq 2 \left.\frac{dV}{dx}\right|_\xi B\left(\operatorname{sign}(K\xi) - \sigma_0(\theta_a(\xi))\right) - \gamma^2 \|\Delta\|^2.$$

Hence, the choice of γ required is bounded over Ω_d.

(3) $\implies$ (1): This is a consequence of Theorem 5.4.1. □

Remark 5.6.9: There are a number of consequences of Theorem 5.6.8 which have been observed by the author. We will briefly state two of the more interesting implications, without formal proof. Firstly, the result implies that if one fixes any linear state feedback which stabilises the local linearisation of the plant, the origin of the resulting nonlinear closed loop (for the plant input saturation problem) will necessarily have a domain of attraction with non-zero measure. When this observation is combined with those in Section 5.7, there will exist some $\mathcal{W}_{2,2}(N)$ set over which the system is finite gain stable. Secondly, if the plant is open loop stable, one can fix an arbitrary compact set Ω in the state space and ensure that the induced $\mathcal{L}_{2+}$ norm over the set Ω equals the $\mathcal{H}_\infty$ norm of the local linearisation, by using (low gain) feedback such that the control action does not saturate over Ω.

Remark 5.6.10: Theorem 5.6.8 is presented as a existence result, which shows the equivalence between state space and input/output stability, with the mathematics associated with the computational heuristic not given prominence. This was a stylistic choice, as this equivalence is important, and the use of a dissipation function of the proposed form provides what is felt to be an elegant method of proof of this equivalence. It should be noted however, that the proof of this equivalence is not believed to be the sole merit of Theorem 5.6.8, as this equivalence has been previously noted by other researchers. In the paper by Liu, Chitour and Sontag [33], the equivalence between state space and $\mathcal{L}_p$ stability was derived from basic analysis of the signal sizes ([33, Section 6.1]). As such, the observation of the equivalence alone cannot be considered a new insight; rather the contribution of the characterisation herein is that the result appears a quite natural consequence of the computational scheme.

Other researchers have examined the necessity for the use of bounded input sets, and semi-global stability results. In the paper by Lin, Saberi and Teel ([31]), it is shown that if the input set to an input saturated plant lies in an $\mathcal{W}_{2,\infty}(N)$ set, a linear state feedback exists which locally stabilises the plant and gives rise to a finite induced $\mathcal{L}_{2+}$ norm over the bounded set. There is an affinity to the result in Theorem 5.6.8, although the analysis found herein takes a different approach. The analysis paradigm in this dissertation is to fix the controller and then *a posteriori* determine over which $\mathcal{W}_{2,q}(N)$ sets the closed loop is finite gain stable, as opposed to fixing the input set *a priori* (as in [31]). (See Section 8.6 for a preliminary synthesis example.)

More concretely, the advocated synthesis technique (towards which the analysis in this dissertation is directed) would take the form of *a priori* fixing γ and determining the controller which ensures that $\|\mathbf{\Upsilon}\|_{i,2} \leq \gamma$ over $\mathcal{W}_{2,q}(N)$, for the maximal value of N, instead of *a priori* fixing $\mathcal{W}_{2,q}(N)$ and then determining a controller which gives the lowest value of $\|\mathbf{\Upsilon}\|_{i,2}$ over that set. This may appear to be a somewhat minor difference, but having γ as the *a priori* chosen design variable may mesh more cleanly with linear $\mathcal{H}_\infty$ design. The linear design methodology is usually developed with sub-optimal values of γ chosen *a priori*, and so the linearisation of the controller can be fixed using the far better understood linear theory (and more computationally efficient, so higher order models can be studied). The nonlinear action "in the large" can then be determined to preserve the sub-optimal γ as best possible. Fixing the input bound *a priori* may seem sensible for a purely nonlinear design, but in practice the synthesis algorithm may be too dependent upon this bound, and the linear design would be driven by the nonlinear constraints. This seems unsatisfactory, as the author feels that the nonlinear tail should not wag the linear design methodology dog.

5.6.1 Motivation of the Heuristics

Although it is believed that the proposed class of dissipation functions will give acceptable algorithmic performance, giving a formal justification for its form appears difficult, however. We will now present a somewhat rough motivation for the form chosen. Firstly, a result will be stated which shows that all valid dissipation functions can be generated by the correct re-definition of the function ψ which appears in the definition of V, and it will then be easier to explain the reasoning behind the choice choice of ψ taken herein.

Proposition 5.6.11: Let $\hat{V}$ be any valid dissipation function which determines a finite induced $\mathcal{L}_{2+}$ norm bound γ over Ω. There then exists a function $\psi_{\hat{V}}: \Omega \mapsto \mathbb{R}_{-}$ such that if V is generated as in equation (5.18) with $\psi_{\hat{V}}$ replacing ψ, then $V(\xi) = \hat{V}(\xi)$ for all $\xi \in \Omega$.
Proof: Let $\psi_{\hat{V}}$ be given by:

$$\psi_{\hat{V}}(\xi) := 2 \left.\frac{d\hat{V}}{dx}\right|_{\xi} (A\xi - B\sigma_0(K\xi)) + \|g(\xi, -K\xi)\|^2 .$$

As $\psi_{\hat{V}}(\xi) = H(\xi, 0)$, we have that $\psi_{\hat{V}} \leq 0$ is implied by $\hat{V}$ being a valid dissipation function. Now fix some $\xi \in \Omega$. Examine $V - \hat{V}$ along $x(\xi)$:

$$\begin{aligned} & 2\left.\frac{d(V-\hat{V})}{dx}\right|_{x(t;\xi)} (Ax(t;\xi) - B\sigma_0(Kx(t;\xi))) \\ = \;& \|g(x(t;\xi), -Kx(t;\xi))\|^2 - \psi_{\hat{V}}(x(t;\xi)) - 2\left.\frac{d(\hat{V})}{dx}\right|_{x(t;\xi)} (Ax(t;\xi) - B\sigma_0(Kx(t;\xi))) \\ = \;& 0. \end{aligned}$$

Thus, by integrating along the path of the induced trajectory, we see that $(V - \hat{V})(\xi) = (V - \hat{V})(x(t;\xi))$ for all t. However, as $\lim_{t\to\infty} x(t;\xi) = 0$, the difference is determined by the boundary condition at 0. Then, as $V(0) = 0$ by construction, $\hat{V}(0) = 0$ by definition, it follows that $V - \hat{V} \equiv 0$. □

We have just shown that all valid dissipation functions may be generated by the path integral (5.18), with the correct re-definition of ψ. The choice for the particular form in (5.17) was taken so that it forces $V = V_l$ locally, and for $|K\xi| \leq 1$, the "slack" matches that for V_l. The choice of ψ is then taken to be the minimal value to satisfy the $\Phi_{B.1}$ expression, and which appears to minimise $\left\|\frac{d(V-V_l)}{dx}\right\|$. The usefulness of minimising $\left\|\frac{d(V-V_l)}{dx}\right\|$ seems to be apparent when one examines the expressions which are used to bound above γ which appear in the proof of Theorem 5.6.8.

5.7 Bounding of the Reachable Set

We have now motivated an approach, to be presented in the following proposition, which allows us to give a simple method of overbounding the reachable set Ω, given that the disturbance lies in a $\mathcal{W}_{2,2}(N)$ set. The analysis for other $\mathcal{L}_{p+}$ norm bounds on the signal set is presumably more difficult, as the present result exploits the way in which a valid dissipation resembles a potential energy function for $\mathcal{L}_{2+}$ signals. If the disturbance signal set is $\mathcal{L}_{2+}$, it appears likely that the reachable set will be similar to that for the open loop plant; i.e., unbounded if unstable, or overbounded by some ellipsoid if the plant is open loop asymptotically stable.

Proposition 5.7.1: Fix a system which is of the form specified by Theorem 5.6.8, which is $\mathcal{L}_2$-detectable. Assume that there exists some compact set $\Omega^o \subset \mathbb{R}^n$ containing the origin over which the induced trajectories (i.e., $w = 0$) all lie in $\mathcal{L}_{2+}$. Then if we fix a matrix $Q > 0$, by the use of Lemma 5.6.2, there exists a function V_l which determines a function ψ which is used to compute a valid dissipation function V (given by equation (5.18)) with a finite induced norm bound γ over Ω^o (by Theorem 5.6.8). Then, if

$$\gamma^2 N^2 < \min_{\xi \in \partial\Omega^o} \{2V(\xi)\}, \tag{5.21}$$

the reachable set Ω_N for the input set $\mathcal{W}_{2,2}(N)$ is such that

$$\Omega_N \subseteq \Omega_N^o = \{\xi : \xi \in \Omega^o, 2V(\xi) \leq \gamma^2 N^2\}.$$

Proof: This is a standard overbounding technique, and is probably quite conservative in most cases. The bound comes about by noting that the dissipation inequality implies that if $x(t) \in \Omega_o$ for all $t \in [0, T]$, then

$$2V(x(T)) \leq \gamma^2 \int_0^T \|w(t)\|^2 \, dt. \tag{5.22}$$

In order to verify that $x(t)$ cannot exit the set Ω_o, assume the converse: i.e., that there exists some w in $\mathcal{W}_{2,2}(N)$ such that for some t, $x(t) \notin \Omega_o$. Let T be the infimal value of the set $\{t : x(t) \notin \Omega_o\}$. As Ω_o is compact, $x(T) \in \partial\Omega_o$. However, as $x(t) \in \Omega_o$ for $t \in [0, T]$, $2V(x(T)) \leq \gamma^2 \int_0^T \|w(t)\|^2 \, dt \leq \gamma^2 \|w\|_2^2$, contradicting (5.21). This contradiction implies that $x(t)$ is confined to Ω_o. This in turn implies that (5.22) holds for all $T \in \mathbb{R}_+$, hence

$$2V(x(T)) \leq \gamma^2 N^2, \forall T \in \mathbb{R}_+ \implies x(T) \in \Omega_N^o, \forall T \in \mathbb{R}_+ \implies \Omega_N \subseteq \Omega_N^o. \quad \square$$

Remark 5.7.2: As the set Ω_N^o may be a strict subset of Ω^o, if one applies the test theorem over Ω_N^o, the indicated value of the induced norm bound may decrease. The same bounding procedure on the reachable set could then by applied, giving rise to a new set which is a bound for Ω_N. If repeatedly applied, the bounding set should presumably converge to some minimal set which is determined by that choice of V (not necessarily Ω_N, however).

Remark 5.7.3: Note that if the system is globally state space stable, and if the candidate function is radially unbounded (which seems most likely), Proposition 5.7.1 would indicate that for all $\mathcal{W}_{2,2}(N)$ sets, there is an induced norm bound. The bound as a function of N may be unbounded, however (see Section 8.4).

Remark 5.7.4: It is not necessary that a bounded reachable set exists when the disturbance set is a $\mathcal{W}_{2,\infty}(N)$ set, even if the system is globally state space stable. For example, take the second order system:

$$\begin{aligned}
\dot{x_1}(t) &= x_2(t),\\
\dot{x_2}(t) &= \sigma_0(w(t) - x_2(t) - \sigma_0(x_1(t))),\\
x(0) &= 0.
\end{aligned}$$

It is easy to see that the $\mathcal{L}_{\infty+}$ signal $w(t) = 2$ drives the state to ∞. However, the analysis given in [50] shows that this system is globally state space stable. The motivation for this system is that this is a globally stabilising state feedback for a double integrator, of a form which is needed to stabilise third order and higher integrator chains (as discussed in Section 4.3).

5.8 Induced Norm Computation for Controller Only Saturation

We will now briefly comment on some related work on the induced norm computation for the case of controller only saturation. This problem configuration gives rise to an "unfair game" between the actions of the control output and external disturbance. For this reason, some of the more interesting properties exhibited by the plant input saturation configuration do not appear.

The state space stability version of this problem has received extensive attention; the induced norm computation much less so. The work closest in spirit to that found herein is that in the paper [22] by J.W. Helton and W. Zhan. In that paper, the main study is of piecewise linear dynamics equations with one "break" in the function definition. The extension to having two "breaks" (such as occurs in the standard saturation) is obviously straightforward.

The analysis in [22] is quite heavily tied to the viscosity solution dissipation theory, although the attempt is made to move away from numerical schemes relying upon discretisations over the state space. It is recognised that the solutions are close to being piecewise quadratic in nature. The analysis is quite complicated, due to the necessity of piecing together the quadratic solutions along the boundary between the different operating regions. The results in [22] are explicitly described as being preliminary in nature. Gaining a better understanding of the computational problem is an ongoing research problem. For example, although a viscosity solution is known to exist from general principles, no guarantee is made that it will fit the class of piecewise quadratic functions given.

Although the author has not fully examined the problem, it is conjectured that the generic existence theorem equivalent to Theorem 5.6.8 could be easily extended for the class of systems studied in [22]. (Of course, this would again only be a sub-optimal characterisation.) In order to adapt the technique, the role of the function ψ would be to attempt to force the solution to appear like the quadratic functions which are the result of the solution of the two appropriate Riccati equations. Once the integration is made along the paths of the induced trajectories, it is likely that the result would no longer be purely piecewise quadratic in form.

The problem of continuity along boundaries, and how the scheme could be computed in an efficient manner, appears to be the problem of most significance in this area of research.

In the short term, it is quite likely that computational schemes could be developed most easily for the controller only saturation configuration, 'as opposed to the plant input saturation configuration. The test condition which results is much simpler, being a straight application of the completion of squares, since the complications associated with the disturbance entering through the nonlinearity is eliminated. As such, it appears the solutions will be close to the quadratic form associated with the solutions of the appropriate Riccati equations. The results may be somewhat unsatisfying in that open loop unstable plants cannot be stabilised by purely bounded control inputs, and so it may become necessary to introduce boundedness assumptions for the disturbance, which will probably introduce similar complications to those which occur with the plant input saturation configuration.

Chapter 6

Induced Norm Computation for Discrete Time Systems

6.1 Introduction

In this chapter, we will turn our attention to stability results for discrete time systems. Some generic results stated for continuous time systems, which are given in terms of operator descriptions, will apply to discrete time systems as well, e. g. the Small Gain Theorem. For linear systems, the ability to apply bilinear transformations to frequency domain descriptions means that many results can be proven in continuous or discrete time, then extended to the other case (which is admittedly not a generally favoured procedure). When analysis is based on state space descriptions, it is often possible (even in the nonlinear case) to apply the same general principles, but there are necessarily larger practical differences.

In this chapter, we will examine areas where there are significant differences in the stability analysis between continuous and discrete time systems. We have seen in Chapter 3 that controllability/observability conditions are essentially the same, and in fact many of the technical issues involving existence/uniqueness of solutions become trivial in discrete time as the solutions are generated as sums and not integrals. We will examine here the extension of Lyapunov stability, dissipation results, and then the extension of the results on induced norm computations for single input systems. We will treat incremental gain results for continuous and discrete time systems together in Chapter 7, as there is a closer affinity between the two cases, although some constructive results are presented only for discrete time due to the apparent complexity of their extension to continuous time.

In Section 6.2, a discrete time version of a Lyapunov stability theorem is given. Section 6.3 gives general discrete time dissipation results, and discusses their application to linear systems.

New results, which are the discrete time extension of the continuous time ones in the previous chapter, are then given in Sections 6.4, 6.5. The generic test is applied to saturating systems, giving a somewhat more specific condition. Of some interest is that it is possible to extend the heuristically determined dissipation function to discrete time, and hence give a non-local, (sub-optimal) existence result. A final note on the determination of the reachable set is given in 6.6.

6.2 Lyapunov Stability For Discrete Time Systems

The extension of Lyapunov stability results to discrete time systems is a fairly straightforward operation. In fact the definitions could be given together for both continuous and discrete time without much loss of clarity. Discrete time versions of the theory have an added advantage in that lack of differentiability of the Lyapunov functions is not an issue; however, it is necessary to examine differences between states which can be widely separated. This is in distinction with the continuous time case, where the comparison is essentially between states arbitrarily close to each other - the information on which is contained in the derivative of V. This makes closed form analysis of discrete time systems generally more difficult. The results will be stated here for shift-invariant systems.

Assumption 6.2.1: The set Ω upon which the state space system evolves is assumed to either be equal to $\mathbb{R}^n$, or a bounded, connected, subset of $\mathbb{R}^n$ with non-zero measure which contains the origin.

Definition 6.2.2: A function $V : \Omega \mapsto \mathbb{R}_+$ is a *candidate Lyapunov function* (discrete time) on $\Omega \subseteq \mathbb{R}^n$ if V meets the requirements listed in in Definition 4.2.2, except for the condition that V is continuously differentiable. It is sufficient that V is continuous.

Theorem 6.2.3: Let a shift invariant nonlinear discrete time state space system be generated by the difference equation,

$$\begin{aligned} x[k+1] &= f(x[k]), \\ x[0] &= \xi_0. \end{aligned}$$

We assume that $\xi_0 \in \Omega$, $f : \mathbb{R}^n \mapsto \mathbb{R}^n$ continuous and with $f(0) = 0$, and with Ω invariant under f. Then, if there exists a candidate Lyapunov function V such that for all $\xi \in \Omega \setminus 0$, we have that

$$\Delta(\xi) := V(f(\xi)) - V(\xi) < 0, \tag{6.1}$$

with

$$\lim_{R\to\infty} \min_{1 \le ||\xi|| \le R} \Delta(\xi) < \kappa_\infty < 0, \tag{6.2}$$

(in the case where Ω is unbounded), then the system is uniformly asymptotically stable (Definition 4.2.4). (In such a case, V is termed a valid Lyapunov function.)
Proof: Found in [45, Chapter 4] (amongst others). □

6.2.1 Linear Systems and the Algebraic Lyapunov Equation

When the system of interest is linear, by assuming a quadratic form for the Lyapunov function, it is possible to give a necessary and sufficient test for uniform asymptotic stability in terms of a matrix algebraic equation. Once again, the second derivative of the function is sufficient to characterise it, and is a constant matrix. This allows stability results to be given in terms of tests on matrices, leaving no need to examine Lyapunov functions pointwise over $\mathbb{R}^n$.

Theorem 6.2.4: Fix a linear discrete time dynamic system with dynamics matrix $A \in \mathbb{R}^{n\times n}$ and initial condition $\xi_0 \in \mathbb{R}^n$. The following statements are equivalent:

1. $\rho(A) < 1$;
2. for all symmetric $Q > 0$ in $\mathbb{R}^{n\times n}$, there exists a symmetric solution $X > 0$ to the *discrete algebraic Lyapunov equation*:
$$A'XA - X = Q; \tag{6.3}$$
3. the system is uniformly asymptotically stable over $\mathbb{R}^n$.

Proof: This is standard; see [62, Chapter 18]. Note that if the conditions hold, $V(\xi) = \xi' X \xi$ is a valid Lyapunov function. □

6.3 Dissipation Results For Discrete Time Systems

We will now summarise the basic dissipation test used for computing induced ℓ_{2+} norms. It is possible to bundle the results here with those for continuous systems given in Section 4.5; however the technical issues in the discrete time case are much less in the continuous time case as differentiability is no longer an issue. That being said, the fact that macroscopic comparisons across the state space are needed in the test makes analysis much more difficult.

Definition 6.3.1: A function $V : \Omega \subset \mathbb{R}^n \mapsto \mathbb{R}_+$ is a *candidate dissipation function* (discrete time) if V is continuous, $V(0) = 0$, and $V(\xi) > 0$ for all $\xi \in \Omega$ for $\xi \neq 0$.

Theorem 6.3.2: Let the system operator $\boldsymbol{\Upsilon} : w \mapsto z$ be generated by the the difference equations,

$$\begin{aligned} x[k+1] &= f(x[k], w[k]), \\ z[k] &= h(x[k], w[k]), \\ x[0] &= 0, \end{aligned}$$

with $x[k] \in \mathbb{R}^n$, $w[k] \in \mathbb{R}^p$, $z[k] \in \mathbb{R}^m$. The functions f, h are assumed to be continuous with respect to x, w, and $f(0,0) = 0$, $h(0,0) = 0$. Fix an input set $\mathcal{W}_{2,\infty}(N)$; it is assumed that the reachable set Ω is bounded. Then, if there exists a candidate dissipation function such that for all $\xi \in \Omega$,

$$\max_{\omega \in [-N,N]} H(\xi,\omega) := \left\{ V(f(\xi,\omega)) - V(\xi) - \gamma^2 \|\omega\|^2 + \left\| h(\xi,\omega)^2 \right\| \right\} \leq 0, \tag{6.4}$$

then $\|z\|_2 \leq \gamma \|w\|_2$ for all $w \in \mathcal{W}_{2,\infty}(N)$.
Proof: This is standard; for example see [30]. □

The theoretical properties of dissipation results are not as well known in the discrete case as in the continuous time case. In the papers [6, 30] by C.I. Byrnes and W. Lin, affine nonlinear systems are examined. In [6], the problem of determining whether a system is lossless is examined, and characterised in terms of a quadratic storage function. In the paper [30], results on dissipativity for affine systems is given. The results are given in terms of the solution of a Hamilton-Jacobi discrete time equation. They are unfortunately local results when the issue of existence arises. When examining linear systems with the standard saturation, such results are essentially trivial as the local properties are those of the equivalent linear system, as will be seen later.

6.3.1 Induced ℓ_{2+} Norm Computation for Linear Systems

When examining the input/output behaviour of linear systems, by use of frequency domain theory, one can quite often convert results given for continuous time systems to equivalent formulations for discrete time. During the development of $\mathcal{H}_\infty$ control theory in the 1980's, it was typical to develop state space formulations first for the continuous time case, and then rely on frequency domain transformations. This was necessitated by the fact that discrete time Riccati equations are somewhat more tedious to analyse, and the algebra was much easier to work through in the continuous time case first.

The tactic of using frequency domain transformations provides little guidance when the extension to nonlinear systems is considered, however. It may be possible to relate the results for linear systems with saturation between discrete and continuous time. That issue is beyond the scope of the present work, however. We will instead tailor our results for discrete time systems, and will need to use the corresponding background results from linear system theory. The text [62] and the thesis by P.A. Iglesias [25] (amongst others) may be consulted for a discussion of the linear theory.

Theorem 6.3.3: Fix a discrete time shift invariant linear system $\boldsymbol{\Upsilon}_L$ with transfer function $G(z) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$, with $[C, A]$ observable. Define $R = \gamma^2 I - D'D$. We assume that $\rho(A) < 1$, and that $\bar{\sigma}(D) < \gamma$. Then, the following statements are equivalent:

1. $\|\boldsymbol{\Upsilon}_L\|_{i,2} < \gamma$;
2. $\|G\|_\infty < \gamma$;
3. there exists a symmetric $X > 0$ with $R - B'XB > 0$ which solves the algebraic Riccati equation,

$$A'XA - X + C'C + (B'XA + D'C)'(R - B'XB)^{-1}(B'XA + D'C) = 0; \tag{6.5}$$

4. there exists a symmetric $X > 0$ such that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}' \begin{bmatrix} X & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} - \begin{bmatrix} X & 0 \\ 0 & \gamma^2 I \end{bmatrix} < 0.$$

Proof: See [25, Section 3.2] or [62, Chapter 18] for details. □

We finally give the time domain consequence of the existence of a solution to the Riccati equation as a dissipation result, being the discrete time version of the completion of squares.

Lemma 6.3.4: For the system $\Upsilon_L\colon w \mapsto z$ as in Theorem 6.3.3, if the solution X exists to the equation (6.5), then the function $V(\xi) = \xi' X \xi$ is a valid global dissipation function for the bound γ, and,

$$V(x[k+1]) - V(x) - \gamma^2 \|w[k]\|^2 + \|z[k]\|^2 = \\ - \left\|(R - B'XB)^{\frac{1}{2}}(w[k] - (R - B'XB)^{-1}(B'XA + D'C)x[k])\right\|^2.$$

Proof: This is taken from [25, Lemma 3.6]. □

6.4 Induced ℓ_{2+} Norm Computation for Linear Systems with Saturation

We will now present the test equivalent to that in Theorem 5.4.1. Unfortunately, the nature of discrete time systems is such that the result does not seem to offer a large apparent advantage over the generic statement given in Theorem 6.3.2. The test still requires a macroscopic optimisation over the state space; however the range which must be searched is reduced. This can be ascribed to the fact that we cannot assume a quadratic form for the candidate function (except locally), and so cannot "complete the squares".

Theorem 6.4.1: Take the (discrete time) system operator $\Upsilon\colon w \mapsto z$, which is generated by the difference equations:

$$\begin{aligned} x[k+1] &= Ax[k] + B\sigma_0(-K(x[k]) + w[k]) =: f(x[k], w[k]), \\ z[k] &= Cx[k] + D(-K(x[k]) + w[k]) =: h(x[k], w[k]), \\ x[0] &= 0. \end{aligned}$$

Here, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times 1}$, $C \in \mathbb{R}^{p\times n}$, $D \in \mathbb{R}^{p\times 1}$, and $K\colon \mathbb{R}^n \mapsto \mathbb{R}$ continuous. We assume that $C'D = 0$. Additionally, we assume that once the input set $\mathcal{W}_{2,\infty}(N)$ is fixed, the reachable set $\Omega \subset \mathbb{R}^n$ is bounded. Then, let V be a candidate function defined on Ω. If for a fixed choice of $\gamma > \|D\|$, we have that for all $\xi \in \Omega$,

$$\max_{\omega \in \mathcal{S}(\xi)} \left\{ H(\xi, \omega) = V(f(\xi, \omega)) - V(\xi) - \gamma^2 \|\omega\|^2 + \|h(\xi, \omega)\|^2 \right\} \leq 0, \tag{6.6}$$

where

$$\mathcal{S}(\xi) = \left\{ [K(\xi) - 1, K(\xi) + 1] \bigcap [-N, N] \right\} \bigcup \left\{ N\sigma_0 \left(\frac{-1}{N} \frac{\|D\|^2}{\gamma^2 - \|D\|^2} K(\xi) \right) \right\},$$

then $\|\Upsilon\|_{i,2} \leq \gamma$.

Proof: Comparison between the tests in inequalities (6.4) and (6.6) shows that the test is sufficient at ξ if $[-N, N] \subseteq \mathcal{S}(\xi)$. Therefore, assume that $[-N, N] \not\subseteq \mathcal{S}(\xi)$. As $[-N, N] \not\subseteq \mathcal{S}(\xi)$, we must have $K(\xi) + 1 < N$ and/or $K(\xi) - 1 > -N$. Without loss of generality, assume that $K(\xi) + 1 < N$ (the case $K(\xi) - 1 > -N$ is symmetric). We will now verify that the condition in (6.6) is sufficient to force $H(\xi, \omega) \leq 0$ for $\omega \in [K(\xi) + 1, N] \bigcap [-N, N]$.

For all $\omega \in [K(\xi)+1, N]$, $f(\xi,\omega) = f(\xi, K(\xi)+1)$. Hence, examination of (6.4) gives us that,

$$\max_{\omega \in [K(\xi)+1,N] \bigcap [-N,N]} H(\xi,\omega) = H(\xi,\omega^*),$$

with ω^* the maximum of the function $u_\xi(\omega) = -\gamma^2 \|\omega\|^2 + \|C\xi + D(-K(\xi)+\omega)\|^2$ over the interval $[K(\xi)+1, N] \bigcap [-N, N]$. The global maximum of this quadratic function (in ω) is found at $\omega = -\frac{\|D\|^2}{\gamma^2 - \|D\|^2} K(\xi)$. And it is then straightforward that the maximum to the constrained problem lies either at the global minimum or at the endpoints of the constraint interval. We have forced ω^* to lie in $\mathcal{S}$, as either $\omega^* = \left\{ N\sigma_0 \left(\frac{-1}{N} \frac{\|D\|^2}{\gamma^2 - \|D\|^2} K(\xi) \right) \right\}$, or $\omega^* = K(\xi)+1$. And so, (6.6) implies that $H(\xi,\omega^*) \leq 0$, implying the required condition holds over the interval of interest. □

Remark 6.4.2: It is not a restriction that we take $\gamma > \|D\|$ in Theorem 6.4.1, as if we take as an input the impulse of magnitude N, $\|z\|_2 \geq \|z[0]\| = N\|D\|$, giving $\|\mathbf{\Upsilon}\|_{i,2} \geq \|D\|$. If the system is observable and has non-trivial dynamics, we will have $z[k] \neq 0$ for some $k > 0$, forcing $\|\mathbf{\Upsilon}\|_{i,2} > \|D\|$. Additionally, in the case of $K(\xi)$ being a memoryless linear system with saturation, the condition that $\gamma > \|D\|$ is implied by Proposition 6.4.3.

The weakness of the pointwise test condition in equation (6.6) provides a justification as to why the result is given with the restriction that the input w does not enter the state equation in a linear fashion. If one were to add a term corresponding to B_1 in (3.5), the search to determine $H^*(\xi)$ needs to be made over the entire interval $[-N, N]$; i.e. there is no advance over Theorem 6.3.2.

6.4.1 Lower Bounds for the Induced ℓ_{2+} Norm

We will now state two results showing that the induced ℓ_{2+} norm of linear systems with the standard saturation is bounded below by the behaviour of the local linearisation. The author is unaware of any specific reference for these results, and so the proof is given in Appendix A for completeness.

Proposition 6.4.3: Fix a feedback loop in the standard feedback configuration (Section 3.3.1), with $\mathbf{G}$,$\mathbf{K}$ well posed linear systems with the standard saturation. Let $\mathbf{G_L}, \mathbf{K_L}$ be their respective local linearisations. Then, the induced ℓ_{2+} norm can be characterised as:

1. if $\begin{bmatrix} I & K_L(z) \\ G_L(z) & I \end{bmatrix}^{-1} \in \mathcal{H}_\infty$ and $\left\| \begin{bmatrix} I & K_L(z) \\ G_L(z) & I \end{bmatrix}^{-1} \right\|_\infty = \gamma_l^*$, then
$$\left\| \begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{G} & \mathbf{I} \end{bmatrix}^{-1} \right\|_{i,2} \geq \gamma_l^*;$$
2. if the operators are all generated by finite dimensional realisations, then if $\begin{bmatrix} I & K_L(z) \\ G_L(z) & I \end{bmatrix}^{-1} \notin \mathcal{RH}_\infty$, then $\begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{G} & \mathbf{I} \end{bmatrix}^{-1}$ does not have a finite induced ℓ_{2+} norm.

Proof: This is to be found in Appendix A. □

6.5 A Heuristic Method to Compute Candidate Dissipation Functions

The heuristic computational method developed in Section 5.6 will now be extended to discrete time. It will once again be presented as an existence condition which interrelates state space and input/output stability under certain conditions. The fact that a (sub-optimal) existence condition can be given in a semi-global form appears to be important; for continuous time systems, such existence results were available almost immediately. The motivation for the form of the function V follows that laid out in Section 5.6.1.

Once again, some further restrictive assumptions are made one the form of the system. The restriction that $B_1 = 0$ (as in (3.5)) would be difficult to eliminate, as noted in the previous section. The assumptions that $K(\xi)$ is linear and that $D = 0$ would be more easily relaxed, although it may be useful to assume that $K(\xi)$ is linear over some ball containing the origin.

Assumption 6.5.1: The operator $\mathbf{\Upsilon}$ is assumed to be generated by the difference equations of the form used in Theorem 6.4.1, with the added assumptions that $D = 0$ and $K(\xi)$ is linear (denoted $K\xi$). We (necessarily) assume that $A - BK$ is stable; i.e., has eigenvalues of modulus strictly less than 1.

In order to define the candidate function, use is made of the linear solution. Fix an arbitrary matrix $Q > 0$, and the scalar γ_l such that

$$\gamma_l > \left\| \left[\begin{array}{c|c} \tilde{A} & B \\ \hline \begin{bmatrix} C \\ Q^{\frac{1}{2}} \end{bmatrix} & 0 \end{array} \right] \right\|_\infty .$$

Let $0 < X \in \mathbb{R}^{n\times n}$ be the solution to the discrete algebraic Riccati equation:

$$\tilde{A}'X\tilde{A} - X + C'C + (B'X\tilde{A})'(R - B'XB)^{-1}B'X\tilde{A} + Q = 0,$$

with $R := \gamma_l^2$, $R - B'XB > 0$, $\tilde{A} := A - BK$ (the solution will exist by Theorem 6.3.3). Define $V_l\colon \mathbb{R}^n \mapsto \mathbb{R}_+$ as

$$V_l(\xi) := \xi'X\xi.$$

Let $\psi\colon \mathbb{R}^n \mapsto \mathbb{R}_-$ be defined:

$$\psi(\xi) := \min\left\{0, V_l(A\xi - B\sigma_0(K\xi)) - V_l(\xi) + \|C\xi\|^2\right\}.$$

If $|K\xi| \le 1$, then by the use of Lemma 6.3.4,

$$\psi(\xi) = -\xi'Q\xi - \left\|(R - B'XB)^{-\frac{1}{2}}(B'X(A - BK))\xi\right\|^2.$$

Definition 6.5.2: Let the *induced trajectory* from $\xi \in \Omega$ (denoted $x[k;\xi]$) be the trajectory generated by:

$$\begin{aligned} x[k+1] &= Ax[k] + B\sigma_0(-Kx[k]), \\ x[0] &= \xi. \end{aligned}$$

It is now possible to give the definition of the proposed candidate function. Let the candidate function $V\colon \bar{\Omega} \mapsto \mathbb{R}_+$ be generated by:

$$V(\xi) := \sum_{k=0}^{\infty} \left\{ \|Cx[k;\xi]\|^2 - \psi(x[k;\xi]) \right\}. \tag{6.7}$$

We will first give some preparatory results on the properties of the induced trajectories and candidate functions before giving the main existence result.

Let X_∞ be the solution to the discrete algebraic Lyapunov equation; i.e.,

$$\tilde{A}' X_\infty \tilde{A} - X_\infty = -I, \tag{6.8}$$

(once again $\tilde{A} = A - BK$). Then, define $\Omega_l \subset \mathbb{R}_n$ as

$$\Omega_l = \{\xi\colon \xi' X_\infty \xi \leq b\},$$

with $b = \min_{\xi:|K\xi|=1} \xi' X_\infty \xi$.

Lemma 6.5.3: For all $\xi \in \Omega_l$, $x[k;\xi] \in \Omega_l$, and $|Kx[k;\xi]| \leq 1$ for all $k \in \mathbb{Z}_+$.
Proof: This proof is a straightforward extension of that for Lemma 5.6.4, and will be omitted. □

Lemma 6.5.4: If $x[\xi] \in \ell_{2+}$ for all $\xi \in \bar{\Omega}$, there then exists a finite constant T_Ω such that for all $\xi \in \bar{\Omega}$, $x[T_\Omega, \xi] \in \Omega_l$.
Proof: (This proof needs to be approached in a slightly different manner to that for Lemma 5.6.6 because of technical differences. Some details will not be presented here, as the proof is otherwise felt to be straightforward.) It is obvious that $x[k;\xi]$ is Lipschitz continuous with respect to ξ, as the mapping $\xi \mapsto x[k;\xi]$ is the composition of $k+1$ Lipschitz continuous mappings on $\mathbb{R}^n$ (counting the identity mapping at $k=0$). Then if we define the function $T(\xi)$ as $T(\xi) = \min_{k\in\mathbb{Z}_+}\{k\colon x[k;\xi] \in \Omega_l\}$, we can write

$$\|x[\xi]\|_2^2 = \sum_{k=0}^{T(\xi)} \|x[k;\xi]\|^2 + x_f' X_\infty x_f,$$

(with $x_f = x[T(\xi)+1;\xi]$), by the use of Lemma 6.5.3. The stability property implied by (6.8) ensures that there exists an open ball surrounding x_f which lies in Ω_l. We can therefore ensure that the difference between $\|x[\xi]\|_2$ and $\|x[\xi+\delta]\|_2$ can be bounded in terms of $\|\delta\|$ over some ball surrounding ξ. This would be done by bounding the difference over the finite sum using a growth constant, and then ensuring that $x[T(\xi);\xi+\delta]$ lies in the open ball surrounding x_f in Ω_l, and so the difference over the infinite "tail" is bounded by the use of the quadratic. Hence, continuity of $\|x[\xi]\|_2$, and the fact that we assume it is finite over $\bar{\Omega}$, implies that

$$\kappa = \max_{\xi\in\bar{\Omega}}\{\|x[\xi]\|_2^2\}$$

exists and is achieved. Then, we know that if $x[k;\xi] \notin \Omega_l$, $x[k;\xi]' X_\infty x[k;\xi] \geq b$. Linear algebra then tells us that $\|x[k;\xi]\|^2 \geq \frac{b}{(\bar{\sigma}(X_\infty))}$. Let $T_\Omega \in \mathbb{Z}_+$ with $T_\Omega > \frac{\kappa(\bar{\sigma}(X_\infty))}{b}$. It then follows that for all $\xi \in \bar{\Omega}$, $x[T_\Omega;\xi] \in \Omega_l$. This is shown by contradiction: assume that there exists some $\xi \in \bar{\Omega}$ such that $x[T_\Omega;\xi] \notin \Omega_l$. Lemma 6.5.3 indicates then that $x[k;\xi] \notin \Omega_l$ for $k = 0, \ldots, T_\Omega$. It then would follow that $\|x[\xi]\|_2^2 \geq T_\Omega \frac{b}{\bar{\sigma}(X_\infty)} > \kappa$, a contradiction. □

Lemma 6.5.5: The candidate function V given in (6.7) is uniformly Lipschitz continuous over $\bar{\Omega}$ if for all $\xi \in \bar{\Omega}$, $x[\xi] \in \ell_{p+}$.

Proof: This is again a straightforward adaptation of Lemma 5.6.7 to discrete time, relying upon the bound given in Lemma 6.5.4. It is felt not to be necessary to repeat the steps, and so the proof is omitted. □

Theorem 6.5.6: Fix a closed loop discrete time system $\mathbf{\Upsilon}: w \mapsto z$, which meets the conditions of Assumption 6.5.1. We assume that the system is ℓ_{2+}-detectable and has a reachable set Ω of compact closure for a fixed input set $\mathcal{W}_{2,\infty}(N)$. Additionally, we assume that for all points along the boundary of Ω, the induced trajectory enters Ω in a finite time. The following statements are then equivalent:

1. for all $w \in \mathcal{W}_{2,\infty}(N)$, $\mathbf{\Upsilon} w \in \ell_{2+}$;
2. for all $\xi \in \Omega$, $x(\xi) \in \ell_{2+}$;
3. the candidate function $V: \bar{\Omega} \mapsto \mathbb{R}_+$ given in (6.7) is in fact a valid dissipation function for a finite choice of $\gamma \geq \gamma_l$.

Proof: (1) $\implies$ (2): This follows directly the equivalent arguments in Theorem 5.6.8 (the proof relies on generic operator manipulations which are not affected by the choice of time axis).

(2) $\implies$ (3): The assumed implication and the assumption regarding behaviour along the boundary implies that $x[\xi] \in \ell_{2+}$ for all $\xi \in \bar{\Omega}$. Lemma 6.5.5 indicates that V is uniformly Lipschitz on Ω. It now remains to show that it determines a finite induced norm bound.

We wish to apply the test in Theorem 6.4.1 in a pointwise fashion over $\bar{\Omega}$, and show that for each point, a finite γ exists which ensures that the Hamiltonian at that point is non-positive. If the means of bounding the choice of γ needed is continuous, compactness of $\bar{\Omega}$ ensures that a finite choice of γ only is needed over the entire set.

Fix a point $\xi \in \bar{\Omega}$. We first note that the test interval $\mathcal{S} \subseteq [K\xi - 1, K\xi + 1] \bigcup \{0\}$. If $\omega = 0$, the construction of V ensures that

$$H(\xi, 0) = V(f(\xi, 0)) - V(\xi) + \|C\xi\|^2 \leq 0.$$

We therefore need only verify that $H(\xi, \omega) \leq 0$ for some γ over the interval $[K\xi - 1, K\xi + 1]$.

Consider:

$$\begin{aligned} H(\xi,\omega) &= V(f(\xi,\omega)) - V(\xi) - \gamma^2\omega^2 + \|C\xi\|^2 & (6.9) \\ &= (V(f(\xi,\omega)) - V_l(f(\xi,\omega))) - (V(f(\xi,0)) - V_l(f(\xi,0))) - (\gamma^2 - \gamma_l^2)\omega^2 \\ &\quad + \left\{V_l(f(\xi,\omega)) - V_l(\xi) - \gamma_l^2\omega^2 + \|C\xi\|^2\right\} \\ &\quad + \{V(f(\xi,0)) - V(\xi) - V_l(f(\xi,0)) + V_l(\xi))\}. & (6.10) \end{aligned}$$

From Lemma 6.3.4, it follows that $V_l(f(\xi,\omega)) - V_l(\xi) - \gamma_l^2\omega^2 + \|C\xi\|^2 \leq 0$. Additionally, examination of the construction of V and ψ gives the inequality:

$$V(f(\xi,0)) - V(\xi) \leq V_l(f(\xi,0)) - V_l(f(\xi,0)) + \xi' Q\xi.$$

Substitution into (6.10) yields:

$$\begin{aligned} H(\xi,\omega) &\leq (V(f(\xi,\omega)) - V_l(f(\xi,\omega))) - (V(f(\xi,0)) - V_l(f(\xi,0))) \quad (6.11)\\ &\quad -\xi' Q\xi - (\gamma^2 - \gamma_l^2)\omega^2 \\ &\leq \kappa|\omega| - \xi' Q\xi - (\gamma^2 - \gamma_l^2)\omega^2, \quad (6.12) \end{aligned}$$

as Lemma 6.5.5 ensures that a uniform Lipschitz constant exists over Ω for V, and hence for $V - V_l$.

It can then be seen that a finite choice of γ is needed to make $H(\xi,\omega) \leq 0$ over Ω. To ensure this, note that $V = V_l$ locally, and so there is a minimum value of $\xi' Q\xi$ which appears in the bounding expression (6.12) at points ξ where the test is of interest.

(3) $\implies$ (1): This is a consequence of Theorem 6.4.1 $\square$

6.6 Bounding of the Reachable Set

It is possible once again to use the dissipation function computed in Section 6.5 to overbound the reachable set for signals lying in bounded input sets. There is some distinction here between continuous and discrete time systems, as in fact many of the technical problems that arose in Section 5.7 do not arise here. The result is correspondingly weaker, however.

Proposition 6.6.1: Assume there exists a compact set Ω^o containing the origin over which the induced trajectories all lie in ℓ_{2+}. Also assume that the choice of V determined by (6.7), defined on Ω^o (only), determines a finite induced ℓ_{2+} norm bound γ over Ω^o for test values lying in the range $[-N, N]$ for some $N > 0$. Then for $n_2 < N$, the reachable set Ω_{n_2} for the input set $\mathcal{W}_{2,2}(n_2)$ is such that

$$\Omega_{n_2} \subseteq \Omega^o_{n_2} = \{\xi : \xi \in \Omega^o, V(\xi) \leq \gamma^2 n_2^2\}.$$

Proof: The problems associated with ensuring that the trajectory cannot escape the set Ω^o cannot occur here, as otherwise the test would fail, as V is not defined outside Ω^o. We can then note that the dissipation inequality indicates that:

$$V(x[k+1]) \leq \gamma^2 \sum_{i=0}^{k} \|w[i]\|^2 .$$

It is then obvious that since we have *a priori* that $x[k] \in \Omega^o$, $x[k] \in \Omega^2_{n_2}$ is implied by $\|w\|_2 \leq n_2$. $\square$

Proposition 6.6.1 is not completely satisfactory in that a key point appears to have been assumed in advance (that the trajectory is confined to Ω^o). However, this is to be expected, as the test for the bound γ requires that such a determination be made. It may be possible to eliminate this precondition in a limited sense by showing that for small enough n_2, there will be some subset of the domain of attraction of the origin (in the ℓ_{2+} sense) where it can be guaranteed that the state trajectory will be confined to this subset.

In the case where the induced trajectory lies in ℓ_{2+} globally, Proposition 6.6.1 will be somewhat less trivial. If the input set is taken to be $\mathcal{W}_{2,\infty}(N)$, the result will be tautologically implied by the norm bound test.

Chapter 7

Incremental Gain Properties

7.1 Introduction

In this chapter, the incremental gain properties of linear systems with saturation will be examined. From a general theoretic viewpoint, the use of the incremental gain in robust control analysis has huge advantages over the induced norm. The stronger version of the Small Gain Theorem can be used, and other analysis is simplified. Examples of such use occur in the papers [20],[18].

Unfortunately, when the incremental gain properties of linear systems with saturation are examined, the analysis appears to be far too conservative. Although the results presented in this chapter are somewhat specialised in nature, they seem to point to the principle that the incremental gain becomes unbounded if the plant is open loop unstable.

In Section 7.2, what are believed to be the first main results on the the continuity properties of linear systems with saturation are surveyed, taken from the paper by Chitour, Liu, and Sontag [9]. These results are given for a more specific class of plants, but the analysis gives a good characterisation of how the continuity of generic saturation functions affects the continuity of the closed loop. Some examination is made of generalisations of the idea of incremental gain in order to make the analysis somewhat less conservative.

The remainder of the chapter consists of new results which accentuate the negative findings of the paper [9]. The nonlinearity has been restricted to be the standard saturation, and for simplicity, the feedback loop has been forced to consist of unity negative feedback. In this case, it has been found that it is possible to bound the incremental gain using a conceptually simple linear matrix inequality problem. As the development of linear matrix inequality software is an important research topic, it would appear that such analysis may be done neatly in a numerically efficient fashion.

A cursory examination of the algebraic problem indicates that the linear matrix inequality used to determine a bound for the incremental gain will be infeasible if the plant is open loop unstable. This is proven not to be an artefact of conservative analysis, as it is found to be possible to generate test signals which display explosive incremental gain properties. Even the use of bounded input sets is not enough to stop this behaviour, as the construction makes use of "small" signals. Also, the construction used would indicate that such drastic loss of continuity would be exhibited by less conservative generalisations of the incremental gain.

7.2 Results on the Continuity and Incremental Gain Properties of a Class of Linear Systems with Saturation

In the paper by Chitour, Liu and Sontag [9], the continuity and incremental gain properties for continuous time neutrally stable systems are investigated. Some of the results are constructive, others rely on using counter-examples to show that postulated properties do not hold in general. The results will now be surveyed, after some preliminary definitions are given.

Definition 7.2.1: A function $g\colon \mathbb{R}_+ \mapsto \mathbb{R}_+$ is a $\mathcal{K}$-function if it is continuous, strictly increasing, and satisfies $g(0) = 0$.

Definition 7.2.2: The operator $\mathbf{\Upsilon}$ satisfies the *generalised incremental gain property* with respect to $\mathcal{L}_{p+}$ (denoted $\mathbf{GIG}_p$) if there exists a $\mathcal{K}$-function g such that for all $u, v \in \mathcal{L}_{p+}$,

$$\|\mathbf{\Upsilon}v - \mathbf{\Upsilon}u\|_p \leq g(\|v - u\|_p).$$

Definition 7.2.3: The operator $\mathbf{\Upsilon}$ satisfies the *semi-global Lipschitz property over $\mathcal{L}_{p+}$* (denoted $\mathbf{SLP}_p$) if there exists a $\mathcal{K}$-function and a constant c such that for all $u, v \in \mathcal{L}_{p+}$,

$$\|\mathbf{\Upsilon}v - \mathbf{\Upsilon}u\|_p \leq (c + g(\|u\|_p))\,\|v - u\|_p\,.$$

Definition 7.2.4: The class $\mathcal{C}_{(0)}$ is the subset of saturation functions which are globally Lipschitz, differentiable at 0 and satisfy

$$\lim_{t,s\to 0, t\neq s} \frac{\sigma(t) - \sigma(s)}{t - s} = \sigma'(0).$$

Definition 7.2.5: The class $\mathcal{C}^{1,+}$ is the set of functions which are continuously differentiable, and satisfy the condition that $\sigma'(x) > 0$ for all x.

Theorem 7.2.6: Fix a (multiple input) linear system of the form:

$$\begin{aligned}
\dot{x}(t) &= Ax(t) + B\sigma(w(t) - Fx(t)), \\
z(t) &= x(t), \\
x(0) &= 0,
\end{aligned}$$

with A neutrally stable, and F the standard stabilising feedback for the passivity condition (see [34] for details, the function σ is the multiple input component-wise saturation function). The following statements describe the continuity properties of this class of systems with respect to $\mathcal{L}_{p+}$.

- If $p < \infty$, we then have that:

1. $\boldsymbol{\Upsilon}$ is continuous with respect to $\mathcal{L}_{p+}$, but does not necessarily satisfy $\mathbf{SLP}_p$.
2. If $\sigma \in \mathcal{C}_{(0)}$, then $\boldsymbol{\Upsilon}$ is $\mathbf{SLP}_p$.
3. $\boldsymbol{\Upsilon}$ is not in general $\mathbf{GIG}_p$.

- If $p = \infty$, we have that:

1. in general, $\boldsymbol{\Upsilon}$ is not continuous with respect to $\mathcal{L}_{\infty+}$, hence it is not $\mathbf{GIG}_p$.
2. If each component of σ is non-decreasing, and $n = 1$, $\boldsymbol{\Upsilon}$ is globally Lipschitz with respect to $\mathcal{L}_{\infty+}$.
3. If $\sigma \in \mathcal{C}^{1,+}$, $\boldsymbol{\Upsilon}$ is $\mathbf{SLP}_p$,

Proof: See [9]. □

The authors also deal with differentiability issues, the results on which are now given.

Theorem 7.2.7: Take the operator $\boldsymbol{\Upsilon}$ as in Theorem 7.2.6. Then,

1. If $\sigma \in \mathcal{C}^{1,+}$, $\boldsymbol{\Upsilon}$ is Fréchet-differentiable with respect to $\mathcal{L}_{\infty+}$.
2. If $1 \leq p < \infty$, $\sigma \in \mathcal{C}^1$ and globally Lipschitz, $\boldsymbol{\Upsilon}$ is Gâteaux-differentiable with respect to $\mathcal{L}_{p+}$.

Proof: See [9]. □

7.3 Bounding the Incremental Gain with the Standard Saturation

By fixing the saturation function to be the standard saturation, it develops that one can bound the incremental gain by the use of the solution of a linear matrix inequality (LMI) problem. One of the more interesting aspects of these results is that it appears that algebraic approaches to norm computation in the presence of this shift invariant sector bounded nonlinearity are in fact mainly useful in the determination of the incremental gain, which is an upper bound for the induced norm. Another nice feature of these results is that the characterisations for both continuous and discrete time are closely related in an intuitive fashion.

Unfortunately, there are presently no analytic characterisations of the conditions which guarantee the existence of a solution to the linear matrix inequalities. For the problem of bounding the incremental gain, we have been able to show sufficient conditions for the non-existence of a finite bound by the tactic of generating test signals which violate any such bound (Section 7.4). These sufficient conditions are quite strong, as one of the conditions is that the plant is open loop unstable. As the results proved in this chapter are mainly negative in nature, the problem configuration has been fixed to be a strictly proper plant along with unity negative feedback. Although the work could be readily generalised, it is not apparent that such an effort is immediately justified.

We will now fix the class of closed loop operators analysed throughout the remainder of this chapter. The system operator $\boldsymbol{\Upsilon}\colon w \mapsto z$ is generated by:

$$\Delta x(k) = Ax(k) + B\sigma_0(w(k) - Cx(k)), \tag{7.1}$$
$$z(k) = Cx(k), \tag{7.2}$$
$$x(0) = 0, \tag{7.3}$$

(i.e., unity negative feedback). Additionally, we assume that $[A, B, C]$ is minimal, with $A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times 1}, C \in \mathbb{R}^{1\times n}$.

7.3.1 Discrete Time Incremental Gain Bounds

It turns out that if dissipation style methods are used to compute an upper bound for the incremental gain of the system, the conditions appear to be the sort that arise from linear systems. As such, it is possible to then assume a quadratic form for the potential function, and use algebraic techniques to find a solution. We will now give the inequalities which arise in the theorem used to determine the bound.

Definition 7.3.1: The matrix P is the solution to the *discrete incremental gain algebraic inequalities* for a gain γ, if:

1. P is such that $P = P'$, and $P > 0$,
2. P is a solution to the matrix inequalities:

$$A'PA - P + C'C \leq 0, \tag{7.4}$$

$$\begin{bmatrix} A - BC & B \\ C & 0 \end{bmatrix}' \begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A - BC & B \\ C & 0 \end{bmatrix} - \begin{bmatrix} P & 0 \\ 0 & \gamma^2 I \end{bmatrix} < 0. \tag{7.5}$$

Theorem 7.3.2: For a system of the form given by (7.1) - (7.3), if a solution to the discrete incremental gain algebraic inequalities exists for a gain γ, then $\|\boldsymbol{\Upsilon}\|_{\Delta,2} \leq \gamma$.
Proof: Take any $w, \tilde{w} \in \ell_{p+}$, along with the corresponding state trajectories $x, \tilde{x}$. Let $\eta := x - \tilde{x}$. It is sufficient to prove that if $V(\xi) = \xi' P \xi$, we have:

$$s[k] := V(\eta[k+1]) - V(\eta[k]) - \gamma^2 \left\|(w - \tilde{w})[k]\right\|^2 + \left\|C\eta[k]\right\|^2 \leq 0, \tag{7.6}$$

for all (finite) k, as (7.6) implies $\gamma \left\|w - \tilde{w}\right\|_2 > \left\|C\eta\right\|_2$, by standard dissipation arguments.

By using the fact that we can write $\sigma_0(a) - \sigma_0(b) = \lambda(a - b)$, for some λ with $0 \leq \lambda \leq 1$, we arrive at:

$$\begin{aligned} \eta[k+1] &= A\eta[k] + B\left[\sigma_0(w[k] - Cx[k]) - \sigma_0(\tilde{w}[k] - C\tilde{x}[k])\right] \\ &= \lambda\left[(A - BC)\eta[k] + B(w - \tilde{w})[k]\right] + (1 - \lambda)A\eta[k] \\ &=: \lambda\zeta_1 + (1 - \lambda)\zeta_2. \end{aligned}$$

If we define $\kappa = \eta'[k]P\eta[k] + \gamma^2 \|(w - \tilde{w})[k]\|^2 - \|C\eta[k]\|^2$, it follows from Lemma 6.3.4 and (7.5) that $\zeta_1' P\zeta_1 \leq \kappa$. Similarly, by applying the Lyapunov inequality (7.4), we arrive at:

$$\zeta_2' P\zeta_2 = \eta'[k]A'PA\eta[k] \leq \eta'[k]P\eta[k] - \|C\eta[k]\|^2 \leq \kappa.$$

By realising that the set $\Omega = \{\xi : \xi' P\xi \leq \kappa\}$ is convex, $\zeta_1, \zeta_2 \in \Omega \implies \eta[k+1] \in \Omega$. Therefore, $\eta'[k+1]P\eta[k+1] \leq \kappa$, implying (from the definition of κ) that $s[k] \leq 0$, verifying the claim. □

Remark 7.3.3: It is conjectured that the existence of the solution to the incremental gain inequalities is a necessary and sufficient condition for γ to be an upper bound for for $\|\mathbf{\Upsilon}\|_{\Delta,2}$; we have just proven sufficiency. As for the necessity of the condition, it is conjectured that a valid dissipation function will exist if $\gamma > \|\mathbf{\Upsilon}\|_{\Delta,2}$, and that it is possible to prove that one need only consider quadratic functions. The importance of the non-existence result in Theorem 7.4.1 is that it gives an intuitive construction of signal pairs which are "destabilising" (incremental gain sense), and shows that the construction can be done with signals with small ∞-norm bounds. A non-constructive non-existence result may not tell us about the signal sizes necessary.

7.3.2 Continuous Time Incremental Gain

The extension of Theorem 7.3.2 to continuous time is straightforward, and will be done next.

Definition 7.3.4: The matrix P solves the *continuous time incremental gain inequalities* for a gain γ, if:

1. P is such that $P = P'$, $P > 0$, and
2. P is a solution of the matrix inequalities:

$$A'P + PA + C'C < 0,$$

$$(A - BC)'P + P(A - BC) + \frac{1}{\gamma^2}PBB'P + C'C < 0.$$

Theorem 7.3.5: For a continuous time system of the given form, if there exists a matrix P which is a solution to the continuous time incremental gain inequalities for a gain γ, then

$$\|\mathbf{\Upsilon}w - \mathbf{\Upsilon}\tilde{w}\|_2 < \gamma \|w - \tilde{w}\|_2,$$

for all $w, \tilde{w} \in \mathcal{L}_{2+}$.

Proof: Once again, take any two inputs $w, \tilde{w}$, and the corresponding state trajectories $x, \tilde{x}$. Let $\eta = x - \tilde{x}$. We fix $V(\xi) := \frac{1}{2}\xi' P\xi$, and define:

$$H := 2\left.\frac{dV}{dx}\right|_{\eta} (A\eta - B(\sigma_0(-Cx + w) - \sigma_0(-C\tilde{x} - \tilde{w})) - \gamma^2 \|w - \tilde{w}\|^2 + \eta' C'C\eta.$$

We then apply the same observation as in the discrete time case, and we are able to bound H as:

$$\begin{aligned} H \;\; &\leq \;\; \lambda\left(2\frac{dV}{dx}\{(A-BC)\eta + B(w-\tilde{w})\} - \gamma^2\|w-\tilde{w}\|^2 + \eta' C' C\eta\right) \\ &\quad +(1-\lambda)\left(2\frac{dV}{dx}A\eta + \eta' C' C\eta\right), \end{aligned}$$

with $0 \leq \lambda \leq 1$. This implies $H \leq 0$, as the two terms in brackets being less than or equal zero is a consequence of P being the solution to the continuous time incremental gain inequalities. □

For some basic examples of the application of Theorem 7.3.5, the reader is directed to the examples in Chapter 8. Of particular interest is the comparison of the bound generated here and the upper bound for the induced norm given by conservative analysis assumptions.

7.4 Non-existence of the Incremental Gain

Although it is convenient that one can bound the incremental gain using linear matrix inequalities, it appears that a solution will exist only under fairly strong stability conditions. We will now show that one can generate "destabilising" test signals with small ∞-norm bounds if the plant is open or closed loop unstable (although the result is obvious if the plant is closed loop unstable).

Additionally, the conditions for non-existence are further enlarged in Theorem 7.4.8. The test condition is that a combination of the open and closed loop dynamics matrices is unstable, but the test signals are no longer guaranteed to be small. The construction of test signals is done solely for discrete time systems for simplicity; it appears that there would be many minor details which would make the continuous time version quite tedious to develop. The extension of Theorem 7.4.1 to continuous time seems more straightforward than would be the case for Theorem 7.4.8. (The problems associated with the extension to continuous time would be issues such as the integrability of the test signals.)

As noted in Remark 7.3.3, it may be possible to show that the solution of the incremental gain inequalities is a necessary and sufficient condition for γ to be an upper bound for $\|\boldsymbol{\Upsilon}\|_{\Delta,2}$, by the use of generic existence arguments. The importance of the present results is that the constructions provide an intuitive feel as to the form of "destabilising test signals" (incremental gain sense), and that one can use small (∞-norm sense) signals to do so.

As to the intuition regarding the form of the constructed signals, the "in the large" signals are chosen so as to ensure the plant input is kept saturated, and then a small difference will follow the open loop dynamics. The main important detail in the construction is to ensure that the difference between the two states does not cause the two test signals to be in different operating regions of the saturation.

7.4.1 Non-existence of the Incremental Gain: Unstable Plant

We will now give a theorem stating conditions under which the incremental gain does not exist, in the case of an open or closed loop unstable plant. In order for the result to be non-trivial, it is necessary to examine the case of a bounded input set over which the induced norm in fact exists. For simplicity, the bound is given in terms of the ℓ_∞ norm of the signal.

Theorem 7.4.1: Assume that $w \in \mathcal{W}_{p,\infty}(N)$, with $N > 1 + \epsilon$ $(0 < \epsilon < 1)$. Then for the system described by (7.1) - (7.3), $\|\mathbf{\Upsilon}\|_{\Delta,p}$ does not exist over $\mathcal{W}_{p,\infty}(N)$ if:

1. $p < \infty$, and A or $A - BC$ has any eigenvalues of modulus greater than or equal to 1;
2. $p = \infty$ and A or $A - BC$ has any eigenvalues of modulus strictly greater than 1.

Proof: It is trivial that if the induced norm does not exist, the incremental gain is also unbounded. This will occur if the closed loop dynamics matrix $A - BC$ is unstable, or if A is unstable and N is chosen large enough. We now examine the case where $A - BC$ is stable, and the induced norm over $\mathcal{W}_{p,\infty}(N)$ is finite.

We can now make the following observations, which are standard consequences of linear system properties.

Proposition 7.4.2: Since $[A, B]$ is controllable, for all $y_0 \in \mathbb{R}^n$, there exists a $u \in \ell_{p+}$ such that,

$$\sum_{i=0}^{n-1}(A - BC)^{(n-1)-i}Bu[i] = y_0,$$

with $u[k] = 0$ for all $k \geq n$.

Proposition 7.4.3: If $p < \infty$, and A is such that it has any eigenvalues of modulus greater than or equal one, the observability of $[A, C]$ implies that there exists some $y_0 \in \mathbb{R}^n$ such that the sequence

$$(S_{M,p})^p := \sum_{i=0}^{M} \left\|CA^i y_0\right\|^p,$$

is unbounded. If $p = \infty$, and A has any eigenvalues with modulus strictly greater than 1, there exists some $y_0 \in \mathbb{R}^n$ such that the sequence $\{S_{M,\infty}\}$ is unbounded, with $S_{M,\infty}$ defined:

$$S_{M,\infty} := \max_{i \in 0,1,\ldots,M} \left\{\left\|CA^i y_0\right\|\right\}.$$

To prove the theorem, we assume the converse; i.e. that there exists some γ such that

$$\|\mathbf{\Upsilon}w - \mathbf{\Upsilon}\tilde{w}\|_p \leq \gamma \|w - \tilde{w}\|_p,$$

for all $w, \tilde{w} \in \mathcal{W}_{p,\infty}(N)$. Fix some y_0 as in Proposition 7.4.3, and u as the appropriate input as in Proposition 7.4.2. We then see that there exists some $M < \infty$ such that

$$(S_{M,p})^p > \gamma^p \|u\|_p^p, \tag{7.7}$$

(for $p < \infty$), or

$$S_{M,\infty} > \gamma \|u\|_\infty. \tag{7.8}$$

Then choose a constant $\alpha > 0$ such that

$$\begin{aligned}\alpha^{-1} &> \max\left\{\max_{i=0,\ldots,n-1}\left\{\left\|u[i] + \sum_{j=1}^{i} C(A-BC)^{j-i}Bu[j-i]\right\|\right\},\right.\\ &\left.\frac{2}{\epsilon}\max_{i\in 0,\ldots,M-1}\left\{\left\|CA^i y_0\right\|\right\}\right\}.\end{aligned}$$

Now define the signal w, which is causally dependent upon the induced state x:

$$w[k] = \begin{cases} 0 & 0 \le k \le n-1, \\ (1+\epsilon)\text{sign}(-Cx[k]) + Cx[k]) & n \le k \le n+M-1, \text{and } \|Cx[k]\| \le 1+\epsilon, \\ 0 & n \le k \le n+M-1, \text{and } \|Cx[k]\| > 1+\epsilon, \\ 0 & n+M \le k, \end{cases}$$

remembering that $\text{sign}(0) = 1$. Then, $\tilde{w}$ is defined as:

$$\tilde{w}[k] = \begin{cases} \alpha u[k] & 0 \le k \le n-1, \\ w[k] & k \ge n. \end{cases}$$

We can readily verify that $w, \tilde{w}$ lie in $\mathcal{W}_{p,\infty}(N)$.

We first note that by their construction, $\|w - \tilde{w}\|_p = \alpha \|u\|_p$. Define $\eta = \tilde{x} - x$. We will now be able to characterise the behaviour η.

Proposition 7.4.4: The state difference is described by:

$$\eta[k] = \alpha A^{k-n} y_0,$$

for k in the interval $n, \ldots, n+M$.

Proof: This will be shown by induction. By the construction of u, α, we can show that $\eta[n] = \alpha y_0$ (by a similar inductive argument, if necessary). Then, by its definition, $\|w[k] - Cx[k]\| \ge 1+\epsilon$. But then, $\tilde{w}[k] - C\tilde{x}[k] = w[k] - C\eta[k] - Cx[k]$ by the definitions of $\tilde{w}, \eta$. Then, using the inductive assumption, $\eta[k] = \alpha A^{k-n} y_0$, and so by the definition of α, $\|C\eta[k]\| < \frac{\epsilon}{2}$. This ensures that $\sigma_0(w[k] - Cx[k]) = \sigma_0(\tilde{w}[k] - C\tilde{x}[k])$, and so examining the state space equations, $\eta[k+1] = A\eta[k] = \alpha A^{K+1-n} y_0$. This verifies the inductive assumption, proving Proposition 7.4.4. $\square$

By applying Proposition 7.4.4, and the definition of $S_{M,p}$, we then have that:

$$\|\mathbf{P}_{n+M}(\mathbf{I} - \mathbf{P}_{n-1})C\eta\|_p = \alpha S_{M,p}.$$

By the use of (7.7) (resp. (7.8)),

$$\frac{\|\mathbf{\Upsilon}\tilde{w} - \mathbf{\Upsilon}w\|_p}{\|\tilde{w} - w\|_p} \ge \frac{\alpha S_{M,p}}{\alpha \|u\|_p} > \gamma.$$

This contradiction proves Theorem 7.4.1.$\square$

Remark 7.4.5: A similar construction is made in [9, Example 5] for a first order continuous time system, but the test signals can have an arbitrarily large $\mathcal{L}_{\infty+}$ norm.

Remark 7.4.6: We have not specifically dealt with the case where $p = \infty$, and $\rho(A) = 1$. In this case, if A has an eigenvalue of modulus 1 which corresponds to a non-trivial Jordan block, the instability proof may be trivially extended to hold for this polynomially unstable mode, at the cost of some notational complexity.

It should be noted that if we chose n_∞ sufficiently small in Theorem 7.4.1, we can force $\|\mathbf{\Upsilon}\|_{\Delta,2} = \|\mathbf{\Upsilon}_L\|_{\Delta,2}$ by ensuring that the internal signal $w - Cx$ has an infinity norm less than 1, and so the nonlinear operator matches the linearisation at the origin. Keeping this in mind, the incremental gain may exhibit interesting behaviour as it goes from this finite value to infinity (by the time $n_\infty > 1$, at least), which will not be examined herein.

7.4.2 Non-Existence of Incremental Gain: Stable Plant

The previously given results for discrete time systems are not completely satisfying in that not all cases are covered. If the plant is open or closed loop unstable, solutions to the incremental gain algebraic inequalities obviously cannot exist. However, if the plant is such that it is open and closed loop stable, there are no simple characterisations as to whether the inequalities are feasible. What will now be given is a test condition which shows non-existence of a solution; it may in fact turn out to be a necessary and sufficient condition for the existence of the incremental gain, but whether it is such is presently unknown to the author. Although it should be relatively straightforward to extend the construction in Theorem 7.4.1 to continuous time, the construction used here would appear to be more difficult to emulate.

Definition 7.4.7: A binary sequence is a sequence whose elements each lie in the set $\{0, 1\}$.

Theorem 7.4.8: Fix a system $\mathbf{\Upsilon}$ of the form given by (7.1) - (7.3), $\|\mathbf{\Upsilon}\|_{\Delta,p}$ does not exist over ℓ_{p+} if there exists a binary sequence of length l denoted $\{\alpha_1, \ldots, \alpha_l\}$ such that

$$P^l_\beta = \prod_{j=1}^{l} (\alpha_{l-j+1} A + (1 - \alpha_{l-j+1})(A - BC)),$$

has an eigenvalue of modulus greater than or equal one ($p < \infty$) or strictly greater than 1 ($p = \infty$).

Proof: This proof is to be found in Appendix B. $\Box$

Remark 7.4.9: Once again, although the proof may be further extended for the case $p = \infty$, and $\rho(P^l_\beta) = 1$ (with an eigenvalue of modulus 1 corresponding to a non-trivial Jordan block), this is not done here for notational convenience.

7.5 Use of Other Disturbance Sets

The topic of the use of generalisations of the incremental gain, or the use of other ℓ_{p+} bounds will now be discussed. It is conjectured that the use of other ℓ_{p+} ($p < \infty$) disturbance bounds would give rise to finite incremental gain bounds, but they would be quite conservative in robust control analysis when compared to the bounds provided by the induced norm.

With the standard saturation present, examination of the test signals in Theorem 7.4.1 shows that they can be made arbitrarily close to each other, which in fact is preferred for analysis. The differences will be within a scaling factor of each other if the difference is small enough, and so the use of a $\mathbf{GIG}_p$ generalisation would not help. The use of an $\mathbf{SLP}_p$ condition, in which the magnitude of the test signals is taken into account, would work out to be similar in practice to the use of a $\mathcal{W}_{2,2}(N)$ set over which the induced norm is finite.

It is conjectured that in the case where the induced $\mathcal{L}_{2+}$ (ℓ_{2+}) norm is finite over a $\mathcal{W}_{2,q}(N)$ set ($q < \infty$), then the incremental gain would also be finite over $\mathcal{W}_{2,q}(N)$. It should be possible to conservatively estimate the period of time the saturation element will in fact saturate, as all signals will lie in $\mathcal{L}_{2+}$ and be bounded. Then, the difference signal for signals lying in some ϵ ball surrounding a test signal can be conservatively overbounded for a finite time, and then asymptotic stability of the closed loop relied upon. The analysis is then easily extended for large difference signals by repeated application of the triangle inequality, using steps of length ϵ.

After some preliminary examination, this work was not fully developed, as it appears that analysis will give extremely conservative results when applied to unstable systems (when compared to the induced norm bound). The technique may be even more conservative in that it would appear to be quite difficult to estimate the time period during which the nonlinear element will be in the saturated mode.

Chapter 8

Analytical Examples

8.1 Introduction

We have heretofore deferred giving examples of the use of the analysis techniques developed in this dissertation. Rather, the results were stated and proven, with some explanatory remarks. The preference was to bundle the various approaches together when examining the properties of a class of systems. In this way, it is hoped that a better intuition can be developed with regards to the operator theoretic properties of linear systems with saturation.

The main part of this chapter is devoted to the examination of the first order (continuous time) plant following the standard saturation, with proportional feedback. This is the simplest possible nontrivial system, but it is believed to display many of the important theoretical peculiarities of this class of nonlinear systems.

The bounds which are derived as a result of the analysis developed in this dissertation are compared to those provided by other methods, and the different behaviour of the induced norm as compared to the incremental gain is made obvious. Additionally, this analysis provides a most impressive justification of the use of bounded input sets, as ill posed behaviour is immediately obvious. Without their use, or a similar remedy, either the analysis will be quite conservative or will lose its initial engineering justification.

An example of discrete time induced ℓ_{2+} norm analysis is given in Section 8.5. Pen and paper analysis of such systems is more difficult as evaluating the derivative of the candidate function is not sufficient, and so examples would be much better dealt with using numerical methods. However, it is interesting that the bound derived is the same as that for the corresponding continuous time system.

The remainder of the chapter gives some additional analysis on related points. The norm computation of a parallel projection operator is given in Section 8.3. In Section 8.4, an interesting example is surveyed, and some comment made on how it illustrates the importance of the reachable set size in the induced norm computation problem. Finally, in Section 8.6, some preliminary comments are made on the application of the results developed in this dissertation to the state feedback synthesis problem. The approach to a simple example is detailed, although it is believed that the general synthesis problem will require an approach far more tailored to linear systems with saturation.

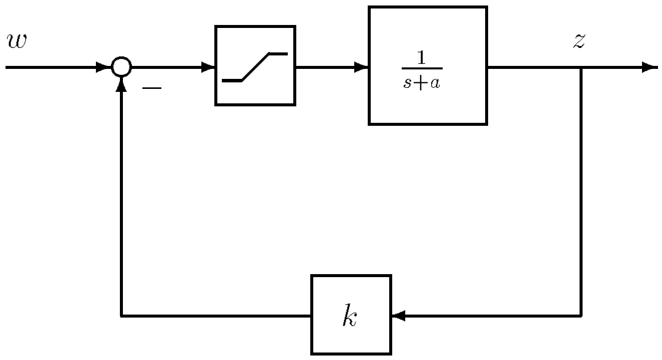

Figure 8.1: Example System

8.2 The First Order Continuous Time Plant

We will now examine the simplest possible non-trivial case, being the first order input saturated plant with proportional feedback. Despite its simplicity, this class of systems exhibits many of the interesting properties which distinguish these nonlinear systems from linear ones.

The system Υ is realised by:

$$\begin{aligned} \dot{x} &= -ax(t) + \sigma_0(-kx(t) + w(t)), \\ z(t) &= x(t), \\ x(0) &= 0, \end{aligned}$$

with $a, k \in \mathbb{R}$. We will now examine various interesting cases which are determined by the values of a, k.

8.2.1 Case $a > 0$, $k < -a$

This is a somewhat unusual case, as the choice of feedback destabilises the stable linear plant. We can easily verify that the induced $\mathcal{L}_{p+}$ norm does not exist for all p (take as a test input $w(t) = \epsilon$ for $0 \leq t \leq \epsilon$, $w(t) = 0$ otherwise). However, it is also simple to verify that any trajectory is confined to $\Omega = (-\frac{1}{a}, \frac{1}{a})$, and so the state is confined to a compact set for all choices of input, and could even be considered to be stable under a loose enough definition of state space stability.

The importance (albeit limited) of this example is that it is possibly one of the simplest which shows that global state space properties are not necessarily determined by the behaviour at the origin, as is the case for linear systems. Additionally, it underlines the fact that the behaviour of the closed loop linear system provides a lower bound for operator gain properties.

8.2.2 Case $a > 0$, $k = 1$

We will now fix $k = 1$, and then examine the stability properties as the plant pole is shifted along the real axis. Examination of this case will provide a basic motivation as to the

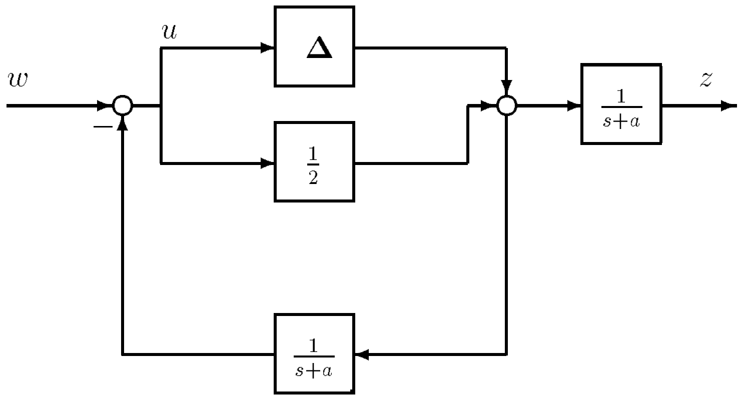

Figure 8.2: The System after Loopshifting

importance of the use of bounded input sets. Additionally, we can compare the results of the analysis herein to that provided by alternative means.

Once again, as $a > 0$, we have that the state is confined to $\Omega = (-\frac{1}{a}, \frac{1}{a})$. The system is locally stable, and the local linearisation has the transfer function $G(s) = \frac{1}{s+a+1}$, which is in $\mathcal{RH}_\infty$, with $\|G(s)\|_\infty = \frac{1}{a+1}$.

We will now examine the different bounds for the operator gain given by the use of the Small Gain Theorem (Circle Criterion), incremental gain analysis, and the result provided by the computational heuristic.

Small Gain Estimate

The first basic approach is to use the estimate provided by the small gain theorem, equivalent to the Circle Criterion. The means by which the estimate is made is via the use of *loop shifting*, a tactic long used in this situation (for example, see [55, Section 6.2]). Loop transformations are the usual means of proving the Circle Criterion, and so the details are of enough interest to be given here.

The first step is to realise that we can express the function σ_0 as

$$\sigma_0(x) = \frac{1}{2}x + \Delta(x),$$

with $\Delta(x)$ given by:

$$\Delta(x) := \begin{cases} \frac{1}{2}x & |x| \leq 1, \\ \text{sign}(x) - \frac{1}{2}x & |x| > 1. \end{cases}$$

The importance of this is that $\frac{1}{2}x$ is the centre of the sector $[0,1]$, and Δ has the minimal norm to cover the sector.

Now examine Figure 8.2. For ease of notation, let the linear operator with transfer function $\frac{1}{s+a}$ be denoted as $\mathbf{G}$, and let $\mathbf{H} = \frac{\mathbf{G}}{\mathbf{I}+\frac{1}{2}\mathbf{G}}$. We then can see that

$$u = -\mathbf{G}\mathbf{\Delta}u - \frac{1}{2}\mathbf{G}u + w.$$

Then,

$$u = (\mathbf{I} + \mathbf{H}\boldsymbol{\Delta})^{-1}(\mathbf{I} + \frac{1}{2}\mathbf{G})^{-1} w,$$

and so,

$$z = \mathbf{G}\boldsymbol{\Delta}(\mathbf{I} + \mathbf{H}\boldsymbol{\Delta})^{-1}(\mathbf{I} + \frac{1}{2}\mathbf{G})^{-1} w.$$

The incremental gain can thus be bounded,

$$\|\boldsymbol{\Upsilon}\|_\Delta \leq \|\mathbf{G}\|_\Delta \|\boldsymbol{\Delta}\|_\Delta \left(\frac{1}{1 - \|\mathbf{H}\|_\Delta \|\boldsymbol{\Delta}\|_\Delta}\right) \left\|(\mathbf{I} + \frac{1}{2}\mathbf{G})^{-1}\right\|_\Delta. \tag{8.1}$$

For the signal space $\mathcal{L}_{2+}$, we arrive at

$$\|\boldsymbol{\Upsilon}\|_{\Delta,2} \leq \frac{1}{2a}\left(1 + \frac{1}{2a}\right). \tag{8.2}$$

(The most difficult part of the transition from from the inequality (8.1) to (8.2) is to verify that

$$\left\|\left(1 + \frac{1}{2}G(s)\right)^{-1}\right\|_\infty = \left\|\frac{s+a}{s+a+\frac{1}{2}}\right\|_\infty = 1,$$

which is true for all $a > -\frac{1}{4}$. This can be verified by the use of the Riccati equation (4.6), with $\gamma^2 = 1 + \epsilon$, for $\epsilon > 0$.)

This analysis agrees with the expected result from graphical analysis that the gain is finite if the system is stable, but becomes unbounded as $a \to 0$.

Incremental Gain Analysis

We would like to now see what the result is if we apply the algebraic test to the system given in Section 7.3.2. The continuous incremental gain algebraic inequalities are particularly easy to solve as they are given in terms of a scalar variable. The inequalities which result are:

$$\begin{aligned} -2ap + 1 &\leq 0, & (8.3)\\ -2(a+1)p + \frac{1}{\gamma^2}p^2 + 1 &\leq 0, & (8.4)\\ p &> 0. & (8.5) \end{aligned}$$

It is an obvious consequence of linear system analysis that the minimal value of γ for which there exists a solution to the inequalities (8.4) - (8.5) is

$$\gamma = \frac{1}{a+1}.$$

The solution p^* which determines this value of γ is given by:

$$p^* = \frac{1}{a+1}.$$

However, from (8.3), it is necessary that $p > \frac{1}{2a}$. Hence, if $a \geq 1$ we can use this value of p^* for solution to (8.3) - (8.4). This implies that the nonlinear system has the same incremental gain as the local linearisation. If $a < 1$, the minimising choice of p is to take $p = \frac{1}{2a}$. In this case, substitution into (8.4) yields:

$$\gamma^2 \geq \frac{1}{4a}, \tag{8.6}$$

which gives an upper bound of $\frac{1}{2\sqrt{a}}$ for γ^*. Once again, the incremental gain becomes unbounded as $a \to 0$, but the bound given for non-zero a is tighter.

Induced $\mathcal{L}_{2+}$ Norm Bounding

We will now apply the heuristic computational method given in Section 5.6 to the class of systems in question. The analysis here gives non-surprisingly the least conservative results for bounding the induced norm. When the analysis is extended to the case where the plant is unstable, the importance of doing such a direct analysis will be underlined.

We will first need the corresponding linear solution. The algebraic Riccati equation which results from (5.15) is:

$$2p(-a-1) + \frac{1}{\gamma_l^2}p^2 + 1 = 0.$$

We have that $\gamma_l^* = \frac{1}{a+1}$, with a corresponding $p = \frac{1}{a+1}$. The actual dissipation function V_l is given by:

$$V_l(\xi) = \frac{1}{2}\frac{1}{a+1}\xi^2.$$

It is then easy to verify that the state is confined to the set $\Omega = (-\frac{1}{a}, \frac{1}{a})$ for all finite time. We now need to see over what subset of Ω the solution $V = V_l$ is a valid candidate function. Referring to Theorem 5.4.1 for the definitions of Φ_A, etc., it is obvious that V_l satisfies $\Phi_A(\xi) \leq 0$ for ξ; we need only verify the region where $\Phi_B(\xi) \leq 0$.

We have that

$$\begin{aligned}\Phi_B(\xi) &= 2\left.\frac{dV_l}{dx}\right|_\xi (-a\xi - \operatorname{sign}(\xi)) + \xi^2, \\ &= -\frac{2}{a+1}(a\xi^2 + |\xi|) + \xi^2.\end{aligned}$$

And so, $\Phi_B(\xi) \leq 0$ if $(1-a)|\xi| - 2 \leq 0$, or $|\xi| \leq \frac{2}{1-a}$. We then have verified that $V = V_l$ is sufficient if $\Omega \subseteq (\frac{-2}{1-a}, \frac{2}{1-a})$, which occurs if $a \geq \frac{1}{3}$. For such cases then, $\gamma^* = \gamma_l^* = \frac{1}{a+1}$. Considering the result of the incremental gain analysis, it is obvious that the norm would not be affected for $a > 1$.

For $a < \frac{1}{3}$, we take $V(\xi)$ to be such that $\Phi_B(\xi) = 0$ for $|\xi| > \frac{2}{1-a}$. Then,

$$\left.\frac{dV}{dx}\right|_\xi = \frac{\xi^2}{2(a|\xi|+1)}. \tag{8.7}$$

As an aside, the anti-derivative of (8.7) is

$$V(\xi) = \frac{1}{2a^3}\left\{\frac{1}{2}(a|\xi|+1)^2 - 2(a|\xi|+1) + \ln(a|\xi|+1)\right\} + \kappa,$$

(although the actual analytic form of $V(\xi)$ is somewhat unimportant - its derivative is the important quantity, and is of a simple form).

As $\Phi_B(\xi) \leq 0$ is automatically satisfied by the construction of V, the Φ_A expression needs to be examined to determine the value of γ. The functions of interest are given by:

$$\begin{aligned}
\theta_a(\xi) &= -\xi + \frac{1}{\gamma^2}\left.\frac{dV}{dx}\right|_{\xi}, \\
&= -\xi + \frac{1}{\gamma^2}\frac{\xi^2}{2(a|\xi|+1)}, \\
\Phi_{A.1}(\xi) &= 2\left.\frac{dV}{dx}\right|_{\xi}(-a\xi + \theta_a(\xi)) - \gamma^2\left\|\frac{1}{\gamma^2}\left.\frac{dV}{dx}\right|_{\xi}\right\|^2 + \xi^2 \\
&= \frac{\xi^2}{a|\xi|+1} - (a+1)|\xi| + \frac{1}{4\gamma^2}\frac{\xi^4}{(a|\xi|+1)^2} + \xi^2, \\
\Phi_{A.2}(\xi) &= 2\left.\frac{dV}{dx}\right|_{\xi}(-a\xi + \mathrm{sign}(\theta_a(\xi))) - \gamma^2\left\|\xi + \mathrm{sign}(\theta_a(\xi))\right\|^2 + \xi^2.
\end{aligned}$$

As $\Phi_{A.1}(\xi) \geq \Phi_{A.2}(\xi)$, we will examine it first. For the expression to be negative (semi)-definite, we need that

$$\gamma^2 \geq \frac{1}{4}\frac{\xi^2}{a\xi^2 + (1-a)|\xi| - 1}.$$

If we express the bounding expression as a function of ξ, basic calculus indicates that the minimum choice occurs at $\xi = \frac{1}{1-a}$, and then monotonically increases. When matched at the beginning of the interval of interest (i.e., $\xi = \frac{2}{1-a}$), the non-surprising result is that a choice of $\gamma = \frac{1}{1+a}$ is indicated.

An interesting case is where we solve the equations for all of $\mathbb{R}^1$, we see that $\lim_{\xi\to\infty}\gamma(\xi)$ is $\gamma^2 = \frac{1}{4a}$. This is precisely the bound determined by the inequality (8.6)! This is not a mere coincidence; this is the behaviour that should be expected in this case. The $\Phi_{A.1}$ equation corresponds directly to the algebraic Riccati equation, and the Φ_B expression to the algebraic Lyapunov equation. Their simultaneous solution corresponds to the continuous incremental gain inequalities.

Reflection upon this example gives the reasons why the more detailed analysis in Chapter 5 is necessary to give less conservative results. Approaches which rely on quadratic properties will correspond to the $\Phi_{A.1}$ expression, and neglect the fact that we can use the less strict $\Phi_{A.2}$ expression if θ_a is larger than 1, which is extremely likely to occur in the large. The other cause for conservatism which can be eliminated is that the limiting properties are used to determine the norm, which neglects the fact that the state will be confined to a bounded set for any fixed initial condition.

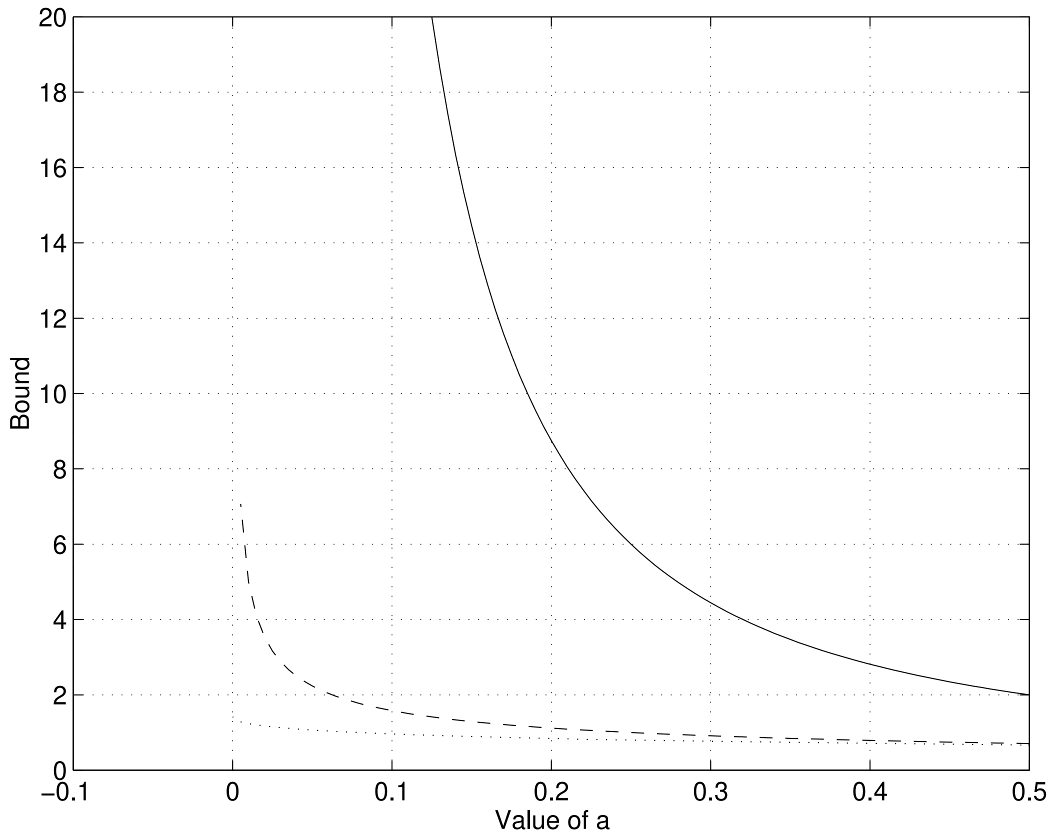

Figure 8.3: Bounds for γ^* as a function of a determined by (from top): Small Gain, incremental gain, induced $\mathcal{L}_{2+}$ norm analysis.

When the full analysis is undertaken, we find that γ is determined by the $\Phi_{A.2}$ expression, and by the fact that $\Omega = (-\frac{1}{a}, \frac{1}{a})$. The full analysis is rather tedious, and in fact will not be given here explicitly. The important point of interest is in fact that $\lim_{a\to 0}\gamma(a) = \sqrt{2}$, which is obviously much better behaved than the prediction given by the small gain or incremental gain analysis, and the limiting behaviour of $\Phi_{A.1}$. The bound determined as a function of a was numerically computed, and is plotted in Figure 8.3, along with the bounds given by the small gain and incremental gain analysis.

8.2.3 Case $a = 0$, $k = 1$

The case where $a = 0$ is of particular interest, being the case of an integrator following the standard saturation with unity negative feedback.

Incremental Gain

In the paper by Chitour, Liu, and Sontag, there is a proof that the incremental gain does not exist for $\mathcal{L}_{2+}$ spaces for this system ([9, Example 5]). However, the construction therein relies on a parameterised pair of signals which become arbitrarily large in infinity norm as the differential gain becomes infinite. It is possible to construct signal pairs which have much smaller infinity norms which display unbounded incremental gain behaviour. The following construction is a continuous time generalisation of the generic construction developed for discrete time systems.

Let $M > 1$ be a free parameter, and define the signals $w(t), \tilde{w}(t)$ for $t \in [0, M]$ as follows:

$$w(t) = \begin{cases} 1 + t - 2T & 2T \le t < 2T + 1, \\ -.1 - t + 2T & 2T + 1 \le t < 2T + 2, \end{cases}$$

$$\tilde{w}(t) = \begin{cases} 0 & 0 \le t < .1, \\ t - .1 & .1 \le t < 1, \\ w(t) & \text{otherwise}, \end{cases}$$

with $T \in \{0, 1, 2, \ldots\}$, and with $w(t) = \tilde{w}(t) = 0$ for $t > M$. We have that $\|w\|_\infty = \|\tilde{w}\|_\infty = 2$. To prove that $\|\Upsilon\|_{\Delta,2}$ does not exist, assume the converse, i.e., that there exists some finite γ such that $\frac{\|\Upsilon w - \Upsilon \tilde{w}\|_2}{\|w - \tilde{w}\|_2} < \gamma$. It can be verified that $\|w - \tilde{w}\|_2^2 = \frac{358}{3000}$, and that $\|\Upsilon w - \Upsilon \tilde{w}\|_2^2 > \frac{M-1}{100}$. Thus, by choosing M such that

$$M \ge \frac{358}{30}\gamma^2 + 1,$$

we contradict the assumption that incremental gain is bounded by γ.

Induced $\mathcal{L}_{2+}$ Norm

Despite the fact that the incremental gain is unbounded, the induced $\mathcal{L}_{2+}$ norm is finite. If the input set is taken to be $\mathcal{L}_{2+}$, we see that the reachable set $\Omega = \mathbb{R}^1$, and that $\gamma = \sqrt{2}$. A candidate dissipation function which determines this value is given by:

$$V(\xi) = \begin{cases} \frac{1}{2}\xi^2 & |\xi| \le 2, \\ \frac{1}{6}\xi^3 + \frac{2}{3} & \text{otherwise}. \end{cases}$$

Once again, the full analysis will be omitted, but the norm bound results by ‘ taking the limiting value of $\Phi_{A.2}$.

It is interesting to note that $\gamma = \sqrt{2}$ is a tight bound; this can be shown by examining the limiting properties of the input signal w_T given by

$$w_T(t) = \begin{cases} t + 1 & 0 \le t \le T, \\ 0 & t > T, \end{cases}$$

as $T \to \infty$. (The response is given by:

$$x(t) = \begin{cases} t & 0 \le t \le T, \\ 2T - t & T \le t \le 2T - 1, \\ e^{-t+2T-1} & t \ge 2T - 1.) \end{cases}$$

In this case, we are able to compare this norm bound with that given in the paper by Liu, Chitour and Sontag [34]. The bound given by their analysis is that $\|\Upsilon\|_{i,2} \le 2(1+1)\|\Upsilon_L\|_\infty = 4$. The analysis in [34] holds of course for general nonlinearities, and so it is not surprising that some deterioration would occur in the bound.

8.2.4 Case $a < 0$, $k = 1$

When the plant is open loop unstable, all signal bounds become undefined, at least over unbounded sets. Once again, the incremental gain displays explosive behaviour, even within bounded input sets. The test signal w given by:

$$w(t) = \begin{cases} 1 + (a+1)t & 0 \leq t \leq -\frac{1}{a} + 1, \\ 0 & t > -\frac{1}{a} + 1, \end{cases}$$

lies in all $\mathcal{L}_{p+}$ spaces, yet gives rise to an output which does not lie in any $\mathcal{L}_{p+}$ space. Therefore, the failure of all analysis techniques is not conservative in this case.

8.2.5 Analysis with Bounded Inputs

The conclusions in Section 8.2.4 may seem somewhat unusual from point of view of robust control theory. When operator theoretic analysis is applied to feedback loops, one expects results to conform to the so-called Meta-Theorem of robust control (stated next).

Theorem 8.2.1 (Meta-Theorem of Robust Control): If a (closed loop) control system is stable (in some input/output sense), then the system will remain stable (in the same sense) for all small (in some sense) perturbations to the system model.

For linear systems, the Meta-Theorem underlies $\mathcal{H}_\infty$ control (for example), as there is a characterisation of classes of perturbations. Such reasoning is exemplified in the Gap Metric approaches to robust control. Much of the underlying belief in the validity of the Meta-Theorem comes from the stability implied by the Small Gain Theorem. This belief gave the author the initial justification for the examination of the gain properties of these system operators.

If we confine ourselves to the use of incremental gain, problems do not occur. The incremental gain increases in a continuous fashion, becoming unbounded as a approaches zero. In fact, we can use the stronger version of the Small Gain Theorem, and so we could develop quite a strong theory. And if the extension of Theorem 7.4.1 to continuous time can be made, the results may not be conservative, even with signals of reasonable size. The only major source of conservatism introduced may be due to that associated with the Small Gain Theorem.

The restriction of theory to the study of the incremental gain does not seem a completely satisfactory option, as this restriction would *a priori* exclude plants with unstable modes which are to be controlled by bounded inputs. From an engineering standpoint, this might seem to largely confine its usefulness to be an academic exercise. This is not to say that there could be no useful results which could be developed, as in areas such as process control, plants may be more likely to be stable. In such a case, the fact that the computation may be done as the solution of linear matrix inequalities may allow easily understood and implemented algorithms to be developed. Although conservative, incremental gain analysis may still be sufficient for many purposes.

However, the preference is to use the induced norm of the systems, as it will give rise to less conservative bounds. The uniqueness/existence difficulties associated with the induced norm version of the Small Gain theorem are likely to be unimportant in non-pathological classes of nonlinear systems, allowing the development of the corresponding robust control theory.

The fact that the induced norm bound behaves in a sensible fashion for $a > 0$ seems encouraging. But the the model seems to display remarkable lack of robustness with $a = 0$. This could be predicted by the explosion in the incremental gain as a goes to zero, but this seems to be a very unsatisfactory situation. The Meta-Theorem seems to have a a counter-example to it, as it seems unreasonable that replacing $\frac{1}{s}$ by $\frac{1}{s-\epsilon}$ is a dramatic perturbation.

We therefore need to re-examine the situation. It would appear something would have to be eliminated: the Meta-Theorem, the class of systems or perturbations, or the theoretical framework. The first two options are unsatisfying, so a need to re-examine the operator theoretic based definitions of stability used. Examination of the destabilising signals indicates that they may need to be pathologically large in order to destabilise the system. Extremely large excursions in the state space are necessary if the plant is "nearly stable". As it is expected that our models/systems would fail long before such excursions became possible, such test inputs seem suspect in terms of an engineering intuition.

The problem is eliminated if we confine the state to bounded subsets of the state space. In order for this to make sense from an operator theoretic standpoint (and to put the means of restriction into a realisation independent form), $\mathcal{W}_{p,q}(N)$ sets are introduced. And, in fact the pathological behaviour exhibited for the induced norm disappears when we examine the bound given in terms of a.

If we take our input set to be $\mathcal{W}_{2,\infty}(N_\infty)$, it is readily verified that the reachable set $\Omega = (-\frac{N_\infty}{a+1}, \frac{N_\infty}{a+1})$, if $\frac{N_\infty}{a+1} < \frac{1}{|a|}$. The basic existence result given by Theorem 5.6.8 then indicates a finite norm bound exists. The analysis was made on a digital computer, and the bound dependence vs. a for various values of N_∞ are given in Figure 8.4. The behaviour of the bound is now sensible, as although there is a value of a at which it becomes unbounded, the failure occurs in a continuous fashion.

8.3 The Parallel Projection Operator

We will now examine the case of the calculation of the norm of the parallel projection operator for the case of an integrator with a saturation element, and with unity negative feedback. The paper by Doyle, Georgiou, and Smith ([13]) defines and discusses the importance of such operators in robust control theory. The parallel projection operator $\mathbf{\Pi_1}$ is defined as the mapping

$$\mathbf{\Pi_1} \begin{bmatrix} d_1 \\ d_2 \end{bmatrix} = \begin{bmatrix} u_1 \\ y_1 \end{bmatrix},$$

for the signals as defined in Figure 8.5. The usefulness of this example is that it ties in with other work in robust control theory, and illustrates the use of the induced norm test when the D matrices are non-zero.

Although the operator has two inputs, we define $w = d_1 - d_2$, and the readily checked claim is that if γ^* is the norm for the thus defined single input system, the parallel projection operator itself has a norm of $\sqrt{2}\gamma^*$.

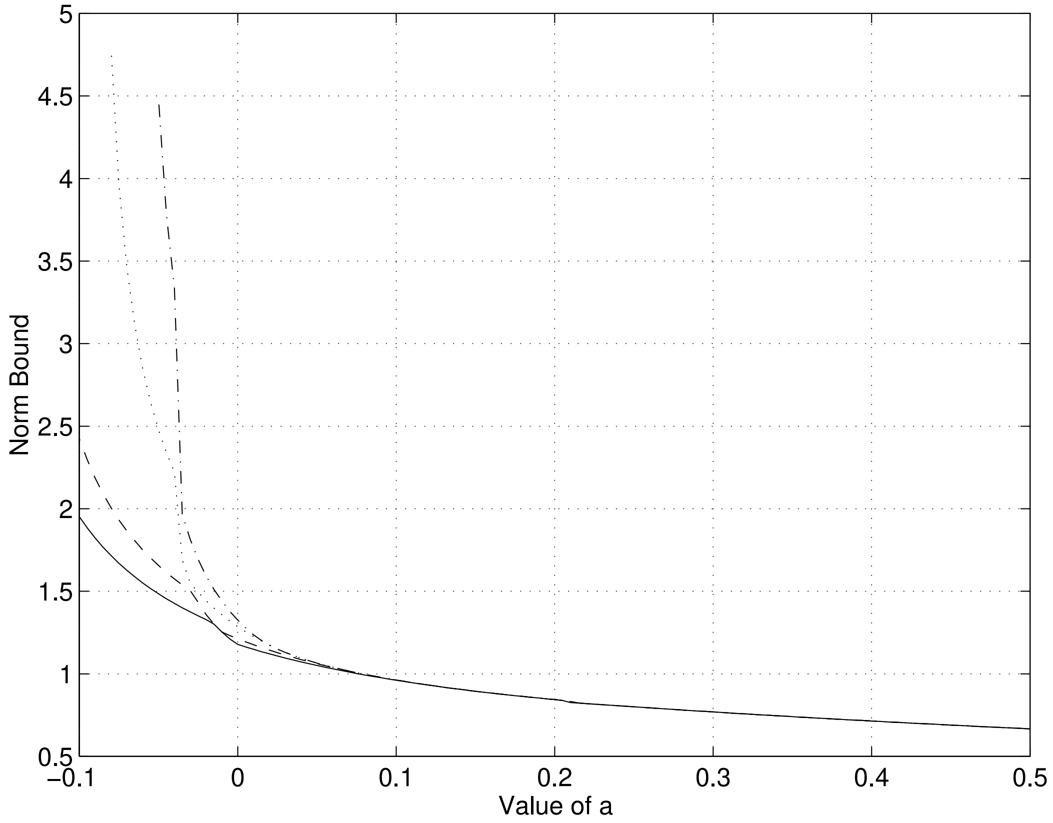

Figure 8.4: Bound for γ^* as a function of a for values of (from top) $N_\infty = 15, 10, 6, 5$.

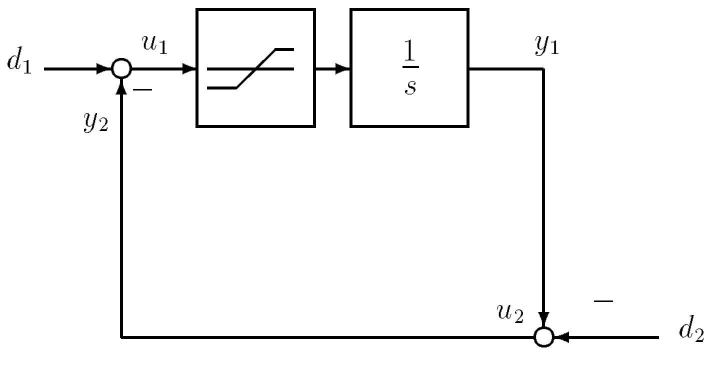

Figure 8.5: Parallel Projection Operator Block Diagram.

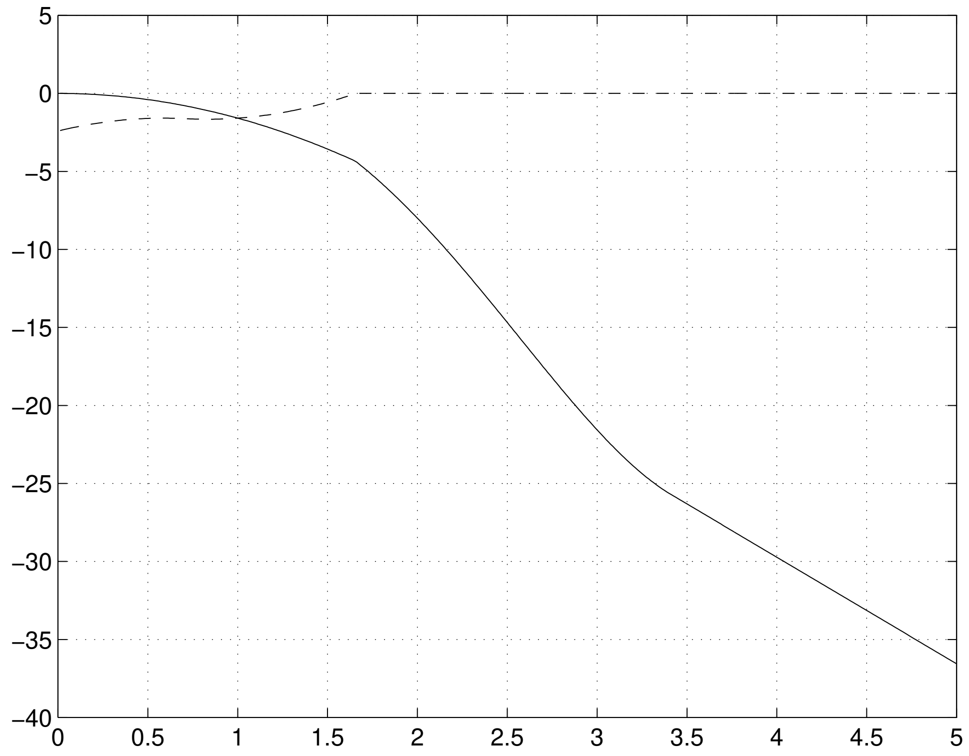

Figure 8.6: Plot of $\Phi_A(\xi)$ (solid) and $\Phi_B(\xi)$ (dashed) vs. $|\xi|$.

The new system has the realisation

$$\begin{aligned} \dot{x}(t) &= \sigma_0(w(t) - x(t)), \\ x(0) &= 0, \\ z &= \begin{bmatrix} w - x \\ x \end{bmatrix}. \end{aligned}$$

In the absence of saturation, the linear system has norm $\gamma_l^* = 1$. With the saturation present, the induced norm rises to $\gamma^* = \sqrt{2+\sqrt{2}}$. This will be first shown by verifying that a candidate $V(x)$ meets the conditions of Theorem 5.4.1 with this upper bound, and then by generating a sequence of test inputs whose limiting output gain tends to the same limit. In this case, the disturbance set is taken to be $\mathcal{L}_{2+}$, and Ω is in fact $\mathbb{R}$.

We take as our candidate dissipation function (with $r = \sqrt{\gamma^{*2} - 1}$):

$$V(\xi) = \begin{cases} \xi^2 & \text{if } |\xi| \le \frac{4}{r}, \\ \frac{r}{6}|\xi^3| + \frac{16}{3r^2} & |\xi| > \frac{4}{r}. \end{cases}$$

The functions Φ_A, Φ_B (defined in Theorem 5.4.1) are plotted versus $|\xi|$ over the interval $[0, 5]$ in Figure 8.6, where it is seen that they are bounded above by 0 . Verification of the analytic form of the functions is omitted due to the tedious nature of the algebra necessary to do so.

As for the asymptotic properties, it is easy to see that if $\xi >> 1$ we have that $\Phi_B(\xi) = 0$. It then remains to be verified that $\Phi_A(\xi)$ is negative over this set. Here,

$$\theta_a(\xi) = -\frac{2+\sqrt{2}}{r} + \frac{1}{2}\xi^2.$$

We will have $|\theta_a(\xi)| > 1$ for $|\xi| \leq 2+\sqrt{2}$. and then,

$$\begin{aligned} \Phi_A(\xi) &= \Phi_{A.2}(\xi) \\ &= \frac{r}{2}\xi^2 - \frac{\gamma^2}{2}(|\xi|+1)^2 + \frac{1}{2}\xi^2 + \frac{1}{2} \\ &= -\frac{\gamma}{2}|\xi| - \frac{r}{2} \\ &\leq 0. \end{aligned}$$

Now, we would like to verify that we can find a sequence of signals in $\mathcal{L}_{2+}$ such that their output gain tends to γ^*. The fact that this class of signals tended to the optimal value, as well as the original motivation for the study of this example, was provided in [19]. Take as the test input $w_{T,N}$ as

$$w_{T,N}(t) = \begin{cases} t+1 & & t \leq T, \\ T+2 & T+2n < & t \leq T+2n+1, \\ -\frac{1}{r}T & T+2n+1 < & t \leq T+2(n+1), \\ 0 & & t > T+2N, \end{cases}$$

for $n = 0, 1, \ldots, N-1$.

We then have that if $T > 2, N > 0$,

$$\frac{\|\boldsymbol{\Upsilon} w_{T,N}\|_2^2}{\|w_{T,N}\|_2^2} = \frac{T^3 + T + \frac{1}{3} + N(4T^2 + (3+\frac{1}{r})T + \frac{10}{3})}{\frac{(T^3+1)-1}{3} + N((1+\frac{1}{r^2})T^2 + 4T + 4)}.$$

In taking the limit, with $T \gg 1, N \gg T$, we see that this lower limit for γ^{*2} tends to $2+\sqrt{2}$.

It may be noticed that the form of the signals $w_{T,N}$ is somewhat unusual. It consists of an initial ramp to drive the state to a large value, and then a high amplitude, relatively high frequency signal which extracts desired behaviour from the feedthrough matrix. Although once again a very simple system, this seems to indicate that the use of pre-filtering or weighting functions may be useful for less conservative analysis. Of course, the addition of extra states due to weighting functions will probably be quite unwelcome for any algorithms developed in the near future.

8.4 The Double Integrator

We will now briefly comment on a result given the paper by Liu, Chitour and Sontag [33]. In that paper, they prove the following Proposition.

Proposition 8.4.1: Consider the system operator generated by:

$$\begin{aligned}\dot{x_1}(t) &= x_2, \\ \dot{x_2}(t) &= -\sigma(ax_1(t) + bx_2(t) + w(t)), \\ x(0) &= 0,\end{aligned}$$

with $a, b > 0$. Then the mapping $w \mapsto x$ is not $\mathcal{L}_{p+}$ stable for all $p \in [1, \infty]$.
Proof: This requires a fairly detailed analysis of test signals, and is done in [33, Section 8]. □

As pointed out in [33], the system in the proposition is globally state space stable. This system provides a very good understanding as to the differences between linear and nonlinear systems when it comes to global properties. By the use of Theorem 5.6.8, the fact that the system is globally state space stable means that for any compact subset of $\mathbb{R}^2$, trajectories confined to that set will exhibit a finite gain property. The result is that the norm bound will deteriorate as the set size increases in an unbounded fashion. This dependence upon the set size may be of extreme importance for numerical computations.

8.5 Discrete Time Case

We will now briefly give an example of the analysis for a simple discrete time system. The analysis is complicated by the fact that it is not sufficient to evaluate the derivative, and so it is likely that pen and paper calculations will be of more limited value.

The system is a "summer" following the standard saturation; the difference equations generating the system operator are given by:

$$\begin{aligned}x[k+1] &= x[k] + \sigma_0(w[k] - x[k]) =: f(x[k], w[k]), \\ x[0] &= 0,\end{aligned}$$

with $x[k] \in \mathbb{R}$.

The induced $\mathcal{L}_{2+}$ norm for the equivalent linear system is 1. When the saturation is included, the norm deteriorates to $\sqrt{2}$, as will be shown below. The fact that these match the values for the continuous time equivalent is somewhat suggestive; this may point to a more general principle. The nature of the saturation would discourage the use of high amplitude high frequency test signals, and so there may be some form of sampling property which could be exploited.

The dissipation function generated by the heuristic is given by:

$$V(\xi) = \begin{cases} \xi^2 & |\xi| \leq 1, \\ \frac{|\xi|(|\xi|+1)(2|\xi|+1)}{6} & |\xi| > 1. \end{cases}$$

The function ξ^2 is the solution to the corresponding (discrete) algebraic Riccati equation, and so for test values of the Hamiltonian which correspond to differences between two points lying in $[-1, 1]$, the test holds with a test value of 1.

If we examine some ξ such that $|\xi| \leq 1$, but $|f(\xi,\omega)| > 1$, we first note that for all values of ω in the test interval, $f(\xi,\omega) = \omega$. Then,

$$\begin{aligned} H(\xi,\omega) &= V(f(\xi,\omega)) - V(\xi) - \gamma^2\omega^2 + \xi^2 \\ &= V(\omega) - \gamma^2\omega^2 \\ &= \frac{1}{3}|\omega|^3 - \frac{3}{2}\omega^2 + \frac{1}{6}|\omega|. \end{aligned}$$

As $|\omega| \leq 2$ over the test interval, we find that $H(\xi,\omega) \leq 0$.

For $|\xi| > 1$, it is easily tested that $H(\xi,0) = 0$. Otherwise, we need to test $H(\xi,\omega)$ for $\xi - 1 \leq \omega \leq \xi + 1$. Without loss of generality, assume that $\xi > 1$. If we write $\omega = \xi + \Delta$, we have that:

$$\begin{aligned} H(\xi,\omega) &= \left[\frac{\xi + \Delta(\xi+\Delta+1)(2\xi+\Delta+1)}{6} - \frac{\xi(\xi+1)(2\xi+1)}{6}\right] \\ &\quad -2(\xi+\Delta)^2 + \xi^2, \\ &= \frac{1}{6}\left((6\Delta - 6)\xi^2 + (6\Delta^2 + 6\Delta - 24)\xi + 2\Delta^3 - 9\Delta^2 + \Delta\right). \end{aligned}$$

The maximising choice of Δ over the interval $[-1,1]$ is $\Delta = 1$, and then we see that

$$H(\xi,\omega) = -2|\xi| - 1.$$

If one verifies the algebra without setting $\gamma = \sqrt{2}$, one will see that the choice of $\gamma = \sqrt{2}$ is the minimal one such that the leading coefficient in the $H(\xi,\omega)$ expression is non-positive.

8.6 A First Order State Feedback Synthesis

We will now examine a simple application which illustrates the potential power of the analysis technique developed herein. The first order integrator with input saturation will once again be examined, but we will now synthesize a general nonlinear (state) feedback, using two different approaches. The first technique uses partial differential inequalities, following the sort of technique developed by A.J. van der Schaft in [51]. The second technique is similar in principle to the most basic anti-windup schemes, as the controller is fixed to be piecewise linear, and then algebraic conditions solved to fix the value of free parameters.

We take the plant to be controlled as:

$$\begin{aligned} \dot{x}(t) &= \sigma_0(-K(x(t)) + w(t)), \\ x(0) &= 0. \end{aligned}$$

8.6.1 Controller Synthesis Using Partial Differential Inequalities

The analysis here follows closely that given in [51], and is of a preliminary nature. It may be possible to solve the state feedback problem using dynamic programming methods, however the preference is once again is to see whether more class specific algorithms may be developed. The somewhat unusual nature of the test condition in Theorem 5.4.1 seems to indicate that the rather naive extension attempted here will not be sufficient.

The controller is a state feedback, and is to be generated by the equation

$$K(\xi) = -\left.\frac{dV}{dx}\right|_{\xi}, \tag{8.8}$$

with $V(\xi)$ a smooth function mapping $\mathbb{R} \mapsto \mathbb{R}$. We would like that the induced $\mathcal{L}_{2+}$ norm for the mapping $w \mapsto z$ to be less than or equal one.

As is typical in this situation, the function V in (8.8) is to be a valid dissipation function for $\gamma = 1$. The first unusual point to note is that the "optimal" input function $\theta_a(\xi)$ is identically zero, as

$$\theta_a(\xi) = -K(\xi) + \frac{1}{\gamma^2}\left.\frac{dV}{dx}\right|_{\xi}.$$

Therefore, V needs to satisfy $\Phi_{A.1}(\xi), \Phi_B(\xi) \leq 0$ for all ξ. The inequalities of interest become:

$$\begin{aligned} -\left\| \left.\frac{dV}{dx}\right|_{\xi} \right\|^2 + \xi^2 &\leq 0, \\ 2\left.\frac{dV}{dx}\right|_{\xi}\left(-\sigma_0\left(\left.\frac{dV}{dx}\right|_{\xi}\right)\right) + \xi^2 &\leq 0. \end{aligned}$$

The constraining equations then become:

$$\begin{aligned} x^2 &\leq \left\| \left.\frac{dV}{dx}\right|_{\xi} \right\|^2, \\ \frac{1}{2}x^2 &\leq \begin{cases} \left(\left.\frac{dV}{dx}\right|_{\xi}\right)^2 & \text{if } \left|\left.\frac{dV}{dx}\right|_{\xi}\right| < 1, \\ \left|\left.\frac{dV}{dx}\right|_{\xi}\right| & \text{otherwise.} \end{cases} \end{aligned}$$

One possible smooth solution is to take $\left.\frac{dV}{dx}\right|_{\xi} = \xi + \frac{\xi^3}{2}$.

The intuition behind the solution is that the controller has unit gain near the origin, giving the correct local properties. The gain is then increased as the output get larger, and reduces the "system sensitivity". The controller is thus able to preserve the local norm bound.

An interesting issue arises as a consequence of the behaviour of θ_a with $\gamma = 1$. The synthesis does not take advantage of the fact that the $\Phi_{A.2}(\xi) \leq \Phi_{A.1}(\xi)$, and hence is a weaker condition. It that if we are extend the methods in [51], this issue will have to be seriously examined if the controller performance is to be satisfactory.

8.6.2 Synthesis of a Piecewise Linear Controller

The control synthesised by the straight use of partial differential inequalities may be unsatisfactory, however. The additional nonlinear terms may complicate linear system analysis, and lead to too complex controllers. One of the important implication of the analysis in Chapter

5 is that linear controllers will preserve their norm properties over a non-negligible region around the origin. As the nonlinear analytic techniques and design intuition are nowhere near as developed as that for linear control theory, the preference is that the nonlinear techniques do not affect linear design processes unnecessarily.

As such, an alternative state feedback will be given which preserves the induced norm, but is in fact piecewise linear. The control is fixed to be of the form

$$K(\xi) = k_1\xi + k_2 M\sigma_0\left(\frac{\xi}{M}\right). \tag{8.9}$$

There are three free parameters (k_1, k_2, M), but as we assume that the linearisation of the controller is fixed by *a priori* linear analysis, there are in fact only two.

In order to motivate setting the problem in this framework, it may be noted that the most basic anti-windup schemes are also piecewise linear, although in that case the controller has dynamics. If this method is generalised, it would then amount to providing a synthesis method for anti-windup schemes using nonlinear $\mathcal{H}_\infty$ methods.

We assume that the linearisation of the controller is fixed to be $K_l = k_1 + k_2 = 1$, as a result of linear analysis (for example, this controller gives the optimal (unweighted) robustness measure with respect to the Gap Metric). The resulting (closed loop) local linearisation has an infinity norm of $\gamma_l^* = 1$. We now wish to find a controller $K(\xi)$ of the form determined by (8.9), with $k_1 + k_2 = 1$, and which ensures that $\|\mathbf{\Upsilon}\|_{i,2} = 1$.

We choose $M = \frac{1}{2}$, by reflection upon the fact that $V_l(\xi) = \frac{\xi^2}{2}$ is a valid dissipation function for the unity feedback controller for $|\xi| \leq 2$. We then see that the controller has one free parameter, and the control form may be expressed as:

$$K(\xi) = \begin{cases} \xi & |\xi| \leq 2, \\ k\xi + 2(1-k) & |\xi| > 2. \end{cases}$$

It is readily apparent that in order to preserve the operator norm of 1, $k \geq 1$. Take a candidate dissipation function V to be:

$$V(\xi) = \begin{cases} \frac{1}{2}\xi^2 & |\xi| \leq 2, \\ \frac{1}{6}\xi^3 + \frac{2}{3} & |\xi| > 2. \end{cases}$$

Examination of the asymptotic value of $\Phi_{A.2}(\xi)$ (with $\gamma = 1$) gives:

$$\Phi_{A.2}(\xi) = 2\xi^2 - (k\xi + 3 - 2k)^2,$$

for large enough ξ ($\theta_a(\xi) \gg 1$ for $\xi \gg 2$). Choosing $k = \sqrt{2}$ is sufficient for the candidate function to be valid for $\gamma = 1$.

In order to generalise this synthesis technique, many computational issues would need to be explored. However, by the tactic of fixing the form of the controller to be piecewise linear with free parameters, it is conjectured that the resulting computations should be algorithmically tractable.

Chapter 9

Concluding Remarks

9.1 Contributions of this Dissertation

We will now summarise what are felt to be the main contributions of this dissertation.

1. In this dissertation, standard dissipation theory has been applied to a class of nonlinear systems (the plant input saturation configuration), to allow the bounding of the induced 2-norm of the system operators. The induced norm computation problem has been examined for both the continuous and discrete time cases. For continuous time systems, the normally required optimisation procedure (pointwise in the state space) over the entire input range has been eliminated by repeated completion of squares, making the test algebraic, as opposed to a dynamic programming problem. (Some simplification occurs in the discrete time case, but a search over a subset of the input range is still necessary.) A heuristic algorithm was developed to generate (suboptimal) valid dissipation functions, with the computation not relying upon the discretisation of the state space. The mathematics associated with the algorithm allows an elegant characterisation of the equivalence between state space and input/output stability.

2. The incremental gain of these nonlinear systems was also examined. Linear matrix inequalities were used to develop a means of computing the incremental gain, in both continuous and discrete time. A general construction of (discrete time) test signals was done herein, which display unbounded incremental gain in the presence of open loop unstable plants.

3. The general first order plant was examined in detail. This examination allows a good qualitative feel of the behaviour of the operator gain bounds generated as a result of the analysis techniques developed herein.

4. The input/output results in this dissertation are developed using bounded input sets, the use of which is advocated herein to analyse systems with saturation nonlinearities. The examples studied provide a good motivation for the use of these bounded input sets.

5. Finally, there are a variety of bridging results which are needed in the main development. There is a good amount of survey material as well as problem framework

development included in this dissertation; this was intentionally done as there have been few recent texts which apply operator theoretic analysis to linear systems with saturation specifically.

9.2 Recommendations for Future Work

We will conclude with a summary of topics of research which arise as a result of the work contained in this dissertation.

1. The area which seems to hold the greatest potential is the study of how one could compute dissipation functions in an algebraic fashion over the state space. The techniques used should also be extensible to the state feedback controller synthesis problem, which is also an obvious topic of interest. It is likely that in the short term, finite difference techniques will need to be used. The objective of such work would be to allow the examination of the output feedback problem.

2. It may be useful as a shorter term project to examine the behaviour of the algorithm given for the induced norm computation, and to compare the results given to those obtained from viscosity theory.

3. If the incremental gain results are to be used, the verification of the conjecture that the solution to the incremental gain inequalities is a necessary and sufficient condition for the gain to be bounded may be useful.

4. If robust control theory based upon operator gains is to be well developed, it is likely that a more thorough examination of the reachability and observability properties of closed loop configurations will be needed, as well as the effect of bounded input sets. Additionally, the perturbations for which stability guarantees can be given will need to be determined.

5. Eventually, the techniques developed in this dissertation will need to be extended to more general configurations, as well as allowing multiple inputs, etc.

Appendix A

A.1 Proof of Propositions 4.6.2 and 6.4.3

In this appendix, the proofs for Propositions 4.6.2 and 6.4.3 will be stated. The results can be summarised as saying that the induced 2-norm of linear systems with the standard saturation is bounded below by that for the local linearisation. Such a result is not surprising, and for nonlinear affine systems, it is shown in the paper by A.J. van der Schaft [51, Proposition 6]. However, there are technical differences present (i.e., we are precisely not affine in the input), and the given proofs here seem quite elegant from an operator theoretic point of view. They may in essence seem obvious, and it is quite likely that they have appeared in the literature in some form, but the importance of the results seem to dictate that the proofs should be given here for completeness.

A.1.1 Proof of Proposition 4.6.2

Before stating the main result, we will first need to state some preliminary lemmas.

Lemma A.1.1: For all $x \in \mathcal{L}_{2+}$, for all $\epsilon > 0$, there exists a constant $M > 0$ such that

$$\|x\|_2 - \left\| M\sigma_0(\frac{x}{M}) \right\|_2 < \epsilon.$$

Proof: By definition,

$$\|x\|_2 = \sqrt{\int_0^\infty x^2(\mu)d\mu},$$

with the integral the Lebesgue Integral. This in fact may be restated as

$$\|x\|_2 = \sqrt{\sup_s \int_0^\infty s^2(\mu)d\mu},$$

with $0 \leq s^2 \leq x^2$, and s a *simple function*, by the definition of the Lebesgue Integral (see [39, Chapter 11]). (A simple function is a function with a finite range, and can be written as

$$s = \sum_{i=1}^{n} \alpha_i \chi_i,$$

with $0 \leq \alpha_i < \infty$, χ_i n non-overlapping characteristic functions.) By the definition of the supremum, for all $\epsilon > 0$, there exists some simple function s such that

$$\sqrt{\int_0^\infty s^2 d\mu} \geq \|x\|_2 - \epsilon.$$

Let $M = \max_{i \in 1,\ldots,n}\{\alpha_i\} < \infty$, with $M > 0$. Then $s^2 \leq M^2\sigma_0^2(\frac{x}{M}) \leq x^2$, hence,

$$\|x\|_2 - \epsilon \leq \sqrt{\int_0^\infty M^2\sigma_0^2(\frac{x}{M})d\mu} \leq \|x\|_2 .$$

This proves the claim. □

Lemma A.1.2: If $P(s)$ is rational, with $P(s) \notin \mathcal{RH}_\infty$, with $P(\infty)$ bounded, there then exists a $u \in \mathcal{L}_{2+} \bigcap \mathcal{L}_{\infty+}$ such that

$$P(s)\hat{u} \notin \mathcal{H}_2.$$

Proof: Fix a minimal realisation $\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$. There then exists some $\xi_0 \in \mathbb{R}^n$ such that $\lim_{t\to\infty} Ce^{At}\xi_0 \to \infty$. As $[A, B]$ is controllable, there exists a bounded u_T on $[0, T]$ $(T > 0)$ such that $x(T) = \xi_0$. Defining u as,

$$u(t) = \begin{cases} u_T(t) & 0 \leq t \leq T, \\ 0 & \text{otherwise,} \end{cases}$$

it then follows that $u \in \mathcal{L}_{2+} \bigcap \mathcal{L}_{\infty+}$, and $P(s)\hat{u}$ does not have a finite 2-norm. □

Lemma A.1.3: If $P(s)$ is rational, then if $u \in \mathcal{L}_{\infty+}$, then $Pu \in \mathcal{L}_{\infty,e}$.
Proof: (The details of this proof will only be sketched, as this is the property which largely justifies the use of extended spaces, although the author is unable to give a specific reference.) Let $\hat{y} = P(s)\hat{u}$. Fix any $0 < T \in \mathbb{R}_+$, and it remains to show that $\mathbf{P}_T y \in \mathcal{L}_{\infty+}$. Fix a realisation of $P(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$, and then note that if the state trajectory is denoted x, then $y(t) = Cx(t) + Du(t)$ over $[0, T]$. However, one can then note the standard results on linear differential equations (as in [45, Section C.5]) give that $x(t)$ is globally defined (as u is essentially bounded), and there exists a time dependent bound (expressed in terms of e^{AT}) for the growth of x. Hence, $x(t)$ is in fact bounded over $[0, T]$, implying that $\mathbf{P}_T y \in \mathcal{L}_{\infty+}$. □

Proposition 4.6.2: Fix a feedback loop in the standard feedback configuration (Section 3.3.1), with $\mathbf{G}$,$\mathbf{K}$ well posed linear systems with the standard saturation (multiple inputs allowed). Let $\mathbf{G_L}, \mathbf{K_L}$ be the respective local linearisations. It is assumed that all linear operators have rational transfer functions. Then, the induced $\mathcal{L}_{2+}$ norm can be characterised as:

1. if $\begin{bmatrix} I & K_L(s) \\ G_L(s) & I \end{bmatrix}^{-1} \in \mathcal{RH}_\infty$ and $\left\|\begin{bmatrix} I & K_L(s) \\ G_L(s) & I \end{bmatrix}^{-1}\right\|_\infty = \gamma_l^*$, then

$$\left\|\begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{G} & \mathbf{I} \end{bmatrix}^{-1}\right\|_{i,2} \geq \gamma_l^*;$$

2. if $\begin{bmatrix} I & K_L(s) \\ G_L(s) & I \end{bmatrix}^{-1} \notin \mathcal{RH}_\infty$, then $\begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{G} & \mathbf{I} \end{bmatrix}^{-1}$ does not have a finite induced $\mathcal{L}_{2+}$ norm.

Proof: To prove 1., we need to show that for all $\epsilon > 0$, there exists some $w \in \mathcal{L}_{2+}$ such that

$$\frac{\left\| \begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{G} & \mathbf{I} \end{bmatrix}^{-1} w \right\|_2}{\|w\|_2} \geq \gamma_l^* - \epsilon.$$

By the definition of the induced $\mathcal{L}_{2+}$ norm, there exists a $\tilde{w} \in \mathcal{L}_{2+}$ such that

$$\frac{\left\| \begin{bmatrix} \mathbf{I} & \mathbf{K_L} \\ \mathbf{G_L} & \mathbf{I} \end{bmatrix}^{-1} \tilde{w} \right\|_2}{\|\tilde{w}\|_2} \geq \gamma_l^* - \frac{\epsilon}{2}.$$

Assume that $\tilde{w} \notin \mathcal{L}_{\infty+}$ (otherwise, the proof becomes trivial). Then, as $\begin{bmatrix} I & K_L(s) \\ P_L(s) & I \end{bmatrix} \in \mathcal{RH}_\infty$, there exists a finite γ_∞ such that

$$\left\| \begin{bmatrix} \mathbf{I} & \mathbf{K_L} \\ \mathbf{P_L} & \mathbf{I} \end{bmatrix}^{-1} \right\|_{i,\infty} = \gamma_\infty.$$

By the use of Lemma A.1.1, there exists some $M > 1$ such that

$$\frac{\left\| M\sigma_0(\frac{\tilde{w}}{M}) - \tilde{w} \right\|_2}{\|\tilde{w}\|_2} < \frac{1}{\gamma_l^*}\frac{\epsilon}{2}.$$

Let $w = \frac{1}{\gamma_\infty}\sigma_0(\frac{\tilde{w}}{M})$. If $v = \begin{bmatrix} \mathbf{I} & \mathbf{K_L} \\ \mathbf{P_L} & \mathbf{I} \end{bmatrix}^{-1} w$, then $\|v\|_\infty \leq \gamma_\infty \|w\|_\infty = 1$. It is then obvious that

$$u = \begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{P} & \mathbf{I} \end{bmatrix}^{-1} w = v.$$

Then,

$$\begin{aligned} \frac{\|u\|_2}{\|w\|_2} &= \frac{\left\| \begin{bmatrix} \mathbf{I} & \mathbf{K_L} \\ \mathbf{P_L} & \mathbf{I} \end{bmatrix}^{-1} \sigma_0(\frac{\tilde{w}}{M}) \right\|_2}{\|w\|_2} \\ &\geq \frac{\left\| \begin{bmatrix} \mathbf{I} & \mathbf{K_L} \\ \mathbf{P_L} & \mathbf{I} \end{bmatrix}^{-1} \left\{ \frac{\tilde{w}}{M} - \left(\sigma_0(\frac{\tilde{w}}{M}) - \frac{\tilde{w}}{M} \right) \right\} \right\|_2}{\left\| \frac{\tilde{w}}{M} \right\|_2} \end{aligned}$$

$$\geq \frac{\left\| \begin{bmatrix} \mathbf{I} & \mathbf{K_L} \\ \mathbf{P_L} & \mathbf{I} \end{bmatrix}^{-1} \frac{\tilde{w}}{M} \right\|_2}{\left\| \frac{\tilde{w}}{M} \right\|_2} - \frac{\left\| \begin{bmatrix} \mathbf{I} & \mathbf{K_L} \\ \mathbf{P_L} & \mathbf{I} \end{bmatrix}^{-1} \right\|_{i,2} \left\| \sigma_0(\frac{\tilde{w}}{M}) - \frac{\tilde{w}}{M} \right\|_2}{\left\| \frac{\tilde{w}}{M} \right\|_2}$$

$$> \gamma_l^* - \frac{\epsilon}{2} - \gamma_l^* \frac{1}{\gamma_l^*} \frac{\epsilon}{2}.$$

Hence the claim that $\frac{\|u\|_2}{\|w\|_2} > \gamma_l^* - \epsilon$ has been verified, proving 1.

To prove 2, take an input $\tilde{w} \in \mathcal{L}_{2+} \bigcap \mathcal{L}_{\infty+}$ such that $\tilde{w} \notin \operatorname{dom} \left\{ \begin{bmatrix} \mathbf{I} & \mathbf{K_L} \\ \mathbf{G_L} & \mathbf{I} \end{bmatrix}^{-1} \right\}$ (such a $\tilde{w}$ exists by Lemma A.1.2). The proof then proceeds by contradiction: assume there exists some constant γ such that

$$\left\| \begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{G} & \mathbf{I} \end{bmatrix}^{-1} \right\|_{i,2} = \gamma.$$

If $\tilde{u} = \begin{bmatrix} \mathbf{I} & \mathbf{K_L} \\ \mathbf{G_L} & \mathbf{I} \end{bmatrix}^{-1} \tilde{w}$, $\tilde{u} \notin \mathcal{L}_{2+}$. However, by Lemma A.1.3, $\tilde{u} \in \mathcal{L}_{\infty,e}$, and so in fact $\tilde{u} \in \mathcal{L}_{2,e}$. Therefore, there exists some $T < \infty$ such that $\|\mathbf{P}_T \tilde{u}\|_2 > \gamma \|\tilde{w}\|_2$ (as the indefinite integral defining $\|\tilde{u}\|_2$ is unbounded). Then, let $\alpha = \frac{1}{\|\mathbf{P}_T \tilde{u}\|_\infty} > 0$. Define $w = \alpha \tilde{w}$, and $u = \alpha \tilde{u}$. As $\|\mathbf{P}_T u\|_\infty < 1$,

$$\mathbf{P}_T u = \mathbf{P}_T \begin{bmatrix} \mathbf{I} & \mathbf{K_L} \\ \mathbf{G_L} & \mathbf{I} \end{bmatrix}^{-1} \mathbf{P}_T w = \mathbf{P}_T \begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{G} & \mathbf{I} \end{bmatrix}^{-1} \mathbf{P}_T w,$$

and so $\frac{\left\| \begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{G} & \mathbf{I} \end{bmatrix}^{-1} w \right\|_2}{\|w\|_2} \geq \frac{\|\mathbf{P}_T u\|_2}{\|w\|_2} \geq \gamma$. This is a contradiction, proving 2. □

A.1.2 Proof of Proposition 6.4.3

Proposition 6.4.3: Fix a feedback loop in the standard feedback configuration (Section 3.3.1), with **G**,**K** well posed linear systems with the standard saturation. Let $\mathbf{G_L}, \mathbf{K_L}$ be the respective local linearisations. Then, the induced ℓ_{2+} norm can be characterised as:

1. if $\begin{bmatrix} I & K_L(z) \\ G_L(z) & I \end{bmatrix}^{-1} \in \mathcal{H}_\infty$ and $\left\| \begin{bmatrix} I & K_L(z) \\ G_L(z) & I \end{bmatrix}^{-1} \right\|_\infty = \gamma_l^*$, then

$$\left\| \begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{G} & \mathbf{I} \end{bmatrix}^{-1} \right\|_{i,2} \geq \gamma_l^*;$$

2. if the operators are all generated by finite dimensional realisations, then if $\begin{bmatrix} I & K_L(z) \\ G_L(z) & I \end{bmatrix}^{-1} \notin \mathcal{RH}_\infty$, then $\begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{G} & \mathbf{I} \end{bmatrix}^{-1}$ does not have a finite induced ℓ_{2+} norm.

Proof: To prove 1., take a sequence $\{w^i\}$ in ℓ_{2+} such that it in the limit achieves a gain of γ_l^* for the linear system $\begin{bmatrix} \mathbf{I} & \mathbf{K_L} \\ \mathbf{G_L} & \mathbf{I} \end{bmatrix}^{-1}$. Let $\{u^i\}$ be the sequence generated by:

$$\begin{bmatrix} u_1^i \\ u_2^i \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{K_L} \\ \mathbf{G_L} & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} w_1^i \\ w_2^i \end{bmatrix}.$$

Then as $w^i, u^i \in \ell_{2+}$, and $\ell_{2+} \subset \ell_{\infty+}$, it is possible find scaling factors α^i such that $\|\alpha^i u_1^i\|_\infty, \|\alpha^i u_2^i\|_\infty < 1$. It is then obvious that:

$$\begin{bmatrix} \alpha u_1^i \\ \alpha^i u_2^i \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{G} & \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \alpha^i w_1^i \\ \alpha^i w_2^i \end{bmatrix},$$

and so we have a test sequence for the the nonlinear system which determines that γ_l^* is a lower bound for the induced norm.

To prove 2., examine a disturbance w such that $w \notin \operatorname{dom}\left\{\begin{bmatrix} \mathbf{I} & \mathbf{K_L} \\ \mathbf{G_L} & \mathbf{I} \end{bmatrix}^{-1}\right\}$. Then as we have added the assumption that the operators are generated by finite dimensional realisations, a finite growth constant exists for the signals u_1, u_2, and so $\mathbf{P}_k \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \in \ell_{2+}$ for all finite k. The proof now proceeds by contradiction: assume that there exists some constant γ such that $\left\| \begin{bmatrix} \mathbf{I} & \mathbf{K_L} \\ \mathbf{G_L} & \mathbf{I} \end{bmatrix}^{-1} \right\|_{i,1} = \gamma$. As $\left\| \mathbf{P}_k \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \right\|_2$ is an unbounded, monotonically increasing sequence, there exists a finite k such that

$$\frac{\left\| \mathbf{P}_k \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \right\|_2}{\|w\|_2} > \gamma.$$

Let $\alpha = \frac{1}{\left\| \mathbf{P}_k \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \right\|_\infty} > 0$. Then,

$$\mathbf{P}_k \begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{G} & \mathbf{I} \end{bmatrix}^{-1} \alpha \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \mathbf{P}_k \begin{bmatrix} \mathbf{I} & \mathbf{K_L} \\ \mathbf{G_L} & \mathbf{I} \end{bmatrix}^{-1} \alpha \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}.$$

It then follows that:

$$\frac{\left\| \begin{bmatrix} \mathbf{I} & \mathbf{K} \\ \mathbf{G} & \mathbf{I} \end{bmatrix}^{-1} \alpha \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \right\|_2}{\|w\|_2} \geq \frac{\left\| \mathbf{P}_k \begin{bmatrix} \mathbf{I} & \mathbf{K_L} \\ \mathbf{G_L} & \mathbf{I} \end{bmatrix}^{-1} \alpha \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \right\|_2}{\|w\|_2} \geq \gamma,$$

proving the claim. □

Appendix B

B.1 Proof of Theorem 7.4.8

We will now prove Theorem 7.4.8, which was deferred due to its somewhat unusual nature. This result, which is needed to show unbounded incremental behaviour is proved in several steps. We will first give some preliminary definitions and results.

Definition B.1.1: The sequence $\beta = \{\beta_1, \ldots, \beta_l\}$, of length l, is a *binary sequence* if for all $i \in 1, \ldots, l$, $\beta_i \in \{0, 1\}$.

Remark B.1.2: When the product notation is used herein for matrices, the order of multiplication is with the lowest index right-most; e.g.,

$$\prod_{i=0}^{m} A_i = A_m A_{m-1} \cdots A_1 A_0.$$

Definition B.1.3: The Euclidean divisor of two integers is denoted div, and the modulo operation mod. For example, $7\mathrm{div}2 = 3$, and $7\mathrm{mod}2 = 1$.

Definition B.1.4: If β is a binary sequence of length l, then define the sequence $\hat{A}_\beta$ on $\mathbb{R}^{n\times n} \times \mathbb{Z}_+$ as follows:

$$\hat{A}_\beta[k] = \begin{cases} I_n & k = 0, \\ A & k > 0, \text{and } \beta[k\mathrm{mod}l + 1] = 1, \\ A - BC & k > 0, \text{and } \beta[k\mathrm{mod}l + 1] = 0. \end{cases}$$

Definition B.1.5: For β, a fixed binary sequence of length l, the associated sequence P_β is the sequence in $\mathbb{R}^{n\times n} \times \mathbb{Z}_+$ given by:

$$P_\beta[k] = \prod_{i=0}^{k} \hat{A}_\beta[i].$$

Remark B.1.6: For β a binary sequence of length l, it is apparent from its construction that

$$P_\beta[m] = \left(\prod_{j=0}^{m\mathrm{mod}l} \hat{A}_\beta[j] \right) P_\beta^{m\mathrm{div}l}.$$

Lemma B.1.7: Assume that there exists a binary sequence of length l (denoted β) such that P_β has an eigenvalue λ such that $|\lambda| \geq 1$ (resp. $|\lambda| > 1$). Let $y_0 \in \mathbb{C}^n$ be the non-trivial eigenvector associated with λ. Then the sequence $S_{M,p}$ for $p < \infty$ (resp. $p = \infty$) is unbounded in M, with $S_{M,p}$ defined:

$$S_{M,p} = \begin{cases} \sqrt[p]{\sum_{i=0}^{M} \|CP_\beta[i]y_0\|^p} & p < \infty, \\ \max_{i \in 0,\ldots,M} \|CP_\beta[i]y_0\| & p = \infty. \end{cases}$$

Proof: Note that $P_\beta[l]y_0 = \lambda y_0$ by the definitions of λ, y_0. And by examining Remark B.1.6, we have that:

$$\begin{aligned} P_\beta[M]y_0 &= \left(\prod_{j=0}^{M \bmod l} \hat{A}_\beta[j]\right) P_\beta^{M \operatorname{div} l} y_0 \\ &= \lambda^{M \operatorname{div} l} \prod_{j=0}^{M \bmod l} \hat{A}_\beta[j] y_0 \\ &= \lambda^{M \operatorname{div} l} P_\beta[M \bmod l] y_0. \end{aligned}$$

Hence, for $p < \infty, m \geq 1$,

$$\begin{aligned} S_{ml-1,p}^p &= \sum_{i=0}^{ml-1} \|CP_\beta[i]y_0\|^p \\ &= \sum_{i=0}^{l-1} \sum_{j=1}^{m} \|CP_\beta[(j-1)l+i]y_0\|^p \\ &= \sum_{i=0}^{l-1} \sum_{j=1}^{m} \left\|\lambda^{j-1} CP_\beta y_0\right\|^p \\ &= \sum_{i=0}^{l-1} \left(\sum_{j=1}^{m} |\lambda|^{p(j-1)}\right) \|CP_\beta[i]y_0\|^p \\ &= \left(\sum_{j=1}^{m} |\lambda|^{p(j-1)}\right) S_{l-1,p}. \end{aligned} \tag{B.1}$$

For $p = \infty, m \geq 1$, we have,

$$\begin{aligned} S_{ml-1,\infty} &= \max_{i \in 0,\ldots,ml-1} \{\|CP_\beta[i]y_0\|\} \\ &= \max_{i \in 0,\ldots,l-1} \left\{\max_{j \in 1,\ldots,m} \{\|CP_\beta[(j-1)l+i]y_0\|\}\right\} \\ &= \max_{i \in 0,\ldots,l-1} \left\{\max_{j \in 1,\ldots,m} \left\{|\lambda|^j \|CP_\beta[i]y_0\|\right\}\right\} \\ &= |\lambda|^m S_{l-1,\infty}. \end{aligned} \tag{B.2}$$

And so, examination of equations (B.1) and (B.2) indicates that $S_{ml-1,p}$ is unbounded (hence $S_{M,p}$, as $S_{M,p}$ is trivially monotonic) if $S_{l-1,p} \neq 0$. This will shown to be necessarily true by contradiction.

Conversely, assume that $S_{l-1,p} = 0$. Examination of equations (B.1) and (B.2) indicates that $S_{ml-1,p} \equiv 0$ for all $m \geq 1$. Choose m such that $ml-1 > n-1$, and then we see $S_{n-1,p} = 0$ by its monotonicity. This necessarily implies that $CP_\beta[i]y_0 = 0$ for all $i \in 0, \dots, n-1$. We can then readily see that $P_\beta[i]y_0 = A^i y_0$. This can be shown by induction. The claim is true by definition for $i = 0$. Then for $i \geq 0$,

$$P_\beta[i+1] = \begin{cases} AP_\beta[i]y_0 & \text{if } \beta[i+1] = 1, \\ (A - BC)P_\beta[i]y_0 & \text{if } \beta[i+1] = 0. \end{cases}$$

But, as we have assumed that $CP_\beta[i] = 0$, the induction is seen to hold.

Hence, the implication is that for $y_0 \neq 0$, we have that

$$\begin{bmatrix} Cy_0 \\ CAy_0 \\ \vdots \\ CA^{n-1}y_0 \end{bmatrix} = 0,$$

which violates the observability assumption on $[C, A]$, and hence we have the required contradiction. □

As one cannot excite complex states with a real system and signals, the result in Lemma B.1.7 is used to derive Lemma B.1.8 (stated next), in which the initial state is a real vector.

Lemma B.1.8: Assume that the conditions regarding P_β in Lemma B.1.7 hold. Then in fact, there exists some $r_0 \in \mathbb{R}^n$ such that the sequence $S^{r_0}_{M,p}$ for $p < \infty$ (resp. $p = \infty$) is unbounded in M, with $S^{r_0}_{M,p}$ defined:

$$S^{r_0}_{M,p} := \begin{cases} \sqrt[p]{\sum_{i=0}^{M} \|CP_\beta[i]r_0\|^p} & p < \infty, \\ \max_{i \in 0,\dots,M} \|CP_\beta[i]r_0\| & p = \infty. \end{cases}$$

Proof: Take y_0 as in Lemma B.1.7. Let $u_0, v_0 \in \mathbb{R}^n$ be the real and imaginary parts of y_0 respectively (i.e., $y_0 = u_0 + iv_0$). It then holds that $P_\beta[i]u_0 = \mathrm{Re}\{P_\beta[i]y_0\}$ (resp. $P_\beta[i]v_0 = \mathrm{Im}\{P_\beta[i]y_0\}$). And by application of the fact that $\|P_\beta[i]y_0\|^2 \leq \|P_\beta[i]u_0\|^2 + 2\|P_\beta[i]u_0\|\,\|P_\beta[i]v_0\| + \|P_\beta[i]v_0\|^2$, $S_{M,p}$ being unbounded implies that $S^{r_0}_{M,p}$ is unbounded for at least one choice of $r_0 = u_0$ or $r_0 = v_0$. □

Theorem 7.4.8: Fix a system $\mathbf{\Upsilon}$ of the form given by (7.1) - (7.3). Then $\|\mathbf{\Upsilon}\|_{\Delta,p}$ does not exist over ℓ_{p+} if there exists a binary sequence of length l denoted $\{\alpha_1, \dots, \alpha_l\}$ such that

$$\hat{A} = \prod_{j=1}^{l} (\alpha_{l-j+1}A + (1 - \alpha_{l-j+1})(A - BC)),$$

has an eigenvalue of modulus greater than or equal one ($p < \infty$) or strictly greater than 1 ($p = \infty$).

Proof: This proof matches in many respects that for Theorem 7.4.1. Note that if $l = 1$, the result is implied by that theorem, and we will therefore assume that $l > 1$. The reason for the separate treatment is that it is possible to show the unbounded behaviour in the case of an unstable plant with signals of small norm, even over $\mathcal{W}_{p,\infty}(N)$ sets for which the induced norm is finite. Here, the construction relies on signals which are bounded, but the bound cannot be set *a priori* and can possibly be quite large.

To prove the theorem, assume the converse; i.e., that there exists some finite γ such that

$$\|\boldsymbol{\Upsilon} w - \boldsymbol{\Upsilon}\tilde{w}\|_p \leq \gamma \|w - \tilde{w}\|_p ,$$

for all $w, \tilde{w} \in \ell_{p+}$. Fix $r_0 \in \mathbb{R}^n$ as in Lemma B.1.8, and let u be a bounded sequence of length n such that

$$\sum_{i=0}^{n-1}(A-BC)^{(n-1)-i}Bu[i] = r_0,$$

with $u[k] = 0$ for all $k \geq n$. Let M be such that $S^{r_0}_{M,p} > \gamma \|u\|_p$. Then, choose α such that

$$\begin{aligned}\alpha^{-1} \ > \ \max\Bigg\{ & \max_{i=0,\ldots,n-1}\left\{\left\| u[i] + \sum_{j=1}^{i} C(A-BC)^{j-i}Bu[j-i]\right\|\right\}, \\ & \max_{i\in 0,\ldots,M-1}\{\|CP_\beta[i]r_0\|\}\Bigg\}.\end{aligned}$$

We can now define our test signals $w, \tilde{w}$, with corresponding induced closed loop responses $x, \tilde{x}$. Let w be defined as follows, being causally dependent upon the trajectory x:

$$w[k] = \begin{cases} 0 & 0 \leq k \leq n-1, \\ Cx[k] & n \leq k \leq n+M-1, \beta[(k-(n-1))\mathrm{mod} l + 1] = 0, \\ Cx[k]+2 & n \leq k \leq n+M-1, \beta[(k-(n-1))\mathrm{mod} l + 1] = 1, \\ 0 & k \geq n+M. \end{cases}$$

Define $\tilde{w}$ as:

$$\tilde{w}[k] = \begin{cases} \alpha u[k] & 0 \leq k \leq n-1, \\ w[k] & k \geq n. \end{cases}$$

Once again, let $\eta = \tilde{x} - x$. The following claim is then made regarding the behaviour of this difference state.

Proposition B.1.9: The state difference η is of the form

$$\eta[k] = \alpha P_\beta[k-n]r_0,$$

for $n \leq k \leq n+M$.

Proof: As before, it can be shown that $\eta[n] = \alpha r_0$. Then, the proposition can be shown if we can verify the general inductive statement that

$$\eta[k+1] = \hat{A}_\beta[k+1-n]\alpha P_\beta[k-n]r_0.$$

If $\beta[(k+1-n)\mathrm{mod} l + 1] = 0$, then $w[k] = Cx[k]$, and so

$$x[k+1] = Ax[k] + B\sigma_0(-Cx[k] + Cx[k]) = Ax[k].$$

Then,

$$\begin{aligned}\tilde{x}[k+1] &= A\tilde{x}[k] + B\sigma_0(-C\tilde{x}[k] + Cx[k]) \\ &= A\tilde{x}[k] + B - \sigma_0(-C\eta[k]).\end{aligned}$$

The construction of α along with the inductive assumption ensures that $|C\eta[k]| < 1$, and so,

$$\begin{aligned}\eta[k+1] &= \tilde{x}[k+1] - x[k+1] \\ &= (A - BC)\eta[k] \\ &= \hat{A}_\beta[k+1-n]\eta[k].\end{aligned}$$

Otherwise, $\beta[(k+1-n)\mathrm{mod} l + 1] = 1$, and $w[k] = Cx[k] + 2$. Similar analysis yields that

$$x[k+1] = Ax[k] + B,$$

and

$$\begin{aligned}\tilde{x}[k+1] &= A\tilde{x}[k] + B\sigma_0(-C\eta[k] + 2) \\ &= A\tilde{x}[k] + B,\end{aligned}$$

(as $2 - C\eta[k] > 1$ by construction). Then,

$$\eta[k+1] = \tilde{x}[k+1] - x[k+1] = A\eta[k] = \hat{A}_\beta[k+1-n]\eta[k].$$

We have thus verified the inductive assumption. $\Box$

The remainder of the proof then follows that of Theorem 7.4.1. We have that

$$\frac{\|\mathbf{\Upsilon}\tilde{w} - \mathbf{\Upsilon}w\|_p}{\|\tilde{w} - w\|} \geq \frac{\alpha S_{M,p}^{r_0}}{\alpha \|u\|_p} > \gamma,$$

the required contradiction. $\Box$

Appendix C

C.1 Proof of Theorem 4.5.3

Although this result is considered to be essentially known, there does not appear to a reference which proves Theorem 4.5.3 in the exact form stated. A proof will therefore be provided, due to the importance of this result.

What we need to show is that theorem, as stated, is equivalent to following result:

Theorem C.1.1: For the operator $\boldsymbol{\Upsilon}$ as stated in Theorem 4.5.3, $\|\boldsymbol{\Upsilon}\|_{i,2} \leq \gamma$ if the candidate function V satisfies the condition

$$\hat{H}(\xi, DV) = \sup_{\omega \in \mathbb{R}^m} \left\{ DV \cdot f(\xi, \omega) - \gamma^2 \|\omega\|^2 + \|g(\xi, \omega)\|^2 \right\} \leq 0,$$

for all $\xi \in \Omega$. Here, $\hat{H}(\xi, DV) \leq 0$ is interpreted as meaning that for all $\phi \in \mathcal{C}^1$, such that $V - \phi$ has a local minimum at ξ,

$$\hat{H}\left(\xi, \left.\frac{d\phi}{dx}\right|_{\xi}\right) \leq 0.$$

Proof: See the papers by M.R. James [26], [27, Theorem 3.1]. (We have added a Lipschitz continuity on V, so we have no need to refer to the lower semi-continuous envelope.) □

We will now show the equivalence. We will first need to add some notation and results on non-smooth analysis. The notation follows the text by F.H. Clarke ([10, Chapter 2]).

Definition C.1.2: The *generalised gradient* of a function $V: \mathbb{R}^n \mapsto \mathbb{R}$ at x, with V Lipschitz at x, is denoted $\partial V(x)$, and is the subset of $\mathbb{R}^{n*}$ given by

$$\left\{ \eta \in \mathbb{R}^{n*} : \limsup_{y \to x, t \downarrow 0} \frac{V(y + tv) - V(y)}{t} \geq \eta' v, \quad \forall v \in \mathbb{R}^n \right\}.$$

For the following results, $V: \mathbb{R}^n \mapsto \mathbb{R}$ is taken to be (at least locally) Lipschitz.

Lemma C.1.3: If V has a local minimum or maximum at x, then $0 \in \partial V(x)$.
Proof: See [10, Prop. 2.3.2]. □

Lemma C.1.4: If for all but at most one of the functions f_i is strictly differentiable, then

$$\partial\left(\sum f_i\right)(x) = \sum \partial f_i(x).$$

Proof: See [10, Prop. 2.3.3, Cor. 1]. $\square$

Lemma C.1.5: If $V\colon \mathbb{R}^n \mapsto \mathbb{R}$ is Lipschitz, with n finite, and is differentiable over the set Ω_d (with the complement of Ω_d necessarily with zero measure), then for $v \in \mathbb{R}^n$

$$\sup_{p\in\partial V(\xi)} pv = \limsup_{y\to\xi, y\in\Omega_d} \left\{ \left.\frac{dV}{dx}\right|_y v \right\}.$$

Proof: See [10, Theorem 2.5.1, Cor.]. $\square$

Lemma C.1.6:
$$\sup_{\phi\in\mathcal{C}^1,(V-\phi)(\xi)\text{is a local minimum}} \left.\frac{d\phi}{dx}\right|_\xi f(\xi,\omega) \leq \sup_{p\in\partial V(\xi)} pf(\xi,\omega). \quad \text{(C.1)} \quad \text{(C.1}$$

Proof: By Lemma C.1.3, $0 \in \partial(V-\phi)(\xi)$, as $V-\phi$ has a local minimum at ξ. Then by Lemma C.1.4, $\partial(V-\phi)(\xi) = \{\partial V(\xi) - \left.\frac{d\phi}{dx}\right|_\xi\}$. Therefore, as $0 \in \partial(V-\phi)(\xi)$, $\left.\frac{d\phi}{dx}\right|_\xi \in \partial V(\xi)$. Thus $\partial V \supseteq \{\left.\frac{d\phi}{dx}\right|_\xi : \phi \in \mathcal{C}^1, (V-\phi)(\xi)$ is a local minimum$\}$, implying (C.1). $\square$

Lemma C.1.7:
$$\sup_{p\in\partial V(\xi)} 2pf(\xi,\omega) - \gamma^2\|\omega\|^2 + \|g(\xi,\omega)\|^2 \leq 0, \quad \forall\xi\in\Omega,$$
is implied by

$$2\left.\frac{dV}{dx}\right|_\xi f(\xi,\omega) - \gamma^2\|\omega\|^2 + \|g(\xi,\omega)\|^2 \leq 0, \quad \forall\xi\in\Omega_d \quad \text{(i. e., almost everywhere).}$$

Proof: By Lemma C.1.5

$$\sup_{p\in\partial V(\xi)} 2pf(\xi,\omega) = \limsup_{y\to\xi, y\in\Omega_d}\left\{2\left.\frac{dV}{dx}\right|_y f(\xi,\omega)\right\}. \tag{C.2}$$

We need to show then that

$$\limsup_{y\to\xi, y\in\Omega_d}\left\{2\left.\frac{dV}{dx}\right|_y f(\xi,\omega)\right\} = \limsup_{y\to\xi, y\in\Omega_d}\left\{2\left.\frac{dV}{dx}\right|_y f(y,\omega)\right\}. \tag{C.3}$$

Assume the converse, that there exists an $\epsilon > 0$ such that

$$\left|\limsup_{y\to\xi, y\in\Omega_d}\left\{2\left.\frac{dV}{dx}\right|_y f(\xi,\omega)\right\} - \limsup_{y\to\xi, y\in\Omega_d}\left\{2\left.\frac{dV}{dx}\right|_y f(y,\omega)\right\}\right| > \epsilon.$$

Then, pick any sequence $\{y_i\} \to \xi$. Choose the subsequence such that $\|y_i - \xi\| < \frac{\epsilon}{2\kappa_V\kappa_f}$, with κ_V, κ_f Lipschitz constants for $\frac{dV}{dx}, f$ respectively. Then for all y_i in this subsequence,

$$\left|2\left.\frac{dV}{dx}\right|_{y_i} f(\xi,\omega) - 2\left.\frac{dV}{dx}\right|_{y_i} f(y_i,\omega)\right| \leq \kappa_V\kappa_f\|y_i - \xi\| < \frac{\epsilon}{2},$$

a contradiction. The theorem statement then follows from (C.2), (C.3). □

To prove Theorem 4.5.3, note that

$$\sup_{\phi\in\mathcal{C}^1,(V-\phi)(\xi)\text{is a local minimum}} 2\left.\frac{d\phi}{dx}\right|_{\xi} f(\xi,\omega)-\|\omega\|^2+\|g(\xi,\omega)\|^2\le 0,\quad \forall\xi\in\Omega,$$

is implied by

$$2\left.\frac{dV}{dx}\right|_{\xi} f(\xi,\omega)-\gamma^2\|\omega\|^2+\|g(\xi,\omega)\|^2,\quad \forall\xi\in\Omega_d\quad \text{(i. e., almost everywhere)},$$

follows from Lemmas C.1.6, C.1.7. Then Theorem 4.5.3 follows from the application of Theorem C.1.1, taking the supremum with respect to ω pointwise over Ω. □

Appendix D

D.1 Proof of Lemmas in Chapter 5

We will now give the proofs of the lemmas from Chapter 5 whose proof had been deferred to avoid breaking the continuity of argument.

Lemma 5.6.2: For all $Q > 0$, there exists a finite $\gamma_l \geq \tilde{\gamma}_l$ such that there is a solution P to the equation (5.15) and with the additional inequality (5.16) also holding.
Proof: It is a consequence of Theorem 4.6.1 that there exists a solution to (5.15) for all $\tilde{\gamma}_l$ such that

$$\tilde{\gamma}_l > \left\| \left[\begin{array}{c|c} A - BK & B \\ \hline \begin{bmatrix} C \\ Q^{\frac{1}{2}} \end{bmatrix} & \begin{bmatrix} D \\ 0 \end{bmatrix} \end{array} \right] \right\|_\infty . \tag{D.1}$$

In order to ensure that the inequality (5.16) holds, it is sufficient that one chooses γ_l such that

$$\gamma_l^2 > \|D_1\|^2 + \|D_1\|^4 \frac{\lambda_{\max}(K'K)}{\lambda_{\min}(Q)}, \tag{D.2}$$

($\lambda_{\max}, \lambda_{\min}$ denoting the maximum and minimum eigenvalues respectively). Therefore it is sufficient to choose γ_l as the maximum of the two expressions in (D.1) - (D.2). □

Lemma 5.6.4: For all $\xi \in \Omega_l$, $x(t;\xi) \in \Omega_l$, and $|Kx(t;\xi)| \leq 1$ for all $t \in \mathbb{R}_+$.
Proof: By its definition, it is easily seen that $|K\xi| \leq 1$ for all $\xi \in \Omega_l$. Denote by $x_l(t;\xi)$ the induced trajectory for the linear system, where $x_l(t;\xi)$ is generated by:

$$\begin{aligned} \dot{x}_l(t) &= (A - BK)x_l(t), \\ x_l(0) &= \xi. \end{aligned}$$

It is easy to see (by examining the Lyapunov equation) that $x_l(t;\xi)'P_\infty x_l(t;\xi)$ is an exponentially decaying function of time, and hence if $\xi \in \Omega_l$, $x_l(t;\xi) \in \Omega_l, \forall t \in \mathbb{R}_+$. And then by noting that if $|Kx(t;\xi)| \leq 1$ for $0 \leq \tau \leq t$, then $x(t;\xi) = x_l(t;\xi)$, and so the claim holds. □

Lemma 5.6.6: If $x(t;\xi) \in \mathcal{L}_{2+}$ for all $x \in \bar{\Omega}$, there then exists a finite constant T_Ω such that $T_\Omega = \max_{\xi \in \bar{\Omega}} T(\xi)$, with $T: \bar{\Omega} \mapsto \mathbb{R}_+$ defined as

$$T(\xi) = \min_{t \in R_+} \{t: x(t;\xi) \in \Omega_l\}.$$

Proof: The details of this proof consist in showing that T is a continuous function, and are only sketched here. If a point ξ in $\bar{\Omega}$ is fixed, by the assumption that $x(t;\xi) \in \mathcal{L}_{2+}$ we know that $T(\xi)$ is finite (as Ω_l contains an open ball at the origin, which $x(t;\xi)$ must enter or the square integral of the path would be unbounded), and by Lemma 5.6.4, the trajectory will remain in Ω_l. We then show that T is continuous at ξ by noting the exponential stability with respect

to $x^T(t)P_\infty x(t)$ will mean that for all $t > T(\xi)$, there will exist an open ball surrounding $x(t;\xi)$ which is contained in Ω_l. And since $x(t,\cdot)$ is known to be continuous by the use of Lemma 5.6.5, there exists an open ball surrounding ξ such that $x(t,\cdot)$ over that ball lies in Ω_l, hence bounding above $T(\cdot)$ over that ball. Similarly, it is possible to bound T arbitrarily close to $T(\xi)$ from below over some ball surrounding ξ, and so continuity can be shown to hold. The fact that a maximum exists then follows from the fact that $\bar{\Omega}$ is compact. □

Lemma 5.6.7: Let $U: \bar{\Omega} \mapsto \mathbb{R}$ be defined:

$$U(\xi) = \int_0^\infty \alpha(x(t;\xi))dt,$$

with $\alpha: \bar{\Omega} \mapsto \mathbb{R}$ having a (uniform) Lipshitz constant k, and with $\alpha(0) = 0$. Then if $x(t;\xi) \in \mathcal{L}_{2+}$ for all ξ in $\bar{\Omega}$, U is uniformly Lipschitz.

Proof: Fix any distinct points ξ, η in $\bar{\Omega}$. To prove that the function is uniformly Lipschitz, examine $|U(\xi) - U(\eta)|$. From Lemma 5.6.6 we have some $T_\Omega < \infty$ such that $x(T_\Omega;\xi), x(T_\Omega;\eta) \in \Omega_l$. By Lemma 5.6.5, there exists some constant c_1 such that

$$\|x(t;\xi) - x(t;\eta)\| < c_1 \|\xi - \eta\|,$$

for $0 \le t \le T_\Omega$. It is a well known fact from linear theory that the state trajectory in response to an initial condition (only) for an asymptotically stable linear shift invariant system lies in $\mathcal{L}_1$; hence it can be seen that there exists another constant c_2 such that $\int_{T_\Omega}^\infty \|x(t;\xi) - x(t;\eta)\| \, dt < c_2 \|x(T_\Omega;\xi) - x(T_\Omega;\eta)\|$. Hence,

$$\begin{aligned} |U(\xi) - U(\eta)| &< k \int_0^{T_\Omega} \|x(t;\xi) - x(t;\eta)\| \, dt + k \int_{T_\Omega}^\infty \|x(t,\xi) - x(t;\eta)\| \, dt \\ &< kc_1 T_\Omega \|\xi - \eta\| + kc_2 c_1 \|\xi - \eta\|. \end{aligned}$$

Hence, some Lipschitz constant exists. □

Bibliography

[1] K.J. Åström and B. Wittenmark. *Computer controlled systems: Theory and design.* Prentice-Hall, Englewood Cliffs, N.J., 1984.

[2] J.A. Ball and J.W. Helton. Viscosity solutions of Hamilton-Jacobi equations arising in nonlinear $\mathcal{H}_\infty$-control. Submitted for journal publication.

[3] J.A. Ball, J.W. Helton, and M.L. Walker. H^∞ control for nonlinear systems with output feedback. *IEEE Trans. Automat. Contr.*, 38(4):546–559, 1993.

[4] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in Systems and Control Theory.* SIAM, Philedelphia, 1994.

[5] R. Brammer. Controllability in linear autonomous systems with positive controllers. *SIAM J. Control*, 10(2):339–353, 1972.

[6] C.I. Byrnes and W. Lin. Losslessness, feedback equivalence, and the global stabilization of discrete-time nonlinear systems. *IEEE Trans. Automat. Contr.*, 39(1):83–98, 1994.

[7] P.J. Campo. *Studies in Robust Control of Systems Subject to Constraints.* Ph.D. thesis, Cal. Inst. Tech., 1990.

[8] C.T. Chen. *Linear System Theory and Design.* Holt, Rinehart and Winston, New York, 1984.

[9] Y. Chitour, W. Liu, and E.D. Sontag. On the continuity and incremental-gain properties of certain saturated linear feedback loops. Submitted for journal publication.

[10] F.H. Clarke. *Optimization and Nonsmooth Analysis.* Wiley-Interscience, New York, 1983.

[11] M.G. Crandall and P.-J. Lions. Viscosity solutions of Hamilton - Jacobi equations. *Trans. Amer. Math. Soc.*, 227(1):1–42, 1983.

[12] C.A. Desoer and J. Wing. The minimal time regulator problem for linear sampled-data systems: General theory. *J. Franklin Inst.*, 272:208 –228, September 1961.

[13] J.C. Doyle, T.T. Georgiou, and M.C. Smith. The parallel projection operators of a nonlinear feedback system. *Systems and Control Letters*, 20:79–85, 1993.

[14] W.H. Fleming and H.M. Soner. *Controlled Markov Processes and Viscosity Solutions.* Springer Verlag, New York, 1993.

[15] D.K. Frederick and A.B. Carlson. *Linear Systems in Communication and Control.* Wiley, New York, 1971.

[16] A.T. Fuller. In-the-large stability of relay and saturating control systems with linear controllers. *Int. J. Control*, 10(4):457–480, 1969.

[17] J.E. Gayek. A survey of techniques for approximating reachable and controllable sets. In *Proc. IEEE Conf. Dec. and Cont.*, 1991.

[18] T.T. Georgiou. Differential stability and robust control of nonlinear systems. *Math.of Controls, Signals and Systems*, to appear.

[19] T.T. Georgiou and M.C. Smith. Unpublished notes.

[20] T.T. Georgiou and M.C. Smith. Metric uncertainty and nonlinear feedback stabilization. Technical report, University of Cambridge, Dept. of Eng., Cambridge, 1994.

[21] M. Green and D.J.N Limebeer. *Linear Robust Control.* Prentice-Hall, New Jersey, 1995.

[22] J.W. Helton and W. Zhan. Piecewise Riccati equations and The Bounded Real Lemma. In *Proc. IEEE Conf. Dec. and Cont.*, pages 196–201, 1993.

[23] D.J. Hill and P.J. Moylan. The stability of nonlinear dissipative systems. *IEEE Trans. Automat. Contr.*, 21:708–711, 1976.

[24] D.J. Hill and P.J. Moylan. Connections between finite-gain and asymptotic stability. *IEEE Trans. Automat. Contr.*, 25:931–936, October 1980.

[25] P.A. Iglesias. *Robust and Adaptive Control for Discrete-Time Systems.* Ph. D. thesis, University of Cambridge, 1991.

[26] M.R. James. Computing the H_∞ norm for nonlinear systems. In *IFAC World Congress*, Sydney, 1993.

[27] M.R. James. A partial differential inequality for dissipative nonlinear systems. *Systems and Control Letters*, 21:315–320, 1993.

[28] E.B. Lee. Mathematical aspects of the synthesis of linear minimum response-time controllers. *IRE Trans. Aut. Cont.*, AC-5(4):283–290, 1960.

[29] E.B. Lee and L. Markus. *Foundations of Optimal Control Theory.* Wiley, New York, 1968.

[30] W. Lin and C.I. Byrnes. Dissipativity, L_2-Gain and H_∞ control for discrete-time nonlinear systems. In *Amer. Control Conference*, pages 2257–2260, 1994.

[31] Z. Lin, A. Saberi, and A.R. Teel. Simultaneous $\mathcal{L}_p$-stabilization and internal stabilization of linear systems subject to input saturation - state feedback case. Submitted for journal publication.

[32] P.-L. Lions. *Generalized solutions of Hamilton-Jacobi equations.* Pitman, London, 1982.

[33] W. Liu, Y. Chitour, and E.D. Sontag. On finite gain stabilizability of linear systems subject to input saturation. Submitted for journal publication.

[34] W. Liu, Y. Chitour, and E.D. Sontag. Remarks on finite gain stabilizability of linear systems subject to input saturation. In *Proc. IEEE Conf. Dec. and Cont.*, pages 1808–1813, 1993.

[35] A.M. Lyapunov. *The General Problem of the Stability of Motion.* Taylor and Francis, 1992. Translated and edited by A. T. Fuller.

[36] P.J. Moylan. Implications of passivity in a class of nonlinear systems. *IEEE Trans. Automat. Contr.*, 19(4):373–381, August 1974.

[37] N.B.O.L. Pettit and P.W. Wellstead. A graphical analysis method for piecewise linear systems. In *Proc. IEEE Conf. Dec. and Cont.*, pages 1122–1127, 1994.

[38] B.G. Romanchuk. On the computation of the induced $\mathcal{L}_2$ norm of single input linear systems with saturation. In *Proc. IEEE Conf. Dec. and Cont.*, pages 1427–1432, 1994.

[39] W. Rudin. *Principles of Mathematical Analysis.* McGraw Hill, New York, 3rd edition, 1964.

[40] I.W. Sandberg. On the L_2-boundedness of solutions of nonlinear functional equations. *Bell Syst. Tech. J.*, 43:1601–1608, 1965.

[41] S. Saperstone and J. Yorke. Controllability of linear oscillatory systems using positive controllers. *SIAM J. Control*, 9:253–262, 1971.

[42] W.E. Schmitendorf and B.R. Barmish. Null controllability of linear systems with constrained controls. *SIAM J. Control and Optimization*, 18(4):327–345, 1980.

[43] R. Schwarzschild, E.D. Sontag, and M.L.J Hautus. Output saturated systems. In *Amer. Control Conference*, pages 2504–2509, 1992.

[44] E.D. Sontag. An algebraic approach to bounded controllability of linear systems. *Int. J. Control*, 39(1):181–188, 1984.

[45] E.D. Sontag. *Mathematical Control Theory.* Springer Verlag, New York, 1990.

[46] E.D. Sontag and Y. Yang. Global stabilization of linear systems with bounded feedback. Technical report, Rutgers Center for Systems and Control, New Brunswick, NJ, 1991.

[47] H.J. Sussman, E.D. Sontag, and Y. Yang. A general result on the stabilization of linear systems using bounded controls. In *Proc. IEEE Conf. Dec. and Cont.*, pages 1802–1807, 1993.

[48] H.J. Sussman and Y. Yang. On the stabilizability of multiple integrators by the means of bounded feedback controls. In *Proc. IEEE Conf. Dec. and Cont.*, pages 70–73, 1991.

[49] W.A. Sutherland. *Introduction to metric and topological spaces.* Oxford University Press, Oxford, 1975.

[50] A.R. Teel. Global stabilization and restricted tracking for multiple integrators with bounded controls. *Systems and Control Letters*, 18:165–171, 1992.

[51] A.J. van der Schaft. L_2-gain analysis of nonlinear systems and nonlinear state feedback $\mathcal{H}_\infty$ control. *IEEE Trans. Automat. Contr.*, 37(6):770–784, 1992.

[52] M. Vidyasagar. *Nonlinear Systems Analysis.* Prentice-Hall, New Jersey, 1978.

[53] M. Vidyasagar. On the use of right-coprime factorizations in distributed feedback systems containing unstable systems. *IEEE Trans. Circuits and Syst.*, 25(11):916–921, 1978.

[54] M. Vidyasagar. *Control System Synthesis: A Factorization Approach.* MIT Press, Cambridge, Massachusetts, 1987.

[55] J.C. Willems. *The Analysis of Feedback Systems.* The M.I.T. Press, Cambridge, Massachusetts, 1971.

[56] J.C. Willems. Dissipative dynamical systems, Part I: General theory. *Arch. Rational Mech. Anal.*, 45:321–351, 1971.

[57] J.C. Willems. Dissipative dynamical systems, Part II: Linear systens with quadratic supply rates. *Arch. Rational Mech. Anal.*, 45:352–393, 1971.

[58] J.C. Willems. The generation of Lyapunov functions for input-output stable systems. *SIAM J. Control*, 9(1):105–134, 1971.

[59] J. Wing and C.A. Desoer. The multiple-input minimal time regulator problem (general theory). *IEEE Trans. Automat. Contr.*, AC-5:125 – 136, 1963.

[60] G. Zames. On the input-output stability of time-varying nonlinear feedback systems, Part I, conditions using concepts of loop gain, conicity and positivity. *IEEE Trans. Automat. Contr.*, 11(2):228–238, April 1966.

[61] G. Zames. On the input-output stability of time-varying nonlinear feedback systems, Part II: Conditions involving circles in the frequency plane and sector nonlinearities. *IEEE Trans. Automat. Contr.*, 11(3):465–476, July 1966.

[62] K. Zhou, J. Doyle, and K. Glover. Robust and optimal control. Draft.

[63] W.P. Ziemer. *Weakly Differentiable Functions.* Springer Verlag, New York, 1989.