MULTI-OBJECTIVE ROBUST CONTROL OF ROTOR/ACTIVE MAGNETIC BEARING SYSTEMS

by

İbrahim Sina Kuseyri

B.Sc. in Mechanical Engineering, Boğaziçi University (1984)

M.Sc. in Mechanical Engineering, Gazi University (1989)

Submitted to the Institute for Graduate Studies in
Science and Engineering in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

Graduate Program in Mechanical Engineering
Boğaziçi University
2011

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Prof. Emre Köse for his invaluable guidance and help during the preparation of this dissertation. I would like to mention his patience, giving me inspiration and hope when I was stuck at death-ends.

I would like to thank Prof. Günay Anlaş and Assoc. Prof. Hilmi Luş for their precious guidance and support throughout my work. I am also grateful to Prof. Eşref Eşkinat, Assoc. Prof. Tankut Acarman, and Asst. Prof. Serdar Soyöz for their valuable comments on my thesis.

ABSTRACT

MULTI-OBJECTIVE ROBUST CONTROL OF ROTOR/ACTIVE MAGNETIC BEARING SYSTEMS

This dissertation considers the application of recent robust linear control techniques to rotors with active magnetic bearings. Rotors suspended with electromagnetic bearings are inherently unstable; therefore feedback control is an essential part of their operation. The purpose of the study is to design, analyze and compare the performance of various stabilizing robust controllers for a model of horizontal rotor with active magnetic bearings. Particular emphasis is placed on the study of time varying and parameter varying linear systems, due to the parameter (rotational speed) dependent structure of rotor dynamics.

Despite the nonlinear form of the actuator (electromagnetic bearing) dynamics, the controlled system is linear and time invariant at constant speed. Control inputs from the actuators are linearized using a constant bias current in order to synthesize a linear stabilizing controller. Main sources of uncertainty in the system stem from the changing spring stiffness (due to different operating conditions) of the electromagnetic bearings, and changing rotor dynamics at different rotational speeds due to gyroscopic effects. Several H_∞ controllers with nominal performance are designed. Simulations are carried out on the actual nonlinear system at speeds as high as 6000 rpm. Robust stability of the system with H_∞ control using an appropriate uncertainty structure is tested and the limits of uncertainty for the uncertain parameters are established. In order to reduce the uncertainty in the system model and to enhance the operation speed with robust stability, a LPV controller is designed with H_∞ performance via scheduling by the rotor speed on-line. Finally, to enhance the performance of controller with respect to possible transients during the operation, a multi-objective LPV controller is designed with generalized H_2 performance, trading-off robustness against performance.

ÖZET

AKTİF MANYETİK YATAKLI ROTORLARIN ÇOK AMAÇLI DAYANIMLI KONTROLU

Bu tez, dayanımlı doğrusal kontrol konusunda aktif manyetik yataklı rotorlara uygulanılabilecek güncel teknikleri ele almaktadır. Elektromanyetik yataklarla taşınan rotorların yapısal olarak kararsız olmaları nedeniyle, geri beslemeli kontrol sistemin önemli bir parçasıdır. Bu çalışmanın amacı, aktif manyetik yataklı yatay bir rotor için kararlılık kazandıran çeşitli dayanımlı kontrol elemanlarının tasarlanması, analiz edilmesi ve performanslarının kıyaslanmasıdır. Rotordinamiğinin parametreye (dönüş devri) bağlı yapısı nedeniyle, "zamanla değişen" ve "parametreye bağlı değişen" doğrusal sistemlerin üzerinde ağırlıklı olarak durulmuştur.

Aktüatörün (elektromanyetik yatak) doğrusal olmayan dinamiği hariç, kontrol edilen sistemin dinamiği, sabit hızda, zaman içinde değişmeyen doğrusal yapıdadır. Kararlılılık sağlayıcı doğrusal bir kontrol elemanı kullanabilmek için aktüatörlerden alınan kontrol girişleri sabit bir önakım kullanılarak doğrusallaştırılmıştır. Sistem dinamiğindeki belirsizliğin ana unsurları elektromanyetik yatakların değişken yay direngenliği (çalışma şartları nedeniyle) ve farklı devirlerde jiroskopik etkiler nedeniyle değişen rotordinamiğidir. Nominal performanslı birkaç H_∞ kontrol elemanı tasarlanmıştır. Doğrusal olmayan asıl sistem üzerinde 6000 rpm'e kadar devirlerde simülasyonlar yapılmıştır. Uygun bir belirsizlik yapısı kurularak sistemin H_∞ kontrolla dayanımlı kararlılığı test edilmiş ve belirsiz parametreler için belirsizliğe ilişkin sınırlar belirlenmiştir. Sistem modelindeki belirsizliği azaltmak ve dayanımlı kararlılığa sahip çalışma devrini arttırmak amacıyla, rotor devrine göre anında ayarlanan ve H_∞ performanslı bir LPV kontrol elemanı tasarlanmıştır. Son olarak, çalışma sırasında karşılaşılabilecek "geçici" şartlara yönelik performansın arttırılması amacıyla, dayanımlılık performans lehine azaltılarak, genelleştirilmiş H_2 performanslı çok amaçlı bir LPV kontrol elemanı tasarlanmıştır.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF SYMBOLS

a	Distance from the rotor center to the radial magnetic bearing on the positive direction
A	System matrix
$\tilde{A}$	Transformed system matrix
$\mathcal{A}$	Closed-loop system matrix
$\mathbf{A}$	Augmented closed-loop system matrix
$\mathbb{A}$	Set of polytopic system matrices
$\tilde{\mathcal{A}}$	Closed-loop system matrix of a LPV system
A_A	Rotor to stator (magnetic bearing) air gap area
A_D	Damping related part of the flexible rotor system matrix
A_e	Output error system matrix
A_{Fe}	Bearing radial area exerting electromagnetic force
A_G	Gyroscopic part of the rotor system matrix
$A_{\mathcal{G}}$	Speed dependent gyroscopic rotor system matrix
A_i	Input disturbance system matrix
A_K	Controller system matrix
A_n	System matrix of the noise
A_q	Uncertain dynamics output matrix
A_r	Rotor system matrix
A_S	Stiffness related part of the rotor system matrix
$A_{\tilde{u}}$	System matrix of the control input weight
A_{uf}	System matrix of the control input filter
$A_{11}, A_{12}, A_{21}, A_{22}$	System matrices of a quasi-LPV model
A_0	Parameter independent part of the system matrix
A_Ω	Rotor speed dependent part of the system matrix
b	Distance from the rotor center to the radial magnetic bearing on the negative direction
B	Input matrix
$\mathcal{B}$	Closed-loop input matrix

$\mathbf{B}$	Augmented closed-loop input matrix
B_A	Magnetic flux density
B_e	Output error input matrix
B_i	Input disturbance input matrix
B_K	Controller input matrix
B_n	Input matrix of the noise
B_p	Uncertain dynamics input matrix
B_r	Rotor input matrix
B_u	Control input matrix
$B_{\tilde{u}}$	Input matrix of the control input weight
B_{uf}	Input matrix of the control input filter
B_w	Disturbance input matrix
$\mathcal{B}_{\mathbf{\Delta}}$	Unit ball
B_1	Disturbance input matrix of multi-objective system
B_2	Control input matrix of multi-objective system
B_{1r}	Rotor disturbance input matrix
B_{2r}	Rotor control input matrix
$\tilde{B}_1, \tilde{B}_2$	Input matrices of a quasi-LPV model
B_0	Parameter independent part of the input matrix
B_Ω	Rotor speed dependent part of the input matrix
co	Convex hull
C	Controllability matrix
C	Output matrix
$\tilde{C}$	Modal output matrix
$\mathcal{C}$	Closed-loop output matrix
$\mathbf{C}$	Augmented closed-loop output matrix
$\mathbb{C}$	Complex numbers
C_e	Output error output matrix
C_i	Input disturbance output matrix
C_K	Controller output matrix
C_n	Output matrix of the noise

C_q	Uncertain dynamics weighted output matrix
C_r	Rotor output matrix
C_s	Sensor output matrix
$C_{\tilde{u}}$	Output matrix of the control input weight
C_{uf}	Output matrix of the control input filter
C_y	Measurement output matrix
C_{y1}, C_{y2}	Measurement output matrices of a quasi-LPV model
C_z	Performance output matrix
C_1, C_2	Output matrices of multi-objective system
$\mathcal{C}_1, \mathcal{C}_2$	Closed-loop output matrices of multi-objective system
d	Rotor diameter
d_D	Diameter of the disk
d_i	Input disturbance signal
$\tilde{d}_i$	Normalized input disturbance signal
d_{mb}	Outer diameter of the magnetic bearing
d_o	Output disturbance signal
$\tilde{d}_o$	Normalized output disturbance signal
d_S	Diameter of the shaft
D	Set of scaling matrices
D	Direct feed-through matrix
$\mathcal{D}$	Closed-loop direct feed-through matrix
$\mathbf{D}$	Augmented direct feed-through matrix
$\mathbb{D}$	Scaling matrix
$\tilde{D}$	Direct feed-through matrix of a quasi-LPV model
D_e	Output error direct feed-through matrix
D_i	Direct feed-through matrix of input disturbance
$D_{i/r}$	Internal and/or rotating damping matrix
$D^*_{i/r}$	Circulatory internal and/or rotating damping matrix
D_K	Controller direct feed-through matrix
$\tilde{\bar{D}}_m$	Modified reduced order modal damping matrix
D_n	Direct feed-through matrix of the noise

D_{qp}	Direct feed-through matrix from uncertainty input to uncertain dynamics
D_{qw}	Direct feed-through matrix from disturbance input to uncertain dynamics
D_s	Structural damping matrix
$\tilde{D}_s$	Modal structural damping matrix
$\tilde{\tilde{D}}_s$	Modified modal structural damping matrix
D_{uf}	Direct feed-through matrix of the control input filter
D_{un}	Direct feed-through matrix of the control input normalizing weight
$D_{\tilde{u}}$	Direct feed-through matrix of the control input weight
$D_{yu}, D_{yw}, D_{zu}, D_{zw}$	Direct feed-through matrix from control and disturbance input to measurement and performance outputs respectively
D_ω	Frequency dependent scaling matrix
$\mathcal{D}_1, \mathcal{D}_2$	Closed-loop direct feed-through matrices of multi-objective system
$D_{11}, D_{12}, D_{21}, D_{22}$	Direct feed-through matrices of multi-objective system
e	Error signal
$\tilde{e}$	Normalized and/or filtered error signal
e_{max}	Maximum eccentricity
e_{ub}	Eccentricity due to unbalance
E	Matrix variable
f	Function
$f_{a1}, f_{a2}, f_{a3}, f_{a4}$	Decoupled magnetic forces in bearing A
$f_{b1}, f_{b2}, f_{b3}, f_{b4}$	Decoupled magnetic forces in bearing B
$f_{A,x}, f_{A,y}, f_{B,x}, f_{B,y}$	Magnetic forces at bearing A and bearing B in orthogonal directions x and y
f_d	Perturbation force
F	Matrix variable
$\mathcal{F}$	Total force acting on the rotor
F_d	Disturbance force in Laplace domain
F_i	Magnetic bearing control force matrix
$\tilde{F}_i$	Magnetic bearing modal control force matrix

F_l	Lower linear fractional transformation
F_M	Magnetic force
$\tilde{F}_m$	Reduced order modal control force matrix
F_r	Magnetic force acting on the rotor at a bearing position
F_s	Magnetic bearing stiffness matrix
F_u	Upper linear fractional transformation
F_+	Magnetic force at positive direction
F_-	Magnetic force at negative direction
g	Gravitational constant
G	Physical plant
G	Matrix variable
$\mathcal{G}$	Gyroscopic matrix
G_a	Power amplifier current gain
G_s	Sensor gain
$\mathcal{H}$	Angular momentum
H_A	Magnetic field strength
i	Coil current
$\hat{\mathrm{i}}$	Peak value of sinusoidal current input to power amplifier
$i_{A,x}, i_{A,y}, i_{B,x}, i_{B,y}$	Control currents at bearing A and bearing B in orthogonal directions x and y
i_b	Bias current
i_c	Control current
$i_{c,max}$	Maximum control current
i_g	Gravity compensation current
i_{in}	Current input to power amplifier
i_{max}	Maximum current that can be provided before saturation
$i_{max,lin}$	Maximum current that can be provided in linear operation
$i_{max(hw)}$	Maximum current limit of the power amplifier
i_{out}	Current output from power amplifier
i_r	Reference current
i_0	Premagnetization current

i_+	Magnetic pole where the control current is in the same direction with the premagnetization current
i_-	Magnetic pole where the control current is in the opposite direction with the premagnetization current
I_p	Polar moment of inertia
I_r	Radial (transverse) moment of inertia
$[I]$	Moment of inertia tensor
$\mathfrak{I}$	Momentum
$\mathfrak{J}$	Mixed objective functional
J	Cost functional
$\tilde{k}$	Stiffness matrix of the rigid rotor
k_i	Current stiffness of electromagnetic bearing
k_M	Magnetic bearing constant
k_s	Spring stiffness of electromagnetic bearing
$k_{s,A}, k_{s,B}$	Isotropic spring stiffness of bearing A and bearing B
k_u	Motion induced voltage coefficient
K	Controller
$\mathcal{K}$	Set of controllers
K_b	Rotor stiffness matrix with mechanical bearings
$\tilde{K}_b$	Modal rotor stiffness matrix with mechanical bearings
K_f	Kalman filter
$\tilde{\bar{K}}_m$	Reduced order modal system stiffness matrix
K_r	Rotor stiffness matrix
$\tilde{\bar{K}}_r$	Modal system (rotor and magnetic bearing) stiffness matrix
K_{reg}	Feedback controller in linear quadratic regulator problem
K_ρ	Parameter dependent controller
l_A	Length of airgap in magnetic circuit
l_{Fe}	Total length of iron in magnetic circuit
L	Minimized matrix function in loop-shaping controller design
$\mathfrak{L}$	Coil inductance
L_D	Disk length
L_G	Open-loop gain

L_{mb}	Length of the magnetic bearing
L_S	Shaft length
$\mathfrak{L}_0$	Nominal coil inductance
m	Mass of the rigid rotor
$\bar{m}$	Mass for the rigid rotor weight acting in a radial axis direction
$\tilde{m}$	Inverse of mass matrix of the rigid rotor
m_{max}	Maximum allowable unbalance mass
$m_{ub,s}, m_{ub,c}$	Static and couple unbalance mass respectively
m_1, m_2, m_3, m_4	Elements of mass matrix of rigid rotor in bearing coordinates
M	Closed-loop nominal system
$\tilde{M}$	Modified closed-loop nominal system with robust performance
$\mathfrak{M}$	Parameter dependent linear matrix inequality
$\tilde{M}_s$	Modified closed-loop nominal system with robust stability
$M_{qp}, M_{qw}, M_{zp}, M_{zw}$	Closed-loop transfer matrices from uncertain dynamics and disturbance inputs to uncertainty and performance outputs
$\mathcal{M}$	Total Moment acting on the rotor
$\tilde{M}_m$	Reduced order modal mass matrix of the flexible rotor
M_r	Mass matrix of the flexible rotor
$\tilde{M}_r$	Modal mass matrix of the flexible rotor
$\mathcal{M}_{xyz}$	Moment evaluated in stationary coordinates
M_Δ	Closed-loop system in linear fractional representation form
M_ρ	Closed-loop system in linear parametrically varying structure
$\mathcal{M}_{\zeta\eta\xi}$	Moment evaluated in rotating coordinates
M_1, M_2	Closed-loop nominal systems for different objectives
n	Noise signal
$\tilde{n}$	Normalized and/or filtered error signal
n_c	Number of coil windings
n_{max}	Maximum amplitude of the noise signal
N	Closed-loop system with uncertain dynamics
O	Observability matrix
p	Input from uncertain dynamics
$\tilde{p}$	Normalized input from uncertain dynamics

p_0	Open-loop poles of the rotor at standstill
p_i	Poles of the flexible rotor
P	Generalized plant
$\tilde{P}$	Generalized plant with weights
P_i	Positive definite matrix variable
P_ρ	Parameter dependent generalized plant
$P_{yu}, P_{yw}, P_{zu}, P_{zw}$	Transfer matrices from control and disturbance inputs to measurement and performance outputs respectively
$\mathcal{P}$	Set of allowable parameters
$\dot{\mathcal{P}}$	Set of allowable rate of change of parameters
$\mathcal{P}^v$	The set of vertices of parameters
q	Uncertainty output
$\tilde{q}$	Normalized uncertainty output
$\tilde{q}_m$	Reduced order modal generalized displacement vector for the flexible rotor
q_r	Generalized displacement vector for the flexible rotor
$\tilde{q}_r$	Modal generalized displacement vector for the flexible rotor
Q	Positive semidefinite weighting matrix
Q_i	Amplitude of rotor displacement at each eigenfrequency
r	Rotor displacement
r_A	Rotor displacement at location bearing A
r_B	Rotor displacement at location bearing B
r_i	Reference input
$\tilde{r}_i$	Normalized and/or filtered reference input
$r_{A,x}, r_{A,y}, r_{B,x}, r_{B,y}$	Rotor displacement in orthogonal directions x and y at bearing locations A and B respectively
R	Positive definite weighting matrix
$\mathcal{R}$	Coil resistance
$\mathbb{R}$	Real numbers
$\mathbb{R}_e$	Extended real numbers
Re	Real part of a complex number
s	Airgap length between the rotor and magnetic bearing

s_A, s_B	Distance from the displacement sensors for bearings A and B to the center of gravity of the rotor respectively
s_0	Nominal airgap between the rotor and magnetic bearings
s_1	Airgap between the rotor and touch-down bearings
S	Matrix variable
$\mathcal{S}$	Sensitivity matrix function
S_i	Input sensitivity matrix function
S_o	Output sensitivity matrix function
$\mathcal{S}_{\boldsymbol{\Delta}}$	Set of block structures for uncertainty
t	Time variable
T	State transformation matrix
$\mathcal{T}$	Complementary sensitivity matrix function
T_i	Input complementary sensitivity matrix function
T_o	Output complementary sensitivity matrix function
T_1, T_2	Coordinate transformation matrices
u	Control input
$\tilde{u}$	Normalized and/or filtered control input
U	Matrix variable
$\mathcal{U}$	Set of unitary matrices
U_{app}	Applied voltage (power amplifier output)
U_{coil}	Coil voltage
U_r	Reference voltage
U_0	Rated voltage
v	Rotor velocity
v	Output signal
v_1, v_2	Perturbations acting on a loop
V	Lyapunov function
$\mathcal{V}$	Power spectral density matrix of measurement noise
w	Disturbance signal
$\tilde{w}$	Normalized disturbance signal
W	Matrix variable
$\mathcal{W}$	Power spectral density matrix of plant noise (disturbance)

W_A	Magnetic field energy
W_a	Weighting function on additive uncertainty
W_e	Weighting function on the error signal
W_i	Weighting function on the input disturbance signal
W_m	Weighting function on multiplicative uncertainty
W_n	Weighting function on the noise signal
W_o	Weighting function on the output disturbance signal
W_p	Weighting function on the uncertainty input
W_q	Weighting function on the uncertainty output
W_S	Weighting function on the sensitivity matrix function
W_T	Weighting function on the complementary sensitivity matrix
W_u	Weighting function on the control input signal
W_{uf}	Weighting function to filter the control input signal
W_{un}	Weighting function to normalize the control input signal
W_w	Weighting function on the disturbance input signal
W_z	Weighting function on the performance output signal
x	Augmented state vector
x_e	Error weighting function state vector
x_i	Input disturbance weighting function state vector
x_K	Controller state vector
x_r	Rotor state vector
x_u	Control input weighting function state vector
x_0	Initial state condition
X	Positive definite matrix variable
$\tilde{X}$	Fundamental matrix of state transition matrix
$\mathcal{X}$	Positive definite matrix variable for the closed-loop system
$\mathbb{X}$	Set of states of a dynamic system
X_c	Controllability gramian
xyz	Stationary rectangular coordinates triple
y	Controller input signal
y_m	Measurement signal

Y	Positive definite matrix variable
$\mathcal{Y}$	Positive definite matrix variable for the closed-loop system
Y_o	Observability gramian
z	Performance output signal
$\tilde{z}$	Normalized performance output signal
z_1, z_2	Performance output signals in a multi-objective system
Z	Positive definite matrix variable
$\mathcal{Z}$	Positive definite matrix variable for the closed-loop system
$\alpha, \alpha_1, \alpha_2, \alpha_3$	Real constant numbers
α_M	Angle between a pole and magnet centerline
β	Generalized H_2 performance index
γ	H_∞ performance index
γ_s	Relaxed H_∞ performance index
Γ	Symmetric performance matrix
δ	Parametric uncertainty
δk	Uncertainty on the bearing spring stiffness
Δ	Block uncertainty structure
$\tilde{\Delta}$	Normalized block uncertainty structure
Δ_e	Extended block uncertainty structure
Δ_f	Fictitious block uncertainty structure
ϵ	Arbitrarily small real constant
ε	Expectation operator
$\zeta\eta\xi$	Rotating rectangular coordinates triple
θ	Rotation angle around y axis
λ	Eigenvalue of a matrix
λ_{max}	Maximum eigenvalue
μ	Structured singular value
μ_A	Magnetic permeability of air
μ_{Fe}	Magnetic permeability of iron
μ_r	Magnetic permeability ratio

μ_0	Magnetic permeability in vacuum
ν	Rate of change of a parameter
$\underline{\nu}, \overline{\nu}$	Lower and upper bounds on the rate of change of a parameter
ρ	Parameter in a a linear parametrically varying system
$\underline{\rho}, \overline{\rho}$	Lower and upper bounds for the value of a parameter
ϱ	Spectral radius
$\overline{\sigma}$	Maximum singular value
Σ	Block diagonal matrix of singular values
τ	Period
ϕ	Rotation angle around z axis
ϕ_i	Mode shapes
ϕ_M	Modal matrix
ϕ_m	Reduced order modal matrix
Φ	State transition matrix
Φ_c	Magnetic flux
φ	Phase angle
χ	State vector
ψ	Rotation angle around x axis
ω	Frequency
ω_n, ω_p	Nutation and precession frequency respectively
ω_0	Cut-off frequency
Ω	Rotational speed of the rotor
Ω_{max}	Maximum rotational speed of the rotor
Ω_n	Nominal rotational speed of the rotor

LIST OF ABBREVIATIONS

AMB	Active Magnetic Bearing
ARE	Algebraic Riccati Equality
BRL	Bounded Real Lemma
CG	Center of Gravity
DLMI	Differential Linear Matrix Inequality
DSP	Digital Signal Processor
FEM	Finite Element Method
IQC	Integral Quadratic Constraint
KYP	Kalman-Yakubovich-Popov
LFR	Linear Fractional Representation
LFT	Linear Fractional Transformation
LHP	Left Half-plane
LMI	Linear Matrix Inequality
LPV	Linear Parametrically Varying
LQG	Linear Quadratic Gaussian
LQR	Linear Quadratic Regulator
LTI	Linear Time Invariant
LTV	Linear Time Varying
MIMO	Multi-input Multi-output
NS	Nominal Stability
PD	Proportional Derivative
PDLF	Parameter Dependent Lyapunov Function
PID	Proportional Integral Derivative
PMB	Passive Magnetic Bearing
PWM	Pulse Width Modulation
QS	Quadratic Stability
RHP	Right Half-plane
RMS	Root Mean Square
rpm	Revolutions Per Minute

RS	Robust Stability
SISO	Single-input Single-output

1. INTRODUCTION

The demands on rotating machinery with respect to speed, efficiency, precision, reliability and safety are continuously rising. However, the potential of rotors with conventional mechanical bearings are limited. Active control of bearing characteristics (i.e., stiffness, damping) is possible with the regulation of oil pressure as in the active hydrostatic bearings or with magnetorheological fluids in hydrodynamic bearings, and with hybrid bearing designs in roller bearings. However, power loss due to the friction between rotor and stator/lubricant fluid is considerably high in passive or active control of rotors supported on mechanical bearings.

In this respect, adaptation of techniques to exploit magnetic levitation to isolate the rotor from the stator or from any other power dissipating media is deemed to be an attractive option. Moreover, possibility of elimination of unbalance forces with free-free boundary conditions of the rotor and prevention of friction between the rotor and stator provide the means for very high speed of operation. Rotational speed of the rotor is constrained solely by the strength of the rotor materials, and especially by the strength of ferromagnetic rotor laminations and their fitting methods.

1.1. Overview

Magnetic bearings are increasingly used in rotating machinery due to the main advantage of contact free operation. As the rotation is frictionless, there is no wear, no lubrication requirement, and very high rotation speeds are possible. The service life of magnetic bearings is higher compared to conventional journal or ball bearings and life cycle cost with respect to maintenance and operation are potentially low. The latter is true despite the requirement of power during the operation, as due to the friction free environment the total power consumption of the active magnetic bearings is much lower than the conventional mechanical bearings.

Two basic magnetic bearing types can be distinguished: The active magnetic bearing (AMB) and the passive magnetic bearing (PMB). With passive magnetic bearings bearing forces result from pairs of permanent magnets bearings with opposite field directions producing mutual repulsion. Although this kind of magnetic support seems to be very simple, there are two major drawbacks when compared with an AMB. Firstly a complete control of six degree of freedom of a rigid body by uncontrolled magnetic forces is impossible. Secondly, the bearing characteristics of a PMB can not be changed easily during operation, as is the case in AMB.

AMBs are operated in closed-loop control system due to the fact that, at constant current, the attractive magnetic forces between the bearings and the rotor increase with decreasing gap nearly in quadratic form. This behavior resembles non-linear spring with negative stiffness which destabilizes states of equilibrium without appropriate control action. The structure of a typical system is shown in Figure 1.1 [33]. The aim is to

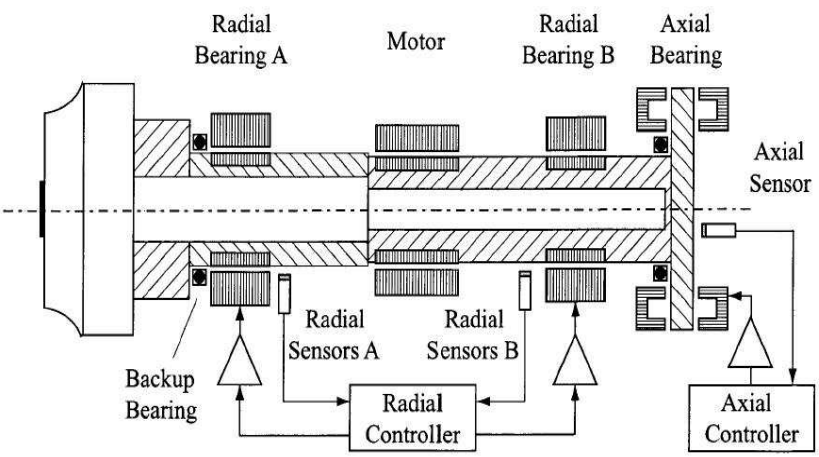

Figure 1.1. Rotor with active magnetic bearings.

hold the rotor in a certain position. The sensors measure the position and transmit the information to controller in form of voltage signal. Sensorless operation is also possible as the inductance of the bearing coils are a function of rotor to bearing radial clearance. A control voltage is then calculated in the controller based on a linear or non-linear control law and passed to the power amplifier which provides current to the electromagnets to apply forces on the rotor.

Active control to allow the rotor to rotate around its axis of inertia is used in unbalance compensation (see Appendix A) schemes to eliminate the unbalance force which might otherwise be transferred to the stator of the machine as mechanical vibrations. Elimination of the unbalance force not only creates a silent and reliable steady state operation, but reduces the resonant behavior of the rotor in the vicinity of bending mode critical speeds.

Usually retainer bearings, also called emergency bearings, auxiliary bearings or touch-down bearings are fitted for emergency operation in case of a problem in power supply, control system failure or temporary overload. These bearings prevent the rotor to contact with magnets or sensors which would otherwise cause severe damage to the machine.

Main advantages of rotor/AMB systems are: i) No mechanical wear and friction, ii) no lubrication therefore non-polluting, iii) low power consumption, iv) high circumferential speeds (more than 300 m/s), v) operation in severe/demanding environments, vi) easily adjustable bearing characteristics (stiffness, damping), vii) active control against disturbances and easy passing of critical speeds, viii) online balancing and unbalance compensation, ix) online system parameter identification, x) online condition monitoring for safe and reliable operation.

Due to the features mentioned above, it can be said that rotors suspended and controlled by AMBs are among typical *mechatronic* systems having interrelated electromagnetic, electronic, mechanical parts, and a controller, providing the system with some level of *artificial intelligence*.

1.2. Applications

Magnetic bearings were first used in 1970s for *satellite flywheels* and *gyrostabilizers*. The particular advantages of AMBs in these applications are low power consumption and long service life. Since these early days, magnetic bearings have become

increasingly popular for use in industrial machinery. One of the most successful applications is in *turbomolecular pumps*. They run at high speeds and produce a high vacuum which is very sensitive to contamination. Turbomachinery is another application area of magnetic bearings. *Turbogenerators, turbocompressors* and *turboexpanders* often run at over-critical speeds. Hence, the rotor has to pass at least one resonance region. Rotor displacement in these resonance regions can become quite large in the conventional passive systems. If magnetic bearings are used, however, active control can be utilized to reduce these high amplitude "vibrations" significantly resulting in increased safety and reliability. A particular interesting area of application is *high speed milling* and *grinding spindles* which often run at speeds in excess of 100,000 RPM. The rotor is not always centered in the stator but controlled to follow certain trajectories. This can be very useful in manufacturing nonsymmetric parts.

Electric motors using the stator coils, both for the generation of torque required for rotation and for the magnetic suspension/control of rotor position, instead of the conventional mechanical bearings, is another developing field of application. Such electric motors are called *bearingless drives*.

Besides the main applications mentioned above, magnetic bearings can also be found in *blood pumps, centrifuges, x-ray devices, computer hard disk drives* and in *gas turbines* in transport applications.

The main disadvantage of operation of rotors with AMBs is the relatively high complexity and the requirement of an integrated control system. Furthermore, there is a certain scepticism in the industry regarding this technology which is proving hard to overcome.

1.3. Current State of Research

Electromechanical design of standard magnetic bearing (homopolar or heteropolar with 8 poles in horseshoe structure) for linear control is well established and offers limited scope for improvement. In order to achieve more performance (increasing spe-

cific load capacity, reducing eddy current losses etc.) more complex designs are being developed. Most research, however, focuses not on the bearing hardware but on the controller algorithms. With intelligent control, the desired bearing characteristics can be optimized for the particular application. Milling and grinding spindles, for example, must be positioned precisely and must be insensitive to disturbances from external loads. Bearings for turbomachinery should have a large stiffness and damping, in order to reduce the whirl amplitudes of the rotor due to wide range of excitations ranging from synchronous unbalance forces to aerodynamic influences. In transport applications base motion must be considered, and reliability in the priority in aerospace systems. For turbomolecular pumps vibrations transmitted to the base are often minimized. This can be achieved by letting the rotor to rotate around its principle axis of inertia (see Appendix A). The same applies to flywheels where power consumption of the bearings is an important design criteria.

Decentralized PD/PID controllers are widely used in commercial systems as they are simple to implement and exhibit sufficient performance. Much research has been undertaken to enhance PD controllers with the objective of minimum displacements, especially in the resonance zones. A common strategy is to apply speed synchronous forces through active magnetic bearings which oppose the unbalance forces. The amplitude and phase of compensation forces is calculated on-line based on a dynamic model of the system and the displacements in the rotating frame. Burrows et al. [7] describe this method.

Another strategy to reduce resonant vibrations during the operation close to critical speeds is based on the fact that resonance regions shift depending on the bearing stiffness (only rigid body modes). Zhang [8] showed how optimization techniques can be used to compute an adaptive controller which adjusts the bearing parameters so that rotor operates in antiresonance zone at all running speeds up to the 4th eigenfrequency (2 distinct eigenfrequencies in case of isotropic bearings) of the system. In order to achieve a force free operation, notch filters are often inserted into the close loop. Herzog et al. [9] presented a design approach and verified the results with experiments in a large turboexpander. Though they have studied the system at a constant rotor speed,

filtering can be accomplished based on the tacho signal for variable speed of operation.

Decentralized PD control has inherent limitations due to limitations on defining the performance criteria for the controllers. Moreover, SISO PD/PID control can not handle the cross-coupling due to gyroscopic effects and it can not stabilize the flexible modes of the rotor. Therefore, current research is concerned with model based centralized controllers and, especially, the application of modern robust control techniques in order to improve reliability and performance of rotor/AMB systems.

The availability of affordable high performance computers was a driver for the development of robust control theory in the 1980s. One of the most popular robust control strategies is based on H_∞ optimization. In their pioneering work, Glover and Doyle [10,11] showed that state-space descriptions of all stabilizing controllers can be found by solving two algebraic Riccati equations. In the 1990s the first H_∞ algorithms were developed in which the design constraints are reformulated using linear matrix inequalities (LMIs). This procedure is applicable to a wider range of plants and, it can be computationally advantageous by utilizing efficient convex optimization algorithms. Explicit solutions to LMI H_∞ problem are given by Apkarian and Gahinet [12,13].

An interesting application of LMI methods is the multi-objective feedback control. This combines the advantages of H_∞ method such as robustness to model uncertainty and explicit frequency domain specifications with the strengths of H_2 method in handling stochastic aspects. Furthermore, transient responses and minimum closed loop damping can be specified by time domain and pole placement constraints. Scherer et al. [14,21] show how these design objectives can be formulated as LMI constraints used to synthesize a controller.

LMI framework allows the derivation of gain scheduled controllers, which increase robustness and/or performance for feedback systems with linear parameter dependent plants. In robust LTI H_∞ design, time varying plant characteristics are treated as uncertainty, which often leads to overly conservative controllers. If the state-space matrices of the plant are fixed functions of time varying physical parameters, which

can be continuously measured, LMI synthesis methods allow the computation of gain scheduled controllers as functions of measured parameters. The derivations of LMI controller synthesis for two classes of parameter varying plants are shown in the papers of Packard, Becker, Apkarian and Gahinet [15-18]. These techniques offer great potential for rotor/AMB systems as their dynamics depend on the rotational speed due to gyroscopic effects. Simulation studies of systems with small rigid rotors have been performed by Sivrioğlu and Nanomi [19] and Tsiotras and Mason [6]. In the standard LMI gain scheduling approach, the time varying parameter(s) is allowed to change at an arbitrary fast rate. In rotordynamics applications, however, the rotor angular acceleration is limited by the available maximum torque of the driver. Tsiotras and Knospe [20] show a priori known bounds of the parameter change rate can be considered in the controller design to improve the closed loop performance using the results of the paper by Apkarian and Adams [5].

In linear control of rotor/AMB systems, linearization of the actuator forces requires bias currents continuously supplied to the bearings, and a considerable part of the consumed power is wasted. Therefore nonlinear control of rotor/AMB systems, without bias current, is considered to be a viable option and attracts the attention of many researchers. Back-stepping techniques, sliding mode control and nonlinear adaptive control in conjunction with zero or very low bias current application has been investigated by several authors [22-24]. The benefit of non-linear control in rotor/AMB systems is significantly reduced power consumption hence, it is a promising approach for the future. However, many issues are still left to be resolved, such as how excitation of weakly damped uncertain flexible modes can be avoided.

Uncertainty in the system dynamics and system parameters have a crucial importance in control system design and they can be explicitly formulated in linear robust control techniques. Most common significant sources of uncertainty in rotor/AMB systems are high frequency dynamics (flexible modes) of the rotor, changing dynamics due to gyroscopic effect and varying displacement (spring) stiffness of the magnetic bearing. Effect of gyroscopic forces on uncertain dynamics can be eliminated by appropriate modeling and control [6,19], but AMB displacement stiffness (which can not

directly be measured) is a major source of uncertainty. This parameter is very sensitive to static load acting on the rotor and also anything that modifies the nominal air gap in the actuator including; manufacturing tolerances, thermal growth, centrifugal growth, and vibration (whirling) due to unbalance [33]. Uncertain parameters considered in the literature for robust control of rotors with active magnetic bearings are the rotor speed and high frequency dynamics. In our approach we will also consider the displacement stiffness of the bearing to be uncertain.

In the literature performance objectives are defined in terms of L_2 gain (H_∞ norm for LTI systems). This norm maintains good robust performance. On the other hand, H_∞ control is mainly concerned with frequency domain performance and does not guarantee good transient response for the closed loop system. H_2 control gives more suitable performance on the behavior of the system for transients. Combining H_2 and H_∞ control would enhance the robustness of the system taking into account both of the requirements on the time and frequency response. We will use mixed performance control, with a signal-based interpretation for LPV systems, to enhance the overall performance of the closed-loop system incorporating the requirements on time and frequency response simultaneously.

1.4. Objective of the Thesis

Industrial applications of rotor/AMB systems are continuously increasing due to their benefits as mentioned in the previous sections, and vibration control of structures such as rotors, using robust control techniques, is one of the main active research areas in control science and engineering.

Uncertain parameters considered in the literature for robust control of rotors with active magnetic bearings are the rotor speed and high frequency dynamics. In our approach we will also consider the displacement stiffness of the bearing to be uncertain. *First objective of the thesis is to analyze and establish the limits of robust stability for the system with the uncertainty structure comprised of the rotor speed and*

displacement stiffness of the bearing.

Generally, time domain specifications like LQG or H_2 control and frequency domain specifications like H_∞ control are mostly considered as valuable criterions in active vibration control. In practice, both specifications are important for a successful design. If there are uncertainties in the model (as in the case for rotor/AMB systems) H_∞ control maintains good robust performance. On the other hand, H_∞ control is mainly concerned with frequency domain performance and does not guarantee good transient response for the closed loop system. H_2 control gives more suitable performance on the behavior of the system for transients. Combining H_2 and H_∞ control will enhance the robustness of the system taking into account both of the requirements on the time and frequency response. However, using H_2 norm in the synthesis makes the controller susceptible to noise from the measurement sensors, especially due to the mechanical and electrical run-out (glitch) on the laminations at the rotor journal surface. Therefore, compensation for the run-out should be considered in the design of the controller, and the corresponding displacements must be subtracted from the measurements before they are fed into the controller.

Mixed H_2/H_∞ control using convex optimization involving LMIs, and multi-objective control using mixed H_2/H_∞ specifications are described in references [14,21]. *Second objective of the thesis is to derive and study the multi-objective H_2/H_∞ output feedback control of the disturbance rejection problem for LPV structured rotor/AMB systems.* Robust stability of the system and minimizing the displacements of the rotor at the measurement locations with minimum control effort is achieved with H_∞ performance. Keeping the H_∞ norm as an additional constraint, another convex optimization problem is simultaneously designed, with possible relaxations on the H_∞ norm, to minimize the rotor displacements with H_2 performance.

2. SYSTEM DYNAMICS

2.1. Actuator Dynamics

2.1.1. Electromagnetic Forces

The AMB model presented in this section is based on zero leakage assumption which says that magnetic flux in a high permeability magnetic structure with small air gaps is confined to iron and gap volumes. Hysteresis effects are neglected and saturation effects are approximated. These simplifications give rise to some errors but results are usually sufficiently accurate in the operating range of the bearing.

The electromagnetic forces of a "horse shoe" magnet can be derived from energy conservation. The magnetic field energy in an air gap with a volume V_A is

$$W_A = \frac{1}{2}\int_{V_A} B_A H_A \, dV_A \,, \tag{2.1}$$

where the magnetic flux density B_A and magnetic field strength H_A can be calculated using magnetic circuit analysis as follows:

$$B_A = \frac{n_c i}{\frac{l_{Fe}}{\mu_{Fe}}\frac{A_A}{A_{Fe}} + \frac{l_A}{\mu_A}} \quad \text{and} \quad H_A = \frac{n_c i}{\frac{\mu_A}{\mu_{Fe}}\frac{A_A}{A_{Fe}} l_{Fe} + l_A} \cdot$$

The magnetic force F_M results from the change of magnetic field energy in the gap $s = \frac{1}{2} l_A$. It is a potential force given by

$$F_M = -\frac{\partial W_A}{\partial s} = \frac{\mu_0 A_A n_c^2}{4} \cdot \frac{(s - \frac{l_{Fe}}{2\mu_r}) i^2}{(s + \frac{l_{Fe}}{2\mu_r})^3} \tag{2.2}$$

where it is assumed that $A_A = A_{Fe}$, $\mu_A = \mu_0 = 4\pi \cdot 10^{-7} \frac{N}{A^2}$, and $\mu_r := \mu_{Fe}/\mu_0$. Equation (2.2) can only be solved if the non-linear relationship between the permeability

and the magnetic field strength of the iron core is known.

If the bearing is operating below the saturation point, however, μ_r is large (>1000) and (2.2) can be simplified to

$$F_M = \frac{\mu_0 A_A n_c^2}{4} \cdot \left(\frac{i}{s}\right)^2 . \tag{2.3}$$

If the magnets in a radial magnetic bearing are arranged as in Figure 2.1, the forces in the z and y directions are almost decoupled and can be calculated separately. Two opposing magnets in two orthogonal directions cause the force

$$F_r = F_+ - F_- = k_M \left(\left(\frac{i_+}{s_0 - r}\right)^2 - \left(\frac{i_-}{s_0 + r}\right)^2 \right)$$

on the rotor where s_0 is the nominal air gap length and r is the displacement of the rotor as shown in Figure 2.1 [33]. The magnetic bearing constant k_M is

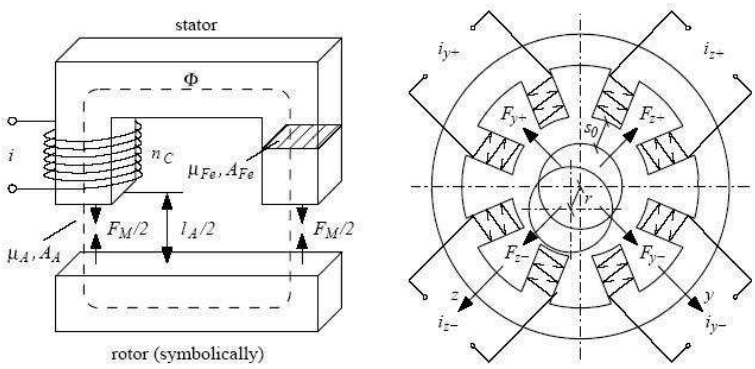

Figure 2.1. Electromagnetic forces.

$$k_M := \frac{\mu_0 A_A n_c^2}{4} \cos \alpha_M$$

with α_M denoting the angle between a pole and magnet centerline.

The non-linearities of the magnetic force are generally reduced by adding a high bias current i_b to the control current i_c, and in some applications the gravitational forces are compensated by a constant current i_g. The resulting actuator currents are, therefore, given by

$$i_+ = i_b + i_g + i_c \quad and \quad i_- = i_b - i_g - i_c \,,$$

which leads to

$$F_r = F_+ - F_- = k_M \left(\left(\frac{i_b + i_g + i_c}{s_0 - r} \right)^2 - \left(\frac{i_b - i_g - i_c}{s_0 + r} \right)^2 \right) \cdot \tag{2.4}$$

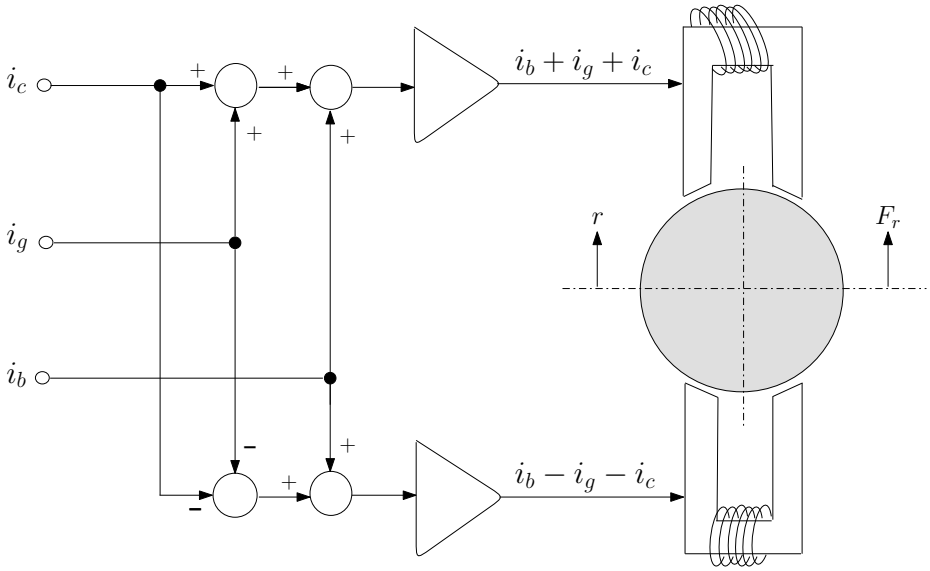

Figure 2.2. Currents in one axis.

The magnetic bearing force on the rotor can be linearized around the operating point to read

$$F_r \cong F_r|_{OP} + \left. \frac{\partial F_r}{\partial i} \right|_{OP} (i_c - i_{cOP}) + \left. \frac{\partial F_r}{\partial r} \right|_{OP} (r - r_{OP}) \cdot$$

Hence, with $i_{cOP} = 0$ and $r_{OP} = 0$, the linearized magnetic bearing force of the

bearing for small currents and small displacements is given by

$$F_{r,lin} = F_{r,g} + k_i i_c - k_s r \,, \tag{2.5}$$

with the actuator gain k_i and the open loop negative stiffness k_s defined to be

$$k_i := 4k_M \frac{i_b}{s_0^2} \quad \text{and} \quad k_s := -4k_M \frac{i_b^2 + i_g^2}{s_0^3} \,.$$

The constant force $F_{r,g}$ vanishes if gravity compensation is not considered separately. It is given by

$$F_{r,g} = 4k_M \frac{i_b i_g}{s_0^2} \,.$$

In principle, the stiffness of each axis can be adjusted individually by means of the bias current i_b and gravity compensation current i_g. However, in almost all practical applications, *premagnetization current* in each axis is designed to be identical to provide the means for constant bearing stiffness in opposing directions. Otherwise (the case of nonlinear and anisotropic stiffness or axially unsymmetrical rotor), backward whirl modes of the rotor can be also be excited by the disturbance forces during the operation. This would create an additional burden for the system as backward whirling must also be suppressed by the damping forces to be created at the magnetic bearings. Therefore, instead of providing constant currents i_b and i_g separately, premagnetization current i_0 is provided identically to opposing directions in all axes, which is large enough to cover both the bias and gravity compensation currents.

Magnetic flux density B_A has a saturation point for ferromagnetic materials. Therefore, magnetic force that can be applied to a rotor has a limit as given by the

well-known Maxwell stress equations

$$F_{M,max} = \frac{B_{A,sat}^2 A_A}{2\mu_0} \cdot$$

Hence, increasing the current beyond this magnetic flux density saturation point will be useless. Any controller output demanding currents above this value will lead to actuator saturation which changes the closed loop system dynamics significantly. Replacing $F_{M,max}$ into (2.3), maximum current at which saturation will take place is approximately

$$i_{max} = \left(\frac{2}{\mu_0}\right)^{\frac{1}{2}} \frac{s_0}{n_c} B_{A,sat} \cdot$$

Usual practice is to limit the rated current output of the amplifier to an artificial hardware limit $i_{max(hw)}$, which is below the reachable maximum current at steady state.

2.1.2. Power Amplifier and AMB

Active magnetic bearings are generally fed by a pulse width modulation (PWM) type half bridge power amplifiers. The half bridge is capable of driving currents in only one direction. This is sufficient because the required magnetic force in the bearing is independent of current direction. Main advantage of PWM over the conventional inverters is the significant reduction on power losses and less non linearity in amplifier dynamics. These amplifiers operate in *voltage mode* or *current mode*. In voltage mode, feedback is used to ensure that applied voltage $U_{app}(t)$ tracks a reference voltage $U_r(t)$. In current mode the voltage $U_{app}(t)$ is adjusted so that coil current $i(t)$ follows a reference current $i_r(t)$. The coil dynamics are derived from Faraday's and Kirchhoff's voltage law. The physical coil is modeled by an ideal coil (no resistance) in series with a resistor that represents the distributed coil resistance. Of course this is a lumped parameter model and one can not actually consider the coil without considering

its resistance as well. Moreover, there is always a voltage drop across the coil and consequently, the voltage U_{coil} that appears across the ideal coil must be inferred through knowledge the applied terminal voltage U_{app} and $i\mathcal{R}$ drop. By Kirchhoff voltage law

$$U_{app} = i\mathcal{R} + U_{coil}\,.$$

Coil voltage is related to the rate of change of flux by Faraday'law, $U_{coil} = n_c\dot{\Phi}_c$. Thus, the coil dynamics are given by

$$U_{app} = i\mathcal{R} + n_c\dot{\Phi}_c\,.$$

Since there is generally a nonlinear relationship between the current through the coil and the generated flux (due to hysteresis), one may choose the current or the flux as the electrical state and eliminate the other variable in the equation. However, in magnetic circuits with airgaps, the airgap reluctance often dominates the reluctance of the core. Consequently, the hysteresis nonlinearity is less dramatic and is well-approximated with a saturation like airgap dependent function $\Phi_c = f(n_c i, s)$. Furthermore, if the core is not excited into its saturation region, an airgap dependent inductance may be assumed and the linear relationship between the current and flux may hold, $n_c\Phi_c = \mathcal{L}(s)i$.

In PWM amplifiers, the demanded set current is created by means of a very quick switching among a positive or negative voltage (d.c.) that is applied to the coil. If the coil current is too small, positive U_0, if it is too large, negative U_0 is applied. The switching takes place at very high frequency (e.g., 80 kHz), which is far above the sampling rate and/or bandwidth of the AMB control. However, this procedure has two consequences: First the set current value is not exactly achieved but the true current oscillates around this value. This causes remagnetization losses in the system. More importantly, due to the finite voltage is used in the amplifier, current can not

be made rise or fall arbitrarily fast. Given the output voltage rating U_0 of the PWM amplifier, and the coil inductance $\mathcal{L}$ and the resistance $\mathcal{R}$ of the magnetic bearing coils, the equation relating voltage and current in the coil is

$$U_0 = \frac{d}{dt}\big(\mathcal{L}(s)i\big) + \mathcal{R}i\,.$$

Neglecting the coil resistance and assuming that inductance of the coil is constant (actually it is a function of air gap), maximum change possible for the current is

$$\left(\frac{d}{dt}i\right)_{max} = \frac{U_0}{\mathcal{L}}\,.$$

Assuming a sinusoidal controller output acting as an input i_{in} for the current amplifier

$$i_{in} = \hat{\mathrm{i}}\sin(\omega t) \quad \Rightarrow \quad \frac{d}{dt}\,i_{in} = \omega\,\hat{\mathrm{i}}\cos(\omega t)\,.$$

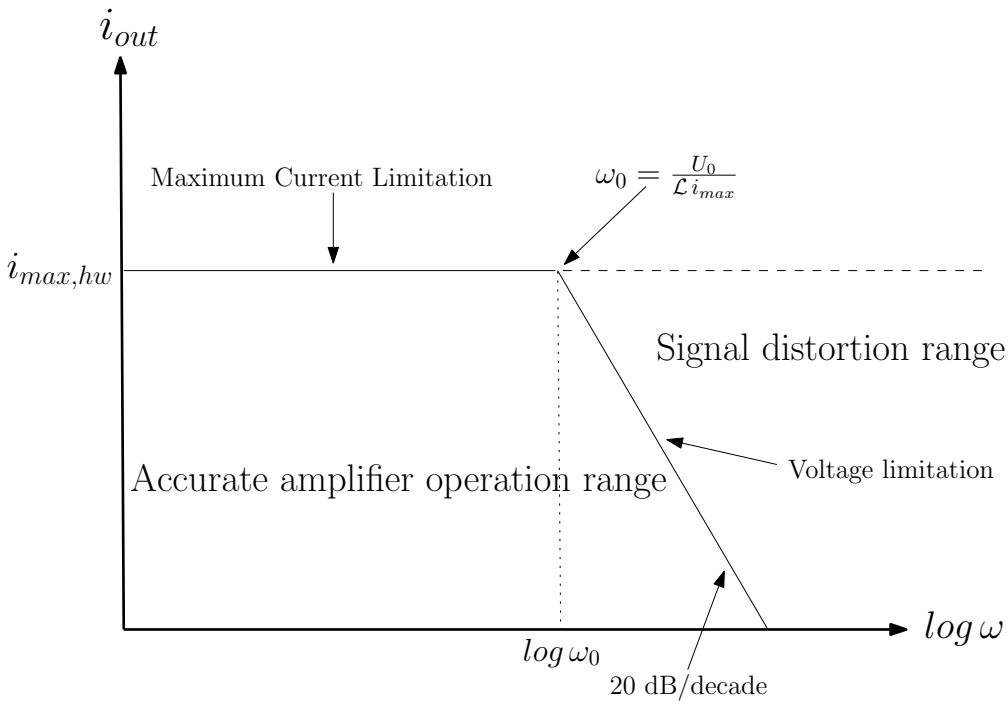

Figure 2.3. Current limits due to amplifier saturation.

The amplifier will only be capable of producing the corresponding current output, i_{out}, if the following condition is met:

$$\omega \hat{\mathrm{i}} \leqslant \frac{U_0}{\mathcal{L}} \quad \Leftrightarrow \quad \hat{\mathrm{i}} \leqslant \frac{U_0}{\omega \mathcal{L}} \,.$$

Hence, in a well defined operation range, the switched type PWM current amplifiers in conjunction with the load of the magnetic bearing work without any remarkable signal distortion nor phase lag, where as, depending on the frequency and magnitude of current signal to be reproduced, important signal distortion may occur.

In order to include the actuator rate saturation effect to the linear force model and avoiding distortion in current signal, gain reduction of the actuator in high frequencies can be approximated as a first order low-pass filter in the *Laplace domain*

$$F_{(Mi,lin-lp)}(s) = F_{(Mi,lin)}(s) \frac{\omega_0}{s + \omega_0}$$

where $F_{(Mi,lin)}$ is given by the second term in (2.5). Cut-off frequency, ω_0, is approximately equal to $U_0/(\mathcal{L}\, i_{max,lin})$, or can be determined by making experiments.

2.2. Rotordynamics

Based on the force model from (2.5) and bearing in mind the limitations of the actuator as discussed in the previous section, magnetic bearing forces can be linearly modeled. Without gravity compensation (premagnetization current i_0 can be made large to also cover the gravity compensation current i_g) and without control current, the linear equation of motion of a single mass supported by magnetic bearing perturbed by an arbitrary linear force f_d is

$$m\ddot{r} = -k_s r + f_d \,.$$

Based on this equation, *Laplace transform* can be used to derive the following one degree of freedom model

$$\mathrm{R}(s) = \frac{(1/m)}{s^2 + (k_s/m)}\mathrm{F}_d(s)\,.$$

This implies that a single mass supported by AMB, in one-dof, is a second order system with the poles and a static gain

$$p_0 = \pm\sqrt{\frac{-k_s}{m}}\ , \quad \text{Gain} = \frac{1}{k_s}\ \cdot$$

From the existence of a pole on the positive real axis (note that k_s is a negative number), it can directly be concluded that the system is unstable.

Dynamic mechanical systems always have more than only one degree of freedom. In fact, such systems have an infinite number of degrees of freedom as they must be considered as a *continuum*. While there are analytic methods for dealing with continuous structures, it turns out to be impractical to deal with infinite number of degrees of freedom. It is a much more suitable and practical approach to first get clarity about how many degrees of freedom must be considered to be able to describe the physical effects one is interested in. Hence, for example, if one is interested in looking at rigid body behavior of a structure, it is enough to consider six degrees of freedom that the rigid body have in space. Considering only the radial (lateral) motion of the rotor, four degrees of freedom is sufficient as the remaining degrees of freedom are controlled by the axial magnetic bearing and the driving motor/turbine.

2.2.1. Rigid Rotor with AMBs

A rigid rotor without control is shown in Figure 2.4. Axial support is not taken into account. This is justified by negligible coupling between radial and axial dynamics due to the fact that, in most of the cases, axial dynamics can be described by the one-dof model as shown above. Radial rotor motion in one plane, without rotation, can be

completely described by a rigid beam model, i.e., by the displacement x at the rotor center of gravity, CG, and the rotation θ of the rotor about an axis through CG.

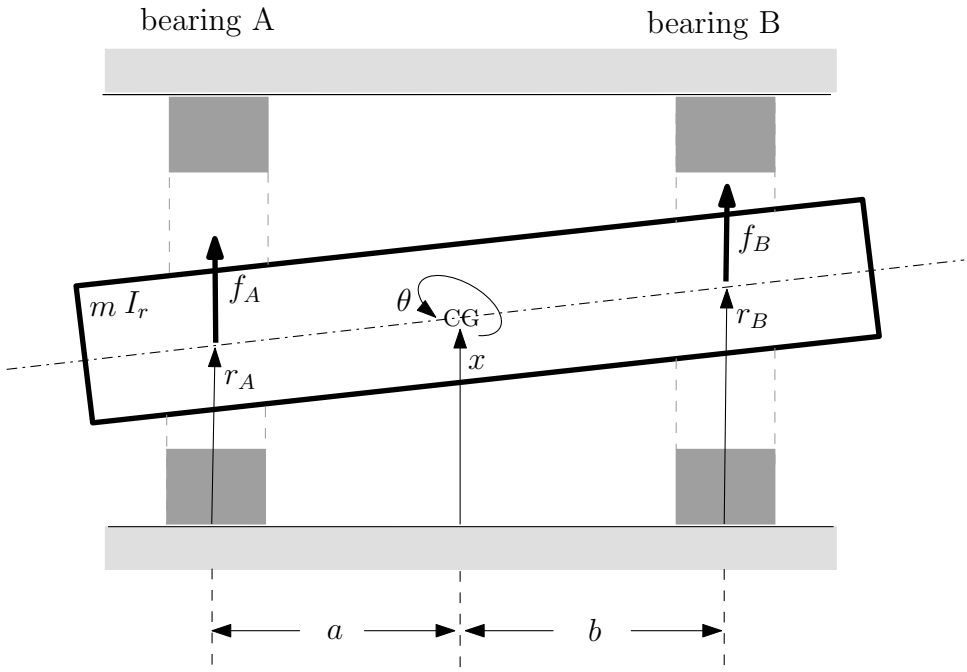

Figure 2.4. Rotor at standstill.

The equation of motion of a non-rotating rotor in two dimensions is

$$\begin{bmatrix} m & 0 \\ 0 & I_r \end{bmatrix} \begin{pmatrix} \ddot{x} \\ \ddot{\theta} \end{pmatrix} = -T_1 \begin{bmatrix} k_{s,A} & 0 \\ 0 & k_{s,B} \end{bmatrix} \begin{pmatrix} r_A \\ r_B \end{pmatrix} + T_1 \begin{bmatrix} k_{i,A} & 0 \\ 0 & k_{i,B} \end{bmatrix} \begin{pmatrix} i_A \\ i_B \end{pmatrix} - \begin{pmatrix} \bar{m}g \\ 0 \end{pmatrix}$$

where i is the rotor mass, I_r is rotor's radial moment of inertia, $k_{s,A}$ and $k_{s,B}$ are the force-displacement factors of bearing A and bearing B, and $k_{i,A}$ and $k_{i,B}$ are corresponding force-current factors. The matrix T_1 corresponds to the transformation matrix of forces generated by the bearings to the center of gravity of the rotor and is given by

$$T_1 := \begin{bmatrix} 1 & 1 \\ -a & b \end{bmatrix}.$$

The term $\bar{m}g$ stands for the gravitational force acting on the rotor. For the vertical plane (xz), $\bar{m}g = mg$, and for the horizontal plane (yz), $\bar{m}g = 0$. If the two planes are rotated by 45°, with respect to horizontal/vertical (typical case), $\bar{m}g = mg/\sqrt{2}$ for

both planes. For vertical rotors $\bar{m}g = 0$ for both radial bearing planes as gravity is acting on the axial direction.

System can completely be transferred to the bearing coordinates using the transformation

$$\begin{pmatrix} x \\ \theta \end{pmatrix} = T_2 \begin{pmatrix} r_A \\ r_B \end{pmatrix}, \qquad T_2 := \frac{1}{a+b} \begin{bmatrix} b & a \\ -1 & 1 \end{bmatrix}.$$

The transformed system is then of the form

$$\begin{bmatrix} m_1 & m_3 \\ m_3 & m_2 \end{bmatrix} \begin{pmatrix} \ddot{r}_A \\ \ddot{r}_B \end{pmatrix} = - \begin{bmatrix} k_{s,A} & 0 \\ 0 & k_{s,B} \end{bmatrix} \begin{pmatrix} r_A \\ r_B \end{pmatrix} + \begin{bmatrix} k_{i,A} & 0 \\ 0 & k_{i,B} \end{bmatrix} \begin{pmatrix} i_A \\ i_B \end{pmatrix} - T_2 \begin{pmatrix} \tilde{m}g \\ 0 \end{pmatrix},$$

where

$$m_1 := \frac{mb^2 + I_r}{(a+b)^2}, \quad m_2 := \frac{ma^2 + I_r}{(a+b)^2}, \quad m_3 := \frac{mab - I_r}{(a+b)^2}.$$

The poles of the system above can be calculated considering only the homogeneous equation (without external forces from the gravity or coil currents). Defining the matrix A and the state variable χ as

$$A = - \underbrace{\begin{bmatrix} m_1 & m_3 \\ m_3 & m_2 \end{bmatrix}}_{\tilde{m}}^{-1} \underbrace{\begin{bmatrix} k_{s,A} & 0 \\ 0 & k_{s,B} \end{bmatrix}}_{\tilde{k}}, \qquad \chi = (r_A \;\; r_B)^T,$$

the state-space form is

$$\frac{d}{dt} \begin{pmatrix} \chi \\ \dot{\chi} \end{pmatrix} = \begin{bmatrix} 0 & I \\ A & 0 \end{bmatrix} \begin{pmatrix} \chi \\ \dot{\chi} \end{pmatrix}. \tag{2.6}$$

Eigenvalues of the matrix A are the same as the matrix $\tilde{A}$ which can be defined as: $\tilde{A} :=$

$-\tilde{m}^{-1/2}\,\tilde{k}\;\tilde{m}^{-1/2}$. Matrix $\tilde{A}$ is symmetric with positive diagonal elements, therefore its eigenvalues are real and positive. Denoting the eigenvalues of matrix $\tilde{A}$ or equivalently the eigenvalues of matrix A as; λ_1 and λ_2, eigenvalues of the rotor/AMB system are $\pm\sqrt{\lambda_1}$ and $\pm\sqrt{\lambda_2}$. Therefore, uncontrolled rotor/AMB system in one plane (xz plane) has four rigid body poles on the real axis of which two are negative and two are positive. Considering the motion at the plane (yz plane) perpendicular to the first one analyzed above with the assumption of axially symmetric rotor and *isotropic* bearing stiffness, there are eight poles in total, each of the four poles to occur twice. Due to the fact that half of the poles are on the right half plane, non-rotating rotor/AMB system modeled as a rigid body, without control, is unstable.

When the rotor starts to rotate, motion in one plane becomes coupled with the motion of the other perpendicular plane due to the gyroscopic effects. Moreover, inherent mass unbalance of the rotor with respect to axis of rotation creates rotating unbalance forces or displacements (if unbalance compensation is used) which must also be taken into account in the synthesis for a stabilizing controller.

Motion of the rotating rotor can be described using the *Newton's Second Law*:

$$\mathcal{F} = \dot{\mathcal{J}} = \frac{d}{dt}(m\mathrm{v})\,, \tag{2.7}$$

and

$$\mathcal{M} = \dot{\mathcal{H}} = \frac{d}{dt}([I]\omega)\,, \tag{2.8}$$

where m is the mass of the rotor, v is the velocity of the mass center, $\mathcal{J}$ and $\mathcal{H}$ are the momentum and angular momentum vectors respectively, $[I]$ is the inertia tensor, ω is the angular velocity of the rotor, and $\mathcal{F}$ and $\mathcal{M}$ are the total external force and moment about the mass center respectively. Equation (2.8) states that; *the instantaneous time rate of change of the rigid body's angular momentum* $\mathcal{H}$ *is equal to the sum of the instantaneous moments* $\mathcal{M}$ *exerted upon the rigid body.* However, due to rotation (inclination) and spin of the rotor, inertia tensor $[I]$ changes in time with respect to

stationary (inertial) coordinates xyz. Therefore, differentiation in stationary coordinates is complicated. Instead, a rotating coordinate system $\zeta\eta\xi$ attached to the center of gravity of the rotor can be used to evaluate (2.8) as follows:

$$\left(\frac{d}{dt}\right)_{\zeta\eta\xi}([I]\omega) + \omega \times [I]\omega = \mathcal{M}_{\zeta\eta\xi}\,. \tag{2.9}$$

Assuming that uneven mass distribution due to unbalance masses is negligible, the

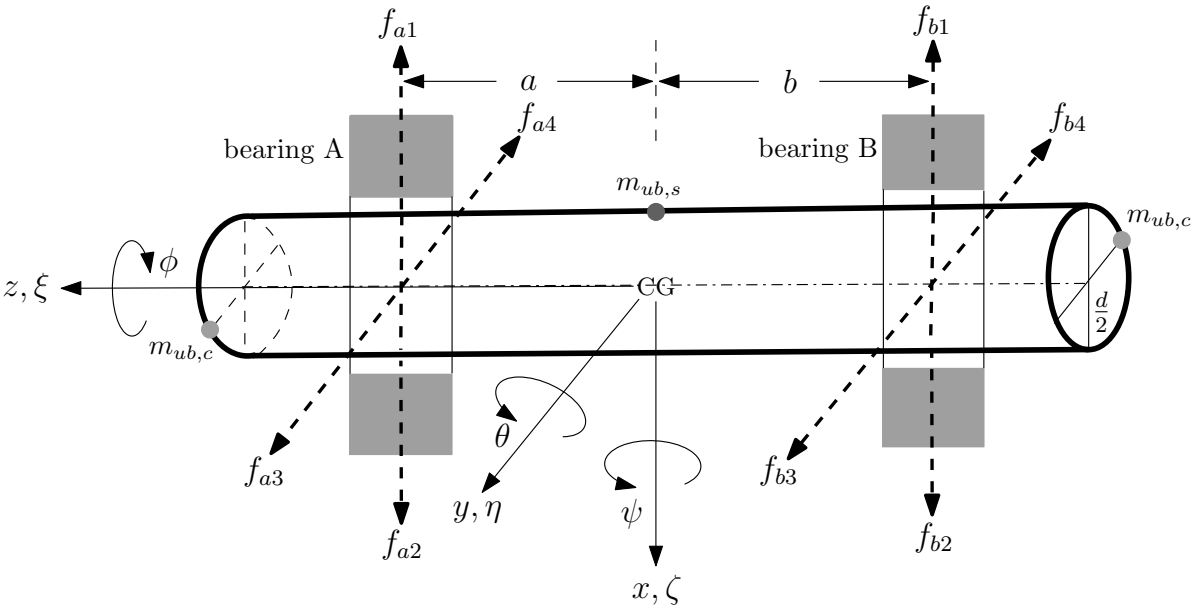

Figure 2.5. Rotating rigid rotor.

inertia tensor in rotating coordinates and angular velocity vector are given by

$$[I]_{\zeta\eta\xi} = \begin{bmatrix} I_r & 0 & 0 \\ 0 & I_r & 0 \\ 0 & 0 & I_p \end{bmatrix} \quad \text{and} \quad \begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} = \begin{pmatrix} \dot{\theta}\sin\phi + \dot{\psi}\cos\phi \\ \dot{\theta}\cos\phi - \dot{\psi}\sin\phi \\ \dot{\phi} + \dot{\psi}\sin\theta \end{pmatrix}, \tag{2.10}$$

where I_r and I_p are the principle radial and polar moment of inertia of the rotor respectively. The angles ϕ, θ and ψ are denoting the orientations of body fixed rotating frame, $\xi\eta\zeta$, about the stationary, zyx, frame respectively (3-2-1 Euler angles).

Since the angular deviations ψ and θ are typically very small, we have $\cos\theta \approx 1$

and $\sin\theta \approx 0$ and (2.10) simplifies to

$$\begin{pmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \end{pmatrix} = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} \dot{\psi} \\ \dot{\theta} \\ \dot{\phi} \end{pmatrix} . \tag{2.11}$$

Similarly, moment vector $\mathcal{M}$ in rotating coordinates is

$$\mathcal{M}_{\zeta\eta\xi} = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix} \mathcal{M}_{xyz} . \tag{2.12}$$

Replacing (2.11) and (2.12) into (2.9), leads to the following equations at the first two rows:

$$\ddot{\psi} = -\Omega\frac{I_p}{I_r}\dot{\theta} + \frac{a}{I_r}(-f_{A,y}) + \frac{b}{I_r}(f_{B,y}) + \frac{(a+b)}{2I_r}m_{ub,c}\,\Omega^2 d\sin(\Omega t + \varphi_c) \tag{2.13}$$

$$\ddot{\theta} = \Omega\frac{I_p}{I_r}\dot{\psi} + \frac{a}{I_r}(f_{A,x}) + \frac{b}{I_r}(-f_{B,x}) - \frac{(a+b)}{2I_r}m_{ub,c}\,\Omega^2 d\cos(\Omega t + \varphi_c) \tag{2.14}$$

where $\Omega = \dot{\phi}$ is the rotational speed of the rotor.

Using (2.7) for the translational motion of the center of gravity of the rotor leads

$$\ddot{x} = \frac{1}{m}[f_{A,x} + f_{B,x} + \bar{m}g + \frac{m_{ub,s}}{2}\,\Omega^2 d\cos(\Omega t + \varphi_s)] \tag{2.15}$$

$$\ddot{y} = \frac{1}{m}[f_{A,y} + f_{B,y} - \bar{m}g + \frac{m_{ub,s}}{2}\,\Omega^2 d\sin(\Omega t + \varphi_s)] \tag{2.16}$$

Displacements at the bearings are the superpositions of translation and inclination

motions which are

$$r_{A,x} = x + a\theta, \qquad r_{B,x} = x - b\theta, \qquad r_{A,y} = y - a\psi, \qquad r_{B,y} = y + b\psi\,.$$

Forces in (2.13-2.16) for translational and tilting motions are

$$f_{A,x} = f_{a2} - f_{a1} = k_{i,A} i_{A,x} - k_{s,A}(x + a\theta), \qquad f_{B,x} = f_{b2} - f_{b1} = k_{i,B} i_{B,x} - k_{s,B}(x - b\theta),$$

$$f_{A,y} = f_{a3} - f_{a4} = k_{i,A} i_{A,y} - k_{s,A}(y - a\psi), \qquad f_{B,y} = f_{b3} - f_{b4} = k_{i,B} i_{B,y} - k_{s,A}(y - a\psi).$$

Dynamic unbalance is represented by static and couple components, $\frac{1}{2} m_{ub,s} d$ and $\frac{1}{2} m_{ub,c} d$ respectively.

The equations of motion in state-space form are as follows:

$$\frac{d}{dt} x_r = \begin{bmatrix} 0 & I \\ A_S & A_{\mathcal{G}}(\Omega) \end{bmatrix} x_r + \Big[B_w\Big] m_{ub} + \Big[B_u\Big] u + g\,, \tag{2.17}$$

where
$x_r = (x\ y\ \psi\ \theta\ \dot{x}\ \dot{y}\ \dot{\psi}\ \dot{\theta}\,)^T, \quad u = (i_{A,x}\ i_{A,y}\ i_{B,x}\ i_{B,y})^T, \quad g = (0\ 0\ 0\ 0\ \frac{\bar{m}g}{m}\ -\frac{\bar{m}g}{m}\ 0\ 0)^T,$
and

$$m_{ub} = \frac{1}{2} d\,(m_{ub,s}\ \ m_{ub,c})^T.$$

The effect of gravity term g can be omitted from (2.17) as premagnetization current can be chosen large enough to cover the gravity compensation current. For vertical rotors, there is no effect of gravity on the radial bearings as the rotor is levitated by the axial bearing.

Matrices A_S, $A_{\mathcal{G}}(\Omega)$, B_w, and B_u are as follows:

$$A_S = \begin{bmatrix} -\frac{k_{s,A}+k_{s,B}}{m} & 0 & 0 & \frac{a\,k_{s,B}-b\,k_{s,A}}{m} \\ 0 & -\frac{k_{s,A}+k_{s,B}}{m} & \frac{b\,k_{s,A}-a\,k_{s,B}}{m} & 0 \\ 0 & \frac{b\,k_{s,A}-a\,k_{s,B}}{I_r} & -\frac{b^2\,k_{s,B}+a^2\,k_{s,A}}{I_r} & 0 \\ \frac{a\,k_{s,B}-b\,k_{s,A}}{I_r} & 0 & 0 & -\frac{b^2\,k_{s,B}+a^2\,k_{s,A}}{I_r} \end{bmatrix},$$

$$A_{\mathcal{G}}(\Omega) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{-\Omega I_p}{I_r} \\ 0 & 0 & \frac{\Omega I_p}{I_r} & 0 \end{bmatrix}, \quad B_w = \begin{bmatrix} 0 \\ F_{ub} \end{bmatrix}, \quad B_u = \begin{bmatrix} 0 \\ F_u \end{bmatrix},$$

with

$$F_{ub} = \begin{bmatrix} \Omega^2 \cos(\Omega t + \varphi_s) & 0 \\ \Omega^2 \sin(\Omega t + \varphi_s) & 0 \\ 0 & \frac{(a+b)}{I_r}\Omega^2 \sin(\Omega t + \varphi_c) \\ 0 & -\frac{(a+b)}{I_r}\Omega^2 \cos(\Omega t + \varphi_c) \end{bmatrix}, \quad F_u = \begin{bmatrix} \frac{k_{i,A}}{m} & 0 & \frac{k_{i,B}}{m} & 0 \\ 0 & \frac{k_{i,A}}{m} & 0 & \frac{k_{i,B}}{m} \\ 0 & -\frac{a\,ki,A}{I_r} & 0 & \frac{b\,k_{i,B}}{I_r} \\ \frac{a\,k_{i,A}}{I_r} & 0 & -\frac{b\,k_{i,B}}{I_r} & 0 \end{bmatrix}.$$

The eigenvalues of the system are determined by the matrices A_S and $A_{\mathcal{G}}(\Omega)$. When the rotor is levitated by a stabilizing controller, poles on the negative and positive (unstable) real axis are pulled through the origin by the proportional (to rotor displacement) gain of the controller. When the proportional gain supplied by the control currents becomes equal to the negative stiffness force applied to the rotor, eight rigid body poles meet at the origin. Further increasing the proportional gain of the controller makes the total stiffness force positive and splits the poles through positive and negative directions at the imaginary axis. This is due to the fact that, symmetric matrix A_S, now, has negative diagonal elements. When rotor starts to rotate, the effect of gyroscopic moments pushes the poles related with the inclination motion through positive imaginary axis. This is due to the skew symmetry of the matrix $A_{\mathcal{G}}(\Omega)$.

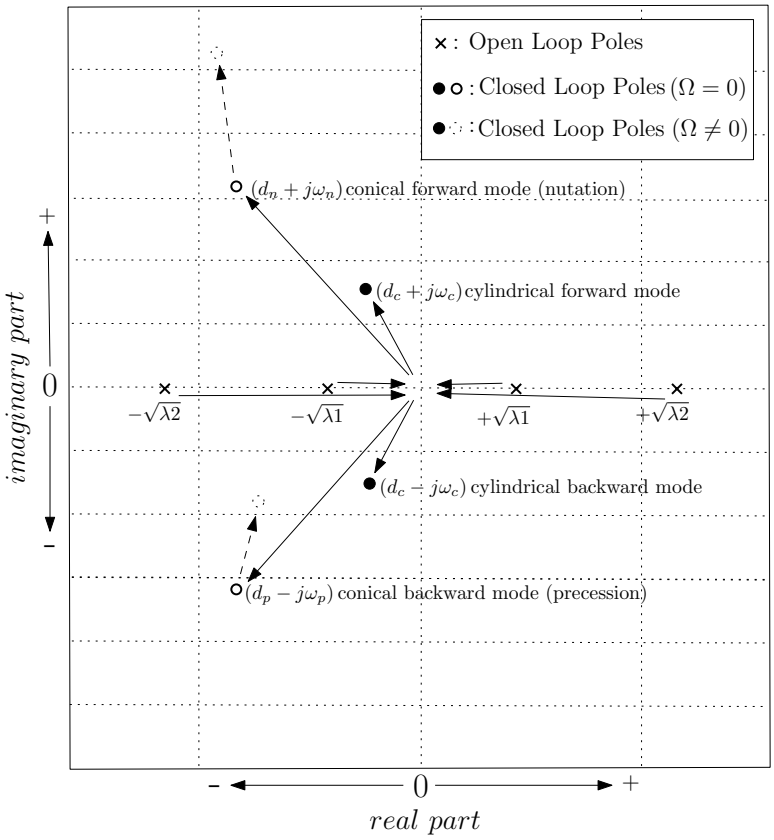

Figure 2.6. Eigenvalue trajectory plot.

Therefore, the rotating rotor modeled as a rigid body has 4 "pole pairs" symmetrically placed on the positive and negative imaginary axis. Positive poles are related with the forward whirl (in direction of rotation), and negative poles are related with the backward whirl of the rotor. Moreover, positive imaginary axis poles are comprised of 2×2 twin poles due to axial symmetry of the rotor. One pair is called *parallel* (eigenvector of translation motion)-or *cylindrical* as the mode shape of the rotor becomes cylindrical due to the rotating unbalance forces-, and the other pair is called *conical* (eigenvector of the inclination motion) as the angular motions are coupled due to gyroscopic effect. Same is valid for the negative imaginary axis poles. Isotropic bearing stiffness created, due to the identical premagnetization currents in both directions, can eliminate the consideration of negative imaginary axis poles. In the cases of *anisotropic* stiffness, nonsymmetric rotor, or rubbing to auxiliary bearings etc., these poles must also be considered in system dynamics and controller design. This phenomenon is also valid for rotors modeled as flexible structures which are analyzed in the next subsection.

2.2.2. Flexible Rotor with AMBs

As stated in the beginning of the section on rotordynamics, dynamic mechanical systems are continuous in space and, therefore, feature an infinite number of degrees of freedom in reality. It is still possible to derive differential equations for such continua, however, these differential equations can not be described by a set of linear second order differential equations with constant coefficients as derived in the previous subsection. In fact, the formulation for the equations of motion for continua always results in partial differential equations, containing not only derivatives with respect to time but also with respect to physical coordinates involved.

Actually, materials used in structures are somewhat elastic, and this is reflected by the deflection of the structure in response to the exerted forces and moments. These forces and moments are often speed dependent for the rotors hence, at high speeds, they deflect more and classified as *flexible rotors* in rotordynamics. In general, flexible rotor classification refers to rotors operating above their first bending critical speed (i.e., first bending mode eigenfrequency). Stiffness of flexible rotors can not considered to be infinite as in the case of rigid rotors, therefore, *flexible modes* of whirling motion due to rotor deflections should also be taken into account.

Differing from rigid structures which are modeled as a single element with maximum six-dof in space, eigenvalue problems for flexible structures are modeled within the scope of continuous systems. However, this approach leads to transcendental equations which, for most of the cases, have to be solved numerically. The effort to evaluate the solutions is rather cumbersome even for very simple models. For realistic systems complexity will be much higher. Therefore, one must think of alternative ways to deal with flexible structures. One possibility is to virtually split up the system into parts with concentrated masses and other concentrated massless parts, e.g., stiffness. Modeling would end up with a multi degree of freedom discrete system and the parts of the system can be linked to each other by using their mutual boundary conditions. This approach leads to the well-known *Finite Element Method* (FEM) used for modeling complex structures with multiple discrete elements. The basic concept is to divide the

structure into particular elements of simple geometry. Between neighboring elements, connecting nodes are introduced. Each node has six-dof about which it can move.

For the case of a rotor with AMBs, the axial displacements are controlled by the axial one-dof AMB or PMB and the axial rotations are controlled by the driving motor/turbine, leaving four-dof to be considered for the lateral motion of each rotating finite element. A comprehensive treatment of FEM modeling for flexible rotors can be found in [1],[2] and [33]. In FEM modeling, linear dynamics of the flexible rotor supported by AMBs is often based on the assumptions: i) Radial displacements are small with respect to rotor dimensions, ii) all system parameters are time invariant, iii) the rate of change of rotor speed is low and can be neglected, iv)rotor torsion and its effect on lateral motion is negligible, v) sensors and actuators can be associated with discrete points on the structure.

Transformation of individual elements to the global coordinate system and connecting the elements at the nodes yields the following equation of motion for the perfectly balanced rotor supported with AMBs [1]:

$$M_r \ddot{q}_r + (D_s + D_{i/r} + \Omega\, \mathcal{G})\, \dot{q}_r + (K + \Omega\, D^*_{i/r})\, q_r = F_s\, q_r + F_i u\,, \tag{2.18}$$

where

M_r : Symmetrical, positive definite rotor mass matrix,

D_s : Symmetrical structural damping matrix,

$D_{i/r}$: Symmetrical internal damping and/or rotating damping matrix,

$D^*_{i/r}$: Skew symmetric "circulatory" internal and/or rotating damping matrix,

$\mathcal{G}$: Skew symmetric semi positive definite gyroscopic matrix,

K_r : Symmetrical, positive semi-definite rotor stiffness matrix,

F_s : AMB stiffness matrix,

F_i : AMB control force matrix,

q_r : Generalized rotor displacement vector,

u : Input (control currents) vector,

Ω : Rotational speed.

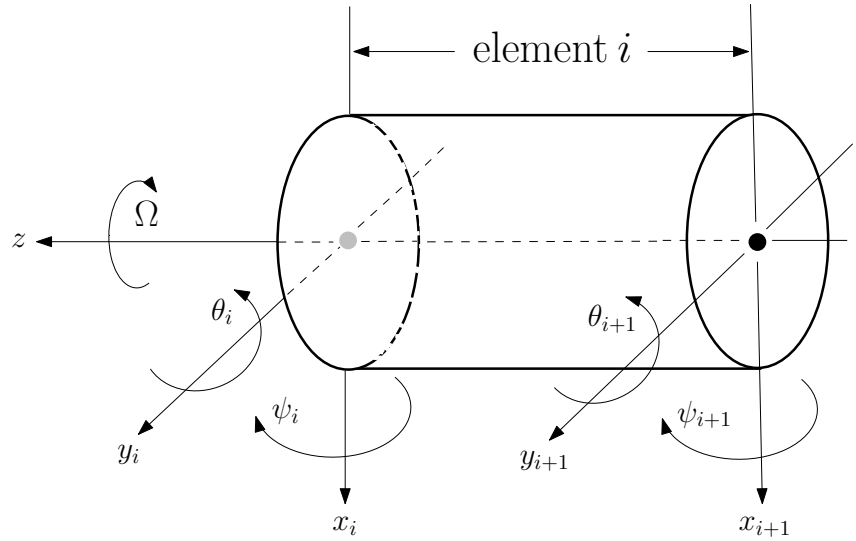

Figure 2.7. Rotor element.

The generalized rotor displacement vector, q_r, consists of displacements and rotations of individual nodes as; $q_r = (q_x \; q_y)^T$, with $q_x = (x_1 \, \psi_1 ..... x_k \, \psi_k)^T$, $q_y = (y_1 \, \theta_1 ..... y_k \, \theta_k)^T$.

The dimension of the system is; $n = 4k$, where k is the number of nodes in the model.

State-space form of (2.18) is

$$\frac{d}{dt}\chi = \begin{bmatrix} 0 & I \\ -M_r^{-1}(K_r - F_s + \Omega \, D_{i/r}^*) & -M_r^{-1}(D_s + D_{i/r} + \Omega \, \mathcal{G}) \end{bmatrix} \chi + \begin{bmatrix} 0 \\ [M_r^{-1} F_i] \end{bmatrix} u \, ,$$

where

$$\chi = (q_r \;\; \dot{q}_r \,)^T \text{ , and } u = (i_{A,x} \;\; i_{A,y} \;\; i_{B,x} \;\; i_{B,y})^T .$$

There is no external damping source like conventional mechanical bearings' viscous damping force. Internal damping can be neglected as its presence requires a relative velocity between rotational and whirling motions of the rotor. One source which can induce a rotating damping with a structure similar to internal friction damping is the rotational motion of a fluid (usually air) at locations of small clearance such as seals.

Dimensions of matrices $M_r, D_s, \mathcal{G}, K_r, F_s$ are square $(n \times n)$ and, the dimensions of matrix F_i is $(n \times 4)$. The rotor stiffness matrix K_r is singular with four vanishing eigenvalues as the rigid body motion of the rotor is unconstrained. Bearing stiffness matrix F_s effects total system stiffness only at the nodes under the action of AMBs.

At standstill, there is no coupling between the orthogonal planes due to gyroscopic effects and (2.18) reduces to:

$$M_r\ddot{q}_r + D_s\dot{q}_r + [K_r - F_s]q_r = F_i u\,, \quad \text{and} \quad \upsilon = Cq_r\,, \tag{2.19}$$

where vector υ is the sensor outputs and C is the $(4 \times n)$ observation (output) matrix.

Defining $\bar{K}_r = K_r - F_s$, solution of (2.19) without structural damping and without control forces from the bearing currents leads to the generalized eigenvalue problem

$$[\bar{K}_r - \omega_i^2 M_r]\phi_i = 0\,,$$

whose $2n$ solutions $(\pm\omega_i)$ and their respective ϕ_is give the information about the systems free response which is a superposition of the components of the form

$$q_r(t) = \sum_{i=1}^{n} \phi_i\, Q_i \cos(\omega_i t + \varphi_i) \quad \text{with} \quad Q_i \text{ and } \varphi_i \in \mathbb{R}\,,$$

which can be evaluated either for the forward whirling $(+\,\omega_i)$ or backward whirling $(-\,\omega_i)$ motion of the rotor. Coefficients Q_i and φ_i are determined by the initial conditions, $q_r(0)$ and $\dot{q}_r(0)$.

For the forward whirling motion, ordering of eigenfrequencies can be such selected that, they are of ascending frequency [3]. Subsequent compilation of the eigenvectors

yields

$$\phi_M = \begin{bmatrix} | & & | \\ \phi_1 & \ldots\ldots & \phi_n \\ | & & | \end{bmatrix} .$$

Matrix ϕ_M is not composed of orthogonal eigenvectors and is rank deficient due to the rigid body modes of the system having 2 eigenfrequencies equal to zero. Instead of using the matrix evaluated from the stiffness matrices of rotor/AMB, a practical solution is to assume that the rotor is suspended by conventional mechanical bearings with their own positive stiffness values. The stiffness of the mechanical bearings will only effect the stiffness related with the rigid body modes of the system. Later during the process of assigning related poles to modal coordinates, the first two poles related with the rigid body modes of the rotor with mechanical bearings can be replaced with the rigid body mode poles of the rotor with AMBs.

Hence, matrix $\bar{K}_r$ will be replaced by the non singular stiffness matrix K_b of the rotor with mechanical bearings and the "ad-hoc" system that will be used to evaluate the dynamics of rotor/AMB is

$$M_r\,\ddot{q}_r + D_s\,\dot{q}_r + K_b\,q_r = F_i\,u\,, \quad \upsilon = C\,q_r\,, \text{ and } \quad [K_b - \omega_i^2 M_r]\phi_i = 0\,.$$

The matrix ϕ_M can be used to transform the (2.19) from physical coordinates to modal coordinates using the transformation, $q_r = \phi_M\,\tilde{q}_r$, and it becomes

$$\tilde{M}_r\,\ddot{\tilde{q}}_r + \tilde{D}_s\,\dot{\tilde{q}}_r + \tilde{K}_b\,\tilde{q}_r = \tilde{F}_i\,u\,, \quad \text{and} \quad \upsilon = \tilde{C}\,\tilde{q}_r\,, \tag{2.20}$$

with $\tilde{F}_i = \phi_M^T F_i$, and $\tilde{C} = C\,\phi_M$. Due to the symmetry of M_r and K_b, the matrices $\tilde{M}_r$ and $\tilde{K}_b$ in (2.20) are of diagonal structure. The modal transformation above is particularly effective when the columns of the transformation matrix ϕ_M are scaled in such a way that matrix $\tilde{M}_r$ becomes the identity matrix. In this case

$$\tilde{M}_r = \phi_M^T M_r\,\phi_M = I\,, \text{ and } \quad \tilde{K}_b = \phi_M^T K_b\,\phi_M = \text{diag}\,(\omega_i^2)\,.$$

In the case of proportional structural damping, i.e., $D_s = \alpha_1 M_r + \alpha_2 K_b$ for $\alpha_1,\ \alpha_2\ \in \mathbb{R}$, the damping matrix is also diagonal as

$$\tilde{D}_s = \phi_M^T D_s\, \phi_M = \mathrm{diag}\,(2\xi_i\omega_i)\,.$$

Thus, system in (2.20) can be coupled into $2n$ one-dof systems with poles

$$p_i = -\xi_i\,\omega_i + j\omega_i\,\sqrt{1-\xi_i^2} \quad \text{and} \quad p_i^* = -\xi_i\,\omega_i - j\omega_i\,\sqrt{1-\xi_i^2}\,, \qquad i = 1, ..., n\,. \tag{2.21}$$

Every pole occurs twice as the system is axially symmetric and isotropic. Each oscillator with an eigenfrequency is associated with one eigenmode of shape ϕ_i. The free damped vibration frequency ($\omega_i\,\sqrt{1-\xi_i^2}$) of each mode shape ($\phi_i$), decays with $e^{-\xi_i\,\omega_i\,t}$. Modal damping values ξ_i are as low as 0.0005 especially at the low frequencies. Note that, actually, there is no structural damping available for the rigid body modes of the rotor suspended by magnetic bearings.

State-space description for the system at standstill can be written in the following form:

$$\frac{d}{dt}\,\chi = \begin{bmatrix} 0 & I \\ A_S & A_D \end{bmatrix} \chi + \Big[B\Big]\, u\,, \tag{2.22}$$

where $\chi = (\tilde{q}_r\ \ \dot{\tilde{q}}_r)^T$, and $u = (i_{A,x}\ i_{A,y}\ i_{B,x}\ i_{B,y})^T$.

As discussed before, the first two diagonal elements a_{11} and a_{22} of the matrix A_S will be replaced with the rigid body poles of the rotor which can be evaluated as shown at the outset of the subsection (2.2.1). Also, the elements a_{11} and a_{22} of the matrix A_D must be zero as there is no damping provided for the rigid body poles of the rotor. Based on the poles obtained from (2.21) and the required modification as explained

above, the matrix blocks in (2.22) are as follows:

$$A_S = \begin{bmatrix} \lambda_1 & 0 & \dots & \dots & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & \dots & \vdots \\ \vdots & 0 & -(p_3 \cdot p_3^*) & 0 & \dots & \vdots \\ \vdots & \vdots & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & \dots & \dots & \dots & 0 & -(p_n \cdot p_n^*) \end{bmatrix}$$

$$A_D = \begin{bmatrix} 0 & \dots & \dots & \dots & \dots & 0 \\ \vdots & 0 & \dots & \vdots & \dots & \vdots \\ \vdots & \vdots & (p_3 + p_3^*) & \vdots & \dots & \vdots \\ \vdots & \vdots & 0 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & \dots & \dots & \dots & 0 & (p_n + p_n^*) \end{bmatrix}$$

$$B = \begin{bmatrix} 0 \\ \tilde{F}_i \end{bmatrix}$$

The matrix A_S is the negative of the stiffness matrix $\tilde{\bar{K}}_r$ and, the matrix A_D is the negative of the modal damping matrix $\tilde{\bar{D}}_s$ which are modified in accordance with the physical requirements of the rotor/AMB system.

For the case of rotating rotor, the motion in two orthogonal planes are coupled with the gyroscopic term $\Omega\mathcal{G}$. Skew symmetric matrix $\mathcal{G}$ obtained from the FEM model can easily be transformed to modal coordinates. Dynamics of the system in the state-space becomes as follows:

$$\frac{d}{dt}\chi = \begin{bmatrix} 0 & I \\ -\tilde{\bar{K}}_r & -(\tilde{\bar{D}}_s + \Omega\,\tilde{\mathcal{G}}) \end{bmatrix} \chi + \begin{bmatrix} 0 \\ \tilde{F}_i \end{bmatrix} u\,, \tag{2.23}$$

where $\tilde{\mathcal{G}} = \phi_M^T \mathcal{G}\, \phi_M$.

System has a total of $(2n-8)$ flexible poles. One half is on the positive frequency plane and the other half is on the negative frequency plane. Moreover, due to the axial symmetry, and the isotropy of the bearings, each pole occurs twice resulting $(n/2-2)$ eigenfrequencies to be considered for the flexible modes with a positive frequency.
The coupling of two planes causes the systems flexible poles to move with increasing Ω from their original positions. The forward whirling poles close to the positive imaginary axis move towards increasing frequencies and, the backward whirling poles close to the negative imaginary axis move towards decreasing frequencies.

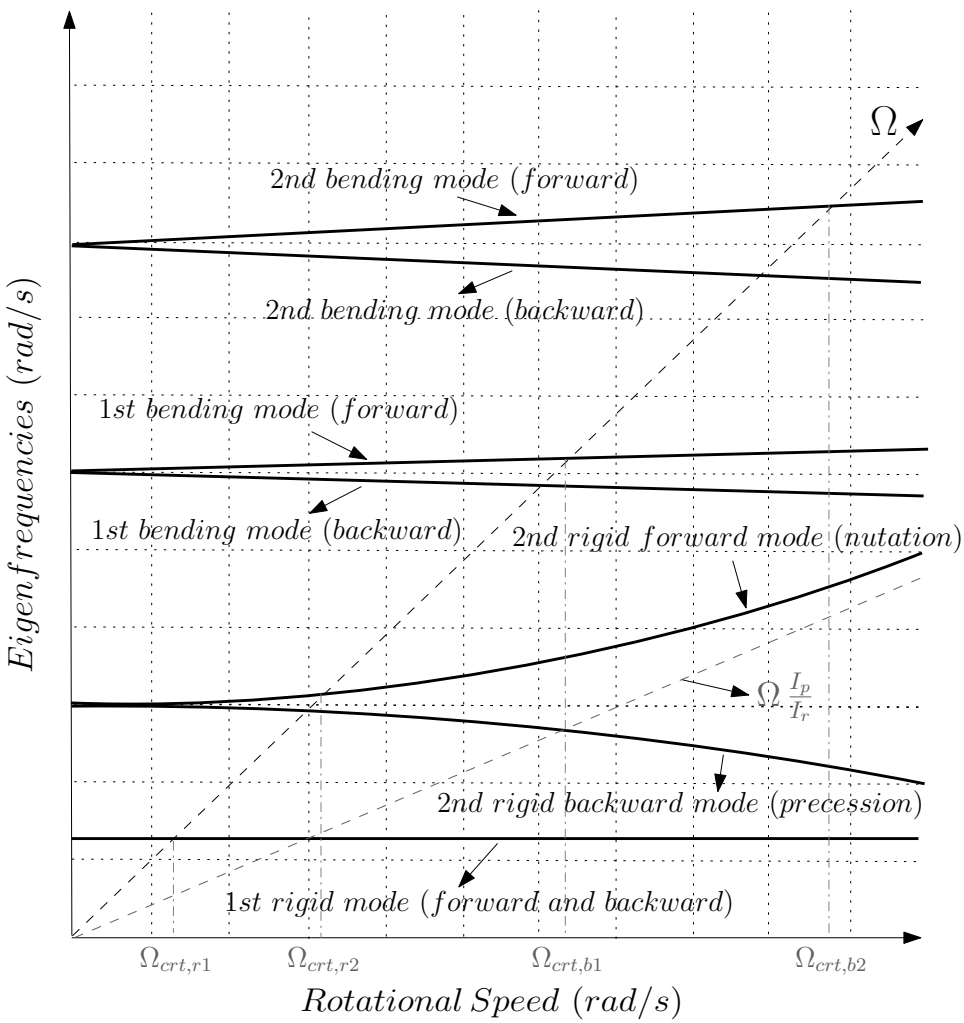

Figure 2.8. Campbell diagram.

The splitting of eigenfrequencies with rotation speed is shown in the *Campbell Diagram*, Figure 2.8. When sped up from standstill, several of the eigenfrequencies are passed

which are called critical speeds of the rotor. At these speeds the unbalance force may cause resonance. Excitation of a resonance at a specific eigenfrequency depends on the axial distribution of the unbalance forces along the rotor. If the rotor is axially symmetric and the bearings are isotopic, only forward modes of the rotor may be excited. Otherwise, backward modes might also be excited.

Application of the stabilizing controller brings the unstable rigid body poles to the left half plane. With increasing controller gain, poles first move through the origin. If the gain is further increased, they split and follow the positive and negative imaginary axis. Using a damping behavior in the controller, the poles can additionally be moved to the left of imaginary axis leading a stable oscillating system. Under rotation, the rigid body poles with conical mode shapes exhibit a similar behavior to flexible modes and they move with increasing Ω from their original positions, which are symmetric with respect to negative real axis, towards positive imaginary axis direction. Damping of backward conical mode will decrease by increasing operation speed if the controller design is based on isotropic bearing stiffness.

The size of the model is determined by the number of nodes in the finite element model. Yet, even simple models have typically 10 nodes with 40 dof and, entailing a state space model of 80. The majority of the single mass oscillators in the modal model have very high eigenfrequencies which can never be excited due to the limited operation band of the rotor and the controller. Moreover, high frequency modes which will not be excited in the normal operation will also be included in the system model and the design will either end up with a controller which is not practical to be built or with a "nervous" controller reacting too much to high frequency transients and sensor/electonic noise, leading to poor performance. Furthermore, besides the fact that FEM merely approximates the behavior of the continuous structure, no structure can be expected to behave linearly up to arbitrarily high frequencies. To reduce the order of the controller designed, minimum degree of freedom for the rotor's radial motion can be decided based on the geometric structure of the rotor to be modeled (number of disks/blades and their axial locations on the rotor). Alternatively, order of the controller can be reduced after modeling and controller synthesis but this approach

is quite complicated and may not lead to sufficient performance. Controller order reduction may also cause stability problems as the open loop system is inherently unstable.

For these reasons and in order to reduce the computational burden in subsequent calculations, one is interested in adapting the model such that it describes the system in a frequency range where this is reasonable. This can be achieved by ordering the columns ϕ_i of the matrix ϕ_M such that corresponding eigenfrequencies are monotonously increasing. Next, the columns of ϕ_i calculated for the eigenfrequencies above a selected frequency level are truncated. If the first m frequencies are relevant, the new matrix ϕ_m has n rows and m columns. Then, the transformation to modal coordinates are made with the matrix ϕ_m instead of ϕ_M. The resulting system then has the same properties as before, but in lower dimensions as

$$\tilde{M}_m \, \ddot{\tilde{q}}_m + \tilde{\bar{D}}_m \, \dot{\tilde{q}}_m + \tilde{\bar{K}}_m \, \tilde{q}_m = \tilde{F}_m \, u \, .$$

Furthermore, the transformation $q_r = \phi_m \, \tilde{q}_m$ allows to interpret the results for the reduced system in the physical coordinates of the real system.

3. ROBUST CONTROL OF ROTOR/AMB SYSTEMS

3.1. Preface

All physical systems are virtually nonlinear and time-varying in nature. Examples of nonlinearities we are often confronted with range from simple nonlinearity, such as saturation and rate limitations of the actuators (as in AMBs), to the inherently nonlinear behavior of physical systems such as robotic manipulators, aircraft, internal combustion engines and chemical process plants. Nevertheless, it is often possible to describe the operation of the physical system by a linear system, under circumstances that the real operation of the physical system does not deviate too much from the nominal operating (or equilibrium) condition. Therefore the analysis and synthesis of linear systems have occupied a major place in systems theory. Consequently many "computationally tractable" analysis and synthesis techniques have been developed.

We often encounter situations where the linearized model around a nominal operating point is inadequate or inaccurate. In this case, a linear controller from linear systems theory may perform well on the linearized model but may not even be stable when implemented on the real physical system. To cope with uncertainties, linear system theory has been extended to robust linear control theory that takes into account these inherent inaccuracies as uncertainties and then provide systematic analysis and synthesis techniques. However introducing an uncertainty in the design process leads to degradation in the performance of a designed controller (note that performance degradation is alternatively expressed by "conservatism" in the literature). Therefore, one of the important issues in the robust control theory is how to efficiently model uncertainties, i.e., to reduce the conservatism in the analysis and synthesis exploiting the nature of the uncertainties such as "structured", "real" or "time invariant".

Main sources of uncertainty in the design of a linear controller for rotor/AMB systems are the high frequency bending modes of the rotor which are not included in the model, changes in the negative spring stiffness of the magnetic bearings, and the

rotor speed as gyroscopic effects significantly change the system dynamics. For the rotor by itself, gyroscopic forces tend to stabilize the rotor motion. However, in an AMB system, instability may occur due to control forces as the AMB system is not passive. Hence, on-line measurement of the rotor speed and scheduling the dynamics of the controller with respect to the measured speed will substantially reduce the magnitude of uncertainty and improve the performance of the robust controller designed.

3.2. Multi-variable Robust Control

3.2.1. Introduction

To control an object or a system means to influence its behavior so as to achieve a desired goal. In order to implement this influence, devices such as controllers are built that incorporate various mathematical techniques. The interactive analysis and synthesis of these devices combined with the systems they control is the subject of *control theory*. *Stability* is usually the basic requirement of a controlled system; however, typically much more is demanded and, in fact, *feedback* is introduced in systems which are already stable, with the objective of improving different aspects of dynamic behavior. Sometimes, if the effect of disturbances on the behavior of the system is known, a *feedforward* path is also incorporated to improve the performance of control. A pictorial representation of a feedback and feedforward controlled system, being influenced by external inputs (such as reference commands) and external disturbances, is shown in Figure 3.1. Here we have a dynamical system with a control law, which takes some actions based on measurement information.

Generally, there have been two main lines of work in control theory, which sometimes have seemed to proceed in different directions but which are, in fact, complementary. One of these is based on the idea that a good model of the object/system to be controlled is available that one wants to somehow optimize its behavior. The techniques used in this approach are closely related with classical calculus of variations and to other areas of optimization theory.

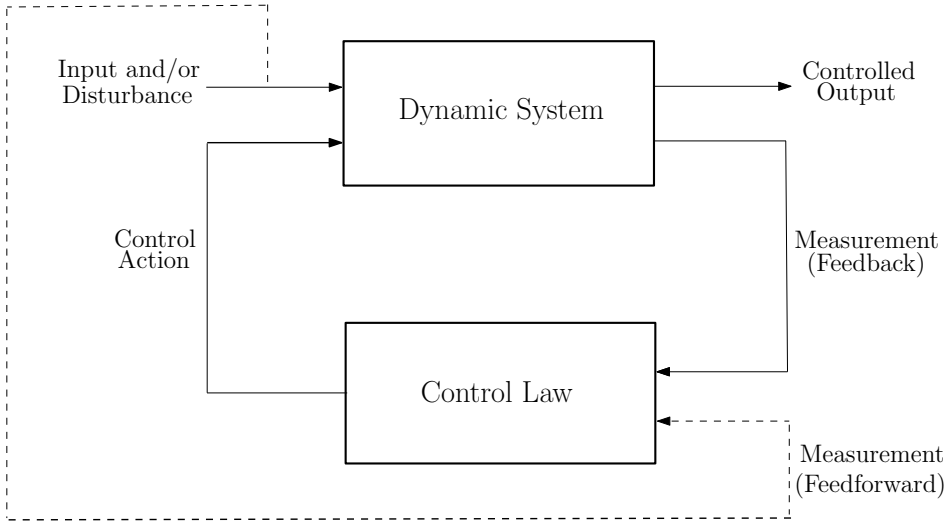

Figure 3.1. Controlled system.

The other main line of work is that based on the constraints imposed by uncertainty about the model or the about the environment in which the object/system operates. In this context, *robust control* refers to the modeling, analysis and design of control systems with *uncertainties*, i.e., control systems for which inexact models are available due to the inevitable difference between models and reality. The central tool here is the use of feedback in order to correct for the deviations from the desired behavior. Mathematically, stability theory, functional analysis and especially the theory of functions of complex variables have had a strong influence on this approach.

The mere assumption of feedback, however, does not guarantee a reduction in uncertainty, and there are many obstacles to achieving the uncertainty-reducing benefits of feedback. In particular, since for any reasonable model representing a physical system uncertainty becomes large and the phase is completely unknown at sufficiently high frequencies, the loop gain must be "small" at those frequencies to avoid destabilizing the high frequency system dynamics. Even worse is that feedback system actually increases uncertainty and sensitivity (with respect to disturbances) of the open loop system in the frequency ranges where uncertainty is significantly large. Often, what can be achieved with intelligent use of feedback is significant reduction of uncertainty for certain signals of importance with a small increase spread over other signals. Thus, the feedback design problem centers around the trade-off involved in reducing the overall impact of uncertainty.

In classical control theory, stability of the single input-single output (SISO) systems with uncertainties due to their simplified models or modeling errors are analyzed using Bode or Nyquist diagrams with respect to their gain and phase margins of instability. Though these methods are satisfactory for simple systems, they do not reflect the true margin of instability for systems with simultaneous gain and phase perturbations. As shown in [26] and [27], gain margin and phase margin do not give the correct indication of system's robustness when the gain and phase of the system are varying simultaneously.

The situation is worse for multi input-multi output (MIMO) systems. In MIMO systems, input and output of the system are not scalars but vectors. Gain of the system from the input to the output is a transfer matrix composed of scalar elements of individual transfer functions. Gain of the system in MIMO structure depends on the direction (relative weight of each input channel) of the input vector. At a certain frequency, the highest possible gain of the system is reflected by the *maximum singular value*, $\overline{\sigma}\big(M(j\omega)\big)$, (which is equal to square root of maximum positive eigenvalue; $\sqrt{\lambda_{max}}$ of $\{M(j\omega)^* M(j\omega)\}$), and lowest possible gain by the *minimum singular value*, $\underline{\sigma}(M(j\omega))$, (which is equal to square root of minimum positive eigenvalue; $\sqrt{\lambda_{min}}$ of $\{M(j\omega)^* M(j\omega)\}$), of the closed-loop complex transfer matrix function $M(j\omega)$. $M(j\omega)^*$ is the complex conjugate transpose (or adjoint) of $M(j\omega)$. Singular values of the matrix $M(j\omega)$ changes with respect to the frequency of input excitations similar to the frequency response of the transfer function for SISO systems. Uncertainty of the system complicates the evaluation of the system transfer matrix and, moreover, uncertainty in the input vector, which can cause the input signal to spread from one input direction to another, makes the control of uncertain plant difficult. Therefore, for MIMO systems it is very useful, if not imperative, to define and model the system with a special uncertainty structure in the synthesis for a stabilizing controller.

3.2.2. Basics of Linear Systems

A *system* is a mapping between inputs and outputs. The parts making up the system may be clearly or vaguely defined. These parts are related to each other through a particular set of variables, called the *states* of the system, that completely determine the behavior of the system at any given time. A *dynamical system* is a system whose state change with time. Specifically, the state of a dynamical system can be regarded as an information storage or memory of past system events. Therefore, the state of a dynamical system at a given time is uniquely determined by the state of the system at the initial time and the present input to the system. Hence, a dynamical system can generate a differential equation of the form

$$\dot{x}(t) = f\big(t, x(t), u(t)\big)\,, \quad x(t_0) = x_0\,, \quad t \in [t_0, t_1]$$

where $x(t) \in \mathbb{X}$ in the time interval $[t_0, t_1]$ is the state of the dynamical system, and $\mathbb{X}$ is an open subset of $\mathbb{R}^n$ with $0 \in \mathbb{X}$.

A *linear system* is a system that satisfies the *principle of superposition* as a transformation from the input to the output, $y = Mu$ (M is the operator that maps the system from the input to the output). Hence it must satisfy the equations

$$M(u_1 + u_2) = Mu_1 + Mu_2 \;\text{ and }\; M(\alpha u) = \alpha M u\,, \text{ with } \alpha \in \mathbb{R}\,.$$

A linear time-invariant (LTI) finite-dimensional system is defined by its state-space representation

$$\begin{aligned} \dot{x} &= Ax + Bu\,, \\ y &= Cx + Du\,. \end{aligned}$$

where $x \in \mathbb{R}^n$ is the state vector, $u \in \mathbb{R}^m$ is the input vector, $y \in \mathbb{R}^p$ is the output vec-

tor, and $\dot{x} = \frac{d}{dt}x$ denotes the time derivative of x. The system matrices $A, B, C,$ and D, of compatible sizes, are fixed in time and describe the behavior of the system. Note that the solution for the state vector x of the LTI state-space system is

$$x(t) = e^{A(t-t_0)}x_0 + \int_{t_0}^{t} e^{A(t-\tau)}Bu(\tau)\,d\tau \quad \text{with} \quad x(t_0) = x_0\,,$$

and the output y of the system is

$$y(t) = Ce^{A(t-t_0)}x_0 + \int_{t_0}^{t} Ce^{A(t-\tau)}Bu(\tau)\,d\tau + Du(t)\,.$$

Given a state-space system (A, B, C, D), the state space system $(TAT^{-1}, TB, CT^{-1}, D)$ has the same input-output behavior for all invertible real matrices T. The operation applied is called a *coordinate transformation* or *similarity transformation.* This implies that state-space descriptions of systems are not unique.

The notation $\mathrm{G}(s) = C(sI - A)^{-1} + D$ is also used for describing a LTI system. The symbol s can be interpreted both as the time derivative operator $\frac{d}{dt}$ and as the argument of *Laplace transform* of the system G. Frequency response of the system can be evaluated at $s = j\omega$. A transfer function is *proper* if $\mathrm{G}(\infty)$ is bounded (if $\mathrm{G}(\infty) = D$). It is *strictly proper* if $\mathrm{G}(\infty) = 0$.

A LTI *continuous* system (systems can also be expressed in *discrete* form, as opposed to their actual structure operating in continuous time, which is particularly useful for digital controller synthesis) is *stable* if A has all its eigenvalues λ_i in the open left half plane $\mathbb{C}^-$, i.e., Re $\lambda_i < 0$.

Controllability is an important property of a control system, and the controllability property plays a crucial role in many control problems, such as stabilization of unstable systems by basic feedback or by optimal/robust control. In most general terms, the concept of controllability denotes the ability to move a system around in its entire configuration space using only certain admissible manipulations. Controllability

matrix for a LTI system is given by

$$\mathrm{C} = [B \;\; AB \;\; A^2B \;\; ... \;\; A^{n-1}B]\,.$$

The system is controllable if and only if the controllability matrix has a full rank.

Observability is a measure for how well the internal states of a system can be inferred by knowledge of its external outputs. A system is observable if, for any possible sequence of state and control vectors, the current state vector can be determined in finite time using only the outputs. Observability matrix for a LTI system is given by

$$\mathrm{O} = [C^* \;\; (CA)^* \;\; (CA^2)^* \;\; ... \;\; (CA^{n-1})^*]^*.$$

The system is observable if and only if the observability matrix has a full rank.

An uncontrollable system is *stabilizable*, if its uncontrollable modes are stable. Similarly, an unobservable system is *detectable*, if its unobservable modes are stable. Of course, all controllable systems are stabilizable and all observable systems are detectable. A state-space *realization* (A, B, C, D) of a proper $\mathrm{G}(s)$ is said to be a *minimal realization*, if matrix A has the smallest possible dimensions. Realization is minimal if and only if (A, B, C, D) is both controllable and observable.

A linear time-varying system has time-varying system matrices and is defined by

$$\dot{x} = A(t)x + B(t)u\,,$$
$$y = C(t)x + D(t)u\,.$$

The solution for the state vector x of the LTV state-space system is

$$x(t) = \Phi(t, t_0)x_0 + \int_{t_0}^{t} \Phi(t, \tau)B(\tau)u(\tau)\,d\tau \quad \text{with} \quad x(t_0) = x_0\,,$$

and the output y of the system is

$$y(t) = C(t)\Phi(t,t_0)x_0 + \int_{t_0}^{t} C(t)\Phi(t,\tau)B(\tau)u(\tau)\,d\tau + D(t)u(t)\,.$$

$\Phi(t,t_0) = \tilde{X}(t)\tilde{X}^{-1}(t_0)$ is called the *state transition matrix*. $\tilde{X}(t)$ is the *fundamental matrix* of the system which can be evaluated as any solution of the matrix differential equation: $\dot{\tilde{X}}(t) = A(t)\tilde{X}(t)$, satisfying $\det\big(\tilde{X}(t)\big) \neq 0$, $\forall\, t \in \mathbb{R}^+$ [29].

A specific description for LTV systems (or even for some nonlinear systems) can be obtained by letting the system matrices depend on (time varying) parameters $\rho(t)$. This model is called parametrically varying linear system (or linear parameter varying system) and is represented by

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

The parameters $\rho \in \mathcal{P} \subset \mathbb{R}^m$ can either be an external input to the system or they can depend on the states of the system. In the latter case, nonlinear effects can be described by an LPV model.

Stability of LTV and LPV systems can not simply be determined as shown for the LTI systems, and it is the subject of subsection (3.4.2).

3.2.3. Controller Synthesis

Nominal (without any kind of uncertainty) generalized plant subject to external reference/disturbance/noise inputs and a control input is shown in Figure 3.2.

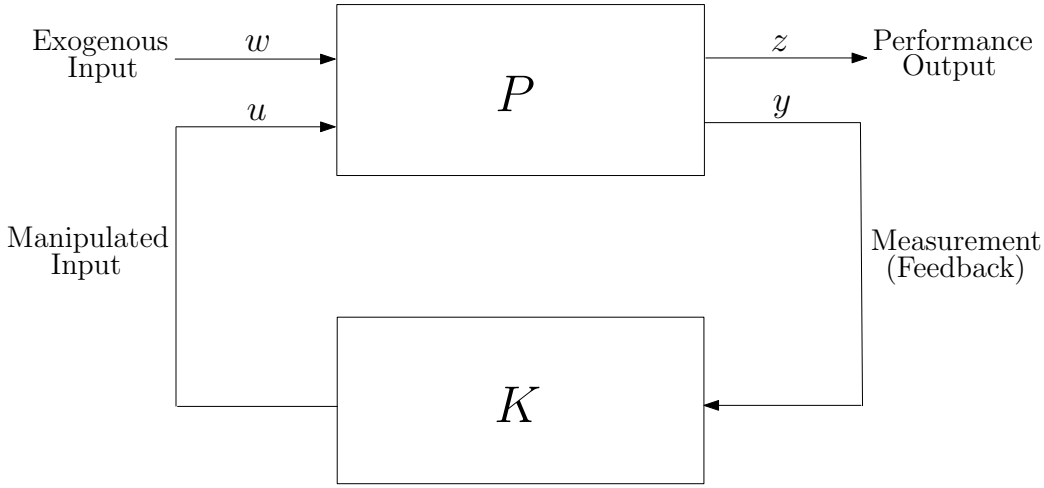

Figure 3.2. Feedback control system.

Controlled system, with order "n", admits a state-space realization

$$\dot{x} = Ax + B_w w + B_u u \tag{3.1a}$$

$$z = C_z x + D_{zw} w + D_{zu} u \tag{3.1b}$$

$$y = C_y x + D_{yw} w + D_{yu} u \tag{3.1c}$$

where $x(t) \in \mathbb{R}^n$, $w(t) \in \mathbb{R}^{m_1}$, $u(t) \in \mathbb{R}^{m_2}$, $z(t) \in \mathbb{R}^{p_1}$ and, $y(t) \in \mathbb{R}^{p_2}$.

In most of the practical applications D_{zw} is zero. Without loss of generality, it can also be assumed that $D_{yu} = 0$.

Considering the controller dynamics in the state space form

$$\dot{x}_K = A_K x_K + B_K y \ , \ u = C_K x_K + D_K y$$

and eliminating the measurement y and substituting the control input u into (3.1) leads to

$$\dot{x} = Ax + B_w w + B_u C_K x_K + B_u D_K C_y x + B_u D_K D_{yw} w \, ,$$

$$z = C_z x + D_{zw} w + D_{zu} C_K x_K + D_{zu} D_K C_y x + D_{zu} D_K D_{yw} w \, .$$

If the state vector is expanded to also include the controller states, $x_{cl} = (x\ x_K)^T$, the closed loop system can be expressed by

$$\begin{pmatrix} \dot{x}_{cl} \\ z \end{pmatrix} = \begin{bmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{bmatrix} \begin{pmatrix} x_{cl} \\ w \end{pmatrix}, \tag{3.2}$$

where the system matrices are given as

$$\left[\begin{array}{c:c} \mathcal{A} & \mathcal{B} \\ \hdashline \mathcal{C} & \mathcal{D} \end{array}\right] = \left[\begin{array}{cc:c} A + B_u D_K C_y & B_u C_K & B_w + B_u D_K D_{yw} \\ B_K C_y & A_K & B_K D_{yw} \\ \hdashline C_z + D_{zu} D_K C_y & D_{zu} C_K & D_{zw} + D_{zu} D_K D_{yw} \end{array}\right].$$

Closed-loop system can also be described by its input w, its output z, and its transfer matrix function $M(s)$ as

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

Equation (3.1) representing the state-space model of the controlled generalized plant in time domain can equivalently be represented in frequency domain as follows:

$$\begin{pmatrix} z \\ y \end{pmatrix} = \begin{bmatrix} P_{zw} & P_{zu} \\ P_{yw} & P_{yu} \end{bmatrix} \begin{pmatrix} w \\ u \end{pmatrix},$$

in which

$$P_{zw} = C_z(sI - A)^{-1}B_w + D_{zw}\,, \qquad P_{zu} = C_z(sI - A)^{-1}B_u + D_{zu}\,,$$
$$P_{yw} = C_y(sI - A)^{-1}B_w + D_{yw}\,, \qquad P_{yu} = C_y(sI - A)^{-1}B_u + D_{yu}\,.$$

Furthermore, eliminating u and y using $u = K(s)\, y$, the transfer matrix function of the

closed loop system from w to z is given by

$$M(s) = P_{zw} + P_{zu}K(I - P_{yu}K)^{-1}P_{yw} =: F_l(P, K)\,. \tag{3.3}$$

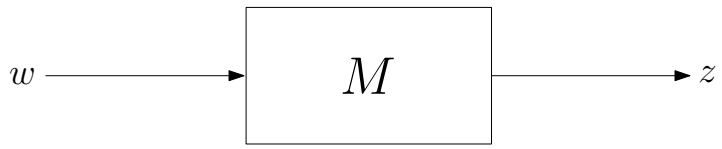

Figure 3.3. Closed-loop nominal system.

Here, $F_l(P, K)$ denotes a lower *linear fractional transformation*, (LFT) of P with K as the parameter.

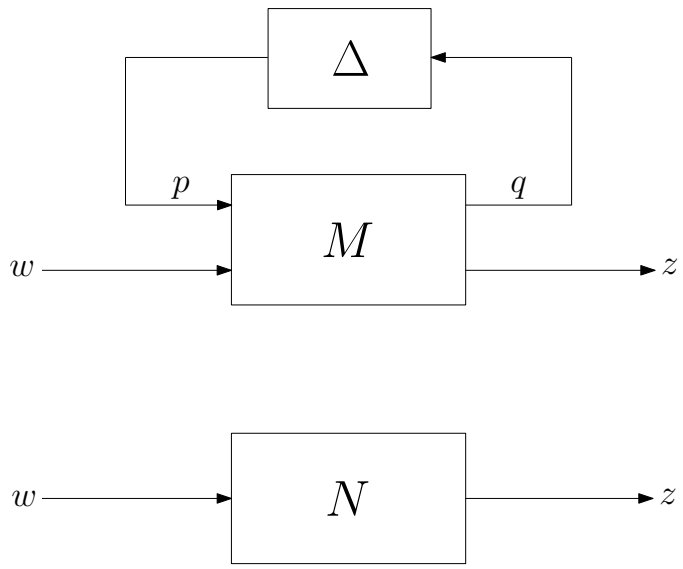

Figure 3.4. Closed-loop system with additional dynamics.

If any other dynamics of the system, denoted Δ, which is not included in the nominal model is drawn above M (Figure 3.4), one needs the upper linear fractional transformation defined as

$$N(s) = M_{zw} + M_{zp}\Delta(I - M_{qp}\Delta)^{-1}M_{qw} =: F_u(M, \Delta)\,, \qquad p = \Delta(s)\,q\,. \tag{3.4}$$

3.2.4. Nominal Stability and Performance

The ultimate goal in every feedback controller design is to achieve a desired behavior of the controlled plant by means of appropriately processing measurements y_m from the physical plant G and feeding back the control signals u as indicated in Figure 3.5. Unfortunately, all of the desired goals can not be achieved to the fullest extent at the same time. Various goals are conflicting and have to be traded against each other.

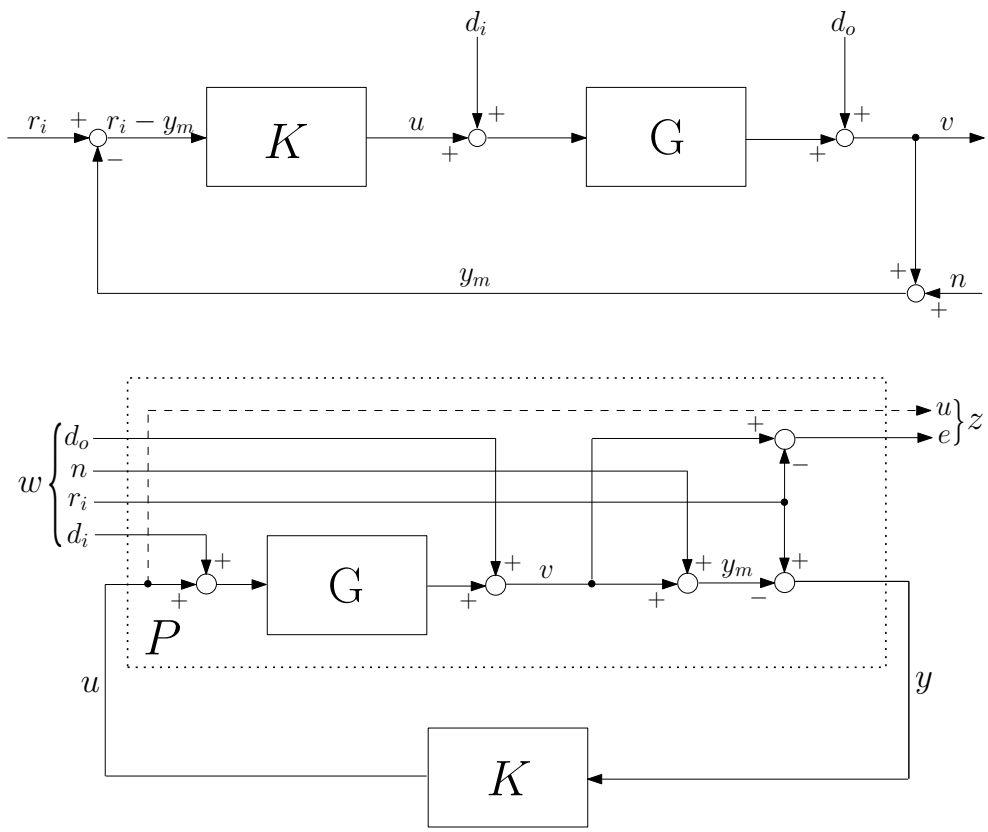

Figure 3.5. Control system configuration and equivalent representation with generalized plant P.

Any linear time-invariant system is stable if and only if all of its poles are on the left half of the complex plane. A MIMO system's stability is equivalent to stability of all of its constituent transfer functions. The term nominal stability refers to the stability of the closed loop made up by the nominal plant (without uncertainty) and the controller. This is achieved if and only if the closed loop is stable in the sense above

and no unstable pole-zero cancelations occur in forming the product KG [27].

Considering the feedback loop in Figure 3.6, the question is: *Are the transfer functions from v_1 and v_2 to any signal in the interconnection stable?* Or, equivalently; *is the system internally stable?* Internal stability is a basic requirement for a practical feedback system. This is because all interconnected systems may be unavoidably subject to some non-zero initial conditions and some small errors, and it can not be tolerated in practice that such errors at some locations will lead to unbounded signals at some other locations in the closed-loop system.

The closed loop system is *internally stable* if it is *well-posed*, and for every initial condition $x(0)$ of P and $x_K(0)$ of K, the limits

$$x(t),\, x_K(t) \to 0 \text{ as } t \to \infty \;\; \text{hold, when}\, w = 0\,.$$

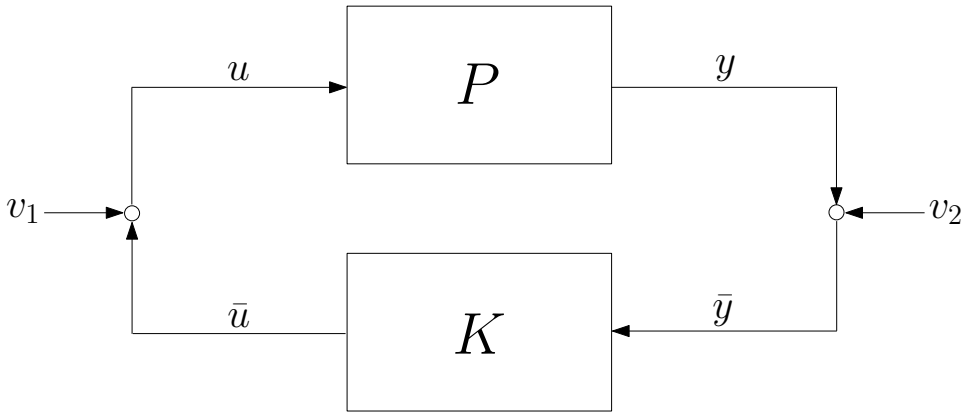

Figure 3.6. Feedback loop with perturbations.

It is clear from Figure 3.6 that

$$\begin{aligned} u &= v_1 + \;\; K\bar{y} \\ \bar{y} &= v_2 + P_{yu}u \end{aligned} \qquad \Rightarrow \qquad \begin{pmatrix} u \\ \bar{y} \end{pmatrix} = \begin{bmatrix} I & -K \\ -P_{yu} & I \end{bmatrix}^{-1} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}. \tag{3.5}$$

Equation (3.5) can be rewritten in the following form (see Appendix B.2):

$$\begin{pmatrix} u \\ \bar{y} \end{pmatrix} = \begin{bmatrix} I + K(I - P_{yu}K)^{-1}P_{yu} & K(I - P_{yu}K)^{-1} \\ (I - P_{yu}K)^{-1}P_{yu} & (I - P_{yu}K)^{-1} \end{bmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \tag{3.6}$$

If the interconnection is well-posed, meaning that all closed-loop transfer matrices are well-defined and proper, $I - P_{yu}K$ must have an inverse for all $K(j\omega)$ and for all $P_{yu}(j\omega)$, $\omega \in \mathbb{R}_e$. This is guaranteed if and only if $I - P_{yu}(\infty)K(\infty) = I - D_{yu}D_K$ is invertible. Most of the system models including the rotor/AMB system are well-posed, as either $D_{yu} = 0$ or $D_K = 0$.

Writing out the state-space equations of the overall system

$$\dot{x} = Ax + B_w w + B_u u\,, \qquad \dot{x}_K = A_K x_K + B_K y\,, \tag{3.7}$$

and

$$\begin{bmatrix} I & -D_K \\ -D_{yu} & I \end{bmatrix} \begin{pmatrix} u \\ y \end{pmatrix} = \begin{bmatrix} 0 & C_K \\ C_y & 0 \end{bmatrix} \begin{pmatrix} x \\ x_K \end{pmatrix} + \begin{bmatrix} 0 \\ D_{yw} \end{bmatrix} w\,,$$

it can easily be seen that the left-hand side matrix is invertible if and only if $I - D_{yu}D_K$ is non-singular. If this holds, one can substitute u, y into (3.7) and find a unique solution to the state equations of the system. Conversely, if $I - D_{yu}D_K$ is singular, when C_y and C_K are slightly perturbed, a linear combination $x(t)$, $x_K(t)$, and $w(t)$ can found to be zero, which means that $x(0)$, $x_K(0)$, and $w(0)$ cannot be chosen arbitrarily.

Hence, the system of Figure 3.6 is internally stable if and only if $I - D_{yu}D_K$ is invertible, and

$$\mathcal{A} = \begin{bmatrix} A & 0 \\ 0 & A_K \end{bmatrix} + \begin{bmatrix} B_u & 0 \\ 0 & B_K \end{bmatrix} \begin{bmatrix} I & -D_K \\ -D_{yu} & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & C_K \\ C_y & 0 \end{bmatrix}$$

is *Hurwitz* (all of its eigenvalues are in the open left half plane, $\mathbb{C}^-$).

Historically, nominal stability was the most important goal of all controller design efforts. Nowadays it is obtained for free, thanks to the H_2 and H_∞ synthesis algorithms, searching for an optimal (or suboptimal) controller from a set of stabilizing controllers. This allows the control engineer to focus on more advanced design goals.

In control design, the performance refers to the closed loop system's behavior. Performance requirements can be formulated in time domain or frequency domain. *Time domain requirements* are often defined in terms of the response behavior to a step input of size one. Typical examples of such requirements are:

i) Limited overshoot (degree to which step value is temporarily exceeded).
ii) Small asymptotic tracking error (how precisely the set value is attained after all transient effects have decayed).
iii) Short rise time (time required to reach a certain percentage of the set value)
iv) Short settling time (time required until the trajectory remains within a certain band around the set value)

Frequency domain requirements are often expressed by demands on *sensitivity matrix function*, $\mathcal{S}$, which is defined as *input sensitivity matrix function* S_i, and *output sensitivity matrix function*, S_o

$$S_i = (I + K\mathrm{G})^{-1} \;\; and \;\; S_o = (I + \mathrm{G}K)^{-1}\,, \tag{3.8}$$

and the *complementary sensitivity matrix function*, $\mathcal{T}$, which is defined to be

$$T_i = (I + K\mathrm{G})^{-1}K\mathrm{G} = K\mathrm{G}(I + K\mathrm{G})^{-1} = K\mathrm{G}S_i \tag{3.9a}$$

$$T_o = (I + \mathrm{G}K)^{-1}\mathrm{G}K = \mathrm{G}K(I + \mathrm{G}K)^{-1} = \mathrm{G}KS_o\,. \tag{3.9b}$$

Note that for SISO systems $S_i = S_o$ and $T_i = T_o$. In the case of positive feedback, input and output sensitivity function matrices are defined as $S_i = (I - K\mathrm{G})^{-1}$ and $S_o =$

$(I - \mathrm{G}K)^{-1}$ respectively. Complementary sensitivity function transfer matrices for positive feedback systems are also defined similarly.

Examining Figure 3.5a the following relations between input/output and closed loop transfer matrices can be defined:

$$v = S_o d_o + \mathrm{G}S_i d_i + T_o r_i - (+)T_o n \tag{3.10a}$$

$$u = -(+)KS_o d_o - (+)T_i d_i + KS_o r_i - (+)KS_o n \tag{3.10b}$$

where (+) sign refers to the case of positive feedback.

In AMB systems, the overall peak (maximum singular value) of output sensitivity matrix function S_o is a measure for robustness of the control system to parameter changes, e.g. due to temperature drifts or component aging [34]. The lower the sensitivity function peak becomes, the more robust to such changes the system performs. The product $\mathrm{G}S_i$ (note that $\mathrm{G}S_i = S_o\mathrm{G}$) constitutes a means for assessing the controlled (closed loop) system's performance, most notably in terms of stiffness and damping of resonant peaks in frequency response. As damping may hide the high gain at resonant frequencies, critical speeds of the rotor in closed loop system can be evaluated by examining the π radian phase shifts in the bode plot of $\mathrm{G}S_i$. Each critical speed occurs at the frequency corresponding to the point of $\pi/2$ radian phase difference within a resonance phase shift interval of π radian.
Typical frequency domain performance requirements comprise:

i) Minimum bandwidth (i.e., T_o must be large up to a certain frequency to achieve a short rise time and good tracking)
ii) Limits on the sensitivity function matrix (i.e., S_i or S_o must be small up to a certain frequency for disturbance rejection)
iii) Limits on some transfer matrices from some inputs to some outputs (e.g. to avoid actuator saturation)
iv) Minimum damping of closed loop poles (i.e., closed loop stability must be assured)

However, there are various performance limitations that prevent achievement of all performance goals at the full extent simultaneously.

A typical performance requirement is the output v to follow reference input r_i. This implies that the output complementary sensitivity matrix function, T_o, should be large. On the other hand, system output should not be affected by the sensor noise. It is obvious from (3.10) that achieving both of the goals over the entire frequency range in contradictory. The designer must trade-off these two goals against each other over frequency.

It can be verified by (3.8) and (3.9) that $\mathcal{S} + \mathcal{T} = I$. Therefore $\mathcal{S}$ and $\mathcal{T}$ can never both be small. This implies that good output disturbance rejection (S_o small) is contradictory to robustness (T_o small), as will become clear in the subsection on *robust stability*. Typical systems have a high level of uncertainty at high frequencies. This demands for T_o being small at high frequencies. S_o, however, must be small at low frequencies in order to have good tracking. In the mid frequency range, a cross-over must take place, and robustness must be traded off against performance.

Trade-offs are not only to made between transfer functions, but are already required when considering sensitivity function $\mathcal{S}$ alone. For rational open loop systems, denoted $L_G(s) = \mathrm{G}(s)K(s)$, with a pole excess of at least two (denominator degree at least by two larger then the numerator degree), the following theorem [32] states that prescriptions for $\mathcal{S}$ in one frequency range have effects on $\mathcal{S}$ at other frequencies.

Theorem 3.1 [27] (Bode Sensitivity Integral Formula) *Given a rational SISO open loop system $L_G(s) = \mathrm{G}(s)K(s)$ with denominator degree at least two larger than its numerator degree (which is a valid assumption for physical systems), with N unstable pole locations p_i, for closed loop stability the following condition must be fulfilled:*

$$\int_0^\infty log|\mathcal{S}(j\omega)|\, d\omega = \pi \sum_{i=1}^{N} \mathrm{Re}(p_i)$$

where $\mathrm{Re}(p_i)$ *denotes the real part of* p_i.

In the case of a stable system the right hand side of the equation is zero, which implies that the area where $\mathcal{S}$ is smaller than one must be compensated by a region of equal (logarithmic) size where $\mathcal{S}$ is larger than one. When the system is unstable, things get worse. Depending on the number and location of unstable poles, the area where $\mathcal{S}$ must be greater than one increases, making good tracking harder for those systems.

The condition imposed by Bode Sensitivity Formula becomes difficult to fulfill due to the fact that vast space of large frequencies can not be used to compensate for low values of $\mathcal{S}$ in the low frequency range. Due to the roll-off of the plant, $\mathcal{T}$ is forced to rapidly approach zero for high frequencies, and since $\mathcal{S} + \mathcal{T} = I$, $\mathcal{S}$ must equally quickly approach one. This leaves only a limited frequency range to fulfill the conditions of Bode Sensitivity Integral, in which the areas of small $\mathcal{S}$ must be balanced with areas of large $\mathcal{S}$. This gives rise to an analogy to a waterbed: Pushing down in one area reduces the water level there, but leads to a rise in water level in another area.

Performance limitations can also be imposed by the open loop system itself. If, for example, the system has right half plane (RHP) zeros, this limits the permissible system gain, since for increasing the gain the closed loop poles move through open loop zeros, and lead to instability. A limitation on gain, however, directly translates into a limit on the frequency up to which the sensitivity function $\mathcal{S}$ can be made small. Unstable poles, on the other hand, require a minimum bandwidth as the stabilization of the poles requires a minimum bandwidth which should cover the frequencies of all poles. This inevitably leads to a minimum value for the frequency at which the complementary sensitivity function $\mathcal{T}$ can start rolling off, which imposes restrictions on the amount of uncertainty the closed loop system will be robust to.

3.2.5. Weighting Functions

In multi-variable controller design, weighting functions are used to limit or bound some performance function matrices (such as S_0) or some signals (such as z and w). The advantage of using weighted performance specifications is obvious. First, some components of a vector signal are usually more important than others. Second, each component of the signal may not be measured in the same units; for example, some signals may be measured in terms of length, and others may be measured in terms of voltage or some others in units of force. Therefore, weighting functions are essential to make these components comparable. Also, we might be primarily interested in rejecting some signals in a certain frequency range, e.g., high frequencies; hence some frequency dependent weights must be chosen.

Each of these signals has some threshold value beyond which the system performance may be said to be unacceptable. If the performance measures are collected into a vector z as in Figure 3.5b, then performance assessment seeks to establish that; $|z_i| \leq z_{max,i}$ simultaneously for all elements of z in response to any foreseeable combination of disturbance inputs, w. Of course, the stipulation of; "any foreseeable combination of disturbance inputs" is crucial. In a linear system if the disturbance signals can be arbitrarily large, then so too will the performance measures. This stipulation implies that some measure of the input signals is bounded. A convenient measure for the signal magnitude is: $|w_i| \leq w_{max,i}$.

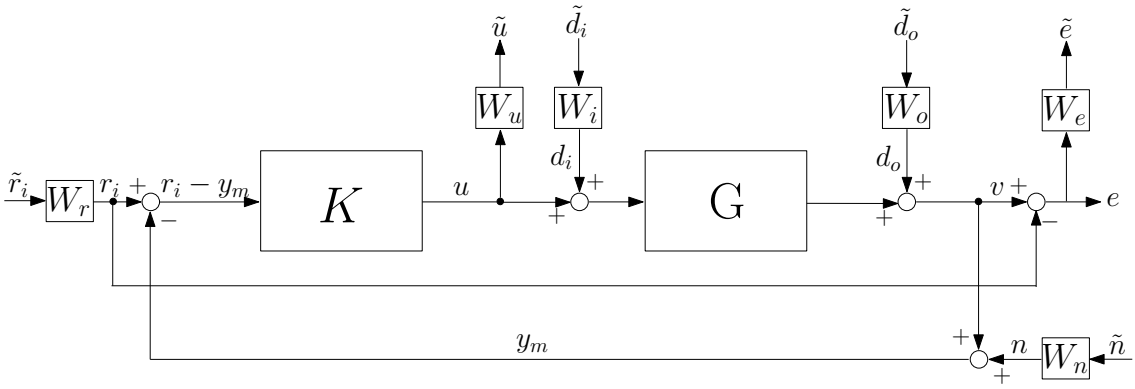

Figure 3.7. Control configuration with weights.

Strict definition of the magnitude of a signal $v(t)$ can take many forms (see Appendix B.1). One choice is the RMS (power) norm

$$\|v\|_{RMS} := \lim_{\tau\to\infty} \sqrt{\frac{1}{\tau}\int_0^\tau v(t)^2\,dt}\,.$$

Using this measure can have important implications for engineering interpretation. In particular, if z_i is measured in this way, then the very largest instantaneous value of z_i could be quite a bit larger than its norm if z_i is a *transient response*. Hence, selection of the RMS norm means that we are focusing on *steady state performance*. This approach is very commonly followed when the frequency response of a system is evaluated. However, in modern control techniques, the signal vectors are mostly measured using the L_2 (energy) norm

$$\|v\|_2 := \sqrt{\int_0^\infty v(t)^T v(t)\,dt} = \sqrt{\frac{1}{2\pi}\int_{-\infty}^\infty \hat{v}(j\omega)^*\hat{v}(j\omega)\,d\omega}\ ,$$

where $\hat{v}(j\omega)$ is the *Fourier transform* of the time signal $v(t)$.

The frequency response view leads to the relatively convenient notion that $|w_i|$ or $|z_i|$ means amplitude of each signal at some particular frequency. Hence, performance analysis seeks to establish that $|z_i| \le z_{max_i}$ for any, including some worst case, $|w_i| \le w_{max_i}$. To exactly establish that each element of a vector z is less than some limit value, normalized vector $\tilde{z}$ can be constructed as

$$\tilde{z} = W_z\, z \ \ :\quad W_z = \mathrm{diag}\,(1/z_{max,i})\,.$$

In a similar fashion, the disturbance signal vector w can also be normalized to read

$$W_w\tilde{w} = w \ \ \Rightarrow\ \ \tilde{w} = W_w^{-1}\,w \ \ :\quad W_w = \mathrm{diag}\,(w_{max,i})\,.$$

In terms of maximum singular value of the closed loop system, if $\overline{\sigma}\,(W_z\,M(j\omega)\,W_w) < \gamma$

$\forall \omega \in \mathbb{R}_e$, the following relations will hold:

$$\tilde{z} = W_z\, M(j\omega)\, W_w\, \tilde{w} : \ \|\tilde{z}\|_2 \leq 1 \ \ \forall \|\tilde{w}\|_2 \leq 1 \quad \Rightarrow \quad \|\tilde{z}\|_2 < \gamma \, \|\tilde{w}\|_2 \, . \tag{3.11}$$

Normalization of the disturbance signals w for the rotor/AMB system using weighting functions can be made as follows:

$$W_i = \mathrm{diag}\,(m_{max}(d/2)\,\Omega^2_{max}) = \mathrm{diag}\,(M_r e_{max}\,\Omega^2_{max})\,, \qquad W_n = \mathrm{diag}\,(n_{max})\,,$$

where m_{max} is the maximum static unbalance mass located at radius $d/2$, M_r is the total mass of the rotor, e_{max} is the maximum eccentricity due to unbalance forces, and n_{max} is the maximum (peak) amplitude of the sensor or any other electrical noise signal. Weighting matrix W_r for the reference signal, which is used to define a trajectory for the rotor (e.g., high speed milling applications) is not considered here as r_i is zero for disturbance rejection problems. There is no exogenous source that can be considered as an output disturbance in rotor/AMB model, therefore, W_o is also redundant. The weighting matrix W_n may be used to model the frequency contents of the sensor/electrical noise, if it is exactly known, at the expense of increasing the order of the system.

Normalization of the performance signal z can be made with the following weighting functions:

$$W_e = \mathrm{diag}\,(1/\mathrm{G}_s s_1)\ , \qquad W_u = W_{un}\, W_{uf} \ : \ W_{un} = \mathrm{diag}\,(1/i_{c,max})\,,$$

where G_s is the sensor gain in (V/m), s_1 is the maximum displacement of the rotor which is the clearance to the touch-down bearings, and $i_{c,max}$ is the maximum control current that can be provided by the power amplifier.

Frequency dependent weight on control input W_{uf} is a kind of low pass filter to limit the *bandwidth* of the controller. If rotor is modeled as a rigid body, band-

width can be limited to its maximum rotational speed. For flexible rotors, bandwidth can be limited to the maximum bending mode eigenfrequency included in the rotor model. Otherwise, the controller may command an actuator with a voltage signal at a frequency close to one of the unmodeled eigenfrequencies of the rotor and may cause stability problems (*spill-over effect*). In some applications, W_e is used to shape the input or output sensitivity matrix function S of the closed loop system in order to reject errors in a certain frequency range (e.g., in low frequency) or to reduce the sensitivity of the system to high frequency disturbances. As this type of weight is frequency dependent, it will increase the order of the generalized plant as well as the order of the controller synthesized. Unfortunately, there is no systematic procedure to select such type of weights, and, in many occasions, this weighting function is chosen purely as a design parameter without any physical basis. Therefore it can be treated as a tuning parameter to achieve the best compromise between the conflicting objectives. Hence, control system design may be regarded as the process of choosing a controller K such that certain weighted signals are made "small" in some sense. Of course, there are different ways to define the smallness of a signal or transfer matrix (see Appendix B.1), and energy norm for signals and maximum singular value for transfer matrices is one of them.

3.2.6. Modeling and Performance Specifications

In Rotor/AMB modeling, the dynamics of the components such as sensors and PWM amplifiers are often neglected as their time constants are too small compared to the rotor and controller dynamics. One approach in modeling is to arrange the open loop system for LFT form, within the framework of Figure 3.5 and Figure 3.7 with positive feedback

$$P(s) = \left[\begin{array}{cc:c} W_e \mathrm{G}\, W_i & 0 & W_e \mathrm{G} \\ 0 & 0 & W_u \\ \hdashline \mathrm{G}\, W_i & W_n & \mathrm{G} \end{array}\right], \text{ with } \quad \tilde{w} = \begin{pmatrix} \tilde{d}_i \\ \tilde{n} \end{pmatrix}, \ \tilde{z} = \begin{pmatrix} \tilde{e} \\ \tilde{u} \end{pmatrix}.$$

Note that, the closed loop system can be described in a lower LFT, $F_l(P, K)$, form

$$\tilde{z} = M(s)\,\tilde{w}\,, \quad \text{where;} \quad M(s) = P_{\tilde{z}\tilde{w}} + P_{\tilde{z}u}K(I - P_{yu}K)^{-1}P_{y\tilde{w}}\,.$$

A state-space realization for the generalized plant P can be obtained by realizing its transfer matrix using a multi-variable realization technique, such as *Gilbert Realization* [26], which yields a minimal realization of the system. However, the direct realization approach is usually complicated. Instead, realization for P can be obtained based on the realizations of each component. G and the other components of the system; W_i, W_n, W_e, W_u have, respectively, the following state-space realizations:

$$\mathrm{G} = \left[\begin{array}{c|c} A_r & B_r \\ \hline C_r & D_r \end{array}\right],$$

$$W_i = \left[\begin{array}{c|c} A_i & B_i \\ \hline C_i & D_i \end{array}\right], \quad W_n = \left[\begin{array}{c|c} A_n & B_n \\ \hline C_n & D_n \end{array}\right], \quad W_e = \left[\begin{array}{c|c} A_e & B_e \\ \hline C_e & D_e \end{array}\right], \quad W_u = \left[\begin{array}{c|c} A_{\tilde{u}} & B_{\tilde{u}} \\ \hline C_{\tilde{u}} & D_{\tilde{u}} \end{array}\right],$$

which is equal to

$$\begin{aligned}
\dot{x}_r &= A_r x_r + [B_{1r}\ B_{2r}]\,(\tilde{w}\ u)^T, & y_r &= C_r x_r + D_r\,(\tilde{w}\ u)^T, \\
\dot{x}_i &= A_i x_i + B_i \tilde{d}_i\,, & d_i &= C_i x_i + D_i \tilde{d}_i\,, \\
\dot{x}_n &= A_n x_n + B_n \tilde{n}\,, & n &= C_n x_n + D_n \tilde{n}\,, \\
\dot{x}_e &= A_e x_e + B_e y_r\,, & \tilde{e} &= C_e x_e + D_e y_r\,, \\
\dot{x}_u &= A_{\tilde{u}} x_u + B_{\tilde{u}} u\,, & \tilde{u} &= C_{\tilde{u}} x_u + D_{\tilde{u}} u\,.
\end{aligned}$$

Weight on input disturbance W_i has no dynamics therefore A_i, B_i, and C_i is zero. Spectral specification of noise signal may not be known in advance, except an estimate

of a bound on its overall amplitude. Therefore, A_n, B_n, and C_n is also zero. D_r can also be set to zero in rotor/AMB applications. Defining a new augmented state vector $x = (x_r\ x_e\ x_u)^T$, a realization for $P(s)$ can be obtained by eliminating the variable y_r

$$\dot{x} = Ax + B_w\tilde{w} + B_u u \tag{3.12a}$$

$$\tilde{z} = C_z x + D_{zw}\tilde{w} + D_{zu} u \tag{3.12b}$$

$$y = C_y x + D_{yw}\tilde{w} + D_{yu} u \tag{3.12c}$$

with the following system matrices:

$$A = \begin{bmatrix} A_r & 0 & 0 \\ B_e C_r & A_e & 0 \\ 0 & 0 & A_{uf} \end{bmatrix}, \quad B_w = \begin{bmatrix} B_{1r} D_i \\ 0 \\ 0 \end{bmatrix}, \quad B_u = \begin{bmatrix} B_{2r} \\ 0 \\ B_{uf} \end{bmatrix},$$

$$C_z = \begin{bmatrix} D_e C_r & C_e & 0 \\ 0 & 0 & D_{un} C_{uf} \end{bmatrix}, \quad D_{zw} = 0\,, \quad D_{zu} = \begin{bmatrix} 0 \\ D_{un} D_{uf} \end{bmatrix},$$

$$C_y = \begin{bmatrix} C_r & 0 & 0 \end{bmatrix}, \quad D_{yw} = \begin{bmatrix} 0 & D_n \end{bmatrix}, \quad D_{yu} = 0\,. \tag{3.13}$$

State-space matrices A_r and B_{2r} of the plant G (which is actually the rotor) are given in (2.17). Considering only the static unbalance force, matrices B_{1r} and C_r can be evaluated as

$$B_{1r} = \begin{bmatrix} 0 \\ \tilde{F}_{ub} \end{bmatrix}, \ \tilde{F}_{ub} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \text{ and } C_r = \mathrm{G}_s\,[C_s\ 0]\ , \ C_s = \begin{bmatrix} 1 & 0 & 0 & s_A \\ 0 & 1 & -s_A & 0 \\ 1 & 0 & 0 & -s_B \\ 0 & 1 & s_B & 0 \end{bmatrix}$$

where G_s is the sensor gain, and s_A/s_B are the distance of displacements sensors (inductive or eddy-current), which are mounted in the vicinity of bearing A and bearing

B, to the origin of the coordinate system (center of gravity of the rotor).

In case W_e is used for the sole purpose of normalizing the rotor position at the bearings, matrices A_e, B_e and C_e will be zero. Hence the matrices in (3.13) shrink to

$$A = \begin{bmatrix} A_r & 0 \\ 0 & A_{uf} \end{bmatrix}, \quad B_w = \begin{bmatrix} B_{1r} D_i \\ 0 \end{bmatrix}, \quad B_u = \begin{bmatrix} B_{2r} \\ B_{uf} \end{bmatrix},$$

$$C_z = \begin{bmatrix} D_e C_r & 0 \\ 0 & D_{un} C_{uf} \end{bmatrix}, \quad D_{zw} = 0\,, \quad D_{zu} = \begin{bmatrix} 0 \\ D_{un} D_{uf} \end{bmatrix},$$

$$C_y = \begin{bmatrix} C_r & 0 \end{bmatrix}, \quad D_{yw} = \begin{bmatrix} 0 & D_n \end{bmatrix}, \quad D_{yu} = 0\,. \tag{3.14}$$

3.2.7. Uncertainty Models

Physical systems can not be modeled with absolute accuracy. As a consequence, a controller that achieves merely nominal stability can not satisfy the control engineer, as no knowledge is available if also the similar but slightly different physical system will be stabilized. Therefore the nominal model is replaced by the concept of a set of models which is obtained by augmenting the nominal model with an appropriate uncertainty description. Robust stability is achieved for such a set of models, if the controller not only stabilizes the nominal model, but all models within the uncertainty specification. If the uncertainty set is cleverly chosen and covers the dynamics of the physical system, this system will be stabilized by all controllers achieving robust stability.
Model uncertainty may result from one or several of the following points:

i) Model Parameter Uncertainty (such as AMB stiffness k_s).
ii) Neglected High Frequency Dynamics (high frequency flexible modes of the rotor).
iii) Nonlinearities (such as hysteresis effects in AMB).

iv) Changing Operating Conditions (as temperature variations may lead to sensor drifts).

v) Neglected Dynamics (such as vibrations of rotor blades).

vi) Parameter Changes Due To Wear (system dynamics may change due to seal wear).

vii) Setup Variations (as a controller for an AMB milling spindle should function with tools of different mass).

viii) Changing System Dynamics (gyroscopic effects change the location of the poles at different operating speeds).

Most common significant sources of uncertainty in rotor/AMB systems are changing dynamics due to gyroscopic effect and varying actuator stiffness, k_s. Effect of gyroscopic forces on the system dynamics can partly be eliminated by appropriate modeling and control (such as LPV), but AMB stiffness k_s is a major source of uncertainty. This parameter is very sensitive to static load acting on the rotor and also anything that modifies the nominal air gap in the actuator including: manufacturing tolerances, thermal growth, and centrifugal growth. It is common to assume that uncertainty in k_s is on the order of 25% [33]. In all cases, it is the responsibility of the control engineer to judge which of the above factors are relevant for the application, and how large the respective uncertainties are. The only information assumed to be known is the maximum size of the expected uncertainty for each frequency ω. This leads to norm bounded perturbations, $\Delta(s)$, that are added to the system, yielding a set with an infinite number possible systems to be considered in the subsequent controller design. While different ways of linking the uncertainty to the system exists, two classes of uncertainty can be distinguished: *Unstructured Uncertainty* and *Structured Uncertainty*.

Unstructured uncertainty is used whenever unknown or neglected system dynamics are to be represented. It is typically represented as additive or multiplicative uncertainty as shown in Figure 3.8.

While $\overline{\sigma}\big(\Delta(j\omega)\big)$ of the delta block is generally considered to be limited to one ($\forall\omega \in \mathbb{R}_e$), the blocks W_a and W_m represent weighting functions that express the size of un-

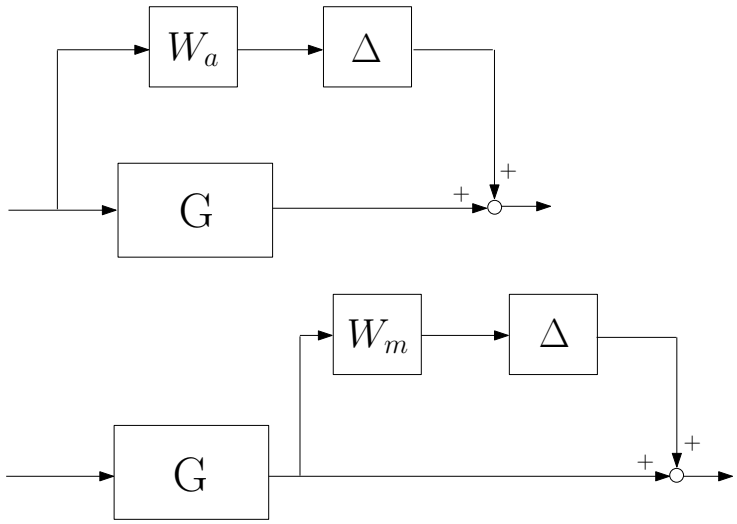

Figure 3.8. Nominal plant with additive and multiplicative uncertainty.

certainty over frequency. Typically uncertainty is small at low frequencies and large and high frequencies. Note that high frequency dynamics of the rotor enters into the system by an additive uncertainty block as high frequency modes are uncoupled from the modes represented in the rotor model G. For multiplicative uncertainty the uncertainty block is always quadratic, while in the MIMO case with additive uncertainty the uncertainty block may have a different number of inputs than outputs, depending on the number of input and output channels of G.

Structured uncertainty occurs whenever more than a single uncertainty is encountered. A special example of structured uncertainty is that of a nominal model with two or more parameters being uncertain. An additional output path is then included in the model for each parameter. The individual delta blocks can be assembled to single delta block similar to that in the unstructured case. The key difference lies in the fact that the delta block encountered here has only entries in its diagonal which makes this a structured uncertainty in contrast to the full delta block in the unstructured uncertainty case.

In practical systems unstructured and parametric uncertainty are often combined, e.g., whenever a system has unmodeled high frequency dynamics and a varying spring stiffness. These blocks can then again be unified in one delta block which again is structured. The resulting delta block is then block-diagonal. Any combination of two

or more blocks of unstructured uncertainty always leads to a structured uncertainty. Although multiple unstructured uncertainties are not used in the uncertainty modeling of practical applications, it is of considerable importance in the synthesis of controller with *robust performance*.

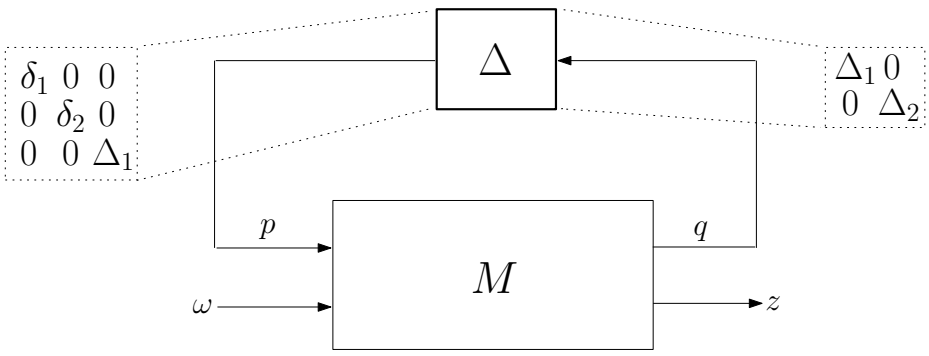

Figure 3.9. LFT representation of uncertainty.

Structured uncertainty can consist of two types of blocks on the diagonal, full uncertainty blocks Δ and scalar blocks δ that may also be repeated (in matrix form; δI). For fixed frequencies ω any structured uncertainty is a complex matrix of the form

$$\mathbf{\Delta} = diag\{(\delta_1\, I_{r1}, ..., \delta_s\, I_{rs}, \Delta_1, ..., \Delta_f) : \ \delta_i \in \mathbb{C}, \ \Delta_j \in \mathbb{C}^{m_j \times m_j}\}$$

and the set of all block diagonal, stable, rational transfer functions that have the same block structure as $\mathbf{\Delta}$ can be written as

$$\mathcal{S}_{\mathbf{\Delta}} := \{\Delta \ : \ \Delta(s) \in \mathbf{\Delta}, \ \forall s \in \mathbb{C}\}\,. \tag{3.15}$$

By adequate algebraic manipulation, the uncertainties can always be pulled out of the system as in Figure 3.9 and the uncertainty model can be written as an upper linear fractional transformation.

The key question in robust stability analysis analysis and design is the question how much uncertainty a nominally stable system can bear before it becomes unstable. For unstructured uncertainty as introduced above, this question is answered by the so-called "Small Gain Theorem".

Theorem 3.2 [26] *Let a stable linear system $M(s)$ is connected to a stable uncertainty system $\Delta(s)$. Then, the closed loop system is internally stable with $\overline{\sigma}\big(\Delta(j\omega)\big) \leq 1/\gamma \;\; \forall\, \Delta$, provided that $\overline{\sigma}\big(M(j\omega)\big) < \gamma, \;\; \forall\, \omega \in \mathbb{R}_e$.*

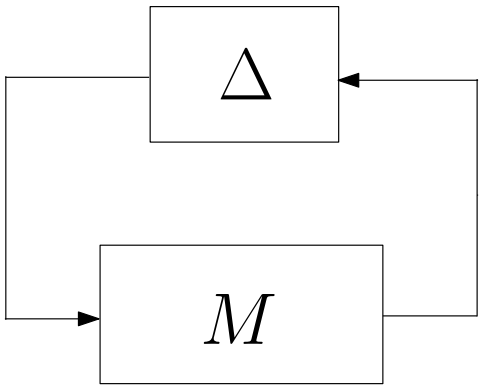

Figure 3.10. Small gain problem.

In order to ensure stability of the closed loop system, the signal gain around the loop must be less than one so that the perturbations introduced into the system get smaller as they circulate around the loop. Hence, the closed loop system can be stable if

$$\overline{\sigma}\big(M(j\omega)\,\Delta(j\omega)\big) < 1 \quad \forall\, \omega \in \mathbb{R}_e\,.$$

As the unstructured uncertainty block Δ is unknown, it is not possible assess this product. However it can easily be established that (see Appendix B.1)

$$\overline{\sigma}\big(M(j\omega)\,\Delta(j\omega)\big) \leq \overline{\sigma}\big(M(j\omega)\big)\;\overline{\sigma}\big(\Delta(j\omega)\big)\,,$$

so that a necessary but not sufficient condition for internal stability is

$$\overline{\sigma}\big(M(j\omega)\big) < \frac{1}{\overline{\sigma}\big(\Delta(j\omega)\big)} \quad \forall\, \omega \in \mathbb{R}_e\,. \tag{3.16}$$

This problem is tractable if it is possible to establish a bound on the maximum singular value of Δ. This is generally possible because it is reasonable to assume that uncertainties in a system are bounded.

The effect of uncertainties on the output sensitivity matrix function S_o is shown in the Figure (3.11). Considering all uncertain dynamics acting on the controlled rotor/AMB system as a lump uncertainty acting on the system, uncertainty perturbations can be modeled as an output disturbance [34]. Referring to Figure (3.11) and equation (3.10), output sensitivity matrix function S_0 for the controlled rotor/AMB system can be defined as

$$q = S_o\, p\,, \quad p = \Delta\, q\,.$$

Note that p affects the system like an output disturbance d_o. According to *small gain theorem*, the overall system can be robustly stable (provided that controlled rotor/AMB and uncertainty system Δ are nominally stable) if

$$\overline{\sigma}\big(S_0(j\omega)\big) < \frac{1}{\overline{\sigma}\big(\Delta(j\omega)\big)} \quad \forall\, \omega \in \mathbb{R}_e\,. \tag{3.17}$$

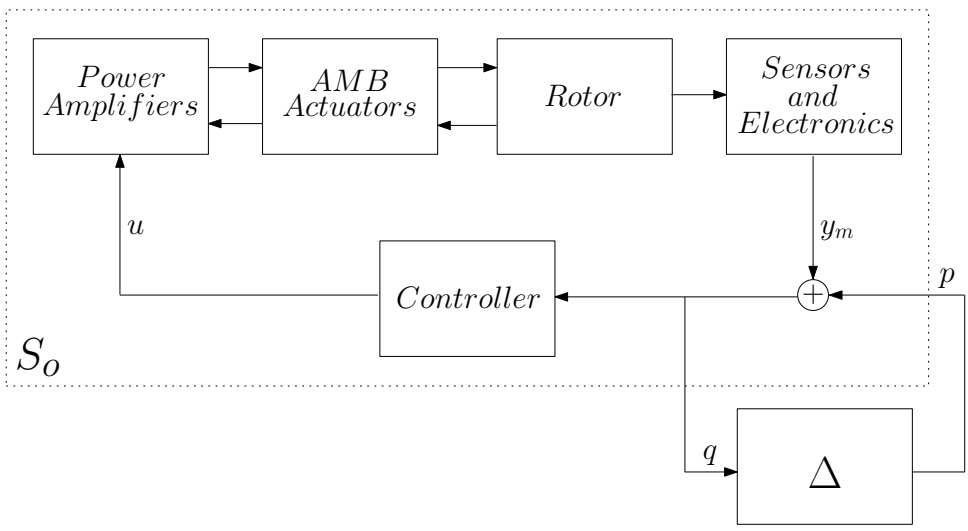

Figure 3.11. Output sensitivity and uncertainty in rotor/AMB system.

If the maximum singular value of S_0 is, for instance, γ_S, then the system can tolerate

a complex gain uncertainty acting in series with the sensors with a maximum singular value of less than or equal to $1/\gamma_S$. Hence, the lower the value $\overline{\sigma}(S_0)$ is, the larger the $\overline{\sigma}(\Delta)$ of lump unstructured uncertainty can be without causing the controlled rotor/AMB system to be unstable.

For structured uncertainty, the above unstructured analysis is conservative. This is due to the fact that the structured uncertainty, which is physically expected, is merely a subset of unstructured uncertainty. To describe the bounds of $\overline{\sigma}\big(\Delta(j\omega)\big)$ is to choose two reasonably simple matrix functions W_p and W_q such that

$$\overline{\sigma}\big(W_p^{-1}(j\omega)\,\Delta(j\omega)\,W_q^{-1}(j\omega)\big) = \overline{\sigma}\big(\tilde{\Delta}(j\omega)\big) \leq 1 \quad \forall\,\omega \in \mathbb{R}_e\,. \tag{3.18}$$

Without the requirement to shape the output sensitivity matrix function S_o, state-space realization to synthesize a robust controller for the rotor/AMB system (for the plant $\tilde{P}$ in Figure 3.12) with two parametric uncertainties, δk and Ω, can be derived in the following form:

$$\begin{aligned}
\dot{x} &= Ax + B_p\tilde{p} + B_w\tilde{w} + B_u u\,,\\
\tilde{q} &= C_q x + D_{qw}\tilde{w}\,,\\
\tilde{z} &= C_z x + D_{zu} u\,,\\
y &= C_y x + D_{yw}\tilde{w}\,,
\end{aligned}$$

$$p = \Delta q\ , \qquad \Delta = \begin{bmatrix} \delta k I & 0 \\ 0 & \Omega I \end{bmatrix}. \tag{3.19}$$

Note from the figure that $\tilde{z} = M(s)\tilde{w}$, where $M(s) = F_l(\tilde{P}, K)$, and $N(s) = F_u(M, \tilde{\Delta})$.

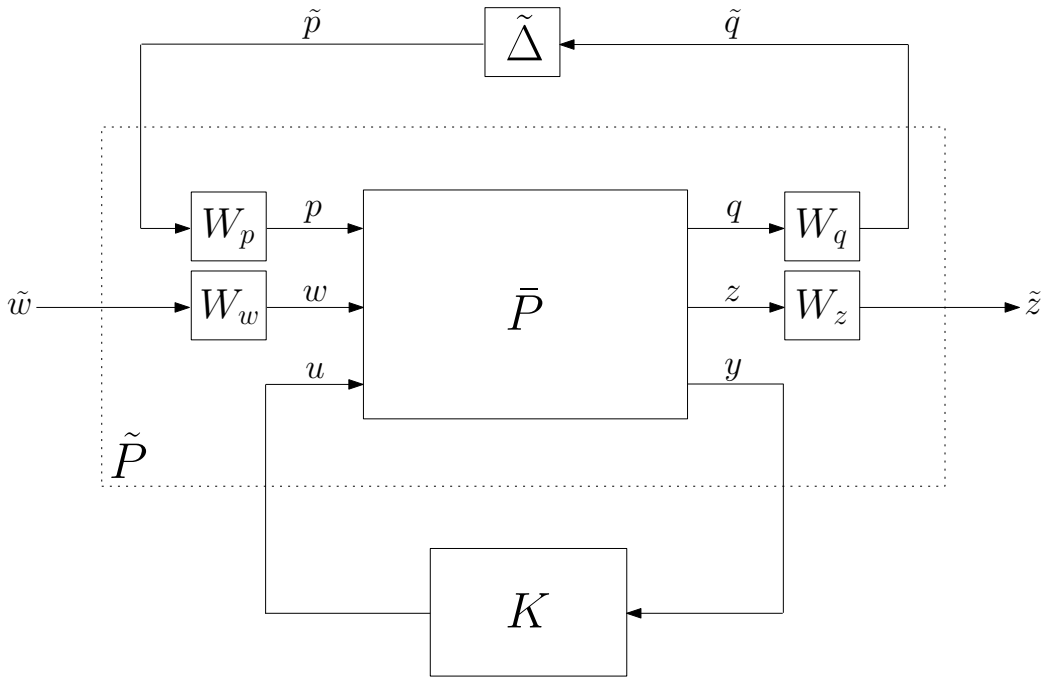

Figure 3.12. Controlled system with uncertainty Δ.

System matrices given in the equations are

$$A = \begin{bmatrix} 0 & I & 0 \\ A_S & \Omega_{max} A_G/2 & 0 \\ 0 & 0 & A_{uf} \end{bmatrix}, \; B_p = \begin{bmatrix} I & I \end{bmatrix}, \; C_q = W_q A_q \,, \; D_{qw} = W_q \begin{bmatrix} 0 \\ 3B_{1r}D_i/4\Omega_{max} \end{bmatrix}. \tag{3.20}$$

Matrix A_q and the weighting matrix functions for the uncertainty block $\tilde{\Delta}$ are

$$A_q = \begin{bmatrix} 0 & A_S & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & A_G/2 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T, \quad W_p = I, \quad W_q = \begin{bmatrix} \delta k_{max} I & 0 \\ 0 & \Omega_{max} I \end{bmatrix}.$$

Stiffness matrix A_S and gyroscopic matrix $A_G = A_{\mathcal{G}}/\Omega$ are given in (2.17). Note that δk is a percentage of nominal negative spring stiffness k_s of the AMB, where Ω is the absolute value of the rotor speed. Augmented state vector x is composed of rotor states and the states of the weighting function W_u selected at the output of the controller, i.e., $x = (x_r \; x_u)^T$. Matrices $B_u, B_w, B_{1r}, C_y, C_z, D_{yw}, D_{zu}$ are given in (3.14). However, B_w of (3.19) is one fourth of B_w of (3.14) as the speed of the nominal system Ω_n is

one half of the maximum operation speed Ω_{max}. Moreover, using (3.2), the closed-loop system $F_l(\tilde{P}, K)$ can be obtained in the form

$$\begin{aligned} M(s) &= \mathcal{C}(sI - \mathcal{A})^{-1}\mathcal{B} + \mathcal{D} \\ &= \begin{bmatrix} M_{\tilde{q}\tilde{p}}(s) & M_{\tilde{q}\tilde{w}}(s) \\ M_{\tilde{z}\tilde{p}}(s) & M_{\tilde{z}\tilde{w}}(s) \end{bmatrix}. \end{aligned} \tag{3.21}$$

3.2.8. Robust Stability and Performance

A control system is robust if it is insensitive to difference between the actual system and the model of the system which was used to design the controller. It should be appreciated that model uncertainty is not the only concern when it comes to robustness. Other considerations include sensor and actuator failures, changes in control objectives, the opening and closing of loops, etc. Furthermore, if a control design is based on optimization, then robustness problems may also be caused by the mathematical objective function not properly describing the real control problem. Also the numerical algorithms themselves may not be robust. However, when we refer to robustness of a system, we usually mean the robustness with respect to model uncertainty.

In robust stability analysis, with a given controller K, one determines whether the system remains stable for all plants in the uncertainty set. If robust stability is satisfied, robust performance analysis is implemented to determine how "large" the transfer function from exogenous inputs w to performance outputs z may be for all plants in the uncertainty set.

In terms of $M\Delta$ structure in Figure 3.4, requirements for stability and perfor-

mance can be summarized as follows:

Nominal Stability (NS)	$\Leftrightarrow$	M is internally stable
Nominal Performance	$\Leftrightarrow$	NS, and $\overline{\sigma}\big(M(j\omega)\big) < \gamma \;\; \forall\omega \in \mathbb{R}_e$
Robust Stability (RS)	$\Leftrightarrow$	NS, and N to be stable $\forall\Delta : \overline{\sigma}\big(\Delta(j\omega)\big) \leq 1 \;\; \forall\omega \in \mathbb{R}_e$
Robust Performance	$\Leftrightarrow$	RS, and $\overline{\sigma}\big(N(j\omega)\big) < \gamma \;\; \forall\Delta : \overline{\sigma}\big(\Delta(j\omega)\big) \leq 1 \;\; \forall\omega \in \mathbb{R}_e$

Suppose that the system M is nominally stable (with $\Delta = 0$), which means that the whole of M, not only M_{zw} (in the expression for N), must be stable. Also assume that Δ is stable. Thus, when we have nominal stability (NS), the stability of the system in (3.4) is equivalent to the stability of the $M\Delta$ structure in Figure 3.4 with $N = M_{zw}$. For nominal performance; $\overline{\sigma}\big(M(j\omega)\big) < \gamma \;\; \forall\omega \in \mathbb{R}_e$, should be valid in addition to the nominal stability condition. Assuming that Δ is also stable, it can also directly be seen that the only source of instability in the closed loop system N in (3.4) can be the feedback term $(I - M_{qp}\Delta)^{-1}$, as all other terms are stable.

For robust stability, the condition $\overline{\sigma}\big(\Delta(j\omega)\big) \leq 1 \;\; \forall\omega \in \mathbb{R}_e$ in (3.18) is equivalent to defining a *unit ball*, $\mathcal{B}_{\mathbf{\Delta}}$, with

$$\mathcal{B}_{\mathbf{\Delta}} = \{\Delta : \; \Delta \in \mathcal{S}_{\mathbf{\Delta}}, \; \overline{\sigma}\big(\Delta(j\omega)\big) \leq 1 \quad \forall\, \omega \in \mathbb{R}_e\}\,.$$

Considering the uncertainty loop in Figure (3.13), the question is: *Are the transfer functions from v_1 and v_2 to any signal in the interconnection stable for all $\Delta \in \mathcal{B}_{\mathbf{\Delta}}$?*
The system in Figure (3.13) can be described in the form

$$\begin{aligned} p &= v_1 + \;\Delta\bar{q} \\ \bar{q} &= v_2 + M_{qp}p \end{aligned} \qquad \Rightarrow \qquad \begin{pmatrix} p \\ \bar{q} \end{pmatrix} = \begin{bmatrix} I & -\Delta \\ -M_{qp} & I \end{bmatrix}^{-1} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}. \tag{3.22}$$

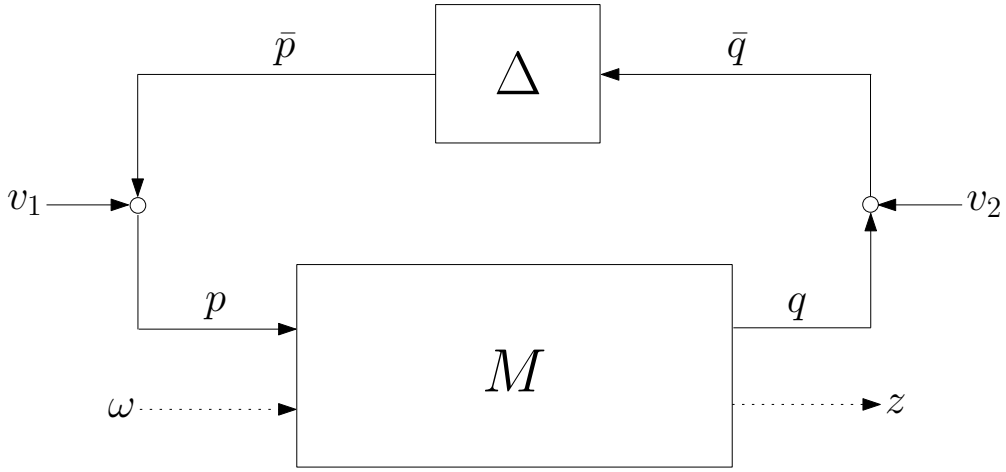

Figure 3.13. Controlled plant with uncertainty perturbations.

Equation (3.22) can be rewritten in the following form (see Appendix B.2):

$$\begin{pmatrix} p \\ \bar{q} \end{pmatrix} = \begin{bmatrix} I + \Delta(I - M_{qp}\Delta)^{-1}M_{qp} & \Delta(I - M_{qp}\Delta)^{-1} \\ (I - M_{qp}\Delta)^{-1}M_{qp} & (I - M_{qp}\Delta)^{-1} \end{bmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}. \tag{3.23}$$

If $(I - M_{qp}\Delta)$ exists, and assuming that the interconnection is well-posed (the sufficient condition to be well-posed is the requirement that the matrix $I - M_{qp}(\infty)\,\Delta(\infty) = I - D_{qp}\Delta$ to be invertible $\forall\Delta \in \mathcal{B}_{\boldsymbol{\Delta}}$. This is guaranteed as, according to (3.19), $D_{qp} = 0$), the only source of instability in (3.23) is $(I - M_{qp}\Delta)^{-1}$ as we have assumed nominal stability for the controlled plant M. This is equivalent to saying that $(I - M_{qp}\Delta)^{-1}$ has no poles in the closed right half plane, $\overline{\mathbb{C}}^{+}$. However, the poles of $(I - M_{qp}\Delta)^{-1}$ are the roots of the polynomial $\det(I - M_{qp}\Delta) = 0$.

Therefore, the system M_{qp} is *robustly stable*, if and only if, $\det(I - M_{qp}\Delta)$ has no zeros in $\overline{\mathbb{C}}^{+}$, $\forall\Delta \in \mathcal{B}_{\boldsymbol{\Delta}}$. If $\det(I - M_{qp}\Delta)$ has no zeros in $\overline{\mathbb{C}}^{+}$, then $(I - M_{qp}\Delta)$ is invertible along extended imaginary axis $\mathbb{C}_e^0$. Hence, for robust stability, it is necessary that

$$\det\big(I - M_{qp}(j\omega)\Delta\big) \neq 0, \quad \forall\Delta \in \mathcal{B}_{\boldsymbol{\Delta}},\ \forall\omega \in \mathbb{R}_e\,.$$

The set of allowable and unallowable uncertainties are defined, respectively, as

$$\Delta \in \mathcal{B}_{\mathbf{\Delta}} \quad \Longleftrightarrow \quad \Delta \in \mathcal{S}_{\mathbf{\Delta}},\ \overline{\sigma}(\Delta) \leq 1$$
$$\Delta \notin \mathcal{B}_{\mathbf{\Delta}} \quad \Longleftrightarrow \quad \Delta \in \mathcal{S}_{\mathbf{\Delta}},\ \overline{\sigma}(\Delta) > 1$$

Therefore, robust stability holds if and only if

$$\inf_{\Delta \in \mathcal{S}_{\mathbf{\Delta}}} \{\overline{\sigma}(\Delta) : \det\big(I - M_{qp}(j\omega)\Delta\big) = 0, \quad \forall \omega \in \mathbb{R}_e\} > 1\,. \tag{3.24}$$

Here we use *infimum* "inf"(greatest lower bound) to define the minimum (smallest) Δ which makes $\det\big(I - M_{qp}(j\omega)\Delta\big) = 0$, attaining $\overline{\sigma}(\Delta) > 1$, as the minimum may only be approached as $\omega \to \infty$ and therefore may not actually be achieved. However, in engineering applications, there is no difference between infimum and minimum.

Inverting (3.24) leads to the definition

$$\mu_{\mathbf{\Delta}}(M_{qp}) := \frac{1}{\inf_{\Delta} \{\overline{\sigma}(\Delta) : \det\big(I - M_{qp}(j\omega)\Delta\big) = 0\}}\,. \tag{3.25}$$

$\mu_{\mathbf{\Delta}}(M_{qp})$ is defined as *structured singular value* of $M_{qp}(j\omega)$ with respect to the uncertainty structure Δ. Robust stability is guaranteed if $\mu_{\mathbf{\Delta}}(M_{qp}) < 1$. Hence, robust stability of nominally stable $M(j\omega)$ is achieved if and only if $\mu_{\mathbf{\Delta}}\left(M_{qp}(j\omega)\right) < 1 \ \ \forall \omega \in \mathbb{R}_e$.

It is well-known that the *computational complexity* of $\mu_{\mathbf{\Delta}}$ is "NP-hard" (see Appendix B.3). However, lower and upper bounds for $\mu_{\mathbf{\Delta}}$, which can be easily computed, can be defined as follows:

Theorem 3.3 [26] *Given a stable system M. The structured singular value $\mu_{\mathbf{\Delta}}$ of M has the lower and upper (crude) bounds; $\varrho(M) \leq \mu_{\mathbf{\Delta}}(M) \leq \overline{\sigma}(M)$, where $\varrho(M)$ is the spectral radius ($max\,\{\,|\lambda_i| : \lambda_i$:eigenvalues of $M\}$), and $\overline{\sigma}(M)$ is the maximum singular value of M.*

See Appendix D for the proof.

Lemma 3.4 [26] *If the n-dimensional uncertainty block is only composed of structured (parametric) uncertainties,* $\Delta = \{\, \delta_i I_n : \delta_i \in \mathbb{C} \,\}$*, then;* $\varrho(M) = \mu_{\mathbf{\Delta}}(M) \leq \overline{\sigma}(M)$*. If the uncertainty block is completely unstructured, , then;* $\varrho(M) \leq \mu_{\mathbf{\Delta}}(M) = \overline{\sigma}(M)$*. for* $\Delta = \mathbb{C}^{n\times n}$

See Appendix D for the proof.

Therefore, if $\overline{\sigma}\big(M_{\tilde{q}\tilde{p}}(j\omega)\big) < 1$ $(\forall\omega \in \mathbb{R}_e)$, it is guaranteed that the uncertain rotor/AMB system will be robustly stable. Considering the closed-loop rotor/AMB model with uncertainty and the partitioned description as in (3.21), $M_{\tilde{q}\tilde{p}}(s)$ can be obtained from the first 8 rows and columns of $M(s)$.

Bounds for $\mu_{\mathbf{\Delta}}(M)$ given in Theorem 3.3 can be refined using transformations on M that do not affect $\mu_{\mathbf{\Delta}}(M)$, but do affect $\varrho(M)$ and $\overline{\sigma}(M)$ [26].
Defining the following two subset of $\mathbb{C}^{n\times n}$

$$\mathcal{U} = \{\, U \in \mathcal{S}_{\mathbf{\Delta}} : U^*U = I_n \,\}$$

$$\mathrm{D} = \left\{ \begin{array}{c} \mathrm{diag}\,[\mathbb{D}_1, ..., \mathbb{D}_S, d_1 I_{m1}, ..., d_{F-1} I_{mF-1}, I_{mF}] : \\ D_i \in \mathbb{C}^{r_i\times r_i},\ \mathbb{D}_i^* = \mathbb{D}_i > 0,\ d_j \in \mathbb{R},\ d_j > 0 \end{array} \right\},$$

the bounds in Theorem 3.3, for $\Delta \in \mathbf{B}_{\mathbf{\Delta}}$, can be tightened to

$$\varrho(M) \leq \max_{U\,\in\,\mathcal{U}} \varrho(UM) \leq \mu_{\mathbf{\Delta}}(M) \leq \inf_{\mathbb{D}\,\in\,\mathcal{D}} \overline{\sigma}(\mathbb{D}M\mathbb{D}^{-1}) \leq \overline{\sigma}(M) \tag{3.26}$$

It has been shown that [26] the tight left inequality $\max \varrho(UM)$ is always an equality. Unfortunately, the maximization of the expression is a non-convex problem with multiple local maxima. The tight upper bound, $\inf_{\mathbb{D}} \overline{\sigma}(\mathbb{D}M\mathbb{D}^{-1})$, however, can be calculated by *convex optimization* through LMIs (see Appendix B.3). Hence, it can be verified that the condition; $\inf_{\mathbb{D}} \overline{\sigma}(\mathbb{D}M\mathbb{D}^{-1}) < 1$ is a sufficient condition for robust stability. However, for necessity, we need to guarantee that $\mu_{\mathbf{\Delta}}(M) = \inf_{\mathbb{D}} \overline{\sigma}(\mathbb{D}M\mathbb{D}^{-1})$ with; $\Delta \in \mathbf{B}_{\mathbf{\Delta}}$.

In some cases, depending on the actual structure of Δ, this is true. For block structures with $2S + F \leq 3$ (S: number of structured (parametric) uncertainty blocks, F: number of full-block (unstructured) uncertainty blocks), $\mu_{\mathbf{\Delta}}(M)$ is equal to the tight upper bound [26], which covers already a large number of application problems.

Stability is not the sole important property of a control system that must be robust to perturbations. While nominal performance is a statement referring to the nominal model that has been used in the controller design, robust performance refers to the set of all similar systems contained in the uncertainty set. Often, there are exogenous disturbances acting on the system (such as unbalance and sensor noise) that result in tracking or regulation errors. Under perturbation, the effect of these disturbances on errors signals can increase. In most cases, long before the onset of instability, the closed loop performance will degrade to the point of unacceptability, hence one needs to test the robustly stable plant's performance under the adverse effects of disturbances. Such a robust performance test will indicate the worst-case level of performance degradation associated with a given level of perturbations.

For robust stability we have the sufficient condition of $\mu_{\mathbf{\Delta}}(M) < 1$ (which is actually $\mu_{\mathbf{\Delta}}(M_{\tilde{q}\tilde{p}}) < 1$). This μ condition is in general not necessary, since cancelations of poles and zeros might occur in the expression; $F_u(M, \tilde{\Delta}) = M_{\tilde{z}\tilde{w}} + M_{\tilde{z}\tilde{p}}\tilde{\Delta}(I - M_{\tilde{q}\tilde{p}}\tilde{\Delta})^{-1}M_{\tilde{q}\tilde{w}}$. If the nominal system has a performance $\overline{\sigma}\big(M(j\omega)\big) < \gamma \ \ \forall\omega \in \mathbb{R}_e$, we have the same performance robustly, if and only if, $\overline{\sigma}\big(F_u(M, \tilde{\Delta})\big) < \gamma \ \ \forall\omega \in \mathbb{R}_e$ as well.

For the rotor/AMB system $F_u(M, \tilde{\Delta})$ can be obtained by using (3.21), and $\tilde{\Delta}$ given as

$$\tilde{\Delta} = \begin{bmatrix} \frac{\delta k}{\delta k_{max}} I & 0 \\ 0 & \frac{\Omega}{\Omega_{max}} I \end{bmatrix}. \tag{3.27}$$

However, standard procedure used in the design of a multi-variable controller with robust performance is based on the tight bound in (3.26) and on the following theorem:

Theorem 3.5 [27] *A stable linear system M with nominal performance γ can have the same performance robustly if $\mu_{\mathbf{\Delta}e}(\tilde{M}) < 1$, where;* $\Delta e = \begin{bmatrix} \Delta & 0 \\ 0 & \Delta_f \end{bmatrix}$,

$\tilde{M} = \begin{bmatrix} M_{\tilde{q}\tilde{p}} & M_{\tilde{q}\tilde{w}} \\ M_{\tilde{z}\tilde{p}}/\gamma & M_{\tilde{z}\tilde{w}}/\gamma \end{bmatrix}$, *with* $\Delta = \mathbb{C}^{n_p \times n_q}$ *and* $\Delta_f = \mathbb{C}^{n_w \times n_z}$.
See Appendix D for the proof.

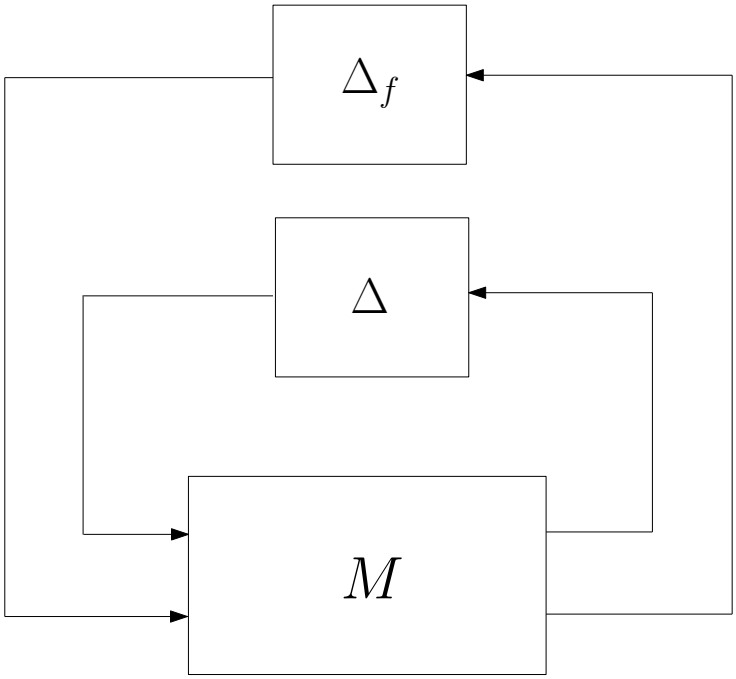

Figure 3.14. Robust performance problem.

Therefore;

$$\overline{\sigma}\big(N(j\omega)\big) = \overline{\sigma}\big(F_u(M, \tilde{\Delta})\big) < \gamma \quad \Longleftrightarrow \quad \mu_{\mathbf{\Delta}e}(\tilde{M}) < 1 \tag{3.28}$$

assuming no pole zero cancelations.

Theorem 3.5 is remarkably useful, as it says that a robust performance problem is equivalent to a robust stability problem with an extended uncertainty structure Δe, incorporating a "fictitious" uncertainty as shown in the Figure 3.14.

3.3. Multivariable LTI Controller Design

3.3.1. LQG Control

Multivariable optimal control based on optimization and optimal filtering theory has reached maturity in 1960's in the form of *Linear Quadratic Gaussian* (LQG) control. In traditional LQG control, it is assumed that the plant dynamics are linear and known, and that the disturbance acting on the plant and measurement noise are stochastic with known statistical properties. Model of the plant is given as

$$\dot{x} = Ax + Bu + w_d$$
$$y = Cx + Du + w_n$$

where w_d and w_n are the disturbance (plant noise) and measurement noise respectively, which are usually assumed to be uncorrelated zero-mean Gaussian stochastic processes with constant power spectral density matrices $\mathcal{W}$ and $\mathcal{V}$.
That is, w_d and w_n are white noise processes with covariances

$$\mathcal{E}\{w_d(t)w_d(\tau)^T\} = \mathcal{W}\delta(t-\tau)$$
$$\mathcal{E}\{w_n(t)w_n(\tau)^T\} = \mathcal{V}\delta(t-\tau)$$

and $\mathcal{E}\{w_d(t)w_n(\tau)^T\} = 0$, $\mathcal{E}\{w_n(t)w_d(\tau)^T\} = 0$, where $\mathcal{E}$ is the expectation operator and $\delta(t-\tau)$ is a delta function.

The LQG control problem is to find the optimal control $u(t)$ which minimizes

$$J = \mathcal{E}\left\{\lim_{T\to\infty}\frac{1}{T}\int_0^T [x^TQx + u^TRu]\,dt\right\},$$

where Q and R are are appropriately chosen constant weighting matrices, such that $Q = Q^* \geq 0$ and $R = R^* > 0$.

The solution to LQG problem consists of first determining the optimal controller for a deterministic *linear quadratic regulator* (LQR) problem, namely, the LQG problem without w_d and w_n. It happens that the solution to this problem can be written in terms of the simple state feedback law; $u = -K_{reg}x(t)$, which is computed as $K_{reg} = R^{-1}B^*X$, where $X = X^* \geq 0$ is the unique positive semidefinite solution of the *algebraic Riccati equation* (ARE)

$$A^*X + XA - XBR^{-1}B^*X + Q = 0\,.$$

The second step is to find an optimal estimate $\tilde{x}$ of the state x, so that $\mathcal{E}\left\{[x-\tilde{x}]^*[x-\tilde{x}]\right\}$ is minimized. The optimal state estimate is given by a *Kalman Filter*. The Kalman filter has the structure of a state estimator or observer with

$$\dot{\tilde{x}} = A\tilde{x} + Bu + K_f(y - C\tilde{x})\,.$$

The optimal choice of K_f, which minimizes $\mathcal{E}\left\{[x-\tilde{x}]^*[x-\tilde{x}]\right\}$, is given by $K_f = YC^*\mathcal{V}^{-1}$, where $Y = Y^* \geq 0$ is the unique positive semidefinite solution of the ARE

$$Y^*A + AY - YC^*\mathcal{V}^{-1}CY + \mathcal{W} = 0\,.$$

The required solution to LQG problem is then found by replacing x by $\hat{x}$, to give

$$u(t) = -K_{reg}\tilde{x}(t)\,.$$

The development of of LQG control coincided with large research programmes on space related problems such as rocket manoeuvring with minimum fuel consumption, which could be well defined and easily formulated as an optimization problem. However, LQG formulation has not proved to be effective on industrial control problems, as accurate plant models are frequently unavailable and the assumption of white noise disturbances is not always relevant or meaningful. The main problem of LQG control

is the fact that; there are no guaranteed stability margins with respect to the perturbations acting on the nominal model of the system and the representation of uncertain disturbances by the white noise processes are often unrealistic [32].

3.3.2. H_2 and H_∞ Control

Motivated by the shortcomings of the LQG control, there was a significant shift in the 1980's towards H_∞ optimization for robust control. H_2 control can be a suitable choice whenever there is information about the spectral content of disturbance w. Therefore, LQG problems can be formulated as a special type of H_2 control problem when w is a stationary noise modeled as stationary random process [27,36]. Moreover, H_2 control is also suitable for minimizing the system gain in transient response.

Consider a proper system $M(s)$, (i.e., $\mathcal{D} \neq 0$ is allowed) in a state-space realization. For the H_∞ norm we use the singular value (induced 2-norm) spatially (for the matrix) and pick out the peak value as a function of frequency

$$\|M\|_\infty := \sup_\omega \overline{\sigma}\big(M(j\omega)\big) \,. \tag{3.29}$$

Here we use *supremum* "sup"(smallest upper bound), as the maximum may only be approached as $\omega \to \infty$ and therefore may not actually be achieved. In terms of performance, H_∞ norm is the peak of the transfer matrix function "magnitude", and by introducing weights, the H_∞ norm can be interpreted as the magnitude of some closed-loop transfer matrix function relative to a specified upper bound.

H_∞ norm also has several time domain interpretations. First, it is the worst-case gain for sinusoidal inputs at any frequency. At a given frequency ω, the gain depends on the direction of $w(\omega)$, and the amplification in the worst-case direction is given by the maximum singular value

$$\overline{\sigma}\big(M(j\omega)\big) = \sqrt{\lambda_{max}\{M(j\omega)^* M(j\omega)\}} \,.$$

The gain also depends on the frequency, and the gain at the worst-case frequency is given by the H_∞ norm as

$$\|M\|_\infty = \sup_\omega \sqrt{\lambda_{max}\{M(j\omega)^* M(j\omega)\}} \ . \tag{3.30}$$

Second, H_∞ norm is equal to the induced (worst-case) 2-norm of any time domain signal

$$\|M\|_\infty = \sup_{w(t)\neq 0} \frac{\|z(t)\|_2}{\|w(t)\|_2} \ . \tag{3.31}$$

This is because the worst input signal $w(t)$ is a sinusoid with frequency ω^* and a direction which gives $\overline{\sigma}\big(M(j\omega^*)\big)$ a maximum gain.

H_∞ norm has also an interpretation as an induced norm in terms of the expected values of the stochastic signals [32]. All these various interpretations make the H_∞ norm very useful in engineering applications.

For the H_2 norm, considering a strictly proper system $M(s)$, (i.e., $\mathcal{D} = 0$) in a state-space realization, *Frobenius norm* (see Appendix B.1) is used spatially (for the matrix) and integrated over frequency

$$\|M\|_2 := \sqrt{\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{trace}\big(M(j\omega)^* M(j\omega)\big)\ d\omega} \ . \tag{3.32}$$

It is clear that $M(s)$ must be strictly proper, otherwise H_2 norm is infinite. H_2 norm can also be given another interpretation. By *Parseval's theorem* (see Appendix B.1), (3.32) is equal to the L_2 norm of the impulse response

$$\|M\|_2 = \|m(t)\|_2 := \sqrt{\int_0^{\infty} \mathrm{trace}\left(\, m(\tau)^T m(\tau)\,\right)\ d\tau} \ . \tag{3.33}$$

Therefore, we have the following deterministic interpretation of H_2 norm, which can

be interpreted as a better performance criterion than H_∞ norm in transient response:

$$\|M\|_2 = \sup_{w(t)=\text{unit impulses}} \|z(t)\|_2 \,. \tag{3.34}$$

Finally, we have a third motivation for the H_2 norm which is of a different nature; here the input is an arbitrary L_2 signal (see Appendix B.1) and has otherwise no known characteristics, but we wish to measure the output signal in the L_∞ norm. That is, our criterion is the induced norm from L_2 to L_∞, which is so-called *generalized 2-norm*

$$\|M\|_{2g} := \|M\|_{L_2 \to L_\infty} = \sup_{\|w\|_2=1} \|z\|_\infty \tag{3.35}$$

also satisfying [42]

$$\|M\|_{2g} = \sqrt{\frac{1}{2\pi}\lambda_{max}\left(\int_{-\infty}^{\infty} M(j\omega)^* M(j\omega)\, d\omega\right)} \,, \tag{3.36}$$

where $\lambda_{max}(\cdot)$ denotes maximum eigenvalue. Notice that when z is scalar valued, the latter expression reduces to the H_2 norm, i.e., for systems with scalar valued variables $\|M\|_{2g} = \|M\|_2$.
In the general case, it can be proved by using *Cauchy-Schwarz inequality* (see Appendix B.1) that; $\|M\|_{2g} \leq \|M\|_2$, hence $\|M\|_2$ is an upper bound to $\|M\|_{2g}$.

To understand the main difference between the H_2 and H_∞ norms, note that (3.32) can be written in terms of singular values

$$\|M\|_2 = \sqrt{\frac{1}{2\pi}\int_{-\infty}^{\infty} \sum_i \sigma_i^2\big(M(j\omega)\big)\, d\omega} \,. \tag{3.37}$$

Comparing (3.29) and (3.37), we see that minimizing the H_∞ norm corresponds to minimizing the peak of the largest singular value, whereas minimizing the H_2 norm corresponds to minimizing the sum of the squares of all the singular values over all frequencies.

The optimal H_2 synthesis problem is stated as follows:
For the system in standard control configuration (Figure 3.2), find a proper, rational, realizable, and stabilizing controller $K(s)$ that leads to a minimal H_2 norm for the closed loop transfer function $M(s)$ from exogenous inputs w to performance outputs z.

Optimal H_2 synthesis problem is based on the solution of *algebraic Riccati equation*

$$A^*X + XA + XRX + Q = 0$$

with the associated *Hamiltonian matrix*:

$$H := \begin{bmatrix} A & R \\ -Q & -A^* \end{bmatrix}.$$

The optimal H_∞ synthesis problem is stated as follows:
For the system in standard control configuration as in Figure 3.2, find the set of all proper, rational, realizable, and stabilizing controllers $\mathcal{K}(s)$ that leads to a minimal H_∞ norm for the closed loop transfer function $M(s)$ from exogenous inputs w to performance outputs z. The optimal H_∞ algorithm leads to a "set of controllers" that achieves minimum worst case amplification of signals in the closed loop system. For practical purposes this theoretical limit is not needed in most of the cases. Furthermore, the search for a "non-unique" optimal controller is numerically and theoretically complicated. For reasons of well-behavedness of the solution and computability, the problem stated above is slightly relaxed in practical applications and the suboptimal problem is solved instead. Moreover, a suboptimal controller may also have better properties (e.g., lower bandwidth) over the optimal ones [27].

The suboptimal H_∞ problem is stated as follows:
Given the set up as the optimal case and a positive scalar γ, find a proper, rational, realizable, and stabilizing controller $K(s)$, if it exists, so that it achieves a closed loop performance $\|M\|_\infty < \gamma$.

General structure of the conventional method, so-called *bisection algorithm* to solve this problem is as follows: Starting out with conservative upper and lower bounds for γ, repeated synthesis attempts are made with the average of the bounds. Depending on whether or not a solution to the problem existed, the upper and lower bound of the problem is adjusted. This process is iterated until the difference between the bounds falls below a predefined threshold value. This procedure is called γ *iteration*. In this way, the minimum is usually not exactly reached but it can be approximated with great precision. In each attempt to solve the suboptimal H_∞ problem two *algebraic Riccati equations* must be solved.

It can be shown that, similar to LQG controllers, H_2 and H_∞ output feedback controllers are composed of two parts; an optimal state observer estimating the system states and a constant feedback of these states [31].
Concerning the order of the controller, it can be said that for a plant with order n (including the weights) an optimal H_∞ controller of order $(n-1)$ exists. Suboptimal controllers are of order n [11].

While there are multitude of design methods in the design of H_2 and H_∞ controllers, they can be split into two main categories: *Loop-shaping based approach*, and *signal based approach*.

Loop-shaping based approach to controller design is somewhat more abstract than the signal based approach. Here the choice of transfer functions to be considered and the size of weighting functions are not mainly based on considerations concerning the size of physical signals encountered. Instead, the scheme is set up in such a way that certain meaningful closed loop transfer functions show a desired behavior in a qualitative way. A well-known method in this approach is the *weighted mixed sensitivity scheme*. As the name suggests, this method aims at minimizing the infinity norm of the sensitivity matrix function S_i (or S_o) and the complementary sensitivity matrix function T_i (or T_o) at the same time. In disturbance rejection problems, minimizing S_i leads to good input disturbance rejection. T_i is the transfer matrix function from the input disturbance d_i to u. Therefore, if avoidance of actuator saturation is critical,

T_i must be minimized. It can be seen from (3.10) that besides the sensitivity and complementary sensitivity matrix functions, there are other transfer matrix functions that may be interesting to minimize. For example, in disturbance rejection problems $\mathrm{G}S_i$ is the transfer matrix function from the input disturbance d_i to the system output v. As a consequence, $\|\mathrm{G}S_i\|_\infty$ determines the worst case effect of input disturbances to the system output. The desired shape of transfer matrix functions to be minimized over frequency, i.e., $\overline{\sigma}(T_i)$ and $\overline{\sigma}(\mathrm{G}S_i)$ versus ω, is achieved by introducing weighting matrix functions $W_T(\omega)$, and $W_S(\omega)$ as shown in the Figure 3.15:

These requirements can be combined into a stacked H_∞ problem

$$\min_K \|L(K)\|_\infty\,, \quad L = \begin{bmatrix} W_U T_i \\ W_S\,\mathrm{G}S_i \end{bmatrix},$$

where K is the stabilizing controller. In other words, we have $z = Lw$ and the objective is to minimize the H_∞ norm from w to z.

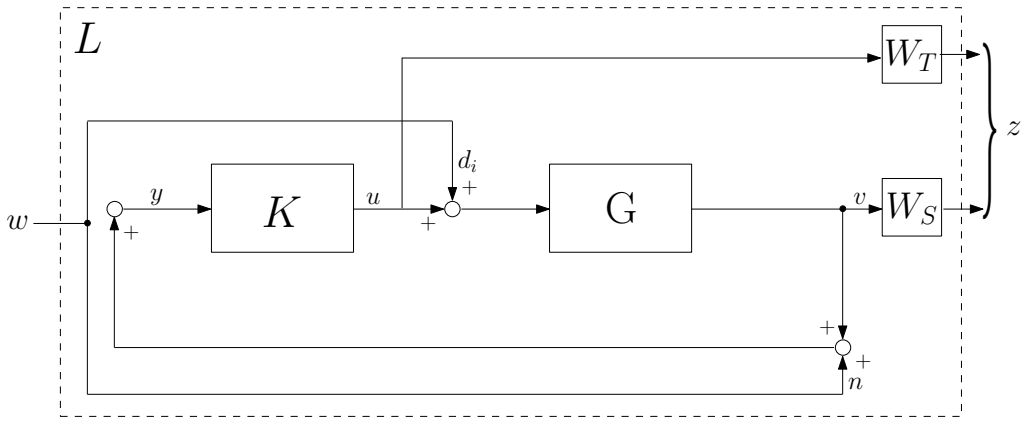

Figure 3.15. Block diagram for (input) disturbance rejection problem in Loop-shaping design.

From Figure 3.15, we get the following set of equations:

$$z_1 = W_S\mathrm{G}u + W_S\mathrm{G}w_1\,, \quad z_2 = W_T u\,, \quad y = \mathrm{G}u + \mathrm{G}w_1 + w_2\,,$$

so that the generalized plant P from $(w\ u)^T$ to $(z\ y)^T$ is

$$P(s) = \left[\begin{array}{cc:c} W_S\mathrm{G} & 0 & W_S\mathrm{G} \\ 0 & 0 & W_T \\ \hdashline \mathrm{G} & I & \mathrm{G} \end{array}\right].$$

Notice the similarity of W_S and W_T with the weighting matrix functions W_e and W_u which are given in the subsection (3.2.6).

Minimization of $\|L\|_\infty$ will include the minimization of $W_T T_i$ and $W_S S_i$ (as G is constant). Due to the *allpass property* of the solution to H_∞ problem [32], the matrix function S_i will be small where the matrix function W_S is large, and similarly T_i will be small where W_T is large. As mentioned in the subsection on Nominal Performance, $S_i + T_i = I$. Therefore, they can not be simultaneously minimized everywhere in the frequency range of operation. However, input disturbance rejection is generally required in low frequency range and minimization of the control input is mainly required in high frequency range (to avoid rate saturation). Due to these considerations W_T is kept small in the low frequencies and W_S is kept small in the high frequencies. The freedom and responsibility of the control engineer lies in deciding how "large" and "small" are to be in each case, where the low frequency range ends and high frequency range starts, and how steep the transition between these areas should be. There are no generic answers to these questions, they depend on the physical system to be controlled and the external signals to be expected during the operation. The application of the H_∞ algorithm and computer simulations gives the information whether design goals are reached or not. In case adjustments are necessary these can be carried out on either one or both of weighting matrix functions W_T and W_S provided that the robustness requirements are also observed.

Signal based approach relies on the information of the physical size of signals that are to be expected or can be accepted. From all possible inputs and disturbances that seem the most important for the present design problem are selected and the corresponding weighting matrix functions are defined to reflect the size of inputs and

the and the inverse of maximum acceptable size of the output signals. The matrix function minimized in this approach is the H_∞ or H_2 norm of the closed-loop transfer matrix function M. Refer to subsections (3.2.5) and (3.2.6) for the application on disturbance rejection problems in rotor/AMB systems.

The advantage of signal based approach lies in the direct mapping of physical reality to the design problem. Moreover, in loop-shaping based approach the uncertainty is modeled in the additive and/or multiplicative form as shown in the Figure 3.8. However, in signal based approach, one can model the system in the structure of Figure 3.12, which is more suitable for further robust analysis and synthesis techniques, such as μ Synthesis.

In the conventional design for optimal H_2 and suboptimal H_∞ controllers based on the Riccati equation theory, the following requirements are assumed to be fulfilled for the system below with standard realization:

$$P(s) = \left[\begin{array}{c|cc} A & B_w & B_u \\ \hline C_z & D_{zw} & D_{zu} \\ C_y & D_{yw} & 0 \end{array}\right]$$

i) (A, B_u) is stabilizable and (C_y, A) is detectable,

ii) $D_{zw} = 0\,, \quad D_{zu}^T D_{zu} \succ 0$ and $D_{yw} D_{yw}^T \succ 0 \quad$ (H_2 synthesis),

iii) $D_{zu} = \begin{bmatrix} 0 \\ I \end{bmatrix}$ and $D_{yw} = \begin{bmatrix} 0 & I \end{bmatrix} \quad$ (H_∞ synthesis),

iv) $\begin{bmatrix} A - j\omega I & B_u \\ C_z & D_{zu} \end{bmatrix}$ has full column rank $\forall\,\omega$,

v) $\begin{bmatrix} A - j\omega I & B_w \\ C_y & D_{yw} \end{bmatrix}$ has full row rank $\forall\,\omega$.

The first assumption is for the stabilizability of P by output feedback. The second assumptions are made to guarantee that H_2 optimal control problem is nonsingular and to make the closed loop system strictly proper to have a finite H_2 norm (synthesizing controller must also have $D_K = 0$). The third assumptions mean that the penalty on the performance output z includes a nonsingular normalized penalty on the control u, that the exogenous signal w includes both plant disturbance and sensor noise, and that the sensor noise weighting is normalized and nonsingular. Relaxation of the third assumption makes the control problem singular. The fourth and fifth assumptions are made for technical reasons on the solvability of Hamiltonian matrices. Relaxation on these assumption make the solution very complicated, and in many cases, leads to some other assumptions to be made [11,26]. The assumption which is implicit in the realization for $P(s)$ as $D_{yu} = 0$ cause no loss of generality as in case $D_{yu} \neq 0$, an equivalent problem can be formed with a linear fractional transformation on the controller $K(s)$, and the controller required can be found as; $K_D(s) = K(I + D_{yu}K)^{-1}$.

Drawback of the Riccati equation theory on H_2 and H_∞ synthesis problems is that it requires unnecessary rank conditions on the generalized plant matrices. Therefore, the ensuing research in removing such restrictions led to the use of *Lyapunov* and *Riccati Inequalities* for *suboptimal H_2 and H_∞ synthesis problems*, which pointed in the direction of LMIs. Moreover, increased performance of computers in 1990's has provided the means to solve LMIs by efficient algorithms through *semidefinite programming* (see Appendix B.3). LMI solution has, however, other advantages beyond removing the unnecessary conditions posed on the plant model. In the first place, the solution methods puts H_∞ synthesis on a common ground with other performance specifications, namely those which involve a closed loop Lyapunov matrix function. This has led to applications to *multi-objective controller design* [14]. LMI formulation has also led to powerful generalizations such as solving problems in LTV and *linear periodic systems* (i.e., $\dot{x}(t) = A(t)x(t) = A(t+\tau)x(t)\,,\ \ \tau > 0$), and the extension to *multidimensional systems* forming the basis of *LPV Control*, an important method for *gain-scheduling* design.

LMI based solutions to H_2 and H_∞ problems use two fundamental theorems in the synthesis for a stabilizing controller, which are stated next.

Theorem 3.6 [36] *The following are equivalent:*
(i) $\|M\|_2^2 < \beta^2$,
(ii) the controllability gramian X_c *satisfies*; $\text{trace}(\mathcal{C}X_c\mathcal{C}^T) < \beta^2$,
(iii) the observability gramian Y_o *satisfies*; $\text{trace}(\mathcal{B}^TY_o\mathcal{B}) < \beta^2$,
(iv) there exists $\mathcal{X} = \mathcal{X}^T \succ 0$ *such that;* $\mathcal{A}\mathcal{X}+\mathcal{X}\mathcal{A}^T+\mathcal{B}\mathcal{B}^T \prec 0$, *and* $\text{trace}(\mathcal{C}\mathcal{X}\mathcal{C}^T) < \beta^2$,
(v) there exists $\mathcal{Y} = \mathcal{Y}^T \succ 0$ *such that;* $\mathcal{A}^T\mathcal{Y} + \mathcal{Y}\mathcal{A} + \mathcal{C}^T\mathcal{C} \prec 0$, *and* $\text{trace}(\mathcal{B}^T\mathcal{Y}\mathcal{B}) < \beta^2$.

LMI formulation to synthesize an H_2 controller can either be based on *(i)-(ii)-(iv)* or *(i)-(iii)-(v)*. In the sequel, we use the former according to the following corollary of Theorem 3.6. For the controllability and observability gramians, see Appendix C.

Corollary 3.7 [36]
Suppose M *is a state-space system with realization* $(\mathcal{A}, \mathcal{B}, \mathcal{C})$. *Then;* $\mathcal{A}$ *is Hurwitz and* $\|M\|_2 < \beta$ *if and only if there exists a symmetric matrix* $\mathcal{X} \succ 0$ *such that*; $\text{trace}(\mathcal{C}\mathcal{X}\mathcal{C}^T) < \beta^2$ *and* $\mathcal{A}\mathcal{X} + \mathcal{X}\mathcal{A}^T + \mathcal{B}\mathcal{B}^T \prec 0$.
See Appendix D for the proof.

Theorem 3.8 [36] (Kalman-Yakubovich-Popov "KYP" Lemma)
Let $\mathbf{M}(s) = \left[\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right]$, *where* $\mathbf{A}$ *has no* $j\omega$*-axis eigenvalues and* $\Gamma = \Gamma^T$ *be given.*

Then the following statements are equivalent:
(i) $\mathbf{M}(j\omega)^*\Gamma\mathbf{M}(j\omega) \prec 0$ *for all* $\omega \in \mathbb{R}_e$.
(ii) There exists a $\mathcal{Y} = \mathcal{Y}^T$ *such that;*

$$\begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \\ I & 0 \end{bmatrix}^T \begin{bmatrix} 0 & 0 & \mathcal{Y} \\ 0 & \Gamma & 0 \\ \mathcal{Y} & 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \\ I & 0 \end{bmatrix} \prec 0\,.$$

Moreover, if $\mathbf{C}^T\Gamma\mathbf{C} \succeq 0$ *then* $\mathbf{A}$ *is Hurwitz if and only if* $\mathcal{Y} = \mathcal{Y}^T \succ 0$.

Note that KYP Lemma establishes the equivalence between a frequency domain domain inequality and a state-space condition in terms of either a Riccati equation or LMI. There are many versions of this result and one may refer to [36] for a proof.

Consider the closed loop system as given in (3.2). For a finite H_2 norm we need $D_{zw} + D_{zu}D_K D_{yw} = 0$. Assume $D_{zw} = 0$ and $D_K = 0$.

Theorem 3.9 [40] *There exists a suboptimal stabilizing controller such that the* H_2 *norm is less than* β *if and only if there exists* $X = X^T$, $Y = Y^T$, $Z = Z^T$, E, F, *and* G *such that*:

$$\begin{bmatrix} AX + XA^T + B_uE + E^TB_u^T & A + G^T & B_w \\ A^T + G & A^TY + YA + FC_y + C_y^TF^T & YB_w + FD_{yw} \\ B_w^T & B_w^TY + D_{yw}^TF^T & -I \end{bmatrix} \prec 0\,,$$

$$\begin{bmatrix} X & I & XC_z^T + E^TD_{zu}^T \\ I & Y & C_z^T \\ C_zX + D_{zu}E & C_z & Z \end{bmatrix} \succ 0\,,$$

$$\mathrm{trace}(Z) < \beta^2\,.$$

If these LMI conditions are satisfied, the required controller is given by

$$C_K = ES^{-1}, \qquad B_K = -Y^{-1}F,$$
$$A_K = (A - B_KC_y)XS^{-1} + B_uC_K - Y^{-1}GS^{-1},$$

where $S := X - Y^{-1}$.

Theorem 3.10 [40] (*Bounded Real Lemma*) *Consider the closed loop system in (3.2). The closed loop system is stable with H_∞ norm less than γ if and only if there exists a $\mathcal{X} = \mathcal{X}^T \succ 0$ such that the LMI*

$$\begin{bmatrix} \mathcal{A}^T\mathcal{Y} + \mathcal{Y}\mathcal{A} & \mathcal{Y}\mathcal{B} & \mathcal{C}^T \\ \mathcal{B}^T\mathcal{Y} & -\gamma I & \mathcal{D}^T \\ \mathcal{C} & \mathcal{D} & -\gamma I \end{bmatrix} \prec 0 \quad \textit{is satisfied.}$$

Lemma 3.11 [40] *There exists a suboptimal stabilizing controller for the system in Theorem 3.5 if and only if there exists $X = X^T$, $Y = Y^T$, E, F, G, and H such that*

$$\begin{bmatrix} AX + XA^T + B_uE + E^TB_u^T & A + B_uGC_y + H^T & B_w + B_uGD_{yw} & XC_z^T + E^TD_{zu}^T \\ A^T + C_y^TG^TB_u^T + H & A^TY + YA + FC_y + C_y^TF^T & YB_w + FD_{yw} & C_z^T + C_y^TG^TD_{zu}^T \\ B_w^T + D_{yw}^TG^TB_u^T & B_w^TY + D_{yw}^TF^T & -\gamma I & D_{zw}^T + D_{yw}^TG^TD_{zu}^T \\ C_zX + D_{zu}E & C_z + D_{zu}GC_y & D_{zw} + D_{zu}GD_{yw} & -\gamma I \end{bmatrix} \prec 0\,,$$

$$\begin{bmatrix} X & I \\ I & Y \end{bmatrix} \succ 0\,.$$

If these LMI conditions are satisfied, the required controller is given by

$$\begin{aligned} D_K &= G, \\ C_K &= (E - D_KC_yX)S^{-1}, \\ B_K &= B_uD_K - Y^{-1}F, \\ A_K &= (A + B_uD_KC_y - B_KC_y)XS^{-1} + B_uC_K - Y^{-1}HS^{-1}, \end{aligned}$$

where $S := Y - X^{-1}$.

See Appendix D for the proofs.

3.3.3. μ Synthesis

According to *small gain theorem*, if the singular value of the closed loop nominal system $\overline{\sigma}\big(M(j\omega)\big)$ is less than γ $\forall\omega \in \mathbb{R}_e$ meaning that $\|M\|_\infty < \gamma$, then the system is robustly stable for unstructured uncertainty with $\|\Delta\|_\infty \leq 1/\gamma$. However, this is a general robustness condition for the case of unstructured uncertainty, and no conclusion can be drawn on the robustness of the closed loop system with structured uncertainty. Actually, H_∞ algorithm merely minimizes the worst case gain for the closed loop nominal system $M(s)$ from exogenous inputs w to performance outputs z (refer to Figure 3.4). If the minimized H_∞ norm γ is greater than one (which is the case for many applications), the transfer matrix function $M_{\tilde{q}\tilde{p}}$ in (3.21) can also be greater one, hence the requirement for the robust stability test. Moreover, the robust performance of the closed loop system with structured uncertainties $N(s)$, can be greater than the H_∞ nominal performance upper bound γ.

The situation would be significantly improved if the H_∞ algorithm could be modified for robust stability and performance *to minimize the structured singular value* of the closed-loop transfer function, i.e., if it would solve the following optimization problem

$$\min_K \sup_\omega \mu_\Delta\big(M(K, j\omega)\big)$$

for a stabilizing controller K. This is referred to as μ Synthesis. Unfortunately, currently no algorithm is available that can achieve this goal directly and guarantee to find the global optimum of the above problem. However, a procedure called *D-K Iteration* yields local optima and constitutes a significant improvement over the classical H_∞ approach for most of the practical control problems. Since μ can not be computed directly (see subsection 3.2.8), the synthesis approach consists of minimizing the upper bound of μ instead of μ itself, and the optimization problem above is relaxed to

$$\min_K \sup_\omega \inf_{D_\omega \in \mathcal{D}} \overline{\sigma}\big(D_\omega M(K, j\omega) D_\omega^{-1}\big)\,.$$

Then the matrices of D_ω which in the above formulation are chosen individually for each ω, are restricted to be parts of $\mathcal{RH}_\infty$ (see Appendix B.1) with the same structure. This simplifies the robust stability and performance problem to

$$\min_K \sup_\omega \inf_{\mathbb{D}\in\mathcal{D}} \|\mathbb{D}\tilde{M}_s(K)\mathbb{D}^{-1}\|_\infty < 1 \iff \text{Robust Stability} \tag{3.38}$$

$$\min_K \sup_\omega \inf_{\mathbb{D}\in\mathcal{D}} \|\mathbb{D}\tilde{M}(K)\mathbb{D}^{-1}\|_\infty < 1 \iff \text{Robust Stability and Performance} \tag{3.39}$$

where $\tilde{M}$ is as given in the subsection (3.2.8). Note that in the synthesis of a controller for only robust stability; the nominal performance γ used in evaluating $\tilde{M}$ can be relaxed to γ_s, and $\tilde{M}_s$ can be constructed accordingly, provided that the resulting performance degradation is acceptable.

The D-K iteration finds *locally optimal solutions* to these problems by the following two step procedure:

K-Step:
Given scaling matrix functions $\mathbb{D} \in \mathcal{RH}_\infty$ (initially identity matrix) and holding these fixed an optimal controller for the problem in (3.38) or (3.39) is calculated by means of the H_∞ algorithm.

D-Step:
Given the controller K from the precious step, matrices D_ω minimizing the expression

$$\inf_{D_\omega\in\mathcal{D}} \overline{\sigma}\big(D_\omega M(K, j\omega) D_\omega^{-1}\big)$$

are computed for each frequency in the grid. Then the optimal matrices D_ω are approximated by a stable, minimum phase transfer matrix function $\mathbb{D}(s)$.

This procedure is iterated until no further improvement of the minimum in (3.38) or (3.39) is reached. Both in the synthesis for robust stability and in the synthesis

for robust stability and performance, if the achieved minimum is less than one then the design is successful. With either K or $\mathbb{D}$ fixed, the *global optimum* in the other variable can be found with a useful convex optimization algorithm. However, the joint optimization of $\mathbb{D}$ and K is not convex (see Appendix B.3) and the global convergence is not guaranteed.

Although fairly complex, well-developed software tools for performing D-K iteration are available [49]. The order of the controllers resulting from the procedure is that of the plant (including weighting matrix functions) plus that of $\mathbb{D}$-scales (twice that of $\mathbb{D}$(s)). Since high order $\mathbb{D}$-scales may be required for efficient minimization of the objective function, this implies a trade-off between closed loop performance and the controller size. Both the choice of maximum scaling matrix function order and of the frequency grid are up to the user. While no guidelines are available in the literature, it can be said that the frequency grid should comprise all eigenfrequencies of the plant model.

The μ synthesis controller design framework is a versatile and effective method for designing robust controllers. Nevertheless, it must be mentioned that the currently available algorithms have considerable problems in dealing with real valued uncertainty. This constitutes an important drawback since this type of uncertainty is often encountered in practical applications. For example uncertain negative spring stiffness of the AMB or uncertain rotational speeds will always be real valued. While unstructured real uncertainty can not be dealt with at all, parametric (structured) real uncertainties can be nominally treated. However, the computation of the μ bounds for real valued uncertainty is not very reliable [26]. As a consequence, convergence problems are often encountered. However, replacing real uncertainty with complex uncertainty often is not a viable alternative since this introduces considerable amount of conservatism (in the complex plane, the parameter then lies on a circular disk rather than on a line). This makes the problem not tractable in some cases as the physically meaningless complex parameter values may lead to equally meaningless systems being considered in the uncertainty model. This results in controllers that have poor performance on the physically relevant systems.

Besides the standard D-K iteration described above, there are some refined algorithms that aim at obtaining controllers with μ values closer to *global optimum.* Among this is the D-G-K iteration which, based on an alternative characterization of μ, tries to perform a *mixed real/complex synthesis* with an additional set of scalings (G scalings) [27]. Mixed real/complex synthesis consists of duplicating the inputs and outputs associated with the real uncertainties, and adding a small, complex uncertainty to these new channels. The "mixed" set-up results in a less conservative uncertainty model (the parameter now lies on a narrow, oval-shaped area around the line) and yields much better convergence of the μ computation. Unfortunately, additional input and output channels also imply a higher plant order and more scaling matrices to be computed in the iteration, resulting in higher order controllers. For analysis purposes, however, this constitutes a considerable improvement over analyzing systems with complex uncertainty.

3.3.4. Controller Order Reduction

Modern multivariable control techniques such as H_∞ and μ-synthesis produce controllers of order at least equal to that of plant, and usually higher because of the inclusion of weights. These control laws may be too complex with regards to practical implementation and simpler designs are normally preferred for some obvious reasons: They are easier to understand and computationally less demanding; they are also easier to implement and have higher reliability since there are fewer things to go wrong in the hardware or bugs to fix in the software. Therefore, a lower-order controller should be sought whenever the performance degradation is kept within an acceptable level. The central problem we address is: Given a high-order LTI stable controller model K, find a low-order approximation K_r such that the closed-loop stability is preserved and the infinity norm i.e., H_∞ or $\hat{L}_\infty$ norm of the difference $\|K - K_r\|$ is small. Note that the error $\|K - K_r\|$ may be unstable and therefore not defined in RH_∞ (see Appendix B.1). Hence we use $\hat{L}_\infty$ norm for the systems in RL_∞ which defines the set of rational functions with no poles on the imaginary axis.

Let (A, B, C, D) be a minimal realization of a stable system $P(s)$, and partition the state vector x, of dimension n, into $[x_1\ x_2]^T$ where x_2 is the vector of $n-k$ states which we wish to remove. With appropriate partitioning of A, B, and C, the state-space equations become

$$\begin{aligned}\dot{x}_1 &= A_{11}x_1 + A_{12}x_2 + B_1 u\,,\\ \dot{x}_2 &= A_{21}x_1 + A_{22}x_2 + B_2 u\,,\\ y &= C_1 x_1 + C_2 x_2 + Du\,.\end{aligned}$$

A k'th order truncation of the realization $P(s) := (A, B, C, D)$ is given by $P_r(s) := (A_{11}, B_1, C_1, D)$. The truncated model K_r is equal to K at infinite frequency, $P(\infty) = P_r(\infty) = D$, but apart from this there is little that can be said in the general case about the relation between P and P_r. If A is in Jordan form (see Appendix B.2) then it is easy to order the states so that x_2 corresponds to high-frequency or fast modes. However, the distance of the eigenvalues λ_i of the diagonalized matrix A from the imaginary axis is, by itself, not a reliable indicator of whether the associated mode should be included in the reduced order model or not [32]. An advantage of model truncation is that the poles of the truncated model are a subset of the poles of the original model and therefore retain any physical interpretation they might have.
Suppose that instead of discarding all the states and dynamics associated with x_2, we simply set $\dot{x}_2 = 0$, i.e., we *residualize* x_2, in the state-space equations. One can then solve for x_2 in terms of x_1 and u, and back substitution of x_2 then gives

$$\begin{aligned}\dot{x}_1 &= (A_{11} - A_{12}A_{22}^{-1}A_{21})x_1 + (B_1 - A_{12}A_{22}^{-1}B_2)u\,,\\ y &= (C_1 - C_2A_{22}^{-1}A_{21})x_1 + (D - C_2A_{22}^{-1}B_2)u\,.\end{aligned}$$

Assuming that A_{22}^{-1} is invertible, the matrices for the reduced system are

$$\begin{aligned}A_r &:= A_{11} - A_{12}A_{22}^{-1}A_{21}\,, \quad & B_r &:= B_1 - A_{12}A_{22}^{-1}B_2\,,\\ C_r &:= C_1 - C_2A_{22}^{-1}A_{21}\,, \quad & D_r &:= D - C_2A_{22}^{-1}B_2\,.\end{aligned}$$

The reduced order model $P_r(s) := (A_r, B_r, C_r, D_r)$ is called a residualization of $P(s) := (A, B, C, D)$. An important property of residualization is that is preserves the steady-state gain of the system, i.e., $P(0) = P_r(0)$. This is in contrast to truncation which retains the system behavior at infinite frequency. Hence it is clear that truncation is to be preferred when accuracy is required at high frequencies, whereas the residualization is better for low-frequeny modelling. Both methods depend to a large extent on the original realization and Jordan form may be used. A better realization, with many useful properties, is the *balanced realization.*

Balanced realization is an asymptotically stable minimal realization minimal realization in which controllability and observability gramians (see Appendix C) are equal and diagonal. Let (A, B, C, D) be a minimal realization of a stable rational transfer matrix function $P(s)$. Then (A, B, C, D) is called balanced if the solutions to the following *Lyapunov equations*

$$AX + XA^* + BB^* = 0$$
$$A^*Y + YA + C^*C = 0$$

are $X = Y = \text{diag}(\sigma_1, \sigma_2, ...\sigma_n) := \Sigma$, where $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_n$. X and Y are the controllability and observability gramians defined also by

$$X := \int_0^\infty e^{At} BB^* e^{A^*t} dt\,,$$
$$Y := \int_0^\infty e^{A^*t} C^* C e^{At} dt\,.$$

Σ is therefore simply referred to as the gramian of $P(s)$. The σ_i's are the ordered *Hankel singular values* of $P(s)$, more generally defined as $\sigma_i := \lambda_i^{1/2}(XY)\,, i = 1, 2, ..., n$. Notice that $\sigma_1 = \|P\|_H$, the *Hankel norm* of $P(s)$.

In a balanced realization the value of each σ_i is associated with a state x_i of the balanced system. And the size of σ_i is a relative measure of the contribution that x_i makes to the input-output behavior of the system.

Therefore if $\sigma_1 > \sigma_2$, then the state x_1 affects the input-output behavior more than x_2, or indeed any other state because of the ordering of σ_i. After balancing a system, each state is just as controllable as it is observable, and a measure of a state's joint observability and controllability is given by its associated singular value.

Let the balance realization (A, B, C, D) of $P(s)$ and the corresponding Σ be partitioned compatibly as

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = \begin{bmatrix} C_1 & C_2 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix},$$

where $\Sigma_1 = \text{diag}(\sigma_1, \sigma_2, ..., \sigma_k)$, $\Sigma_2 = \text{diag}(\sigma_{k+1}, \sigma_{k+2}, ..., \sigma_n)$ and $\sigma_k > \sigma_{k+1}$. The reduced order model given by (A_{11}, B_1, C_1, D) is called a *balanced truncation* of the full-order system. A balanced truncation is also a balanced realization and the H_∞ norm of the error between $P(s)$ and the reduced order system is bounded by twice the sum of the last $n - k$ Hankel singular values [32].

In a balanced truncation above, we discarded the least controllable and observable states corresponding to Σ_2. In *balanced residualization* we simply set to zero the derivatives of all these states. The resulting balanced residualization of $P(s)$ is (A_r, B_r, C_r, D_r) which has been derived above. It has been shown that [32] balanced residualization enjoys the same error bound as balanced truncation. Note that balanced realization preserves the steady-state gain of the system and balanced residualization is related to balanced truncation by the bilinear transformation $s \to s^{-1}$.

An alternative approach is the *optimal Hankel norm approximation* which can be defined as finding a reduced order model $P_r(s)$ of degree k such that the Hankel norm of the approximation error, $\|P(s) - P_r(s)\|_H$, is minimized. Therefore, in the optimization we seek an error which is in some sense closest to being completely unobservable and completely uncontrollable.

Let $P(s)$ be factorized as $P(s) = M^{-1}(s)N(s)$, where $M(s)$ and $N(s)$ are stable

left coprime factors [26] of $P(s)$. Approximating $[N\ M]$ of degree n by $[N_r\ M_r]$ of degree $k < n$ using balanced truncation, balanced residualization or optimal Hankel norm, a reduced order transfer matrix function $P_r(s) = M_r^{-1}N_r$ of degree k can be realized.

3.4. LPV Control

3.4.1. Introduction

Modeling of uncertainty region is a major design step in controller synthesis. Designer, very often, faces with the problem of constraining uncertainty regions of any size into a general uncertainty set model, though this is quite complicated if the system behaves differently at different operating points. As the uncertainty region can change, the governing equations can also change, thus the modeled system diverges from the actual physical system. Hence, some physical systems, albeit linear, can not always be represented by a single LTI model.

The idea of having multiple LTI models for a general representation of parametrically varying system is considered in [30]. Essentially, the entire plant dynamics are divided in several operating stations with a certain region size, where the union of these cover the full operating range. This technique is known as *gain-scheduling.* Although a wide variety of methods are described as gain-scheduling approaches, these are usually linked by a divide and conquer type of design procedure whereby the LTV or even nonlinear control task is decomposed into a number of LTI problems.

The first appearing algorithms are somewhat limited in an analytical sense though they are validated by experimental analysis. This is actually the most appearing drawback, because there is no guarantee of having any of the performance and robustness in the transition region between these discrete operating points which have certain closed loop robustness and stability specifications. Thus, the designer can not assess *a priori* stability and robustness of the system unless validation actions such as extensive

simulations or physical experiments are undertaken. As an alternative or complement, robustness analysis can be used to verify performance.

Robustness and gain scheduling are closely related. Stability and performance of gain scheduling controllers can be analyzed using robustness criteria if the parameters used for gain scheduling are treated as uncertain parameters available to the controller in real time.

The demand for systematic, theoretically rigorous techniques for the gain scheduling method has stimulated a great deal of research on linear parameter-dependent systems where system matrices are matrix functions of time varying parameters in the form

$$\begin{pmatrix} \dot{x} \\ z \\ y \end{pmatrix} = \begin{bmatrix} A(\rho) & B_w(\rho) & B_u(\rho) \\ C_z(\rho) & D_{zw}(\rho) & D_{zu}(\rho) \\ C_y(\rho) & D_{yw}(\rho) & D_{yu}(\rho) \end{bmatrix} \begin{pmatrix} x \\ w \\ u \end{pmatrix},$$

where the parameter(s) $\rho(t)$ takes on values in a known *compact* (closed and bounded) set $\mathcal{P} \in \mathbb{R}^m$, and its rate of change $\dot{\rho}$ is known. However, time variation of the parameters is not known in advance and they are measured in real-time with sensors for control. Hence controller is also parameter-dependent, using the available real-time information of the parameter variation.

Associated with parameter-dependent systems there are two main analysis and design approaches: *Dissipative systems theory* based on LPV models, and *linear robust systems theory* based on linear fractional representation LFR *loop transformation* models. Distinction draws from:

i) The allowable dependence that the state-space data has on parameters,
ii) the extent to which information about the parameter's variations are exploited in the analysis,
iii) the different techniques that are used to analyze systems.

In the dissipative systems framework for parameter dependent systems, the following structure, in LPV form, is used to describe the system

$$\begin{pmatrix} \dot{x}(t) \\ z(t) \end{pmatrix} = \begin{bmatrix} \mathcal{A}\big(\rho(t)\big) & \mathcal{B}\big(\rho(t)\big) \\ \mathcal{C}\big(\rho(t)\big) & \mathcal{D}\big(\rho(t)\big) \end{bmatrix} \begin{pmatrix} x(t) \\ w(t) \end{pmatrix},$$

where $\rho(t) \in \mathcal{P}$ is a trajectory (called an *allowable* ρ trajectory), which satisfies for each ρ_i $(i = 1, ..., m)$, and for all t; $\underline{\nu}_i \leq \dot{\rho}_i(t) \leq \overline{\nu}_i$. Lower and upper bounds for the rate of change of parameter ρ_i; $\underline{\nu}_i$ and $\overline{\nu}_i$, are functions that map $\mathcal{P}$ to $\mathbb{R}^m$.

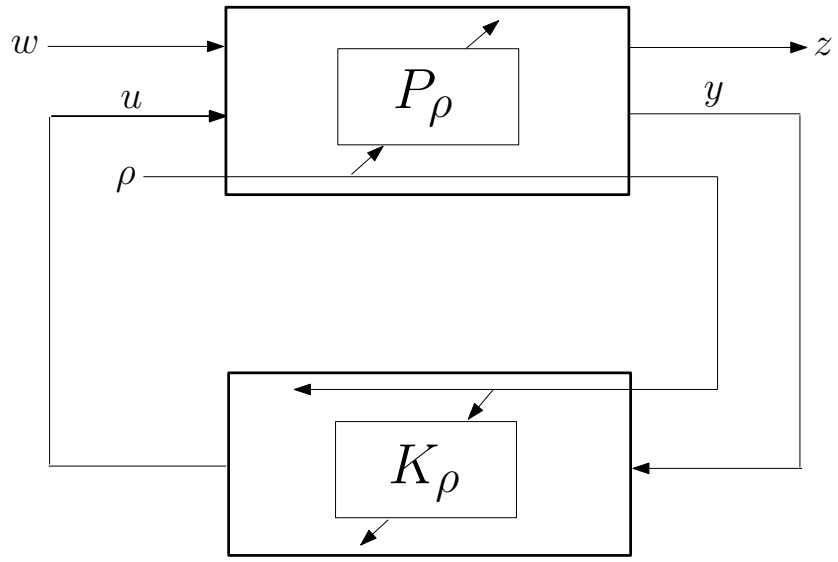

Figure 3.16. LPV representation of parameter dependent systems.

In linear robust systems framework, the system is modeled in a way similar to the shown in Figure 3.4 using the closed loop transfer matrix function $M(s)$, but with time varying uncertainty structure

$$\begin{pmatrix} z \\ q \end{pmatrix} = \begin{bmatrix} M_{zw} & M_{zp} \\ M_{qw} & M_{qp} \end{bmatrix} \begin{pmatrix} w \\ p \end{pmatrix},$$

with

$$p = (p_P \; p_K)^T, \;\; q = (q_P \; q_K)^T, \;\; \text{and} \;\; p(t) = \Delta(t) q(t)\,.$$

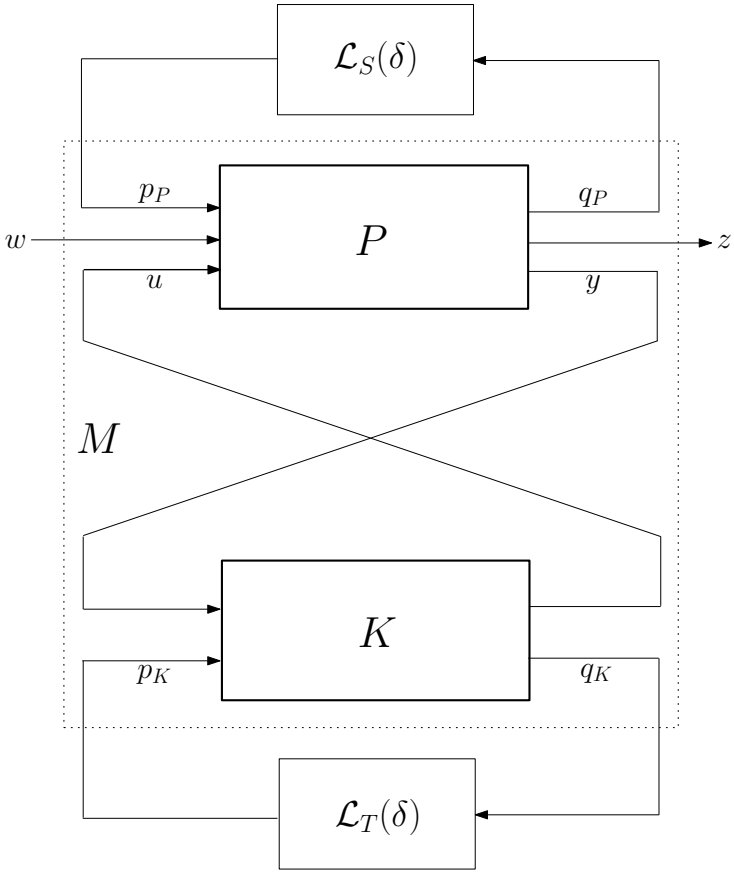

Figure 3.17. LFR structure for parameter dependent systems.

where vector of integers S describe the plant's parameter structure, and vector of integers T describe the controller's parameter structure. Notice that these classes of systems are closed under interconnection, i.e., an LFT of an LFT is an LFT.
For systems modeled in LPV form stability definitions are based on *parameter-dependent Lyapunov functions* (PDLFs), whereas for LFR systems it is based on *structured small gain theorems*. In the analysis (and possibly synthesis) of LFR structured systems, uncertainty descriptions are based on *integral quadratic constraints* (IQCs) [46].

In order the optimize the performance of the systems with the controller K, usually the "worst-case" (over allowable parameter trajectories) induced L_2 gain (H_∞

norm) from $w \to z$ is used as follows:

$$\text{LPV Performance Metric:} \quad \max_{\rho} \max_{w,\, w\neq 0} \frac{\|M_\rho w\|_2}{\|w\|_2}\,.$$
$$\text{LFR Performance Metric:} \quad \max_{\delta} \max_{w,\, w\neq 0} \frac{\|M_\Delta w\|_2}{\|w\|_2}\,.$$

Furthermore, some nonlinear systems can also be represented in parameter-dependent structure, which is so-called *quasi-LPV* form. Consider the nonlinear state equations

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

where $x(t) \in \mathbb{R}^n$ is the state, $u(t) \in \mathbb{R}^{n_u}$ is the control input, $\delta(t) \in \mathbb{R}^{n_\delta}$ is the exogenous time-varying parameter(s), and $y(t) \in \mathbb{R}^{n_y}$ is the output.

Suppose that the state x is partitioned into two groups as; $x = (x_1\ x_2)^T$, and assume that $x_1 \in \mathbb{R}^{n_1}$ and δ are available for scheduling purposes, though $x_2 \in \mathbb{R}^{n_2}$ is not. Then the nonlinear system can be reduced to a quasi-LPV form if there exist;

i) well-defined functions $A_{ij}, \tilde{B}_i, C_{yi}$, and $\tilde{D}$ without singularities,
ii) a function $\tilde{u} : \mathbb{R}^{n_1} \times \mathbb{R}^m \times \mathbb{R}^{n_u}$, written $\tilde{u}(x_1, \delta, u)$ that is invertible with respect to its third argument, such that the governing equations can be rewritten as:

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ y \end{pmatrix} = \begin{bmatrix} A_{11}(x_1,\delta) & A_{12}(x_1,\delta) & \tilde{B}_1(x_1,\delta) \\ A_{21}(x_1,\delta) & A_{22}(x_1,\delta) & \tilde{B}_2(x_1,\delta) \\ C_{y1}(x_1,\delta) & C_{y2}(x_1,\delta) & \tilde{D}(x_1,\delta) \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \tilde{u}(x_1,\delta,u) \end{pmatrix}.$$

Two facts are important:

i) $\tilde{u}(x_1, \delta, u)$ enters the dynamical equations linearly,
ii) there exists a function $\tilde{v} : \mathbb{R}^{n_1} \times \mathbb{R}^m \times \mathbb{R}^{n_u} \to \mathbb{R}^{n_u}$, such that for all $x_1 \in \mathbb{R}^{n_1}$, $\delta \in \mathbb{R}^m$ and $\check{u} \in \mathbb{R}^{n_u}$, such that; $\tilde{u}\big(x_1, \delta, \tilde{v}(x_1, \delta, \check{u})\big) = \check{u}$ holds.

Hence, we can treat $\check{u}$ as the input, leaving quasi-LPV equations

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ y \end{pmatrix} = \begin{bmatrix} A_{11}(x_1,\delta) & A_{12}(x_1,\delta) & \tilde{B}_1(x_1,\delta) \\ A_{21}(x_1,\delta) & A_{22}(x_1,\delta) & \tilde{B}_2(x_1,\delta) \\ C_{y1}(x_1,\delta) & C_{y2}(x_1,\delta) & \tilde{D}(x_1,\delta) \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \check{u} \end{pmatrix},$$

with control input $\check{u}$ designed using a suitable feedback design approach. The actual input u is computed via scheduled inverse; $u(t) := \tilde{v}\big(x_1(t), \delta(t), \check{u}(t)\big)$.
A model with LPV assumption ignores the fact that $x_1 = x_1$, letting $\rho \in \mathbb{R}^{n_1+m}$ be an independent exogenous parameter (covering x_1 conservatively along with δ)

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ y \end{pmatrix} = \begin{bmatrix} A_{11}(\rho) & A_{12}(\rho) & \tilde{B}_1(\rho) \\ A_{21}(\rho) & A_{22}(\rho) & \tilde{B}_2(\rho) \\ C_{y1}(\rho) & C_{y2}(\rho) & \tilde{D}(\rho) \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \check{u} \end{pmatrix}.$$

Therefore, LPV structured models for a nonlinear system is conservative as this introduces additional behavior that the real system can not actually exhibit.

Decision on the selection of the methodology to design a controller for parameter-dependent systems is mainly based on engineering judgement. Nevertheless, it can be said that for a system with few uncertain parameters to consider and relevant rate bounds, LPV form is suitable. However for systems with many quickly varying parameters and/or nonlinearities, LFR structure (LFT with multipliers) [47] can be a better choice. Concerning the rigid body model of the rotor/AMB system, where the uncertainty structure is composed of one measurable (rotor speed) and one unmeasurable parameter (negative spring stiffness of the AMB), LPV form is appropriate.

3.4.2. Stability of LTV/LPV Systems

In this section axioms on *Lyapunov stability* of LTV (and LTI) systems and on *quadratic stability* of LPV systems will be presented, complying with the definitions

in [29]. On the *existence and uniqueness* of solutions, one can also consult [29] or any good book on linear systems theory, such as [28].

Consider the following *autonomous* LTV system with a unique solution $\vartheta(t, t_0, x_0)$ at time t, starting from x_0 at time t_0

$$\dot{x}(t) = \mathcal{A}(t)x(t)\,. \tag{3.40}$$

Definition 3.12 [29]
The equilibrium point $x = 0$ of the LTV system is said to be;

i) stable if and only if for all $t_0 \geq 0$ and for all $t \geq t_0$;
$\forall \epsilon > 0 \;\; \exists \delta(\epsilon, t_0) > 0 : \; \|x_0\| < \delta(\epsilon, t_0) \;\Rightarrow \|\Phi(t, t_0)x_0\| < \epsilon\,,$

ii) uniformly stable if and only if in (i) δ does not depend on t_0 ,

iii) uniformly attractive if and only if for all $t_0 \geq 0$;
$\exists \eta > 0 : \|x_0\| < \eta \Rightarrow \lim_{t\to\infty} \|\Phi(t, t_0)x_0\| = 0$ *uniformly with respect to t_0 and x_0 ,*

iv) uniformly asymptotically stable if and only if it is uniformly stable and uniformly attractive,

v) unstable if and only if it is not stable.

The above definitions apply to LTV systems as well as to nonlinear systems in general. On the other hand, when linear systems are dealt with, the equilibrium point $x = 0$ is attractive if and only if it is *globally attractive.* That is, for linear systems, definition *(iii)* holds for all $x_0 \in \mathbb{R}^n$. From this fact it follows that property of uniform asymptotic stability, when possessed by a linear system, is always global.

Note that, in most of the engineering applications we need uniform asymptotic stability rather than the general notion stability as defined in item (i) of Definition 3.7. The next theorem is a necessary and sufficient condition for uniform asymptotic stability in terms of the *state transition matrix* $\Phi(t, t_0)$ which is defined as: $x(t) = \Phi(t, t_0)x_0$, for the solution of (3.40) with the initial condition $x_0 = x(t_0)$.

Theorem 3.13 [29] *System* (3.40) *is uniformly asymptotically stable if and only if both of the following conditions hold:*

(i) There exists a scalar $\alpha > 0$ *such that for all* $t_0 \geq 0$ *and for all* $t \geq t_0$ *;*
$\|\Phi(t, t_0)\| \leq \alpha$,

(ii) $\lim_{t\to\infty} \|\Phi(t, t_0)\| = 0$, *uniformly with respect to* t_0.

From the definitions we have that exponential stability implies uniform asymptotic stability; however, when the the system is linear, it can be shown that the properties of uniform asymptotic stability and exponential stability are equivalent.

Definition 3.14 [29] *System* (3.40) *is said to be exponentially stable if and only if there exists positive scalars* α_1 *and* α_2 *such that for all* $t_0 \geq 0$ *and for all* $t \geq t_0$*;*
$\exists \mu > 0 : \|x_0\| < \mu \Rightarrow \|\Phi(t, t_0)x_0\| \leq \alpha_1 e^{-\alpha_2(t-t_0)}\|x_0\|$.

Theorem 3.15 [29] *System* (3.40) *is exponentially stable if and only if there exists positive scalars* k *and* α *such that for all* $t_0 \geq 0$ *and for all* $t \geq t_0$; $\|\Phi(t, t_0)\| \leq \alpha_1 e^{-\alpha_2(t-t_0)}$.

Theorem 3.16 [29] *System* (3.40) *is exponentially stable if and only if it is uniformly asymptotically stable.*

To use the stability definitions given above, one should compute the solutions of system (3.40), which is often difficult or even impossible task, unless the the system we are considering is time-invariant. In this case, we can evaluate analytically the solution of the system and its stability is defined by the following well known theorem.

Theorem 3.17 [28] *Let* $\mathcal{A}$ *in equation* (3.40) *be constant, that is* $\mathcal{A}(t) = \mathcal{A} \in \mathbb{R}^{n\times n}$. *Then system* (3.40) *is exponentially stable if and only if all of the eigenvalues of* A *are located in the open left half of the complex plane.*

Recall that a matrix $\mathcal{A}$ having all eigenvalues in the open left half of the complex plane is said to be *Hurwitz*. In the same way, the statement "the linear time-invariant

system $\dot{x}(t) = \mathcal{A}x(t)$ is Hurwitz stable" means that $\mathcal{A}$ is a Hurwitz matrix, i.e., the system is exponentially stable. Therefore, the stability analysis for LTI systems is reduced to the computation of system eigenvalues. Unfortunately, this kind of analysis can not be extended to LTV systems. In general, if the system matrix is time dependent, it is not possible to analytically express the *state transition matrix* $\Phi(t, t_0)$ of the system. To perform the stability analysis we have to use the *Lyapunov Stability Theorem*, which is stated next.

Theorem 3.18 [29] *Consider system* (3.40) *and assume there exists a continuous function;* $V(t, x) : \mathbb{R}^+ \times \mathbb{R}^n \to \mathbb{R}$, *and three positive constants* α_1, α_2 *and* α_3 *such that;*

(i) $V(t, 0) = 0, \quad \forall t \in \mathbb{R}^+$,

(ii) for all $x \in \mathbb{R}^n$ *and for all* $t \in \mathbb{R}^+$

$\alpha_1 \|x\|^2 \leq V(t, x) \leq \alpha_2 \|x\|^2$,

(iii) the time derivative of V *along the system trajectories;*

$\dot{V}(t, x) = \frac{\partial V(t,x)}{\partial t} + \left(\frac{\partial x}{\partial t}\right) \frac{\partial V(t,x)}{\partial x}$

satisfies for all $x \in \mathbb{R}^n$ *and for all* $t \in \mathbb{R}^+$ *the condition;*

$\dot{V}(t, x) \leq -\alpha_3 \|x\|^2$.

Then the system (3.40) *is exponentially stable.*

Often, as a candidate *Lyapunov function*, we use the quadratic form

$$V(t, x) = x^T \mathcal{X}(t) x \,, \tag{3.41}$$

with a continuous and differentiable matrix valued function $\mathcal{X}(t) : \mathbb{R}^+ \to \mathbb{R}^{n \times n}$ which is positive definite and bounded, and satisfies; $\alpha_1 I \leq \mathcal{X}(t) \leq \alpha_2 I$, $\forall t \in \mathbb{R}^+$ for some positive numbers α_1 and α_2 . If V satisfies the condition *(ii)* it is said to be *positive definite* and *decrescent*. If $\dot{V}$ satisfies *(iii)* it is said to be *negative definite*. When V is continuously differentiable, the derivative along the solutions becomes

$$\dot{V}(t, x) = \frac{d}{dt} V(t, x) = \frac{\partial}{\partial t} V(t, x) + \left(\frac{\partial x}{\partial t} \left(\frac{\partial}{\partial x} \right) \right) V(t, x) = x^T (\dot{\mathcal{X}} + \mathcal{A}^T \mathcal{X} + \mathcal{X}\mathcal{A}) x \,.$$

With the choice (3.41), we can guarantee negative definiteness of $\dot{V}$ at a given point t by imposing

$$\dot{\mathcal{X}}(t) + \mathcal{A}^T(t)\mathcal{X}(t) + \mathcal{X}(t)\mathcal{A}(t) \preceq -\alpha I\,, \quad \text{for some } \alpha > 0\,.$$

Conversely, at the points where $\mathcal{X}(t)$ is not differentiable, to guarantee negative definiteness of $\dot{V}$ we must have

$$\frac{\mathcal{X}(t+\triangle t) - \mathcal{X}(t)}{\triangle t} + \mathcal{A}^T(t)\mathcal{X}(t+\triangle t) + \mathcal{X}(t+\triangle t)\mathcal{A}(t) \preceq -\alpha I$$

for all $\triangle t$ in the neighborhood of $\triangle t = 0$.

In general, for the sake of simplicity, we use the following inequality to cover both cases on differentiability of $\mathcal{X}(t)$ as follows

$$\dot{\mathcal{X}} + \mathcal{A}^T\mathcal{X} + \mathcal{X}\mathcal{A} \prec 0\,. \tag{3.42}$$

Notice that we use the notation "$\prec 0$" for negative definite matrix inequalities, whereas the notation "< 0" is used for strictly negative scalar inequalities.

Corollary 3.19 [29] *System* (3.40) *is exponentially stable if there exists a positive definite, bounded and continuous matrix valued function* $\mathcal{X}(t)$ *satisfying* (3.42).

Based on Theorem 3.16 we can derive a number of *sufficient* conditions guaranteeing the stability of a LTV system.

Theorem 3.20 [28] *Consider the system* (3.40) *and assume that* $\lim_{t\to\infty} \mathcal{A}(t) = \overline{\mathcal{A}} \in \mathbb{R}^{n\times n}$. *Then system* (3.40) *is exponentially stable if and only if all the eigenvalues of* $\overline{\mathcal{A}}$ *are located in the open left half of the complex plane.*

Theorem 3.21 [29] *Consider the system* (3.40) *with a continuous matrix-valued function* $\mathcal{A}(t)$. *Moreover, assume that;*

(i) the matrix function $\mathcal{A}(t)$ is bounded; there exists a positive number m such that $\|\mathcal{A}(t)\| \leq m \;\; \forall t \in \mathbb{R}^+$,

(ii) the eigenvalues of $\mathcal{A}(t)$ are located in the open left half of the complex plane,

(iii) there exists a scalar $\epsilon > 0$ such that
$\|\dot{\mathcal{A}}(t)\| \leq \epsilon \;\; \forall t \in \mathbb{R}^+$, *or for the points where $\mathcal{A}(t)$ is not differentiable;*
$\limsup_{\triangle t \to 0} \|\frac{\mathcal{A}(t+\triangle t) - \mathcal{A}(t)}{\triangle t}\| \leq \epsilon \;\; \forall t \in \mathbb{R}^+$.

Then if ϵ is sufficiently small, system (3.40) *is exponentially stable.*

Theorem 3.22 [29] *Consider the system* (3.40) *with a continuous matrix-valued function $\mathcal{A}(t)$. If there exists a scalar $\alpha > 0$ such that $\lambda_{max}(t)\{(\mathcal{A}(t)+\mathcal{A}(t)^T)\} \leq -\alpha \;\; \forall t \in \mathbb{R}^+$* ($\lambda_{max}$ denotes the maximum eigenvalue), *then the system is exponentially stable.* See Appendix D for the proof.

An autonomous linear system subject to parametric uncertainties can be described in the following LPV form

$$\dot{x}(t) = \mathcal{A}(\rho)x(t)\,, \;\; t \in [0, +\infty) \tag{3.43}$$

where $x(t) \in \mathbb{R}^n$, $\rho(t) \in \mathcal{P} \subset \mathbb{R}^m$ is the vector of uncertain parameters, and $\mathcal{A}(\rho)$ is continuous. We assume that the set $\mathcal{P}$ is a *hyper-box*. That is a set in the form

$$\mathcal{P} := [\underline{\rho}_1, \overline{\rho}_1] \times [\underline{\rho}_2, \overline{\rho}_2] \times \cdots \times [\underline{\rho}_m, \overline{\rho}_m]\,. \tag{3.44}$$

Note that, a hyper-box is a particular *polytope* (see Appendix B.3) with 2^m vertices. We denote the set of the vertices of $\mathcal{P}$ by $\mathcal{P}^v$.

Obviously, equation (3.43) represents a collection of infinite number of systems. For any given $\rho \in \mathcal{P}$ (3.43) yields a system of differential equations with constant coefficients. Conversely, if ρ is a vector-valued function of time belonging to certain functional space $\mathcal{S}$ for any $\rho(t) \in \mathcal{S}$, system (3.43) defines a system of differential equations with time-varying coefficients.

A straightforward application of the previous results on Lyapunov stability conditions shows that, if the parameters are time-varying Hurwitzness of the system matrix $\mathcal{A}(\rho)$ for any $\rho \in \mathcal{P}$ no longer guarantees exponential stability of system (3.43).
In this case, we need a different approach based on *quadratic Lyapunov functions.* Hence, we start with the definition of *quadratic stability.*

Definition 3.23 [29] *System* (3.43) *is said to be quadratically stable (QS) in* $\mathcal{P}$*, if and only if, there exists a positive definite matrix* $\mathcal{X} \in \mathcal{R}^{n\times n}$ *such that for all* $\rho \in \mathcal{P}$;

$$\mathcal{A}(\rho)^T\mathcal{X} + \mathcal{X}\mathcal{A}(\rho) \prec 0 \quad \textit{holds.} \tag{3.45}$$

As shown in the next theorem, quadratic stability implies exponential stability of the system (3.43) for all admissible time realizations of the parameters. However, it must be emphasized that *quadratic stability is not a necessary condition for exponential stability.* One may find a Lyapunov function other than the standard quadratic form, $x^T\mathcal{X}x$, and prove exponential stability.

Theorem 3.24 [29] *Assume that the system* (3.43) *QS. Then for any continuous function* $\rho(t)$ $\mathbb{R}^+ \to \mathcal{P}$ *the linear time-varying system* $\dot{x}(t) = \mathcal{A}\big(\rho(t)\big)x(t)\,,\ \ t \in [0, +\infty)$ *is exponentially stable.*
See Appendix D for the proof.

Since $\mathcal{P}$ is compact, (3.45) implies the existence of a positive definite matrix $\mathcal{X}$, with

$$\tilde{\mathcal{A}}^T\mathcal{X} + \mathcal{X}\tilde{\mathcal{A}} \prec 0\,,$$

where $\tilde{\mathcal{A}}(t) := A\big(\rho(t)\big)$. The assumption on piecewise continuity of parameters covers almost all cases of interest in the engineering applications. From a mathematical point of view, it is possible to extend the admissible parameters realizations to the set of *Lebesque measurable* vector-valued functions (see Appendix B.1).

It is readily seen that behind Definition 3.23 and the proof of Theorem 3.24 is the use of a "constant" Lyapunov function of the form $V(x) = x^T \mathcal{X} x$. Time invariance is the key point which allows to deal with arbitrarily varying parameters. Therefore in the presence of further information about the parameters behavior, it is more convenient to use quadratic Lyapunov functions which depends on parameters. Such Lyapunov functions, along a given parameter vector time realization, are "time-varying" Lyapunov functions in the form (3.41).

From Definition 3.23, it follows that system (3.43) is QS if and only if the following problem admits a feasible solution:

Find a symmetric matrix $\mathcal{X} \in \mathbb{R}^{n \times n}$, such that

$$\begin{aligned} \mathcal{X} &\succ 0 \\ \mathcal{A}(\rho)^T \mathcal{X} + \mathcal{X}\mathcal{A}(\rho) &\prec 0 \qquad \forall \rho \in \mathcal{P}\,. \end{aligned} \tag{3.46}$$

The following results shows that the quadratic stability of system (3.43) is equivalent to quadratic stability of the *dual system* which shall be used in subsection (3.4.3).

Lemma 3.25 [29] *System* (3.43) *is QS, if and only if, there exists a positive definite matrix* $\mathcal{Y}$*, such that for all* $\rho \in \mathcal{P}$ *the following holds:*

$$\mathcal{A}(\rho)\mathcal{Y} + \mathcal{Y}\mathcal{A}(\rho)^T \prec 0\,. \tag{3.47}$$

See Appendix D for the proof.

Note that, if $\mathcal{Y}$ is a positive definite matrix satisfying the hypothesis of Lemma 3.25, system (3.43) is QS, and a suitable Lyapunov function is $x^T \mathcal{Y}^{-1} x$.

Problem (3.46) is a feasibility problem in the matrix variable $\mathcal{X}$ subject to an *infinite* number of Linear Matrix Inequalities. In order to solve the problem, it must be

converted into *finite* number of LMIs. The kind of parameter dependence specified in the next assumption allows to reduce (3.45) into a finite number of inequalities.

Assumption 3.26 [29] *The system matrix* $\mathcal{A}(\rho) : \mathcal{P} \to \mathbb{R}^{n\times n}$ *is the ratio of a multi-affine matrix-valued function of* ρ *and a multi-affine polynomial of* ρ;

$$\mathcal{A}(\rho) = \frac{N_{\mathcal{A}}(\rho)}{d_{\mathcal{A}}(\rho)} = \frac{\sum_{i_1,\ldots,i_m=0}^{1} \mathcal{A}_{i_1\ldots i_m}\, \rho_1^{i_1} \cdots \rho_m^{i_m}}{\sum_{i_1,\ldots,i_m=0}^{1} a_{i_1\ldots i_m}\, \rho_1^{i_1} \cdots \rho_m^{i_m}}$$

where $d_{\mathcal{A}} \neq 0$ for all $\rho \in \mathcal{P}$ and $N_{\mathcal{A}}(\rho) \in \mathbb{R}^{n\times n}$.

The parameter dependence considered in Assumption 3.26 is quite general and recovers, as particular cases, the usual affine and multi-affine dependence.

Notice that, the requirement on the entries of $A(\rho)$ to be ratio of multi-affine polynomials does not guarantee the satisfaction of Assumption 3.26. For example, the elements of the matrix function

$$\mathcal{A}(\rho) = \begin{bmatrix} 0 & 1 \\ \frac{\rho_1}{\rho_2} & \frac{\rho_2}{\rho_3} \end{bmatrix},$$

where $\rho_2, \rho_3 \neq 0$, are the ratio of multi-affine polynomials, but $\mathcal{A}(\rho)$ does not satisfy Assumption 3.26 since

$$\mathcal{A}(\rho) = \frac{\begin{bmatrix} 0 & \rho_2\rho_3 \\ \rho_1\rho_3 & \rho_2^2 \end{bmatrix}}{\rho_2\rho_3}.$$

Theorem 3.27 [29] *System* (3.43) *where* $\mathcal{A}(\rho)$ *satisfies Assumption 3.26 is QS, if and only if, there exists a positive definite matrix* $\mathcal{X} \in \mathbb{R}^{n\times n}$ *such that:*
$\mathcal{A}(\rho_i)^T\mathcal{X} + \mathcal{X}\mathcal{A}(\rho_i) \prec 0\,, \quad i = 1, ..., 2^m,$ *where* ρ_i *is the i-the vertex of* $\mathcal{P}$.
Proof: See Theorem B.28 in Appendix B.3.

Consider the LTV system (3.40) with $\mathcal{A}(t) \in \text{co}\{\mathcal{A}_{(1)}, \mathcal{A}_{(2)}, ..., \mathcal{A}_{(l)}\} =: \mathbb{A}\,,\ t \in [0, +\infty)$, where $\mathcal{A}(t)$ is piecewise continuous. A system in this form is called *polytopic* (see Appendix B.3), since it represents a family of time-varying systems whose system matrix attains values into the polytope with vertices $\mathcal{A}_{(i)}$, $i = 1, 2, ..., l$. The polytopic system is said to be QS if and only if there exists a positive definite matrix $\mathcal{X}$ such that

$$\mathcal{A}^T\mathcal{X} + \mathcal{X}\mathcal{A} \prec 0\,,\ \ \forall \mathcal{A} \in \mathbb{A}$$

which is equivalent to require the existence of a positive definite matrix $\mathcal{X}$ such that

$$\mathcal{A}_{(i)}^T\mathcal{X} + \mathcal{X}\mathcal{A}_{(i)} \prec 0\,,\ \ i = 1, 2, ..., l\,. \tag{3.48}$$

It is simple to recognize that quadratic stability of the polytopic system guarantees the exponential stability of any linear time-varying system $\dot{x} = \mathcal{A}(t)x(t)$. In a sense, polytopic systems can be seen as a sub-class of systems satisfying Assumption 3.24. Assume, for example, that the polytopic system we deal with is obtained by the *convex hull* (see Appendix B.3) of two matrices

$$\dot{x} = \mathcal{A}(t)x(t)\,,\ \ \mathcal{A}(t) \in \text{co}\{\mathcal{A}_{(1)}, \mathcal{A}_{(2)}\}\,,\ t \in [0, +\infty)\,.$$

Then consider the uncertain system

$$\dot{x} = \mathcal{A}(\rho)x(t) = \big(\mathcal{A}_0 + \mathcal{A}_1(\rho)\big)x(t)\,,\ \ \text{where } \rho(t) \in [\rho_{min}, \rho_{max}]. \tag{3.49}$$

Obviously we have that $\mathcal{A}(\rho_{min}) = \mathcal{A}_{(1)}$ and $\mathcal{A}(\rho_{max}) = \mathcal{A}_{(2)}$.

Since the image of an affine matrix function defined over a hyper-box (an interval in this case) is equal to the convex hull of the image of vertices (see Theorem B.27 in Appendix B.3) we have that; $\mathcal{A}([\rho_{min}, \rho_{max}]) = \text{co}\{\mathcal{A}_{(1)}, \mathcal{A}_{(2)}\}$. Therefore the uncertain system (3.49) is equivalent (from the point of view of stability analysis) to the polytopic LTV system. Notice that, system (3.49) reflects the dynamics of the closed-loop

rotor/AMB system without the disturbance w.

In general, there always exists an uncertain system in the form of equation (3.43) satisfying $\mathcal{A}(\mathcal{P}) \supseteq \mathbb{A}$. Obviously, when $\mathbb{A}$ is a strict subset of $\mathcal{A}(\mathcal{P})$, quadratic stability of system (3.43) is a sufficient but not necessary condition of the quadratic stability of the polytopic LTV system.

As stated in Theorem 3.24, quadratic stability of system (3.43) guarantees exponential stability for all time behaviors of parameters which are of interest. In particular, exponential stability is guaranteed for discontinuous parameters which exhibit an unbounded rate of variation at the discontinuity points. Conversely, in many engineering applications, the uncertain parameters are a continuous and slowly varying functions of time (e.g., in rotor/AMB applications, acceleration of the rotor is limited by the maximum available torque that can be supplied by the driving motor/turbine). In those cases, the quadratic stability approach may result an extremely conservative tool to test the system stability.

Assume that $\rho(t)$ mapping $\mathbb{R}^+ \to \mathcal{P}$ is a continuous function, and $\mathcal{P}$ is the hyper-box defined in (3.44), and that a bound on the rate of variation is known as

$$\underline{\nu}_i \leq \dot{\rho}_i(t) \leq \overline{\nu}_i \,, \;\; i = 1, ..., m \,, \;\; t \in \mathbb{R}^+. \tag{3.50}$$

Note that condition (3.50) implies that time derivative parameter vector $\dot{\rho}(t)$ belongs for all t to the hyper-box $\dot{\mathcal{P}}$ defined as

$$\dot{\mathcal{P}} := [\underline{\nu}_1, \overline{\nu}_1] \times [\underline{\nu}_2, \overline{\nu}_2] \times \cdots \times [\underline{\nu}_m, \overline{\nu}_m] \,, \tag{3.51}$$

which is centered at the origin of the parameter derivative space. We denote by $\dot{\mathcal{P}}^v$ the set of the 2^m vertices of $\dot{\mathcal{P}}$ and by $\nu_{(j)}$ the j-th vertex of $\dot{\mathcal{P}}$. To take into account the information on the rate of variation of parameters we have to use a quadratic Lyapunov function depending on parameters.

Definition 3.28 [29] *System* (3.43) *is said to be quadratically stable via parameter dependent Lyapunov functions in* $\mathcal{P} \times \dot{\mathcal{P}}$ *if and only if there exists a continuously differentiable positive definite matrix valued function* $\mathcal{X}(\rho) : \rho(t) \in \mathcal{P} \rightarrow \mathcal{X}(\rho)$ *such that*

$$\mathcal{A}^T(\rho)\mathcal{X}(\rho) + \mathcal{X}(\rho)\mathcal{A}(\rho) + \sum_{i=1}^{m} \frac{\partial \mathcal{X}(\rho)}{\partial \rho_i} \nu_{i(j)} \prec 0 \tag{3.52}$$

for all $\rho \in \mathcal{P}$ *and for all* $j = 1, ..., 2^m$ *(*$\nu_{i(j)}$ *denotes the i-th component of vector* $\nu_{(j)}$*).*

Behind Definition 3.28 is the use of a parameter dependent quadratic Lyapunov function in the form $V(\rho, x) = x^T \mathcal{X}(\rho) x$. Such Lyapunov function, along a given vector time realization, is a time-varying Lyapunov function in the form $x^T \tilde{\mathcal{X}}(t) x = x^T \mathcal{X}(\rho(t)) x$. Now we can state the following fundamental result.

Theorem 3.29 [29] *Assume that system* (3.43) *is quadratically stable via parameter dependent Lyapunov functions in* $\mathcal{P} \times \dot{\mathcal{P}}$*. Then the system* (3.43) *is exponentially stable for all vector valued continuous functions* $\rho(t)$ $\mathbb{R}^+ \rightarrow \mathcal{P}$ *satisfying* (3.50).
See Appendix D for the proof.

3.4.3. Multi-objective Controller Synthesis

Given an LPV system, the synthesis goal is to construct a controller which not only uses a measured output but, in addition, the on-line measured actual parameters as information in order to exponentially stabilize the system and to provide good performance properties. The performance objective can be specified as an L_2 (see Appendix B.1) disturbance attenuation problem with the standard interpretation such as guaranteeing robust stability or good reference tracking. This is so-called H_∞ problem for LPV systems.

In principle, an LPV system can be viewed to be time-varying, and any design technique which is available for a specific choice of the performance measure can be

used for the controller construction. However, the actual parameter curve is not known in advance and many existing synthesis techniques, such as H_2 and H_∞ control of LTV systems, instantaneously require the knowledge of the parameter values over the whole time interval of interest. For LPV systems and H_∞ problem it is not difficult, though conservative, to propose a way out of this dilemma. Assume that the parameters are contained in a set, given a priori as in (3.44). Then replace the time-varying solutions of *differential Riccati equations* along the actual parameter curve by constant solutions of *algebraic Riccati inequalities* over the whole set of possible parameters. Under additional hypotheses on the structure of the parameter set (polytopic), and on the dependence of the system on the parameters (complying with Assumption 3.26), the verification of the existence of a suboptimal controller and its construction can be reduced to solving a set of linear matrix inequalities.

Consider the LTV system

$$z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = Mw = \begin{bmatrix} M_1 \\ M_2 \end{bmatrix} w = \left[\begin{array}{c|c} \mathcal{A} & \mathcal{B} \\ \hline \mathcal{C}_1 & \mathcal{D}_1 \\ \mathcal{C}_2 & \mathcal{D}_2 \end{array}\right] w\,. \tag{3.53}$$

We can interpret $w \to z_1$ as the robustness (and performance) channel. We can assume that the uncertainty of the system is described by $w = \Delta z_1$, where Δ comprises the set of (possibly nonlinear) operators $L_2 \to L_2$ with incremental gain not larger than $1/\gamma$. If $\mathcal{A}$ is exponentially stable, small gain theorem implies that stability is preserved if $\|M_1\|_\infty < \gamma$, where

$$\|M_1\|_\infty := \sup_{\|w\|_2=1} \|M_1 w\|_2 \tag{3.54}$$

defines the operator norm of M_1 induced by the signals in L_2. In the LTI case, there is a well-known test on the state-space matrices which characterizes stability of $\mathcal{A}$ and $\|M_1\|_\infty < \gamma$, the so-called *Bounded Real Lemma* which is given in Theorem 3.10. Following theorem paves the way for its generalization to LTV systems. Recall from Lyapunov stability analysis that the time functions are bounded, and that a symmetric

valued time function $\mathcal{X} = \mathcal{X}^T$ is strictly positive ($\mathcal{X} \succ 0$), if there exists an $\epsilon > 0$ with $\mathcal{X}(t) \succeq \epsilon I$ for all $t \geq 0$.

Theorem 3.30 [29] *Consider system (3.53). The following statements are equivalent*:
i) The closed loop system M_1 is exponentially stable and $\|M_1\|_\infty < \gamma$.
ii) $\|\mathcal{D}_1(t)\|_\infty < \gamma$ for all $t \in \mathbb{R}^+$ and there exists positive definite continuous matrix-valued function $\mathcal{X}(t) : \mathbb{R}^+ \to \mathbb{R}^{n\times n}$ such that

$$\dot{\mathcal{X}} + \mathcal{A}^T\mathcal{X} + \mathcal{X}\mathcal{A} + \gamma^{-2}\mathcal{C}_1^T\mathcal{C}_1 + (\mathcal{X}\mathcal{B} + \gamma^{-2}\mathcal{C}_1^T\mathcal{D}_1)(I - \gamma^{-2}\mathcal{D}_1^T\mathcal{D}_1)^{-1}(\mathcal{B}^T\mathcal{X} + \gamma^{-2}\mathcal{D}_1^T\mathcal{C}_1) \prec 0\,.$$

Using *Schur complement formula* (see Appendix B.2) it is not difficult to show the following result:

Lemma 3.31 [14] (*LTV Bounded Real Lemma*)
The closed loop system M_1 is exponentially stable with $\left\| \left[\begin{array}{c|c} \mathcal{A} & \mathcal{B} \\ \hline \mathcal{C}_1 & \mathcal{D}_1 \end{array}\right] \right\|_\infty < \gamma$ *if and only if there exists a $\mathcal{X}(t) = \mathcal{X}(t)^T \succ 0$ such that*;

$$\begin{bmatrix} \dot{\mathcal{X}} + \mathcal{A}^T\mathcal{X} + \mathcal{X}\mathcal{A} & \mathcal{X}\mathcal{B} & \mathcal{C}_1^T \\ \mathcal{B}^T\mathcal{X} & -I & \mathcal{D}_1^T \\ \mathcal{C}_1 & \mathcal{D}_1 & -\gamma^2 I \end{bmatrix} \prec 0 \quad \textit{is satisfied.} \tag{3.55}$$

Hence an H_∞ norm bound can be characterized by the existence of strictly positive solutions to a strict *differential linear matrix inequality* (DLMI). If M_1 is LTI, it causes no loss of generality to confine $\mathcal{X}$ to be constant, which is the case in Theorem 3.10.

The channel $w \to z_2$ is used solely for performance specifications. This channel can be used to generalize certain H_2 norm like criteria for LTI systems to the LTV system M_2 [21]. Among the several possibilities, we pay special attention to the deterministic criterion of assessing performance by the largest amplitude of z_2 for all w of finite and bounded energy. This is particularly useful if z_2 (or its components) are interpreted

as tracking errors. To quantify the gain of M_2 mapping L_2 to L_∞, we use the induced norm (generalized 2-norm) as given in (3.36):

$$\|M_2\|_{2g} = \sup_{\|w\|_2=1} \|z\|_\infty = \sup_{\|w\|_2=1} \|M_2 w\|_\infty \tag{3.56}$$

Suppose $\mathcal{A}$ is exponentially stable, and let $\mathcal{Y}$ denote the bounded solution of the initial value problem

$$\dot{\mathcal{Y}} = \mathcal{A}\mathcal{Y} + \mathcal{Y}\mathcal{A}^T + \mathcal{B}\mathcal{B}^T, \quad \mathcal{Y}(0) = 0\,. \tag{3.57}$$

If $\mathcal{D}_2 = 0$, one has $\|M_2\|_{2g}^2 = \sup_{t\geq 0} \|\mathcal{C}_2\mathcal{Y}\mathcal{C}_2^T\|_{2g}$ [14]. This allows to prove the following analogue to Theorem 3.30.

Theorem 3.32 [21] *The system M_2 is exponentially stable with* $\left\| \left[\begin{array}{c|c} \mathcal{A} & \mathcal{B} \\ \hline \mathcal{C}_2 & \mathcal{D}_2 \end{array}\right] \right\|_{2g}^2 < \beta$ *if and only if $\mathcal{D}_2 = 0$ and there exists a smooth $\mathcal{Z}$ with*

$$\mathcal{Z} \succ 0\,, \quad \dot{\mathcal{Z}} + \mathcal{A}^T\mathcal{Z} + \mathcal{Z}\mathcal{A} + \mathcal{Z}\mathcal{B}\mathcal{B}^T\mathcal{Z} \prec 0\,, \quad \mathcal{C}_2\mathcal{Z}^{-1}\mathcal{C}_2^T \prec \beta I\,. \tag{3.58}$$

In defining the size of the amplitude of z_2 with the spatial norm $\max_j |x_j|$, the squared gain of M_2 equals; $\sup_{t\geq 0} \max_j d_j[\mathcal{C}_2\mathcal{Y}\mathcal{C}_2^T]$, where d_j denotes the j-th diagonal element of the matrix inside the brackets, and $\mathcal{Y}$ solves (3.57). Theorem 3.32 remains valid for this norm after replacing $\mathcal{C}_2\mathcal{Z}^{-1}\mathcal{C}_2^T \prec \beta I$ by; $\max_j d_j(\mathcal{C}_2\mathcal{Z}^{-1}\mathcal{C}_2^T) < \beta$.

An alternative measure arises with a stochastic interpretation. If w is white noise, we recall that $E(z_2\, z_2^T) = \mathcal{C}_2\mathcal{Y}\mathcal{C}_2^T$, due to $x_0 = 0$ [14]. Then;

$$\|M_2\|_2^2 := \sup_{t\geq 0} E(z_2^T z_2) = \sup_{t\geq 0} \operatorname{trace}[\mathcal{C}_2\mathcal{Y}\mathcal{C}_2^T] \tag{3.59}$$

defines the maximal output variance and is a generalization of the genuine H_2 norm to LTV systems. Theorem 3.32 persists to hold for the H_2 norm with $\mathcal{C}_2\mathcal{Z}^{-1}\mathcal{C}_2^T \prec \beta I$ replaced by; $\operatorname{trace}(\mathcal{C}_2\mathcal{Z}^{-1}\mathcal{C}_2^T) < \beta$.

Even for the LTI systems, the synthesis problem of optimizing $\|\cdot\|_{2g}$ or $\|\cdot\|_2$ of over all stabilizing controllers which keep a bound on the norm $\|\cdot\|_\infty$ for a different channel seems very hard [21]. The difficulty arises from the fact that $\mathcal{X}$ in (3.55) and $\mathcal{Z}$ in (3.58) have to be taken different. However, at the expense of introducing conservatism, we consider the problem of minimizing an upper bound of $\|M_2\|_{2g}$ under the constraint $\|M_1\|_\infty < \gamma$. If some solution of $\mathcal{X}$ of (3.55) satisfies $\mathcal{C}_2\mathcal{X}^{-1}\mathcal{C}_2^T \prec \beta I$, then this Lyapunov matrix function can also satisfy (3.58) and we can conclude that $\|M_2\|_{2g}^2 < \beta$. This leads to defining the *mixed H_2/H_∞ objective functional* as follows:

$$\mathfrak{I}(M) := \inf\left\{\beta \,|\, \exists \text{ a smooth time function } \mathcal{X} \text{ satisfying (3.55), and } \mathcal{C}_2\mathcal{X}^{-1}\mathcal{C}_2^T \prec \beta I\right\}.$$

We use the convention $\mathfrak{I}(M) = \infty$ if no solution to (3.55) exists. Notice that $\mathfrak{I}(M) < \beta$ implies $\|M_1\|_\infty < \gamma$, and $\|M_2\|_{2g}^2 < \beta$. Setting $(\mathcal{C}_1\ \mathcal{D}_1) = 0$ implies that the solution sets of DLMIs in (3.55) and (3.58) are identical and we infer $\|M_2\|_{2g}^2 = \mathfrak{I}(M)$ that recovers the generalized H_2 norm. Similar conclusions hold for the other norms of M_2.

Suppose a specific control task leads to the generalized LTV plant

$$\begin{pmatrix} \dot{x} \\ z_1 \\ z_2 \\ y \end{pmatrix} = \begin{bmatrix} A & B_1 & B_2 \\ C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \\ C & D & 0 \end{bmatrix} \begin{pmatrix} x \\ w \\ u \end{pmatrix}, \tag{3.60}$$

with the LTV controller $u = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right] y$, closed-loop system can be described as

$$z(t) = M\big(K(t)\big)w(t) = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} M_1(K) \\ M_2(K) \end{pmatrix} w = \left[\begin{array}{c|c} \mathcal{A} & \mathcal{B} \\ \hline \mathcal{C}_1 & \mathcal{D}_1 \\ \mathcal{C}_2 & \mathcal{D}_2 \end{array}\right] w\,,$$

where $$\left[\begin{array}{c|c} \mathcal{A} & \mathcal{B} \\ \hline \mathcal{C}_1 & \mathcal{D}_1 \\ \mathcal{C}_2 & \mathcal{D}_2 \end{array}\right] = \left[\begin{array}{cc|c} A + B_2 D_K C & B_2 C_K & B_1 + B_2 D_K D \\ B_K C & A_K & B_K D \\ \hline C_1 + D_{12} D_K C & D_{12} C_K & D_{11} + D_{12} D_K D \\ C_2 + D_{22} D_K C & D_{22} C_K & D_{21} + D_{22} D_K D \end{array}\right].$$

The controller $K(t)$ is stabilizing if $\mathcal{A}$ is exponentially stable. We intend to minimize $\mathfrak{I}\big(M(K)\big)$ over all LTV controllers $K(t)$. Recall that $\mathfrak{I}\big(M(K)\big) < \infty$ automatically implies that $K(t)$ is stabilizing. It is standard to approach this problem via a suboptimality test: Characterize whether there exists a $K(t)$ with $\mathfrak{I}\big(M(K)\big) < \beta$ or, equivalently, whether there exists a $K(t)$ and a smooth function $\mathcal{X}(t)$ such that

$$\mathcal{X} \succ 0\,, \quad \begin{bmatrix} \dot{\mathcal{X}} + \mathcal{A}^T\mathcal{X} + \mathcal{X}\mathcal{A} & \mathcal{X}\mathcal{B} & \mathcal{C}_1^T \\ \mathcal{B}^T\mathcal{X} & -I & \mathcal{D}_1^T \\ \mathcal{C}_1 & \mathcal{D}_1 & -\gamma^2 I \end{bmatrix} \prec 0\,, \quad \mathcal{C}_2\mathcal{X}^{-1}\mathcal{C}_2 \prec \beta\,. \tag{3.61}$$

This not only guarantees robust stability against perturbations with incremental gain of at most $1/\gamma$, but also a generalized H_2 performance level of β.

BRL inequality in Theorem 3.10 used in the H_∞ synthesis problems for LTI systems is, for a fixed $\mathcal{X}$, linear in controller parameters. Hence one can eliminate these parameters and solve for the controller matrices. However in the mixed problem (either LTV or LTI) we have to fulfill three inequalities what makes it impossible to eliminate all controller parameters from the final characterization. Instead, we intend to keep as many (transformed) controller parameters as possible such that, still, matrix inequalities result which are linear in parts of $\mathcal{X}$ and $\mathcal{X}^{-1}$ and in the remaining transformed controller parameters. A central step in the proof is the following simple explicit result for the solvability of a specially structured inequality.

Lemma 3.33 [21] *Let Q(t) be a symmetric (partitioned) time function and consider*

the inequality

$$\begin{bmatrix} Q_1 & Q_{21}^T & Q_{31}^T + X^T \\ Q_{21} & Q_2 & Q_{32}^T \\ Q_{31} + X & Q_{32} & Q_3 \end{bmatrix} \prec 0 \tag{3.62}$$

in the time function $X(t)$. This inequality has a solution $X(t)$ if and only if

$$\begin{bmatrix} Q_1 & Q_{21}^T \\ Q_{21} & Q_2 \end{bmatrix} \prec 0\,, \quad \textit{and} \quad \begin{bmatrix} Q_2 & Q_{32}^T \\ Q_{32} & Q_3 \end{bmatrix} \prec 0\,. \tag{3.63}$$

If (3.62) *is solvable, one particular solution is given by:*

$$X = Q_{32} Q_2^{-1} Q_{21} - Q_{31}. \tag{3.64}$$

See Appendix D for the proof.

Theorem 3.34 [21] *There exists a controller* $K := \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right]$ *with* A_K *of size* $k \leq n$ *which satisfies* $\mathfrak{I}\big(M(K)\big) < \beta$ *if and only if there exists* $\epsilon > 0$*, time functions* $X(t), Y(t)$*, and* $Z(t)$ *with* $X \succ 0\,,\ Y \succ 0\,,\ X - Y^{-1} = ZZ^T,\ Z$ *of size* $n \times k\,,\ Z^T Z \succ 0\,,$ *and time functions E(t), F(t), G(t), where* $G(t) = D_K(t)$*, such that*

$$\operatorname{rank} \begin{bmatrix} E & GC \\ Y & I \\ I & X \end{bmatrix} = k + n\,, \quad \operatorname{rank} \begin{bmatrix} B_2 G & Y & I \\ F & I & X \end{bmatrix} = k + n\,,$$

and the following DLMIs are satisfied:

$$\begin{bmatrix} \dot{X} + A^T X + XA + FC + (FC)^T & XB_1 + FD & (C_1 + D_{12}GC)^T \\ (XB_1 + FD)^T & -I & (D_{11} + D_{12}GD)^T \\ C_1 + D_{12}GC & D_{11} + D_{12}GD & -\gamma^2 I \end{bmatrix} \prec 0\,, \tag{3.65}$$

$$\begin{bmatrix} -\dot{Y} + AY + YA^T + B_2E + (B_2E)^T & B_1 + B_2GD & (C_1Y + D_{12}E)^T \\ (B_1 + B_2GD)^T & -I & (D_{11} + D_{12}GD)^T \\ C_1Y + D_{12}E & D_{11} + D_{12}GD & -\gamma^2 I \end{bmatrix} \prec 0\,, \tag{3.66}$$

$$\begin{bmatrix} \beta - \epsilon I & C_2Y + D_{22}E & C_2 + D_{22}GC \\ (C_2Y + D_{22}E)^T & Y & I \\ (C_2 + D_{22}GC)^T & I & X \end{bmatrix} \succeq 0\,. \tag{3.67}$$

See Appendix D for the proof.

Due to the H_2 nature of the performance specification, one might wish to include the requirement $D_{21} + D_{22}GD = 0$ on G. This puts another linear restriction on G without destroying the structure of the DLMIs.

If the system and the controller are LTI, it causes no loss of generality to confine $\mathcal{X}$ in (3.61) to be constant. Hence Theorem 3.34 remains valid by specializing to constant X, Y, Z, and E, F, G. Note that we do not require to include the hypotheses that "(A, B_2) stabilizable and (A, C) detectable", since they are obvious necessary conditions for the existence of positive definite solutions of (3.65) and (3.66).

Let us now specialize Theorem 3.34 to the case without any a priori restriction on the controller size.

Theorem 3.35 [21] *There exists a controller K with $\mathfrak{I}\big(M(K)\big) < \beta$ if and only if there exists smooth time functions X(t), Y(t), and time functions E(t),F(t),G(t) such that DLMIs (3.65) and (3.66) and*

$$\begin{bmatrix} \beta I & C_2Y + D_{22}E & C_2 + D_{22}GC \\ (C_2Y + D_{22}E)^T & Y & I \\ (C_2 + D_{22}GC)^T & I & X \end{bmatrix} \succeq 0\,. \tag{3.68}$$

are satisfied. If existing, the parameter A_K in K can be chosen of size n. Hence, suboptimality is characterized in terms of the solvability of two differential and one algebraic matrix inequalities where all unknowns X, Y, E, F, G enter linearly.

Given X and Y, *Cholesky factorize* $X - Y^{-1} = UU^T$ such that U and U^{-1} are smooth and bounded. Motivated by $XY + UW^T = I$ (see the proof of Theorem 3.34), define $W = (I - YX)U^{-T}$ which is smooth and has a bounded inverse.
Defining the controller parameters as

$$C_K := (E - GCY)W^{-T}, \qquad B_K := U^{-1}(F - XB_2G)\,. \tag{3.69}$$

Finally, with

$$Q_{21} = \dot{X}Y + \dot{U}W^T + (A + B_2GC)^T + X(A - B_2GC)Y + FCY + XB_2E\,,$$

$$[Q_{31}\ \ Q_{32} \mid Q_3] = \left[\begin{array}{cc|cc} (B_1 + B_2GD)^T & (XB_1 + FD)^T & -\gamma I & (D_{11} + D_{12}GD)^T \\ C_1Y + D_{12}E & C_1Y + D_{12}GC & D_{11} + D_{12}GD & -\gamma I \end{array}\right],$$

a suitable A_K is given by

$$A_K = U^{-1}(Q_{32}^T Q_3^{-1} Q_{31} - Q_{21})W^{-T}. \tag{3.70}$$

Note that we could as well start with $U := X - Y^{-1}$ implying $W = -Y$ or, dually, $W := Y - X^{-1}$ implying $U = -X$.

LTV system in (3.60) can be written in LPV form with a specific dependence on time if it is possible to specify continuous and bounded functions

$$\begin{bmatrix} A(\rho) & B_1(\rho) & B_2(\rho) \\ C_1(\rho) & D_{11}(\rho) & D_{12}(\rho) \\ C_2(\rho) & D_{21}(\rho) & D_{22}(\rho) \\ C(\rho) & D(\rho) & 0 \end{bmatrix}$$

defined on a set $\mathcal{P} \subset \mathbb{R}^m$, and letting the system be described by some unknown smooth parameter curve $\rho(t) \in \mathcal{P}$ with a possible restriction on the rate of variation given as $\dot{\rho} \in \dot{\mathcal{P}} \subset \mathbb{R}^m$. The available a priori information consists of the sets $\mathcal{P}$, $\dot{\mathcal{P}}$, and the on-line information is the actual parameter value $\rho(t)$ and its derivative $\dot{\rho}(t) = \nu(t)$ at the time instant t. This structure comprises linear systems with time-varying measurable parameters and gain scheduling structures for possibly nonlinear systems.

The goal is to guarantee a bound on the mixed H_2/H_∞ functional along each possible parameter curve. Hence, we are lead to searching for controller parameters $A_K(\rho, \nu)$ as a function on $\mathcal{P} \times \dot{\mathcal{P}}$, $B_K(\rho), C_K(\rho), D_K(\rho)$ as functions on $\mathcal{P}$ and a smooth function $\mathcal{X}(\rho)$ on $\mathcal{P}$ such that (3.61) holds on $\mathcal{P} \times \dot{\mathcal{P}}$ after replacing $\dot{\mathcal{X}}$ by

$$\acute{\mathcal{X}}(\rho, \nu) := \sum_{i=1}^{m} \frac{\partial \mathcal{X}(\rho)}{\partial \rho_i} \nu_i \,. \tag{3.71}$$

Along any parameter curve, the *chain rule* implies; $\frac{d}{dt}\big(\mathcal{X}(\rho)\big) = \acute{\mathcal{X}}(\rho, \nu)$ such that (3.61) holds along this curve and guarantees robust stability of level $1/\gamma$, and robust H_∞ and generalized H_2 performance levels of γ and β respectively. With literally the same proof, Theorem 3.35 persists to hold, again after replacing $\dot{X}$, $\dot{Y}$ by $\acute{X}$, $\acute{Y}$, and viewing the inequalities over $\mathcal{P} \times \dot{\mathcal{P}}$.

Theorem 3.36 [21] *There exists functions $A_K(\rho, \nu)$, $B_K(\rho)$, $C_K(\rho)$, $D_K(\rho)$, and a smooth function $\mathcal{X}(\rho)$ satisfying (3.61) on $\mathcal{P} \times \dot{\mathcal{P}}$ if and only if there exists smooth functions $X(\rho)$, $Y(\rho)$ and functions $E(\rho)$, $F(\rho)$, $G(\rho)$ which satisfy (3.65), (3.66), and (3.68). If existing, the size of $A_K(\rho, \nu)$ can be chosen as that of $A(\rho)$.*

Note that our explicit formulas for the controller system matrix A_K reveal that we need to let A_K depend on both ρ and ν. Hence, for the controller implementation, the derivative $\dot{\rho}(t) = \nu(t)$ has to be available.

The inequalities (3.65), (3.66), and (3.68) (and their LPV counterparts after replacing $\dot{X}$, $\dot{Y}$ with $\acute{X}$, $\acute{Y}$) consist of convex but infinite-dimensional optimization problem in

terms of the unknowns $X(\rho)$, $Y(\rho)$, $E(\rho)$, $F(\rho)$, $G(\rho)$, γ, β, and the parameters ρ and ν. One obvious but computationally expensive way to solve this problem is to grid the parameter space in terms of ρ and solve the system of equations simultaneously. The inequalities (3.65), (3.66), and (3.68) then reduce to the solution of finite-dimensional optimization problem, parameterized by ρ and ν. For dense enough grid, this procedure will provide close to optimal non-conservative solution to these inequalities. This approach has the drawback, however, that it does not provide a natural way of scheduling for the unknown matrices. The second approach, used here, is to "postulate" a fixed parameter structure for the unknown matrices X, Y, E, F and G. Letting for example;

$$X(\rho) = X_0 + \rho X_1 \,, \quad Y(\rho) = Y_0 + \rho Y_1 \,, \tag{3.72}$$

and

$$E(\rho) = E_0 + \rho E_1 \,, \;\; F(\rho) = F_0 + \rho F_1 \,, \;\; G(\rho) = G_0 + \rho G_1 \,, \tag{3.73}$$

inequalities (3.65), (3.66), and (3.68) become a series of LMIs with linear dependence on ν and linear and quadratic dependence on ρ.

If ρ and ν appeared in these inequalities in an affine way (actually ν does) then, according to Theorem B.28 in Appendix B.3, one only need to check these matrix inequalities at the vertices of the polytope defined by $\tilde{\mathcal{P}} = [\underline{\rho}_i, \overline{\rho}_i] \times [\underline{\nu}_i, \overline{\nu}_i]$ in the $(\mathcal{P}, \dot{\mathcal{P}})$ space. In order to address the non-affine (quadratic) dependence on the parameter ρ, one can impose a *multiconvexity* requirement on the LMIs (3.65), (3.66), and (3.68). The advantage of this approach lies in the fact that one needs to solve the inequalities only at the vertices of the polytope $\tilde{\mathcal{P}}$.

Suppose we are trying to satisfy an LMI condition $\mathfrak{M}(\rho) \prec 0$ for all $\rho \in \mathcal{P}$. If

$$\frac{\partial^2 \mathfrak{M}}{\partial \rho_i^2} \succeq 0 \quad \text{for all } \rho \in \mathcal{P} \quad \text{and} \quad i = 1, ..., m \,, \tag{3.74}$$

we need only to check $\mathfrak{M}(\rho) \prec 0$ for $\rho \in \mathcal{P}^v$. But the derivative condition places

undesirable restrictions on the structure of the unknowns. However, if P_i is a positive semidefinite matrix (see Appendix B.2) for each $i = 1, ..., m$, then

$$\mathfrak{M}(\rho) + \sum_{i=1}^{m} \rho_i^2 P_i \prec 0 \quad \text{implies} \quad \mathfrak{M}(\rho) \prec 0\,, \tag{3.75}$$

and the multiconvexity requirement for the expression on the left is

$$\frac{\partial^2 \mathfrak{M}}{\partial \rho_i^2} + 2P_i \succeq 0 \quad \forall i = 1, ..., m\,, \tag{3.76}$$

which is more relaxed than $\frac{\partial^2 \mathfrak{M}}{\partial \rho_i^2} \succeq 0$. Moreover, since P_is appear linearly, the search over suitable matrices can be incorporated to the existing LMI conditions.

Conversely, when the LMI condition is $\mathfrak{M}(\rho) \succ 0$ for all $\rho \in \mathcal{P}$, for the multiconvexity requirement we need to satisfy

$$\frac{\partial^2 \mathfrak{M}}{\partial \rho_i^2} + 2P_i \preceq 0 \quad \forall i = 1, ..., m\,. \tag{3.77}$$

The multi-objective LPV model for rotor/AMB systems can be constructed in the following form:

$$\begin{pmatrix} \dot{x} \\ z_1 \\ z_2 \\ y \end{pmatrix} = \begin{bmatrix} A(\Omega) & B_1(\Omega^2) & B_2 \\ C_1 & 0 & D_{12} \\ C_2 & 0 & 0 \\ C & D & 0 \end{bmatrix} \begin{pmatrix} x \\ w \\ u \end{pmatrix}. \tag{3.78}$$

System has parameter dependence to $\Omega(t)$ due to gyroscopic effects and to $\Omega^2(t)$ due to unbalance forces. We desire a robustly stabilizing LPV output feedback controller, for all unstructured uncertainties with H_∞ norm not larger than $1/\gamma$. The controller should also have an H_∞ performance of less than γ, and a generalized H_2 performance of less than β. The structure of such a controller depending on $\Omega(t)$ and $\dot{\Omega}(t)$ can be

taken as

$$u = \left[\begin{array}{c|c} A_K(\rho,\dot{\rho}) & B_K(\rho) \\ \hline C_K(\rho) & 0 \end{array}\right] y \tag{3.79}$$

.

For the open-loop LPV system (3.78), there exists an LPV controller of the form given by (3.79) that provides *global asymptotic stability* and an L_2 gain (H_∞ performance) smaller than γ for the closed-loop system, whenever there exists parameter dependent matrix functions $X(\Omega)$, $Y(\Omega)$, $E(\Omega)$, and $F(\Omega)$ satisfying the following infinite dimensional LMIs in (3.80)-(3.81) to be negative definite, and satisfying the semi-positive definiteness of infinite dimensional LMI (3.82).

$$\begin{bmatrix} \dot{\Omega}\frac{\partial X}{\partial \Omega} + A^T X + XA + FC + (FC)^T & XB_1 + FD & C_1^T \\ (XB_1 + FD)^T & -I & 0 \\ C_1 & 0 & -\gamma^2 I \end{bmatrix} \prec 0\,, \tag{3.80}$$

$$\begin{bmatrix} -\dot{\Omega}\frac{\partial Y}{\partial \Omega} + AY + YA^T + B_2E + (B_2E)^T & B_1 & (C_1Y + D_{12}E)^T \\ B_1^T & -I & 0 \\ C_1Y + D_{12}E & 0 & -\gamma^2 I \end{bmatrix} \prec 0\,, \tag{3.81}$$

$$\begin{bmatrix} Y & I \\ I & X \end{bmatrix} \succeq 0\,. \tag{3.82}$$

Introducing the additional constraints from the multiconvexity requirements given in (3.76) and (3.77) which are parameter independent due to the selection of unknown matrices according to (3.72) and (3.73), imposes convex constraints on the variables. Owing to this convexity, it is sufficient for the LMIs (3.76) and (3.80)-(3.81) to be satis-

fied at the vertices of the convex set. This leads to a minimization problem in γ in the matrix variables $X_0, X_1, Y_0, Y_1, E_0, E_1, E_2, F_0$, and F_1. In the case of multi-objective LPV controller synthesis, (3.82) is replaced by

$$\begin{bmatrix} \beta I & C_2 Y & C_2 \\ (C_2 Y)^T & Y & I \\ C_2^T & I & X \end{bmatrix} \succeq 0\,, \tag{3.83}$$

and $L_2 \to L_\infty$ gain β of the closed-loop system M_2 is minimized with an acceptable $L_2 \to L_2$ gain γ_s for the closed-loop system M_1, such that $\gamma_s \geq \gamma$.

A LPV controller can be constructed using (3.69) and (3.70) with the parameter (rotor speed) dependent matrix functions X, Y, E and F found from the solution of the LMIs above. Notice from the structure of (3.70) that the derivative of rotor speed (angular acceleration) is also required during on-line scheduling of the controller.

4. CONTROLLER DESIGN AND SIMULATIONS FOR A ROTOR/AMB SYSTEM

Rotor/AMB model and associated data selected for the design of various robust controllers is shown in Figure 4.1 (not to scale) and Table 4.1.

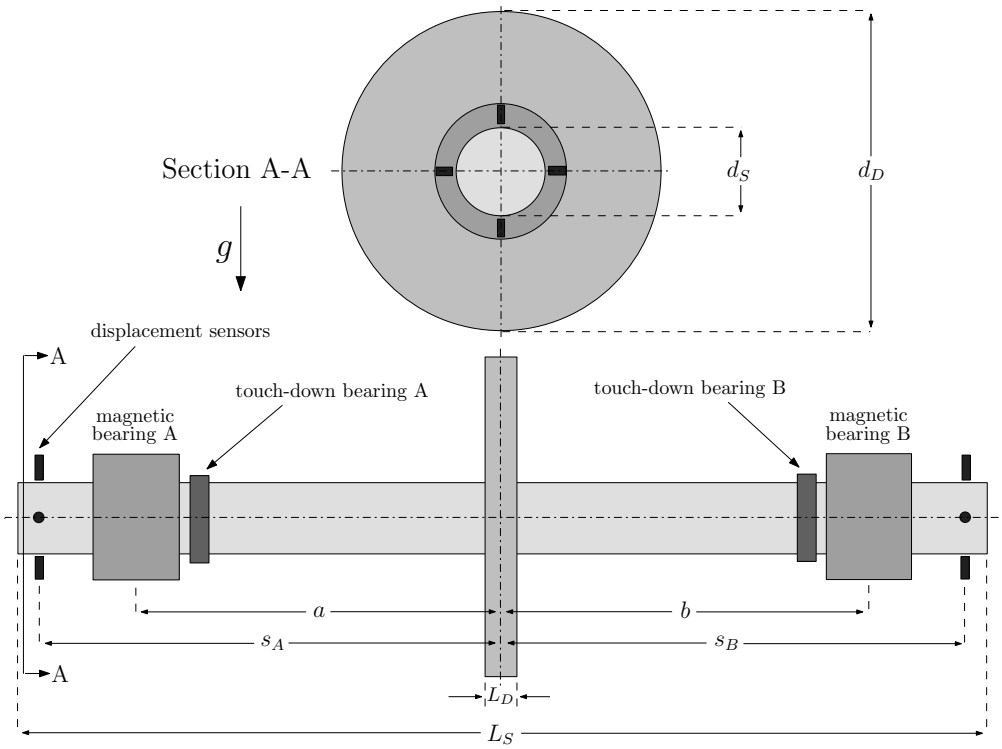

Figure 4.1. Horizontal rotor with radial AMBs.

Table 4.1. Rotor/AMB data.

Symbol	Value	Unit	Symbol	Value	Unit	Symbol	Value	Unit
M_S	85.90	kg	L_S	1.50	m	s_0	$2.0 \cdot 10^{-3}$	m
M_D	77.10	kg	L_D	0.05	m	s_1	$0.5 \cdot 10^{-3}$	m
I_r	17.28	kg·m^2	d_S	0.10	m	i_0	3.0	A
I_p	2.41	kg·m^2	d_D	0.50	m	k_M	$7.8455 \cdot 10^{-5}$	N·m^2/A^2
a	0.58	m	s_A	0.73	m	k_s	$-3.5305 \cdot 10^5$	N/m
b	0.58	m	s_B	0.73	m	k_i	235.4	N/A

Note that the axial magnetic bearing is not shown in the Figure 4.1. A PMB or a one-channel PID controlled AMB can be used to control the axial motion of the rotor.

In the evaluation of magnetic bearing constant k_M, it is assumed that the maximum force F_M which must be provided by the AMBs is three times the rotor weight in the vertical direction. Selection of the premagnetization current i_0 is made by trial end error during the simulations of the closed-loop system. In most of the standard literature and majority of the applications of operation of AMBs with a premagnetization current half the size of maximum amplifier output current is considered ($i_0 = i_{max(hw)}/2$). Reduction of the premagnetization current increases the linearization error as the control currents should not be too large with respect to the bias current. Linear operation can be extended to be valid for the coil currents (which are $i_0 \mp i_c$) in the interval; $|i_0 \mp i_c| \leq (i_{max(hw)} - i_0)$ [25]. For horizontal rotors premagnetization current can not be smaller than the gravity compensation current which is required to suspend the rotor against gravity. Further constraint on the minimization of the premagnetization current is the fact that coil currents to levitate the rotor from the touch-down bearings should not exceed the maximum current limit $i_{max(hw)}$ of the power amplifier. For an airgap s_0 of 2.0 mm between the rotor and the AMBs, maximum levitation distance is found to be 0.5 mm which is set as the gap s_1 between the rotor and touch-down bearings.

A 3 kW power amplifier is chosen with a voltage rating U_0 of 300 V and a maximum current limit $i_{max(hw)}$ of 10 A, with the nominal coil inductance $\mathcal{L}_0$ of 157 mH to yield a cut-off frequency ω_0 around 273 rad/s ($\mathcal{L}_0 \cong k_i^2/k_s$ [33]). Other parameters of interest in the design of a stabilizing controller are the sensor gain G_s which is taken as 7071 V/m, and the gain of the (current) power amplifier G_a which is taken as unity (1 A/V). Further specifications on the mechanical and electromagnetic design of the magnetic bearings are out of the scope of the thesis. One may refer to [33], which is a comprehensive guide on the theory and design of magnetic bearings.

Two H_∞ controllers are designed with nominal performance using the nominal system (3.12) with their respective system matrices given in (3.13) and (3.14). Nominal

speed and the nominal bearing stiffness chosen for both of the controllers are 3000 rpm ($\approx$ 314.2 rad/s), and $-3.5305 \cdot 10^5$ N/m respectively. System given in (3.13) has two frequency dependent weighting matrix functions in order to limit the controller bandwidth and to incorporate integral control action for the error signal. Selected functions used for all channels are given in Figure 4.2. Related transfer matrix functions

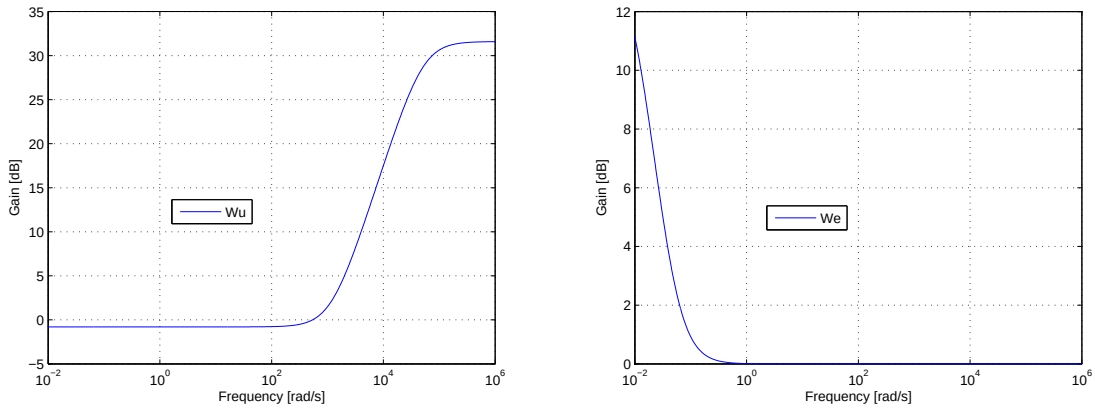

Figure 4.2. Controller output and error weighting functions.

are

$$W_u = \text{diag}\left(38\,\frac{s+1200}{s+50000}\right), \qquad W_e = \text{diag}\left(\frac{s+0.05}{s+0.01}\right). \tag{4.1}$$

System (3.14) uses only the weighting matrix function W_u in controller synthesis.

Suboptimal H_∞ controllers are synthesized using Lemma 3.11 via semidefinite programming [50-51]. The order of the controllers are the same as their respective open-loop augmented system (with weighting functions) states. Hence controllers K1 and K2 have sixteen and twelve states respectively. Frequency response of the synthesized controllers in terms of their singular values are as given in Figure 4.3. Singular values of the closed loop systems with controller K1 and controller K2 are given in Figure 4.4.

H_∞ norm (maximum singular value) γ of the nominal closed-loop system with K1 and with K2 at maximum speed of 3000 rpm and maximum mass center displacement of $0.50 \cdot 10^{-3}$ are 99.24 and 21.05 respectively.

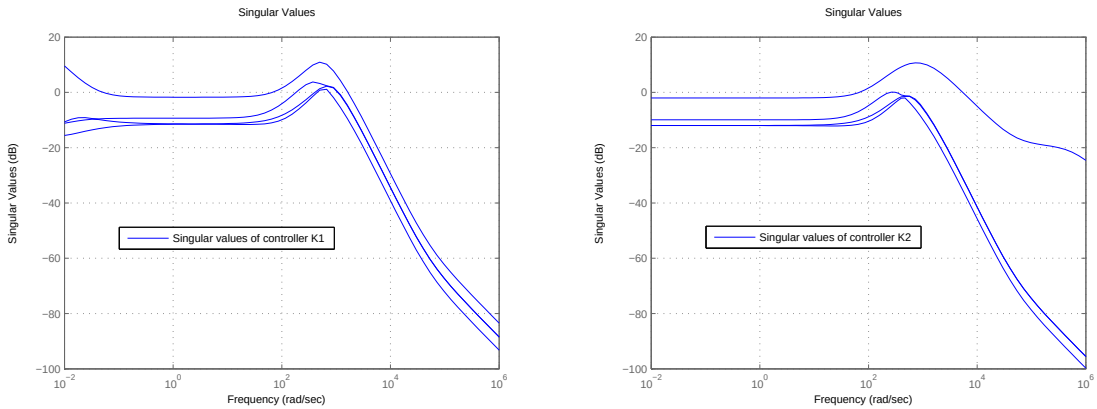

Figure 4.3. Singular values of K1 and K2.

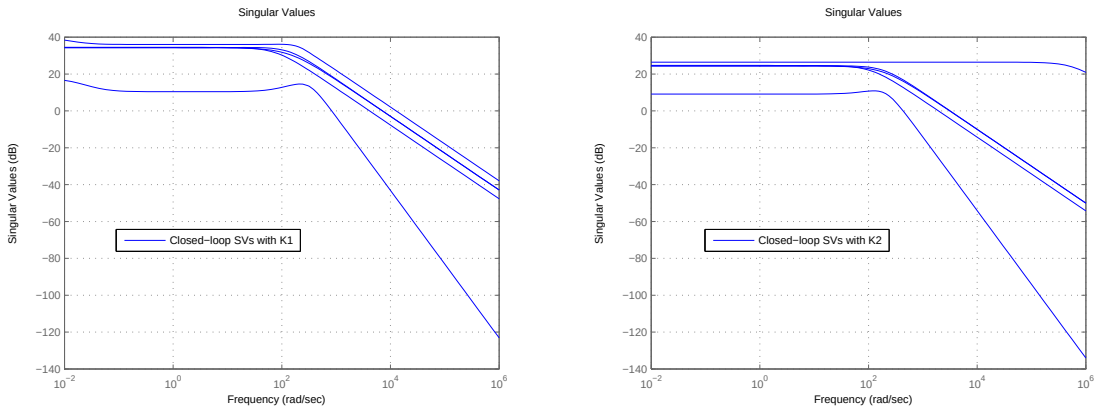

Figure 4.4. Closed-loop singular values with K1 and K2.

Table 4.2. H_∞ performance with K1 for different design parameters.

Maximum speed (RPM)	Maximum mass center displacement (m)	γ
1500	$0.25{\cdot}10^{-3}$	70.96
3000	$0.25{\cdot}10^{-3}$	97.06
6000	$0.25{\cdot}10^{-3}$	99.81
1500	$0.50{\cdot}10^{-3}$	89.57
3000	$\mathbf{0.50{\cdot}10^{-3}}$	**99.24**
6000	$0.50{\cdot}10^{-3}$	100.07

Table 4.3. H_∞ performance with K2 for different design parameters.

Maximum speed (RPM)	Maximum mass center displacement (m)	γ
1500	$0.25{\cdot}10^{-3}$	11.41
3000	$0.25{\cdot}10^{-3}$	15.42
6000	$0.25{\cdot}10^{-3}$	31.77
1500	$0.50{\cdot}10^{-3}$	12.62
3000	$\mathbf{0.50{\cdot}10^{-3}}$	**21.05**
6000	$0.50{\cdot}10^{-3}$	52.01

Apart from the strong effect of the weight W_e used for integration in the controller, the H_∞ norm depends on the nominal operation speed and maximum allowable unbalance (mass center displacement) of the rotor. Change in the performance of the closed-loop systems based on the selection of different nominal parameters are shown in Table 4.2 and Table 4.3 for the system with controller K1 and for the system with controller K2 respectively.

According to *small gain theorem* (Theorem 3.1), robustness of the closed-loop nominal system with respect to unstructured uncertainty is inversely proportional to its H_∞ norm. Therefore the system with K2 is more robust to unstructured uncertainty. Hence K2 is a better choice with respect to robustness to unstructured uncertainty. Moreover, its implementation is easier as it has less number of states (twelve) than the controller K1 having sixteen states. However, as can be seen from the Figure 4.4, system with K2 is more prone to amplification of high frequency signals such as sensor noise. K1 would be a better choice in the case of static loads (d.c.) acting on the system and integral control action is required to keep the rotor at the bearing center and/or when there exists considerable high frequency noise in the system.

Choice of the controller has also an affect on the critical speeds of the rigid body whirling modes of the rotor in the closed-loop system. A controller with high gain on displacement feedback signal shifts the critical speeds to higher frequencies than a controller with lower gain. Critical speeds (eigenfrequencies) of the rotor in open-loop

and in closed-loop with controllers K1 and K2 are shown in Figure 4.5 and Figure 4.6.

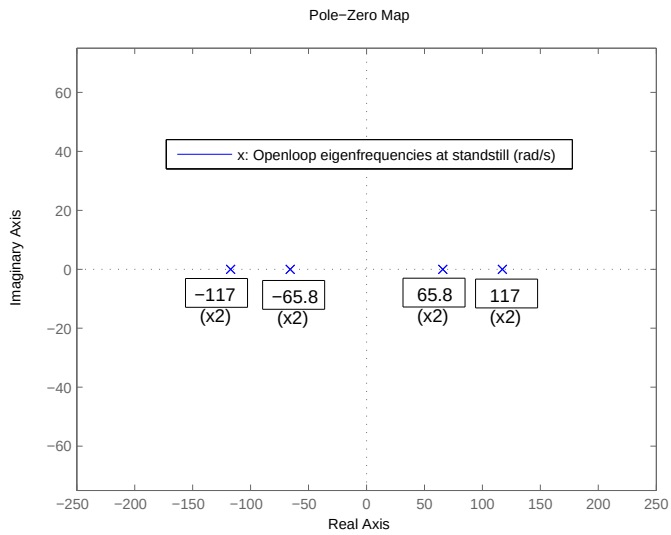

Figure 4.5. Open-loop eigenfrequencies of the rotor at standstill.

When the rotor passes the critical speed a phase shift of 180 degrees occurs due to the fact that the impedance of the rotor to unbalance excitation shifts from the bearing stiffness to the inertia of the rotor (see Appendix A). Critical speed is when the phase shifts exactly 90 degrees, where the impedance of the rotor to unbalance forces is solely provided by the damping forces within the magnetic bearings. Note that the critical speed of the conical mode of the rotor (second rigid mode eigenfrequency) does not appear due to the fact that the unbalance force is modeled to be completely static, acting through the mass center of the axially and radially symmetric rotor.

Note also that closed-loop system with K1 has an additional phase shift (lag) in the low frequency range due to the integral action of the controller. As shown in the Figure 4.6 below, critical speed of the closed-loop system with the controller K1 and K2 are approximately 120 rad/s ($\cong$ 1146 rpm) and 150 rad/s ($\cong$ 1433 rpm) respectively. Therefore both of the controllers are designed to control the system in the supercritical as well as the subcritical range of operation.

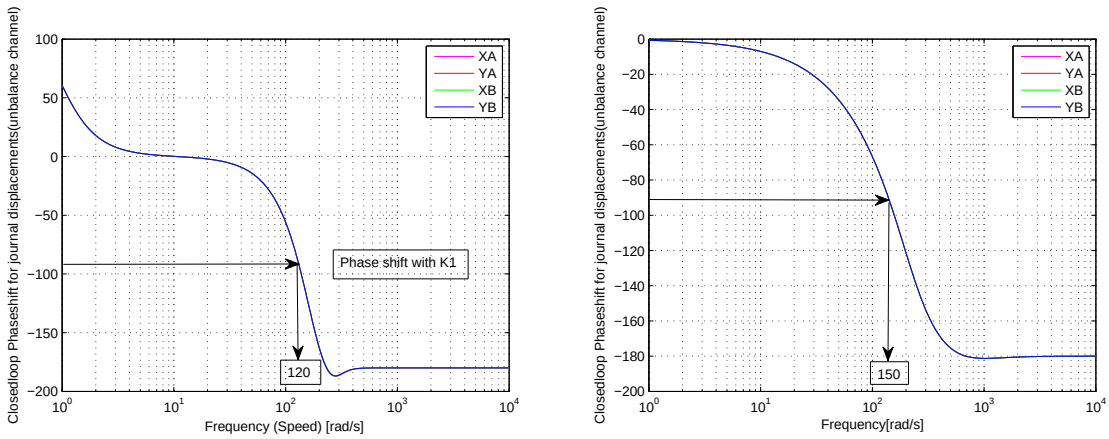

Figure 4.6. Closed-loop parallel mode eigenfrequency of the rotor with K1 and K2.

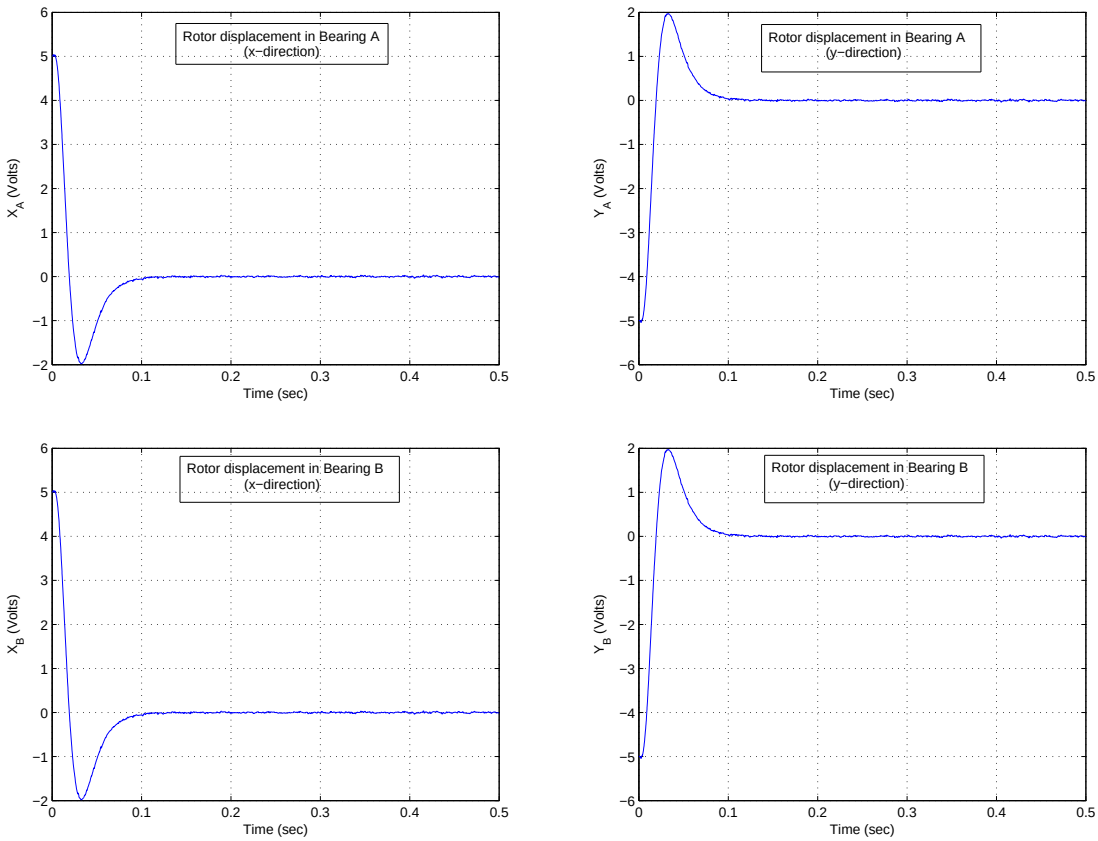

Figure 4.7. Rotor position in the bearings during the start-up.

Due to its simpler structure, we analyze the H_∞ performance of the system using the controller K2 in the simulations. Figure 4.7 and Figure 4.8 show the position of the rotor in the bearings during the initial lift from the touch-down bearings and the respective control currents provided by the controller. Note that it takes only

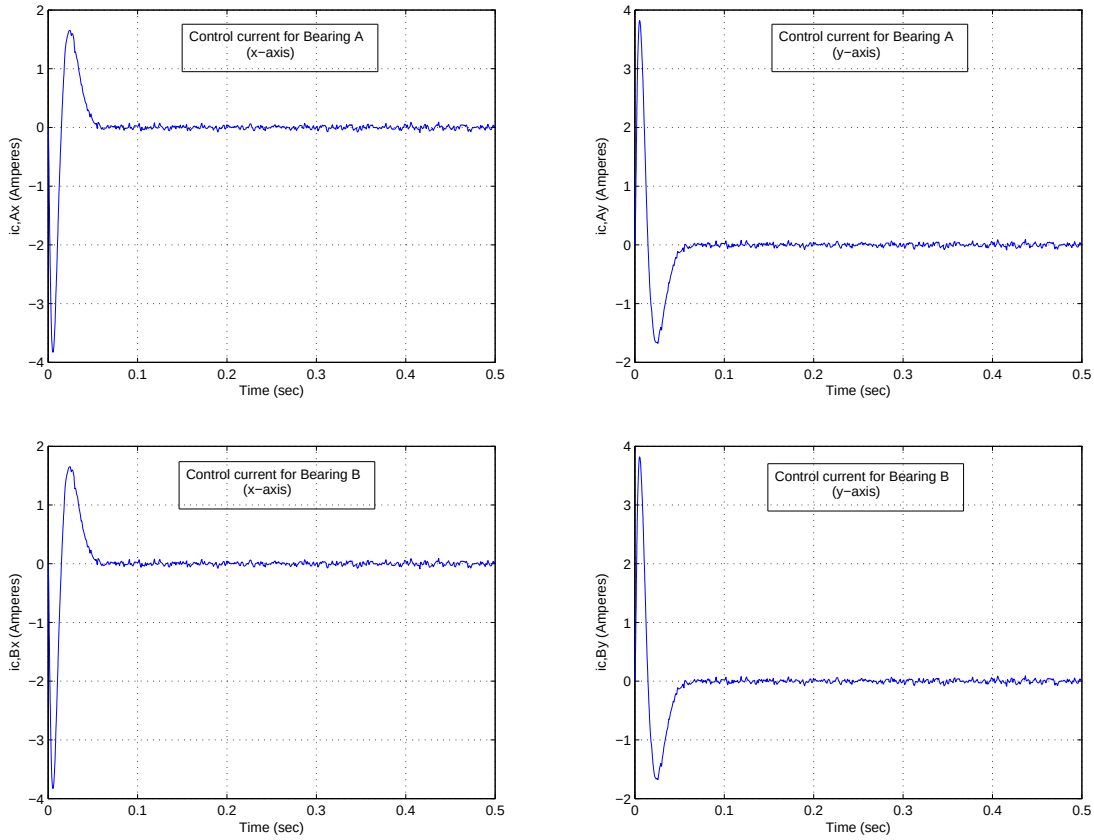

Figure 4.8. Control currents during the start-up.

0.1 second for the bearings to lift the the rotor from the touch-down bearings to the center of rotation with a moderate level of overshoot. As shown in Figure 4.8, control currents do not exceed 4 A which is the limit for linear operation of the actuators $\big(|i_0 \mp i_c| \leq (i_{max(hw)} - i_0)\big)$.

Result of the simulations at steady-state (when the rotor reaches its steady operation speed; 3000 rpm) are shown in Figure 4.9 for bearing A. Results for bearing B are similar. Disturbance acting on the system, i.e., unbalance force and sensor/electronic noise are shown in Figure 4.10. Mass center displacement (eccentricity) due to unbalance of the rotor is assumed to be $0.25 \cdot 10^{-3}$ m in the simulations. Note that the controller substantially suppresses the whirl amplitudes due to unbalance. It is evident

from Figure 4.9 that the peak value of the whirl is approximately 0.1 V, corresponding to $14 \cdot 10^{-6}$ m. Hence, the H_∞ controller K2 reduces of unbalance whirl amplitude of the rotor around 95%. Moreover, as shown in Appendix E, a pulse signal with 1 V amplitude and 0.025 seconds duration is injected to the loop for all channels at the input of the controller. It is clear from the simulation results that it does not cause stability problems hence verifying the internal stability of the feedback connection.

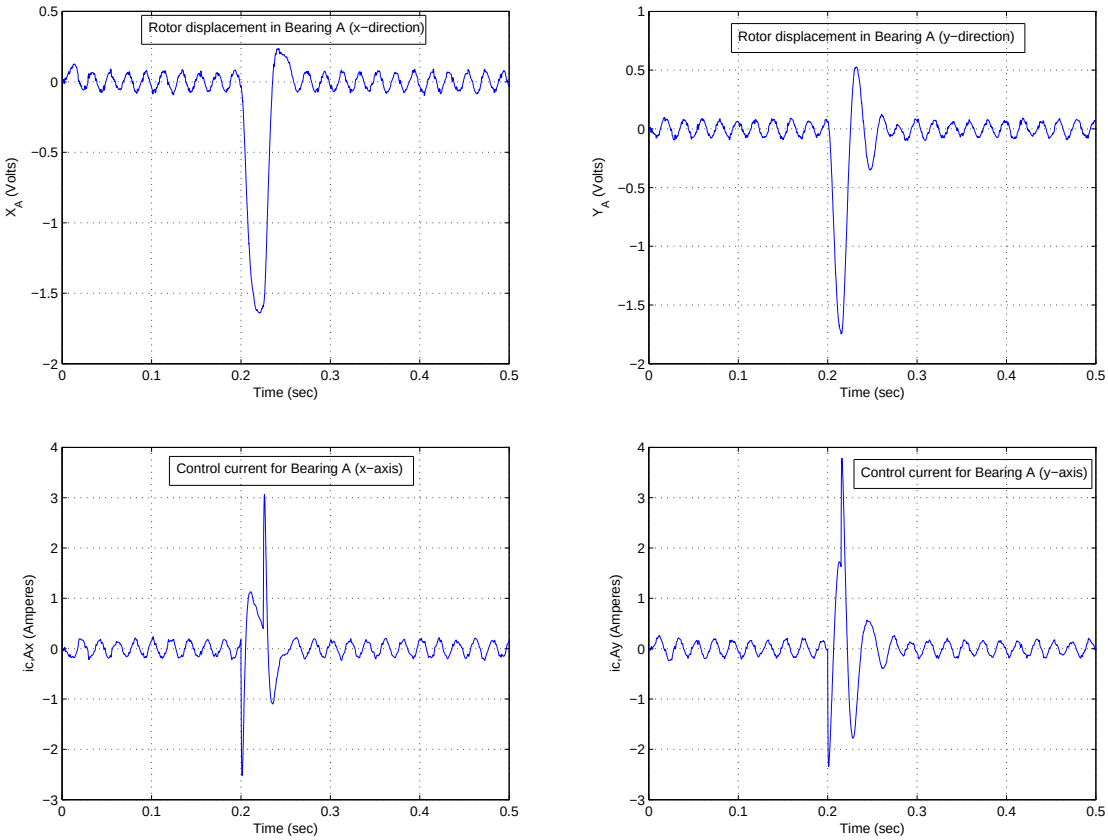

Figure 4.9. Rotor position and control currents for bearing A at 3000 RPM.

Recall that controllers K1 and K2 have orders of sixteen and twelve respectively. If their weakest four Hankel singular values of coprime factors shown in the Figure 4.11 are eliminated, reduced order controllers with twelve and eight states respectively can be constructed. Singular values of the reduced order controllers K1r, K2r and their respective closed-loop singular values with K1r and K2r are shown in the Figure 4.12 and Figure 4.13 respectively. H_∞ norm γ of the closed-loop system with K1r and K2r increases from 99.24 to 529.55 and from 21.05 to 62.07 respectively.

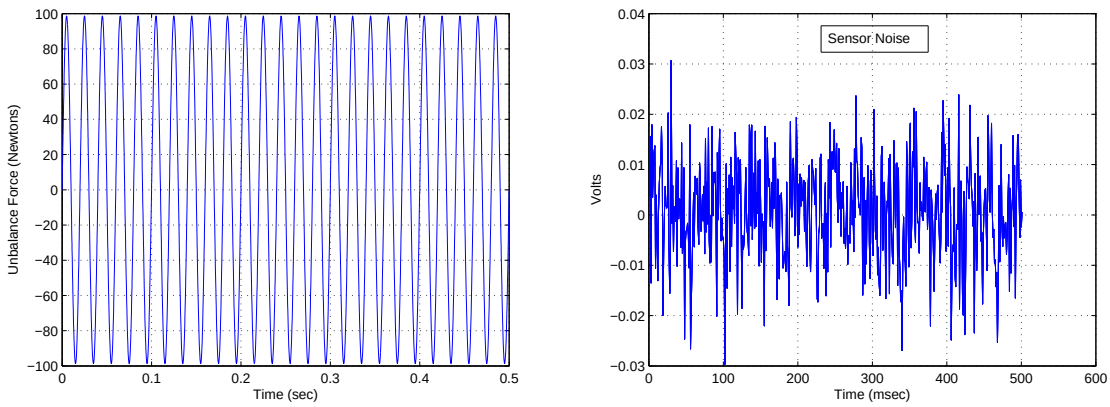

Figure 4.10. Unbalance force and noise signal.

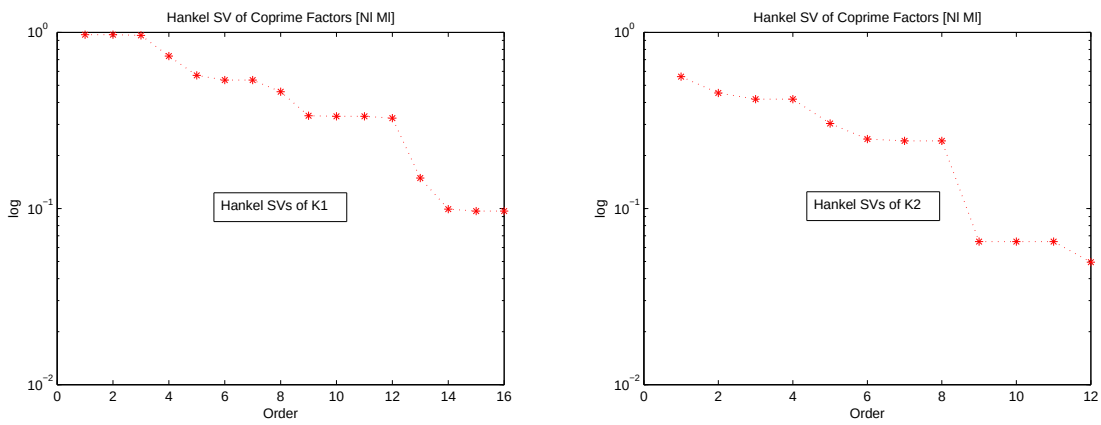

Figure 4.11. Hankel singular values of coprime factors of K1 and K2.

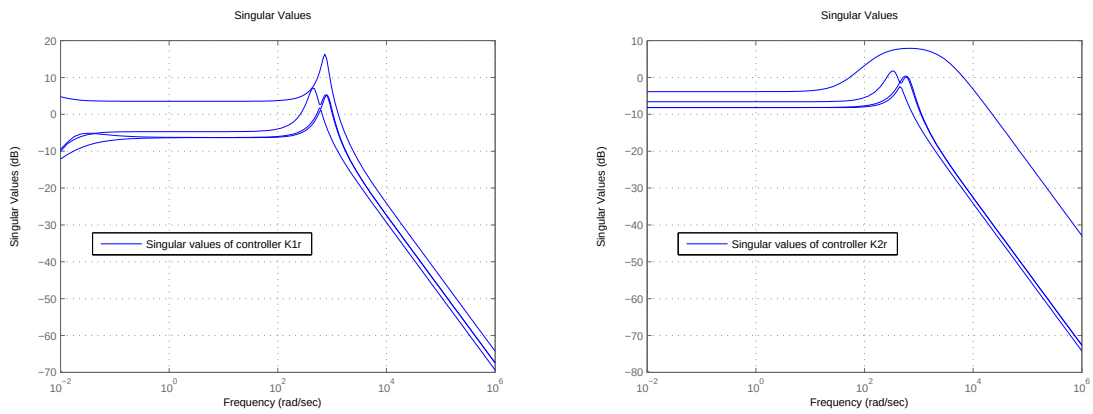

Figure 4.12. Singular values of the reduced order controllers K1r and K2r.

Critical speed of the closed-loop system also shifts with the reduced-order controllers. Phase shifts during the pass from critical speed is shown in the Figure 4.14 respectively.

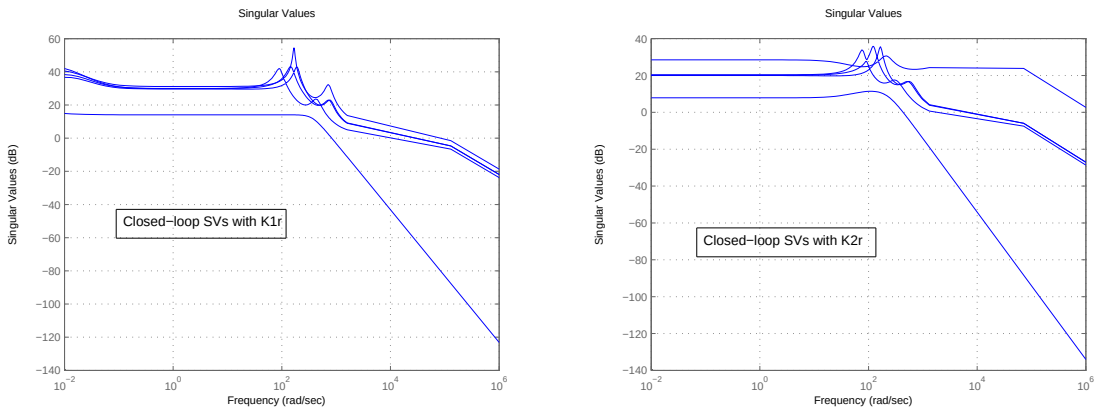

Figure 4.13. Singular values of the closed-loop system with K1r and K2r.

An H_∞ controller (K3) with parametric uncertainty structure is designed according to model (3.19) using the system matrices from (3.14) and (3.20). Nominal bearing stiffness is set to $-3.5305 \cdot 10^5$ N/m as given in Table 4.1 with 25% uncertainty. Nominal speed is selected as half of the maximum speed of operation. H_∞ performance γ of the system for $\Omega_{max} = 3000$ rpm is found to be 38.65.

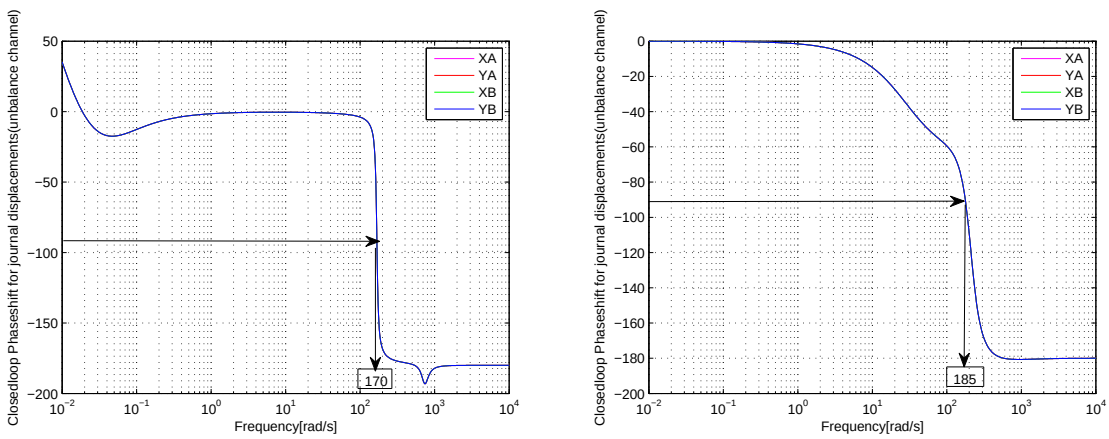

Figure 4.14. Closed-loop parallel mode eigenfrequency of the rotor with K1r and K2r.

Keeping the uncertainty on the bearing stiffness constant (25%), robust stability of the closed-loop system is tested for several maximum operating speeds with μ-analysis according to (3.25) and (3.26) using the robust control toolbox for MATLAB®

[49]. Results are shown in Figure 4.15. It can be seen from the Figure 4.15 (left) that robust stability is lost for rotor speeds exceeding 4085 rpm, where $\mu \geq 1$. Uncertainty in bearing stiffness also has a substantial effect on robust stability. It is shown in Figure 4.15 (right) for constant (3000 rpm) maximum rotor speed.

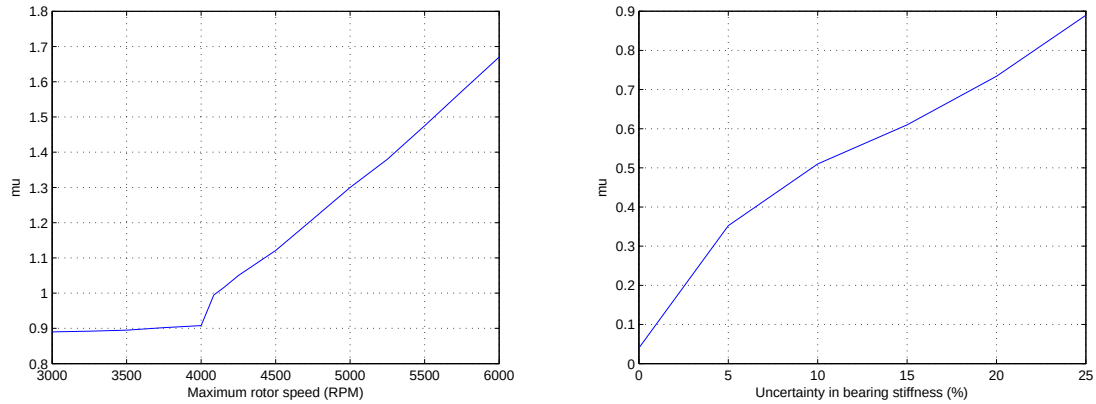

Figure 4.15. Structured singular value with respect to maximum rotor speed (left) and with respect to uncertainty in bearing stiffness (right).

Singular values of the controller and the closed-loop system for a maximum operating speed of 4085 rpm ($\cong$ 418 rad/s) are shown in Figure 4.16. H_∞ performance γ of the closed-loop system for Ω_{max} = 4085 rpm is 47.86.

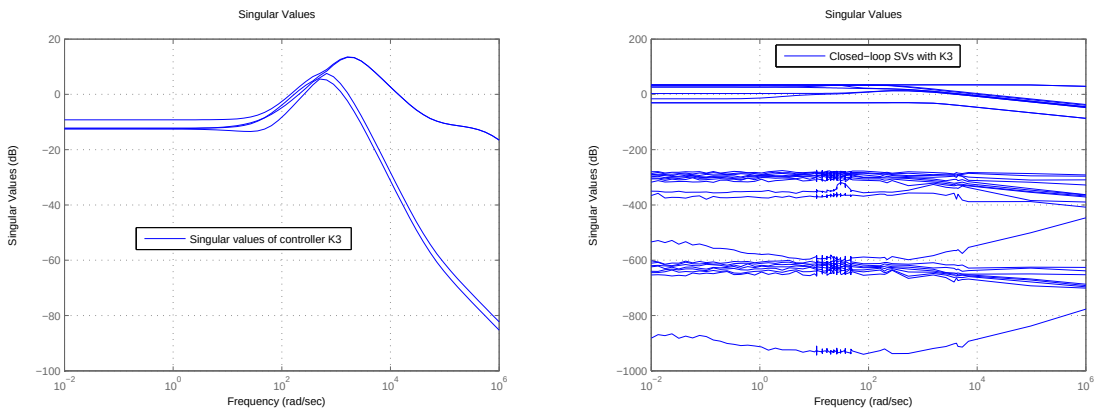

Figure 4.16. Singular values of the controller and closed-loop system with K3.

LPV controller for the parameter (rotor speed) dependent rotor/AMB system (3.78) can be designed via semidefinite programming (50-51) satisfying the LMIs (3.80-3.82), (3.76), and (3.77) at all the vertices of the convex hull (i.e., $\Omega_{min}, \Omega_{max}, \dot{\Omega}_{min}, \dot{\Omega}_{max}$),

with the predefined parameter dependent structure for the variable matrix functions according to (3.71-3.73). System matrices in (3.78) are similar to the matrices for the model in (3.12), except the decomposition of the system matrix $A(\Omega)$ and exogenous input matrix $B_w(\Omega^2)$ into two parts, such as; $A(\Omega) = A_0 + \Omega A_\Omega$ and $B_w(\Omega^2) = B_{w0} + \Omega^2 B_{w\Omega}$. Construction of the full-order controller can be realized on-line as shown in Appendix E using (3.69-3.70).

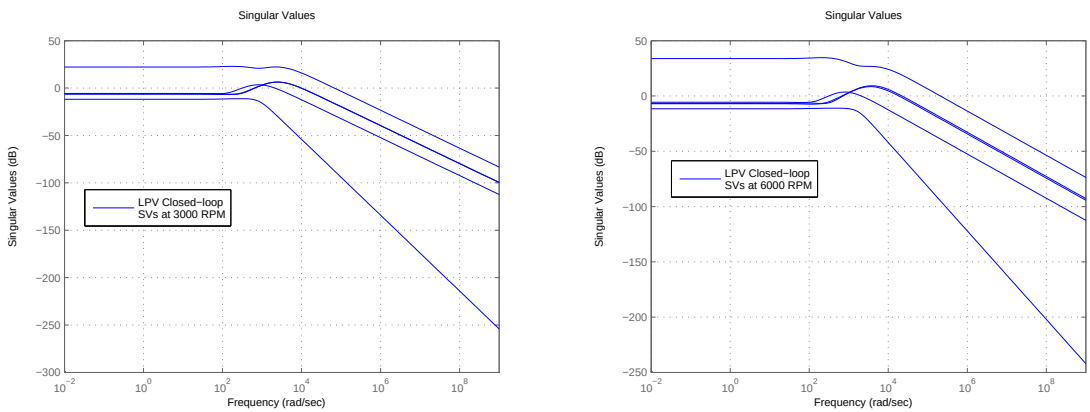

Figure 4.17. Singular values of the closed-loop LPV system.

Singular values of the closed-loop system and the LPV controller at two different speeds; 3000 and 6000 rpm are shown in Figure (4.17) and Figure (4.18) respectively.

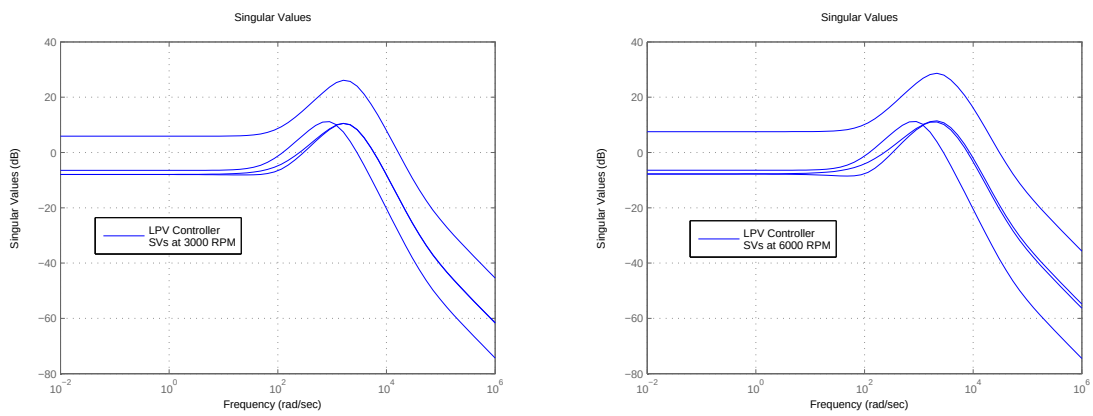

Figure 4.18. Singular values of the LPV controller.

Controller to robustly stabilize the system with H_∞ performance is synthesized inside the convex hull described above within the rotational speed range of 0 rad/s to

Table 4.4. H_∞ performance of LPV closed-loop systems at 3000 rpm.

Structure of X and Y		γ	Controller Form
$X = X_0 + \Omega X_1$	$Y = Y_0 + \Omega Y_1$	15.92	(4.2)
$X = X_0$	$Y = Y_0 + \Omega Y_1$	19.13	(4.3)
$X = X_0$	$Y = Y_0$	27.56	(4.3)

Table 4.5. H_∞ performance of LPV closed-loop systems at 6000 rpm.

Structure of X and Y		γ	Controller Form
$X = X_0 + \Omega X_1$	$Y = Y_0 + \Omega Y_1$	56.31	(4.2)
$X = X_0$	$Y = Y_0 + \Omega Y_1$	65.42	(4.3)
$X = X_0$	$Y = Y_0$	102.29	(4.3)

614 rad/s (6000 rpm), and angular acceleration range of -15 rad/s^2 to 15 rad/s^2. H_∞ performance γ of the closed-loop LPV system at the maximum instantaneous speed 6000 rpm is 56.31. Note that this performance is achieved with a controller of the form

$$u = \left[\begin{array}{c|c} A_K(\Omega, \dot{\Omega}) & B_K(\Omega) \\ \hline C_K(\Omega) & 0 \end{array}\right] y\,, \tag{4.2}$$

If the matrix function X is assumed to be time-invariant, than the controller matrices will not depend on the angular acceleration of the rotor due to the controller formulation given in (3,70) and the controller will be of the form

$$u = \left[\begin{array}{c|c} A_K(\Omega) & B_K(\Omega) \\ \hline C_K(\Omega) & 0 \end{array}\right] y\,. \tag{4.3}$$

H_∞ performance of the LPV controllers and the closed-loop system for different structures on matrix variables X and Y at operating speed of 3000-6000 rpm are given in Table 4.4 and Table 4.5 respectively.

Comparing the performance of case 1 and case 2, it can be said that there is virtually no loss of performance if the controller is constructed without the information on angular acceleration of the rotor. The controller synthesized with $X = X_0$ and $Y = Y_0$ (case 3) guarantees robust stability even in the arbitrarily fast changes in Ω. However the rate of change of angular speed Ω, i.e., angular acceleration, is always limited by the maximum available torque of the driver. Hence, this bound can be used and the controller can be constructed accordingly to have a comparatively better performance of case 2 over case 3.

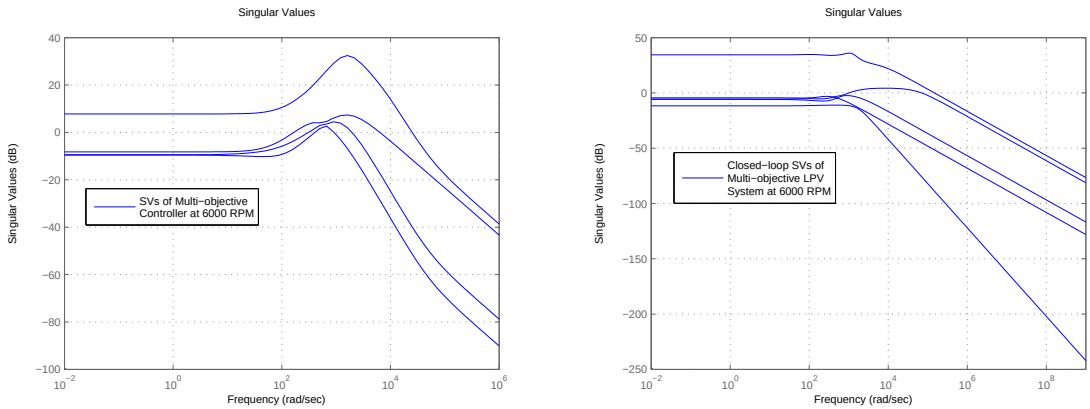

Figure 4.19. Singular values of the multi-objective LPV controller (left) and the closed-loop system (right).

A multi-objective LPV controller with mixed performance specification is synthesized within the same convex hull as the single objective LPV controller for the case 2 with a maximum operating speed of 6000 rpm. Generalized H_2 performance β of the multi-objective LPV controller is found to be 364.4, for a reduced H_∞ performance level γ_s of 72.12 at 6000 rpm. Figure 4.19 shows the singular values of the multi-objective controller and the closed-loop system at 6000 rpm.

A pulse (transient) signal with 1 V amplitude and 0.025 seconds duration and is injected at 0.2 seconds of simulation time into the loop at the input of the controller (see Appendix E). Control current and rotor position at bearing A in y-axis for LPV control with H_∞ performance and with mixed performance is shown in Figure 4.20 and 4.21 respectively. Comparing Figure 4.20 and Figure 4.21 it is clear that the peak values

of both the control current and rotor position are reduced in the closed-loop system with the multi-objective controller. Hence the mixed H_2/H_∞ control for rotor/AMB systems provides additional flexibility with respect to the transient effects during the operation.

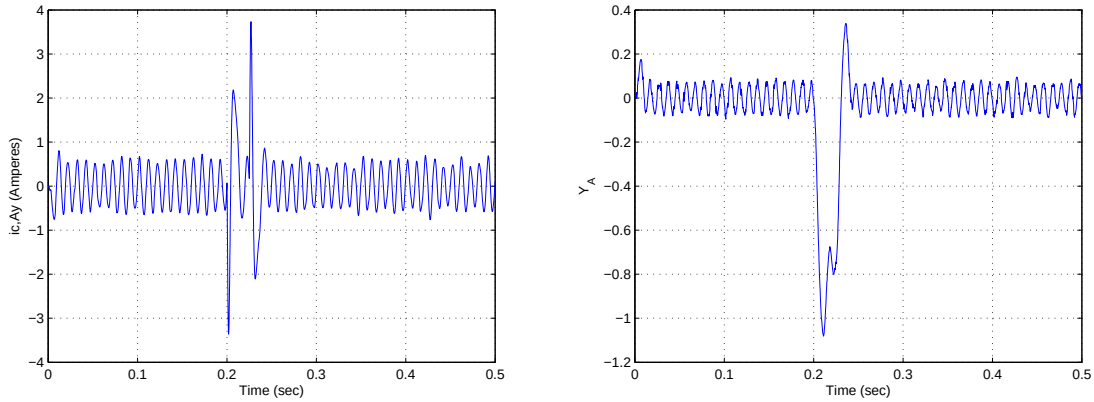

Figure 4.20. Control current and rotor displacement in y-axis of bearing A for LPV H_∞ control.

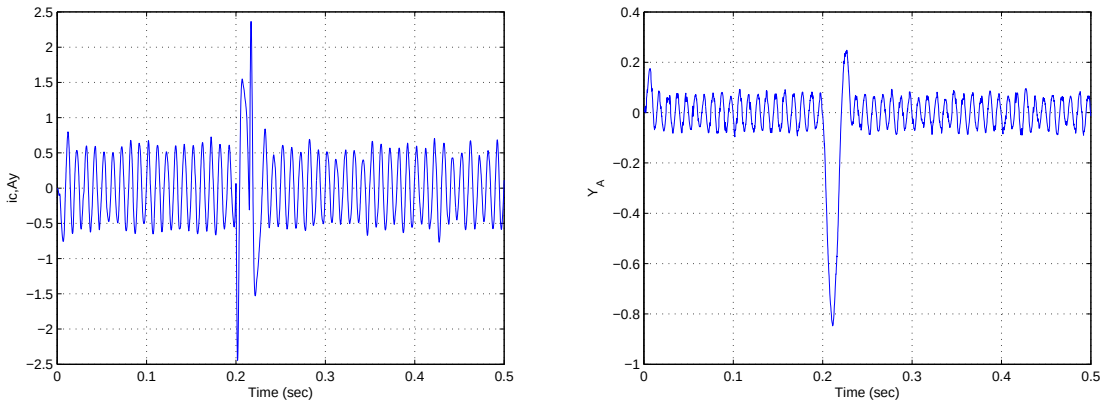

Figure 4.21. Control current and rotor displacement in y-axis of bearing A for LPV "mixed" control.

5. DISCUSSION AND RECOMMENDATIONS FOR FUTURE WORK

Robust linear control techniques for rotor/AMB systems such as, H_∞ and LPV control, can yield to robustly stable high performance controllers. However, linearization of the actuator forces to design a linear controller requires bias currents to be provided to the bearings, and considerable amount of power used in the stabilization the rotor is wasted. Therefore, minimizing the bias currents is favorable for efficient power consumption which is one of the main advantages of magnetic bearings. Moreover, high bias currents pushes the rigid body poles far away from the origin and requires additional effort from the controller to move the poles back and then on to the imaginary axis. Instead of a constant bias current, a promising option can be to provide a variable bias current to bearings in proportion to the variable control currents during the operation. For horizontal rotors, gravity compensation current can be used as the constant part of the bias current and the bias current required for linearization can be calculated and provided to the bearings as half of the instantaneous control current. One important issue is keeping the bearing stiffness isotropic to achieve acceptable rotor dynamics (to prevent backward modes and the separation of the twin poles). Therefore, bias currents for both axes of a bearing should be identical, and equal to the half of the higher control current. For vertical rotors, as gravity compensation is the task of the axial controller, constant part of the bias current will be zero.

The performance of the AMB is measured in terms of bearing stiffness, power consumption and power losses. Bearing stiffness refers to the rate of change of the closed-loop actuator force with respect to a change in position. It is closely related to the controller's ability to regulate the rotor against disturbances. The power consumption of the bearing is measured as the electrical power supplied to the bearing $U_{app} \cdot i$ where U_{app} is the voltage applied across the coil and i is the current through the coil. This power supplied is consumed in three ways: (1) it produces the control forces, (2) it implements the bias flux through the bias current, or (3) gets dissipated as heat

by the power loss mechanisms. The loss mechanisms include resistive power dissipation, eddy-current core loss, eddy-current drag, and hysteretic core loss [33]. Note that eddy-currents generated in the rotor lead to rotor heating, and consequently, thermal expansion of the rotor and degeneration of the nominal airgap. As the bias increases, bearing stiffness increases, the power required to implement the bias increases, and the loss mechanisms dissipate more power. However, the corresponding increase in bearing stiffness reduces the amount of power required to implement the control forces. Thus, the total power consumption may actually decrease as the bias increases. Therefore, a trade-off exists between the bearing stiffness, the power consumption, and power losses. Moreover, minimum bias current may not always produce minimum power consumption under dynamic loading such as that caused by unbalance disturbances. Magnetic bearings already suffer from a limited force capacity, and low bias current operation decreases the dynamic load capacity significantly and render the bearings vulnerable to changes in operating conditions.

Though we have used a weighting function to penalize very quick (high frequency) control input requirements, it has been a tacit assumption in our work that the magnetic bearing coil current (power amplifier output) instantly follows the command signal from the controller (power amplifier input). This assumption has been motivated by the fundamental law of linearized bearing force in (2.5), which expresses the electromagnetic force as a function of the current. However, the inductance of the magnetic bearing coil will resist any sudden change in current, hence fast current changes can only be achieved by a suitably high internal amplifier voltage. In other words, the coil current i is a *system state* and contributes to the overall system dynamics. Therefore, the basic magnetic bearing model derived in Section 2.1 must be expanded by taking into account the electrical properties of the bearing magnet and the power amplifier, i.e., the coil inductance $\mathcal{L}$ and its resistance $\mathcal{R}$ as well as the amplifier output voltage U_{app}. The inductance (or "self-inductance") $\mathcal{L}$ varies with the rotor position r. For a linearized description, however, $\mathcal{L} = \mathcal{L}_0$ is defined for an assumed constant position and its value is considered in the operating point $r = 0$. The rotor motion in the magnetic field of the bearing magnet also generates a voltage across the bearing coil, similar to the case of an electric motor. Hence the total voltage of the power ampli-

fier is used for overcoming the coil inductance and resistance and the motion induced voltage coefficient k_u as [33]:

$$U_{app} = \mathcal{R}i + \mathcal{L}_0 \frac{d}{dt} i + k_u \frac{d}{dt} r \,. \tag{5.1}$$

Based on the theory of electromechanical energy conversion, it can be shown that the coefficient k_u is theoretically equal to the bearing current stiffness k_i [4]. Moreover, it can be shown that [33] the magnetic bearing constants k_i and k_s and the coil inductance $\mathcal{L}$ are interdependent quantities; $\mathcal{L} = k_i^2/|k_s|$. The reason for this is the fact that the magnetic bearing is a device that can transform electrical into mechanical energy back and forth, similar to electric motors and generators. However this energy transformation is not conservative, since losses occur from eddy currents, flux leakage, magnetic hysteresis and other nonlinear sources, all of which compromise this ideal equivalence of k_u and k_i.

Following from the important statement about the applied voltage U_{app} being the "true" system input variable rather than the coil current i, the complete set of basic linearized model equations additionally comprises the AMB's voltage current dynamics (5.1), together with the force/current relation (2.5) and the equations of motion (2.13-2.17) of the mechanical part of the system. Consequently, the power amplifier can no more be considered a voltage-to-current amplifier. More precisely we have to speak of a voltage-to-voltage amplifier when addressing the AMB system's power amplifier. This control scheme is called "voltage control" rather than "current control" which is the method used in the thesis as well. A most important difference between current and voltage controlled AMB systems is the location of open-loop eigenvalues. As shown in the the Section 2.2, current control yields $\mp\sqrt{\lambda_1}$ and $\mp\sqrt{\lambda_2}$ twin poles (8 poles in total) for the open-loop eigenvalues, therefore results in unstable open-loop system having four eigenvalues in the right half of the complex plane. On the other hand, for the coil resistance $\mathcal{R} = 0$ and $k_u = k_i$, it has been shown that the voltage control yields all open-loop eigenvalues located at zero [33], which also represents an open-loop unstable system.

The advantage of voltage control over current control can be summarized as follows:

i) Higher overall system robustness since the plant model is more accurate.
ii) Weaker open-loop instability (no eigenvalue on the right half of complex plane).
iii) Simpler power amplifier structure (voltage-to-voltage).
iv) Possibility to benefit from "two-way" property of electromechanical transducers ("self-sensing bearing"[33]).

The main disadvantage is the increase in the order of the controlled plant, which becomes 12 (addition of 4 coil currents) in the case of voltage control instead of 8 as in the case of current control.

Lately, a certain trend back from current to voltage control is perceivable in AMB technology, which, as a matter of fact, is already state-of-the-art in modern electric motor control, a technology similar to that of AMB systems. The present trend in AMB technology is also facilitated by modern digital signal processors (DSPs) which provide all the peripherals necessary to directly generate the appropriate pulse width modulated (PWM) output voltage command signals for control of the bearing currents or bearing forces respectively. Voltage control with digitally generated PWM command signals also allows for implementing modern robust control techniques in order to reach a much more linear bearing behavior compared to conventional current control, even in the presence of large rotor displacements and large forces or bearing currents.

Uncertainty structure of the rotor/AMB system can be separated into two parts, comprised of a periodic (bearing spring stiffness) and a time-varying (rotor speed) perturbation acting on the nominal system as shown in the Figure 5.1. It can be induced from Figure 4.15 that if there is no uncertainty on the operation speed of the rotor, robust stability of the system is preserved with a maximum uncertainty of 25% on the bearing stiffness. On-line measurement of rotation speed and scheduling the controller with respect to operation speed (and angular acceleration) can provide robust stability for any speed, provided that a controller can be synthesized with nominal stability and

assuming that the performance of the closed-loop system is acceptable.

Unfortunately, it has been shown in [43] that output feedback control of LPV systems with partly measured parameters is a nonconvex problem (see Appendix B.3). However, casting the problem in the form of full-state feedback, leads to a convex problem. Hence measuring the state parameters given in (2.17) (if voltage amplifiers are used instead of current amplifiers as discussed above, coil currents must also be available for the controller) and feeding them back to the gain scheduled controller can yield to robust stability at any speed provided that the nominal stability is granted at the maximum speed of operation.

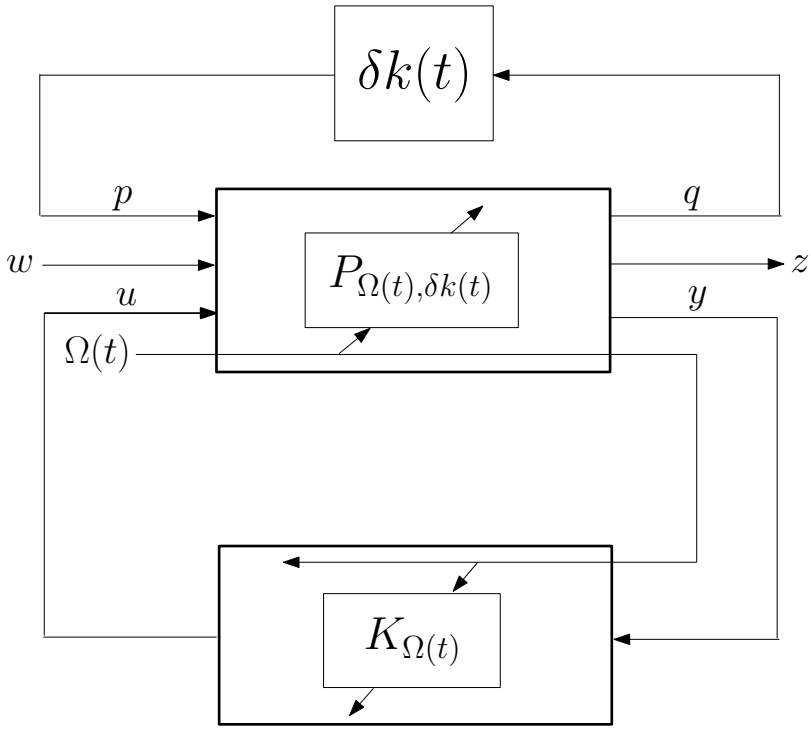

Figure 5.1. LPV model of rotor/AMB system with partly measured parameter structure.

APPENDIX A: ROTOR UNBALANCE COMPENSATION

One of the most important and unique features AMBs is the provision of active control of the system's response due to unbalance forces, a concept not possible with conventional ball, air or fluid film bearings. This additional control facility allows the rotor to either spin around its inertial axis through the mass center - provided that the air gap between the rotor and magnetic bearing is sufficiently large, which is most often the case in practice - or to compensate for the residual unbalance force so that the rotor is forced to rotate around its axis of geometry (i.e., center of rotation coincides with the center of geometry of the rotor).

Unbalance force acting on the rotor can be described as (in direction x)

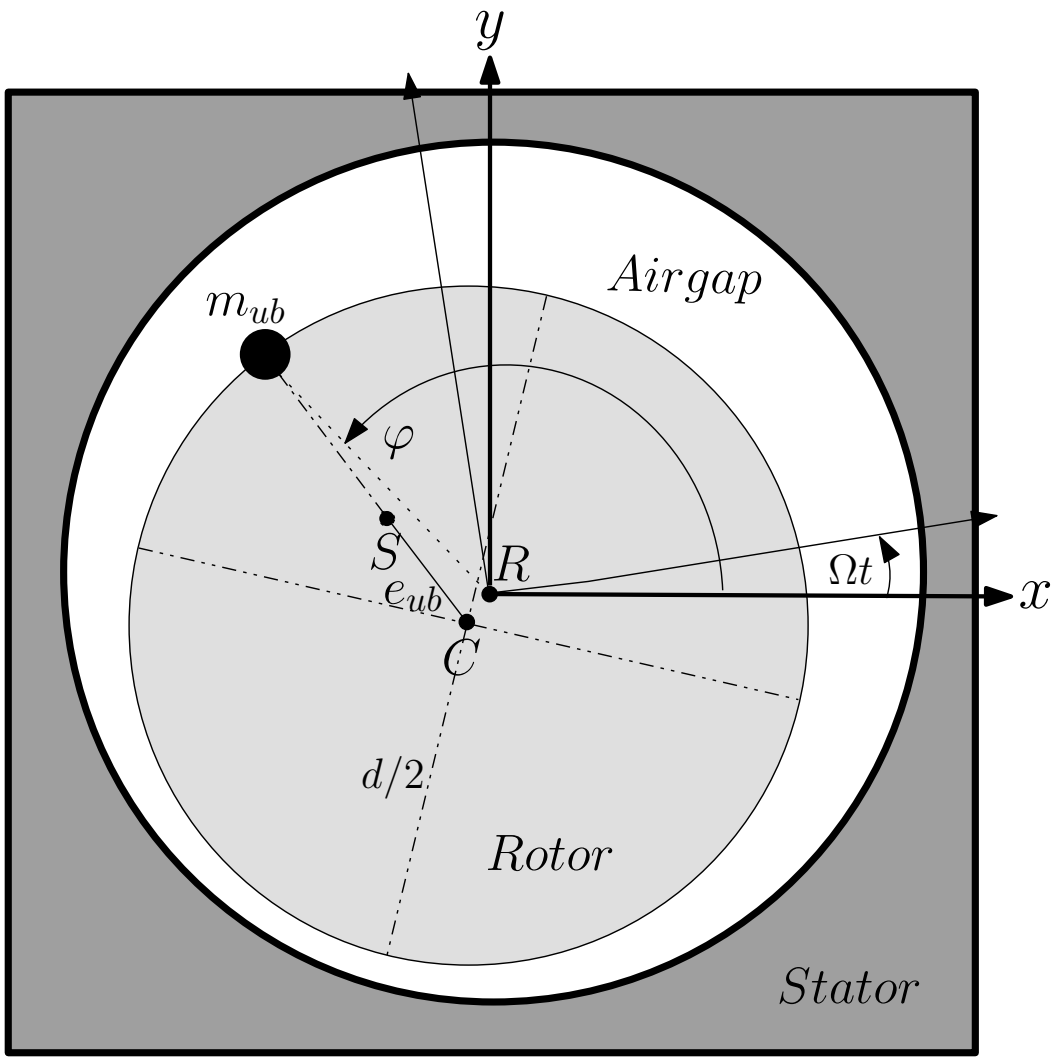

Figure A.1. Axes of rotation.

$$m_{ub}\,\Omega^2(d/2)\cos(\Omega t+\varphi) = M_r\,e_{ub}\,\Omega^2\cos(\Omega t+\varphi) \quad\Longrightarrow\quad e_{ub} = \frac{m_{ub}\,d}{2M_r}$$

Eccentricity (mass center displacement) e_{ub} can be tolerated up to 10 percent of the air gap within the linear operation of AMBs [25]. Assuming that the rotor system is symmetric, and therefore the translational motions are decoupled from angular ones, then the equations of motion for the rotor geometric center due to a static unbalance force can be described in stationary coordinates, with origin at R, as follows:

$$\begin{pmatrix}\ddot{x}_C\\ \ddot{y}_C\end{pmatrix} + \begin{bmatrix}\omega_n^2 & 0\\ 0 & \omega_n^2\end{bmatrix}\begin{pmatrix}x_C\\ y_C\end{pmatrix} = \begin{pmatrix}e_{ub}\,\Omega^2\cos(\Omega t+\varphi)\\ e_{ub}\,\Omega^2\sin(\Omega t+\varphi)\end{pmatrix}, \qquad \omega_n := \sqrt{k/M_r}$$

Note that the positive stiffness k in the equation is provided by the controller to suppress the negative spring stiffness of the open-loop system. Actually, controller also provides damping force through the bearings, but this is neglected for simplification.

If the assumed solutions

$$x_C(t) = X(\Omega)\cos(\Omega t+\varphi), \qquad y_C(t) = Y(\Omega)\sin(\Omega t+\varphi)$$

is introduced into the equations, solution for the whirl amplitudes can be obtained as

$$X(\Omega) = Y(\Omega) = e_{ub}\frac{(\Omega/\omega_n)^2}{1-(\Omega/\omega_n)^2}\,.$$

Geometric center C of the rotor moves in a forward whirl on a circular trajectory with the radius

$$r_C(\Omega) = \sqrt{x_C^2+y_C^2} = e_{ub}\frac{(\Omega/\omega_n)^2}{1-(\Omega/\omega_n)^2}\,.$$

The mass center S, too, moves on a circle with radius

$$r_S(\Omega) = \sqrt{x_S^2 + y_S^2} = \frac{e_{ub}}{1 - (\Omega/\omega_n)^2}$$

with

$$x_S(t) = x_C(t) + e_{ub}\cos(\Omega t + \varphi) = \frac{e_{ub}}{1 - (\Omega/\omega_n)^2}\cos(\Omega t + \varphi)\,,$$
$$y_S(t) = y_C(t) + e_{ub}\sin(\Omega t + \varphi) = \frac{e_{ub}}{1 - (\Omega/\omega_n)^2}\sin(\Omega t + \varphi)\,.$$

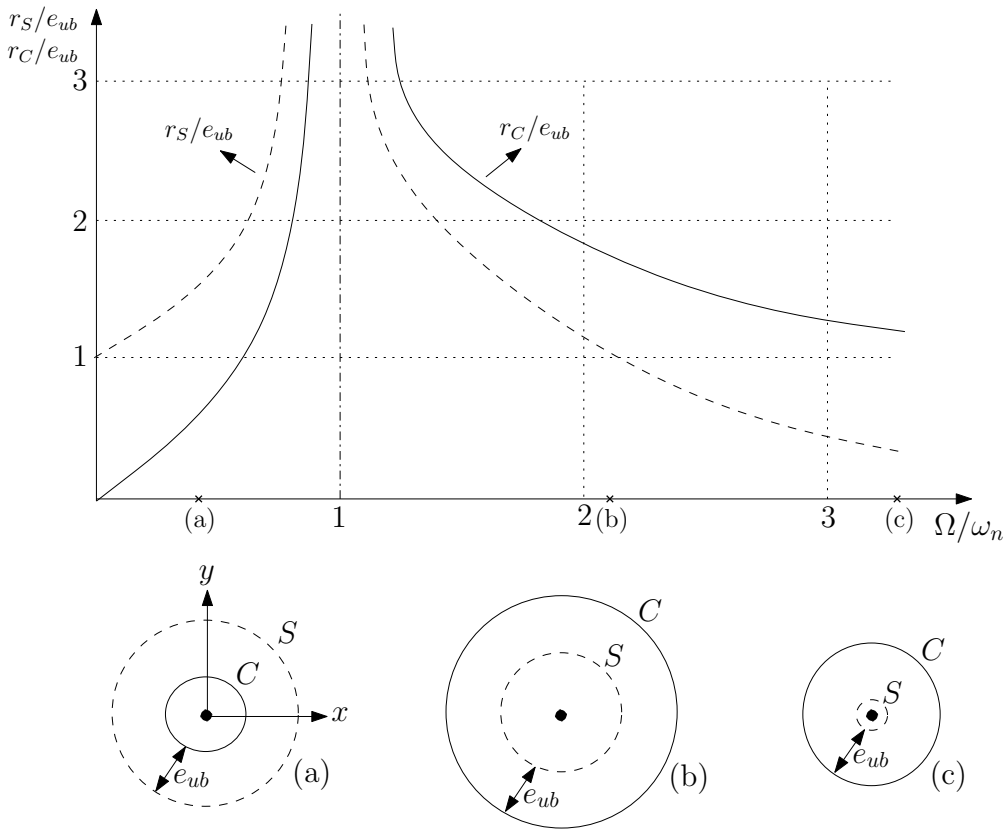

Figure A.2. Resonance curves and trajectories of C and S around R.

Figure (A.2) illustrates the switching from subcritical to supercritical range of rotor speed Ω. For low speeds (location (a) on the axis), S whirls on the outer trajectory, and at high speeds more than location (c), the rotor tends towards spinning about an axis through S, thus centering itself. The phase jump of 180° occurs while passing through the critical speed ($\Omega/\omega_n = 1$), and reflects the fact that impedance of rotor/AMB system to unbalance force shifts from stiffness to inertia.

Therefore, after passing through the first rigid mode critical speed, if the rotor is let to rotate freely around its axis of inertia, no unbalance force will be generated as rotor's mass is evenly distributed around its mass center S. Hence, using a method to filter out speed synchronous components from the control currents fed to AMBs can lead to "vibration free" operation of the stator. (Note that this is not possible with conventional bearings, such as hydrodynamic or even actively controlled hydrostatic bearings, as the rotor's radial motion due to whirl around its axis of inertia will always be resisted by the stiffness and damping force of the fluid film in the bearings.) Moreover, as the synchronous current outputs from the controller/amplifier will be redundant, magnetic saturation of the AMBs due to maximum current rate (or *force slew rate*) will be prevented.

Unbalance control and compensation methods in the literature feature the same physical input-output behavior when considering them as "black boxes". All methods can be understood and mathematically treated as *generalized notch filters*, since they all feature a very narrow band transfer characteristic $N_f(s)$ in order to generate the appropriate rotation synchronous injection I_1 or alternatively I_2 out of the available broadband sensor signal V_s of the AMB system.

The term "generalized notch filter" is motivated by the fact that, differing from a classical notch filter, the open-loop pole location p of such a filter $N_f(s)$ as described by the equation below can be allocated freely, which enables stabilization of the resulting closed-loop system, including unbalance control, over virtually the entire speed range, hence also within the rigid body critical speeds.

In the SISO case a typical representation of a generalized notch filter can be obtained by the following transfer function [35]:

$$N_f = \frac{s^2 + \Omega^2}{(s-p)(s-p^*)} \quad \text{with} \quad p = j\Omega + re^{j\phi}\,.$$

Note that p^* is the complex conjugate of p, and r, ϕ must yield closed loop stability.

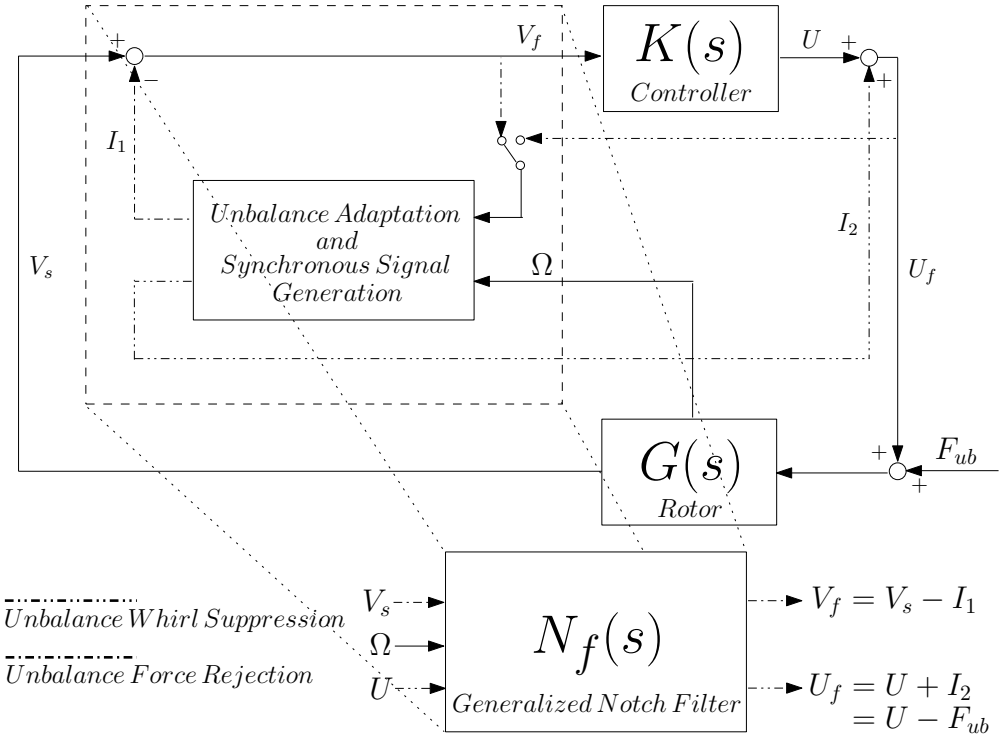

Figure A.3. Schematic diagram of unbalance control.

In the MIMO case the entire transfer function matrix $N_f(s)$ can be obtained by setting up the following block diagonal form

$$N_f = \mathrm{diag}(N_1 ... N_5)\,.$$

Although radial control of the magnetic bearings requires only four filters in a block diagonal implementation, it must be mentioned that an unbalance compensation scheme according to the form above makes it possible to address rotation synchronous signal components even in the axial control channel, a feature which can be very useful in the practice.

The injection point for the rotation synchronous compensation signal can be either at the control input I_1 or at its output I_2, without loss of generality. In *unbalance force rejection* schemes, which are most often applied in industrial rotor/AMB systems, signal injection at the controller input can have advantages, especially in the case of digital control with fixed-point processors [33]. In *unbalance whirl suppression* schemes harmonic signal injection usually takes place at the controller output.

In modern multivariable control techniques, such as H_2 and H_∞ control, whirl suppression can be a standard feature of the controller itself and no additional filter is required, as opposed to the applications with either decoupled or coupled PD/PID controllers.

APPENDIX B: MATHEMATICAL PRELIMINARIES

B.1. Norms on Linear Spaces

At the outset we set some basic concepts and notation. The real numbers are denoted by $\mathbb{R}$, and the complex numbers by $\mathbb{C}$; we use $j := \sqrt{-1}$ for the imaginary unit. Also, given a complex number $z = x + jy$ with $x, y \in \mathbb{R}$; $z^* = x - jy$ is defined as the *complex conjugate*, $|z| = \sqrt{x^2 + y^2}$ as the complex magnitude, and $x =$Re(z) as the real part. The field $\mathbb{F}$ is defined to be the real numbers $\mathbb{R}$,or the complex numbers $\mathbb{C}$.

Definition B.1 *Suppose $\mathcal{V}$ is a nonempty set and $\mathbb{F}$ is a field, and that operations of addition among its elements and multiplication by a scalar are defined in the following way:*
(a) For every pair $u, v \in \mathcal{V}$, a unique element $u + v \in \mathcal{V}$ is assigned called their sum,
(b) for each $\alpha \in \mathbb{F}$ and $v \in \mathcal{V}$ there is a unique element $\alpha v \in \mathcal{V}$ called their product.
Then $\mathcal{V}$ is a linear space (or a vector space) if the following properties hold for all u, v, $w \in \mathcal{V}$, and for all α, $\beta \in \mathbb{F}$:

i) There exists a zero element in $\mathcal{V}$, denoted by 0, such that $v + 0 = v$,
ii) there exists en element $-v$ in $\mathcal{V}$, such that $v + (-v) = 0$,
iii) the association $u + (v + w) = (u + v) + w$ is satisfied,
iv) the commutativity relationship $u + v = v + u$ holds,
v) scalar distributivity $\alpha(u + v) = \alpha u + \alpha v$ holds,
vi) element distributivity $(\alpha + \beta)v = \alpha v + \beta v$ is satisfied,
vii) the associative rule $(\alpha\beta)v = \alpha(\beta v)$ for scalar multiplication holds,
viii) for the unit scalar $1 \in \mathbb{F}$ the equality $1v = v$ holds.

Notice that both $\mathcal{V}$ and $\mathbb{F}$ contain the zero element, which is denoted by "0"regardless of the instance. Given two linear spaces $\mathcal{V}_1$ and $\mathcal{V}_2$, with the same associated scalar field, we use $\mathcal{V}_1 \times \mathcal{V}_2$ to denote the linear space formed by their Cartesian product. Thus every element of $\mathcal{V}_1 \times \mathcal{V}_2$ is of the form (v_1, v_2), where $v_1 \in \mathcal{V}_1$ and $v_2 \in \mathcal{V}_2$.

A *subspace* of a linear space $\mathcal{V}$ is a subset of $\mathcal{V}$ which is also a linear space with respect to the same field and operations; equivalently, it is a subset which is closed under the operations on $\mathcal{V}$. Given elements $v_1, ..., v_m$ in a linear space we denote their *span* by $\text{span}\{v_1, ..., v_m\}$, which is the set of all elements v that can be written as; $v = \alpha_1 v_1 + ... + \alpha_m v_m$ for some scalars $\alpha_k \in \mathbb{F}$. The span always defines a subspace.

If for some elements we have; $\text{span}\{v_1, ..., v_m\} = \mathcal{V}$, we say that linear space $\mathcal{V}$ is *finite dimensional*. If no such finite set of elements exists we say the linear space is *infinite dimensional*. Finite dimensional spaces are comprised of vectors and matrices. Infinite dimensional spaces are composed of functions (of time or of time/space) such as signals and transfer matrices. Note that digital signals in the form of; $v[k]$ with $k = 0, 1, ..., n$ are finite dimensional.

If a linear space $\mathcal{V}$ is finite dimensional we define its *dimension*, $\dim(\mathcal{V})$, to be the smallest number n such that there exists elements $v_1, ..., v_n$ satisfying $\text{span}\{v_1, ..., v_n\} = \mathcal{V}$. In that case we say that the set $\{v_1, ..., v_n\}$ is a *basis* for $\mathcal{V}$.

Definition B.2 *Norm is a real number which gives an overall measure of the size of a vector, a matrix or a function. It is denoted either by $|\cdot|_{\mathcal{V}}$ (for finite-dimensional spaces) or $\|\cdot\|_{\mathcal{V}}$ (for infinite-dimensional spaces), and it satisfies the following properties* (though the notation used is for infinite dimensional case, the required properties are for both finite and infinite dimensional spaces):

(i) $\|v\|_{\mathcal{V}} \geq 0$ *for all* $v \in \mathcal{V}$,
(ii) $\|v\|_{\mathcal{V}} = 0$, *if and only if,* $v = 0$,
(iii) $\|\alpha \cdot v\|_{\mathcal{V}} = |\alpha| \cdot \|v\|_{\mathcal{V}}$ *for all scalars* $\alpha \in \mathbb{F}$,
(iv) $\|u + v\|_{\mathcal{V}} \leq \|u\|_{\mathcal{V}} + \|v\|_{\mathcal{V}}$ *for all* u *and* $v \in \mathcal{V}$.

More precisely, if v is an element in a linear space $\mathcal{V}$ over the field $\mathbb{F}$, than the properties above must be satisfied $\forall v \in V$ and $\forall \alpha \in \mathbb{F}$, to admit $\|\cdot\|_{\mathcal{V}}$ as a norm on v. A linear space together with a norm is called a *normed space* and is denoted $(\mathcal{V}, \|\cdot\|_{\mathcal{V}})$.

Every normed space is a *metric space*, in which notions such as *convergence* can be defined. A sequence $\{v(k)\} \in \mathcal{V}$ is said to be convergent if there exists a $v \in \mathcal{V}$ such $\|v(k) - v(l)\| \to 0$ when $k, l \to \infty$. Other standard concepts of interest include compactness and total boundedness. Let $\mathcal{S}$ be a subset in the normed space $\mathcal{V}$. For any $\delta > 0$, a subset $\mathcal{S}_\delta$ is said to be a δ-net of $\mathcal{S}$ if $\mathcal{S}_\delta$ is finite and for each $v_1 \in \mathcal{S}$ there is a $v_2 \in \mathcal{S}$ such that $\|v_1 - v_2\| \leq \delta$. If for each $\delta > 0$, $\mathcal{S}$ contains a δ-net, $\mathcal{S}$ is said to be *totally bounded*. A sequence $\{v(k)\}$ is called a *Cauchy sequence* if for any $\epsilon > 0$, there exists an integer $n > 0$ such that

$$\|v(k) - v(l)\| \leq \epsilon\,, \;\; \forall\, k, l \geq n.$$

Every convergent sequence is a Cauchy sequence. However, the converse need not to be true. The elements in a Cauchy sequence eventually cluster around each other, so that they are "trying to converge". If every Cauchy sequence in $\mathcal{V}$ is convergent, then $\mathcal{V}$ is said to be *complete*. This is the case with Banach spaces. A *Banach space* is complete normed space. A set S in a Banach space $\mathcal{V}$ forms a subspace if

$$v_1, v_2 \in \mathcal{S} \Rightarrow v_1 + v_2 \in \mathcal{S}\,, \text{ and } v \in \mathcal{S},\, \alpha \in \mathbb{C} \Rightarrow \alpha v \in \mathcal{S}.$$

A subspace $\mathcal{S}$ is *closed* if every convergent sequence in $\mathcal{S}$ has a limit in $\mathcal{S}$. For a finite dimensional space $\mathcal{V}$, every subspace $\mathcal{S} \subseteq \mathcal{V}$ is closed. In general, however, a subspace need not be closed.

Definition B.3 *An inner product, denoted $\langle \cdot\,, \cdot \rangle_{\mathcal{V}}$, on a vector space $\mathcal{V}$ is a function mapping $\mathcal{V} \times \mathcal{V} \to \mathbb{F}$ so that:*

(i) the inner product $\langle v\,, v \rangle_{\mathcal{V}} \geq 0$ for all $v \in \mathcal{V}$,

(ii) $\langle v\,, v \rangle_{\mathcal{V}} = 0$, if and only if, $v = 0$,

(iii) if $v \in \mathcal{V}$, then $\langle v\,, \alpha_1 u_1 + \alpha_2 u_2 \rangle = \alpha_1 \langle v\,, u_1 \rangle + \alpha_2 \langle v\,, u_2 \rangle$ for all $u_i \in \mathcal{V}$ and scalars α_i ; i.e., the mapping $u \to \langle v\,, u \rangle$ is linear on $\mathcal{V}$,

(iv) $\langle u\,, v \rangle_{\mathcal{V}}$ is the complex conjugate of $\langle v\,, u \rangle_{\mathcal{V}}$ for all u and $v \in \mathcal{V}$.

Thus, the properties of an inner product are more restrictive than those of norm. The standard example of an inner product space, from which much intuition about inner product spaces is gained, is *Euclidean space.* That is, $\mathbb{R}^n$ or $\mathbb{C}^n$ with the inner product

$$\langle x\,,y\rangle := x^*y = x_1^*y_1 + ... + x_n^*y_n\,,$$

where $*$ denotes complex conjugate transpose.

Geometrically the inner product captures the idea of angle between two elements of $\mathcal{V}$ in the same way that the so-called "dot product" does in Euclidean space. A vector space $\mathcal{V}$ together with an inner product $\langle\cdot\,,\cdot\rangle_{\mathcal{V}}$ is called an *inner product space.* A complete inner product space is called a *Hilbert space.*

Theorem B.4 [37] (Cauchy-Schwarz inequality) *If $\langle\cdot\,,\cdot\rangle_{\mathcal{V}}$ is an inner product on a linear space $\mathcal{V}$ over the field $\mathbb{F}$, then*;

$$|\langle u\,,v\rangle|^2 \leq \langle u\,,u\rangle\langle v\,,v\rangle \quad \text{for all } \; u,v \in \mathcal{V}\,.$$

Equality occurs, if and only if, u and v are linearly dependent, that is, $u = \alpha v$ or $v = \alpha u$ for some $\alpha \in \mathbb{F}$.

Proof: Let $u, v \in \mathcal{V}$ be given. If $v = 0$, the assertion is trivial, so we may assume that $v \neq 0$. Let $t \in \mathbb{R}$ and consider

$$\begin{aligned} p(t) := \langle u+tv\,,u+tv\rangle &= \langle u\,,u\rangle + t\langle v\,,u\rangle + t\langle u\,,v\rangle + t^2\langle v\,,v\rangle \\ &= \langle u\,,u\rangle + 2t\,\mathrm{Re}\langle u\,,v\rangle + t^2\langle v\,,v\rangle\,, \end{aligned}$$

which is a real quadratic polynomial with real coefficients. Because of item *(i)* in Definition B.3, we know that $p(t) \geq 0$ for all real t, and hence $p(t)$ can have no real simple roots. The discriminant of $p(t)$ must therefore be nonpositive;
$(2\,\mathrm{Re}\langle u\,,v\rangle)^2 - 4\langle v\,,v\rangle\langle u\,,u\rangle \leq 0$, and hence; $\mathrm{Re}\langle u\,,v\rangle)^2 \leq \langle v\,,v\rangle\langle u\,,u\rangle$.

Since this inequality must hold for any pair of $u, v \in \mathcal{V}$, it must hold if v is replaced by $\langle u\,, v\rangle v$, so we also have the inequality; $(\mathrm{Re}\langle u\,, \langle u\,, v\rangle v\rangle)^2 \leq \langle u\,, u\rangle\langle v\,, v\rangle|\langle u\,, v\rangle|^2$.
However, $\mathrm{Re}\langle u\,, \langle u\,, v\rangle v\rangle = \mathrm{Re}\overline{\langle u\,, v\rangle}\langle u\,, v\rangle = \mathrm{Re}|\langle u\,, v\rangle|^2 = |\langle u\,, v\rangle|^2$.
Therefore, $|\langle u\,, v\rangle|^4 = \langle u\,, u\rangle\langle v\,, v\rangle|\langle u\,, v\rangle|^2$. If $\langle u\,, v\rangle = 0$, then the statement of the theorem is trivial. If not, then we may divide the last equality by $|\langle u\,, v\rangle|^2$ to obtain the desired inequality of the theorem.
Because of item *(ii)* of Definition B.3, $p(t)$ can have a real (double) root only if $u+tv = 0$ for some t. Thus, equality can occur in the discriminant condition, if and only if, u and v are linearly dependent. $\square$

Furthermore, it can be verified by using the Cauchy-Schwarz inequality that

$$\|v\| = \sqrt{\langle v\,, v\rangle}$$

satisfies the properties of norm.

We say that two elements u, v in $\mathcal{V}$ are *orthogonal* if $\langle u\,, v\rangle = 0$. The notation $u \perp v$ is sometimes used to indicate this relationship.

In the sequel, we consider the norms of four different objects:

i) v is a constant vector.
ii) v is a constant matrix, which is denoted as M.
iii) v is a time dependent signal $v(t)$, which at each fixed t, is a constant scalar or vector.
iv) v is a system; a transfer (matrix) function $G(s)$ or impulse response $g(t)$, which at each fixed s or t, is a constant scalar or matrix.

Cases 1 and 2 involve *spatial norms* and cases 3 and 4 involve *function norms* or *temporal norms*. Notice that the first two are finite-dimensional norms and denoted by $|\cdot|$, while the latter two are infinite dimensional and denoted by $\|\cdot\|$.

Consider a vector v with n elements; that is the linear space $\mathcal{V} \in \mathbb{F}^n$. We will consider three norms which are special cases of the vector p-norm (or *Hölder norm*)

$$|v|_p := \left(\sum_i |v_i|^p \right)^{1/p} .$$

For example, vector 1-norm is defined as

$$|v|_1 = \sum_i |v_i| .$$

A commonly used vector norm is *Euclidean norm* or 2-norm, which is the shortest distance between two points

$$|v|_2 = \sqrt{\sum_i |v_i|^2} .$$

Another example is vector ∞-norm, which is the largest-element magnitude in the vector

$$|v|_\infty = |v|_{max} = \max_i |v_i| .$$

Notice that p-norm is defined for $1 \leq p \leq \infty$. If p is less than 1, property *(d)* (which is the *triangle inequality*) is violated.

Note also that, for $1 \leq p \leq q \leq \infty$ and $v \in \mathbb{F}^n$

$$|v|_\infty \leq |v_q| \leq |v_p| \leq |v_1| .$$

Equality holds if v is a scalar. Norm of a scalar $v \in \mathbb{F}$ is known as its *magnitude* or its *absolute value*.

Since various vector norms only differ by constant factors, they are often said to be equivalent. It is a well-known fact that $\mathbb{R}^n$ and $\mathbb{C}^n$ are both Banach spaces with any

norm; this can be easily observed in the case of p-norm as shown above, and hence the statement follows in general because all norms are equivalent for a finite dimensional space.

A norm on a matrix $|M|$ is a *matrix norm* if, in addition to the four norm properties in the definition of the norm, it also satisfies the *submultiplicative property*

$$|MN| \leq |M| \cdot |N| \,.$$

This property is very important when combining systems, and forms the basis for the *small gain theorem.* The norms on matrices that do not satisfy the submultiplicative property are called *generalized matrix norms.*
Three matrix norm examples which are the direct extensions of vector p-norms are:

Sum matrix norm, which is the sum of element magnitudes

$$|M|_1 = |M|_{sum} := \sum_{i,j} |m_{i,j}| \,.$$

Frobenius matrix norm, which is the square root of the sum of the squared element magnitudes:

$$|M|_2 = |M|_F := \sqrt{\sum_{i,j} |m_{i,j}|^2} = \sqrt{\langle M, M \rangle} = \sqrt{\mathrm{trace}(M^* M)} \,,$$

where the operator *trace* is the sum of the diagonal elements, and M^* is the complex conjugate transpose or adjoint of M (see Appendix B.2).

Maximum element norm, which is the largest-element magnitude

$$|M|_\infty = |M|_{max} := \max_{i,j} |m_{i,j}| \,.$$

This norm is not a matrix norm as it does not satisfy the submultiplicative property.

However, $\sqrt{kl}\,|M|_{max}$ is a matrix norm [43], where k and l are the number of rows and columns of M respectively.

Induced matrix norms are important because of their close relationship to signal amplification in systems. Considering the equation $z = Mv$, we may think of v as the input vector and z as the output vector. At a certain time or frequency the amplification or gain of the matrix M is given by the ratio $|z|/|v|$.

Definition B.5 *Suppose $\mathcal{V}$ and $\mathcal{Z}$ are Banach spaces. A mapping, M, from $\mathcal{V}$ to $\mathcal{Z}$ is called a linear, bounded operator if*
i) $M(\alpha_1 v_1 + \alpha_2 v_2) = \alpha_1 M(v_1) + \alpha_2 M(v_2)$,
ii) *there exists a scalar $\kappa \geq 0$, such that $|Mv|_{\mathcal{Z}} \leq \kappa \cdot |v|_{\mathcal{V}}$ for all $v \in \mathcal{V}$.*

The space of all linear bounded operators mapping $\mathcal{V}$ to $\mathcal{Z}$ is denoted by; $\mathcal{L}(\mathcal{V}, \mathcal{Z})$. We define the the *induced matrix norm* on this space by

$$|M|_{\mathcal{V}\to\mathcal{Z}} := \max_{v\in\mathcal{V},\, v\neq 0} \frac{|Mv|_{\mathcal{Z}}}{|v|_{\mathcal{V}}}$$

and it can be verified that it satisfies the properties of a norm. Notice that $|M|_{\mathcal{V}\to\mathcal{Z}}$ is the smallest number κ that satisfies the boundedness of the operator M as given in the definition above. It is possible to show that $\mathcal{L}(\mathcal{V}, \mathcal{Z})$ is a complete space whenever $\mathcal{Z}$ is complete [42].

In practical terms, we are looking for a direction of the vector v such that the ratio $|z|_p/|v|_p$ is maximized. Thus, the induced norm gives the largest possible "amplifying power" of the matrix. The following equivalent definition is also used:

$$|M|_{i,p} := \max_{v\neq 0} \frac{|Mv|_p}{|v|_p} = \max_{v=1} |Mv|_p \,.$$

For the induced 1-, 2- and ∞-norms the following identities hold:

$$|M|_{i,1} = \max_j \left(\sum_i |m_{ij}| \right) \quad \text{“maximum row sum”} ,$$

$$|M|_{i,2} = \overline{\sigma}\,(M) = \sqrt{\varrho(M^*M)} \quad \text{“maximum singular value or spectral norm”} ,$$

$$|M|_{i,\infty} = \max_i \left(\sum_j |m_{ij}| \right) \quad \text{“maximum column sum”} .$$

where the spectral radius $\varrho(M) = \max\{\, |\lambda_i| : \lambda_i = \text{eigenvalues of M}\,\}$. Induced 2-norm of a matrix is equal to its maximum singular value, and is often called *spectral norm.*

Proposition B.6 *All induced norms $|M|_{i,p}$ are matrix norms, therefore they satisfy the submultiplicative property*

$$|MN|_{i,p} \leq |M|_{i,p} \cdot |N|_{i,p} .$$

Proof: Consider the set of equations; $z = Mv\,,\ v = Nw \ \Rightarrow\ z = MNw\,.$
From the definition of the induced norm, we get by first introducing $v = Nw$, then by multiplying the numerator and denominator by $|v|_p \neq 0$, and finally maximizing each term involving w and v independently, that

$$|MN|_{i,p} := \max_{w\neq 0} \frac{|MNw|_p}{|w|_p} = \max_{v,w\neq 0} \left(\frac{|Mv|_p}{|v|_p} \cdot \frac{|Nw|_p}{|w|_p} \right) \leq \max_{v\neq 0} \frac{|Mv|_p}{|v|_p} \cdot \max_{w\neq 0} \frac{|Nw|_p}{|w|_p}$$

and the proposition follows from the definition of the induced matrix norm. $\square$

For induced matrix norms the submultiplicative property $|MN| \leq |M| \cdot |N|$ holds for matrices M and N of any dimension as long as the product MN exists. In particular, it holds if we choose M and N as vectors. Thus, choosing N to be a vector, i.e., $N = w$, for any matrix norm we have

$$|Mv| \leq |M| \cdot |v| .$$

We say that; "matrix norm $|M|$ is compatible with its corresponding vector norm $|v|$". Similarly, for the Frobenius norm is compatible with the vector 2-norm, since when v is a vector $|v|_F = |v|_2$.

From the inequality above we also get for any matrix norm that

$$|M|_p \geq \max_{v \neq 0} \frac{|Mv|_p}{|v|_p} .$$

Note that, the induced matrix norms are defined such that the equality holds in the inequality above. The property $|M|_F \geq \overline{\sigma}(M)$ then follows since $|v|_F = |v|_2$. An important property of the Frobenius norm and the spectral norm (induced 2-norm) is that they are invariant with respect to unitary transformations [37], i.e., for unitary matrices U_i, satisfying $U_i^* U_i = I$. Therefore, we have

$$|U_1 M U_2|_F = |M|_F ,$$
$$\overline{\sigma}(U_1 M U_2) = \overline{\sigma}(M) .$$

From *singular value decomposition* (see Appendix B.2) of the matrix $M = U \Sigma V^*$ we then obtain an important relationship between the Frobenius norm and the singular values, $\sigma_i(M)$, in the form of

$$|M|_F = \sqrt{\sum_i \sigma_i^2(M)} .$$

Given a function v mapping $\mathbb{R}^+$ to $\mathbb{C}^n$. We say $v \in L_p^n[0, \infty)$, if the following definitions are valid for $t \geq 0$:

$$\|v\|_p := \left(\int_0^\infty |v(t)|^p \, dt \right)^{1/p} < \infty$$
$$\|v\|_\infty := \text{ess} \sup_{t \in \mathbb{R}} |v(t)| < \infty$$

where $1 \leq p < \infty$. The superscript n is called the *spatial dimension* of L_p^n. In most

cases, we drop the spatial dimension and simply use the notation L_p.

In case v is a digital signal, it can be represented by; $v[k],\ k = 0, 1, ..., n$.
Notice that digital signals are finite dimensional, hence their magnitudes are evaluated within the scope of vector norms. Often, the space related with a digital signal norm is denoted by ℓ_p.

The integral involved in the definition of $\|v\|_p$ is a *Lebesque integral* not a *Riemann integral*. That is, function values on a *measure zero* do not contribute to the integral. A subset $\mathcal{S}$ of $\mathbb{R}^n$ is said to have measure zero if for each $\epsilon > 0$ there is a countable union of balls B_1, B_2,... of volume ϵ_i such that

$$\sum_{i=1}^{\infty} \epsilon_i < \epsilon\,, \quad \text{and} \quad \mathcal{S} \subseteq \bigcup_{i=1}^{\infty} B_i\,.$$

Two functions $f, g\,:\,O \to T$ from some subset $O \subseteq \mathbb{R}^n$ into some other set T are said to be *equal almost everywhere* if the set

$$\{x \in O \,|\, f(x) \neq g(x)\}$$

has zero measure. In general, a property is said to hold almost everywhere provided that it only fails in a set of measure zero. In fact, the "essential supremum" means the supremum except for a set of measure zero.

All of the L_p spaces defined are complete and, in particular, L_2 is an example of Hilbert space. Obviously, if $v \in L_p$, we define $\|v\|_p$ as its L_p norm.
For $p = 2$ and $n = 1$, the L_2 norm or *energy norm* of v is defined as

$$\|v\|_2 := \sqrt{\int_0^{\infty} v(t)^2\, dt}\ .$$

Also the L_2-RMS norm or *power norm* of the signal $v(t)$ is defined as

$$\|v\|_{RMS} := \lim_{T\to\infty} \sqrt{\frac{1}{T}\int_0^T x(t)^2\, dt}\ .$$

Notice that RMS norm of a signal $v(t)$ may not satisfy the property *(b)* given in the definition of the concept of norm. Thus, RMS norm of the signals is accepted as a semi-norm. Note also that, in order to $v \in L_p$ for $1 \leq p < \infty$, it must satisfy; $\lim_{t\to\infty} v(t) = 0$.

If $v(t)$ is vector valued, i.e., $v : \mathbb{R}^+ \to \mathbb{C}^n,\ n > 1$ we define

$$\|v\|_2 := \sqrt{\int_0^\infty v(t)^T v(t)\, dt}\ .$$

For $T = [0, \infty)$, the set of all such functions bounded in the L_p norm forms a linear space, the space $L_p(T)$. In general, $L_p(T)$ is a Banach space, and $L_p(T) \subseteq L_q(T)$ whenever $1 \leq p \leq q$. Moreover, one can distinguish between $L_2[0, \infty)$ the future, $L_2(-\infty, 0]$ the past or $L_2(-\infty, \infty)$ all times, as the domain of the L_2 space.

Signal and system norms also have interpretations in the frequency domain. The *complex inner product space* $\hat{L}_2(j\mathbb{R})$, which consists of functions mapping the imaginary axis $j\mathbb{R}$ to $\mathbb{C}^n$ is defined with the inner product

$$\langle \hat{u}\,, \hat{v}\rangle_{\hat{L}_2} := \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{u}(j\omega)^* \hat{v}(j\omega)\, d\omega\,.$$

Thus a function $\hat{v} : j\mathbb{R}$ to $\mathbb{C}^n$ is in $\hat{L}_2(j\mathbb{R})$ if $\langle \hat{v}\,, \hat{v}\rangle_{\hat{L}_2} = \|v\|_2^2 < \infty$. Here we use the same notation for the norm and inner product of $\hat{L}_2$ in frequency domain, as we did for L_2 in time domain.

The *Fourier transform* of a function $v : \mathbb{R} \to \mathbb{C}^n$ is defined to be

$$\hat{v}(j\omega) := \lim_{T\to\infty} \int_{-T}^{T} v(t)\, e^{-j\omega t}\, dt\,.$$

It can be shown [40] that; when $v(t) \in L_2(-\infty, \infty)$ this limit exists for almost all ω. For convenience we denote this operation by the map Θ, and write simply $\hat{v} = \Theta v$ to indicate the above relationship.

Given a function $\hat{v} : j\mathbb{R} \to \mathbb{C}^n$, we define the *inverse Fourier transform* of $\hat{v}$ by

$$v(t) := \lim_{T \to \infty} \frac{1}{2\pi} \int_{-T}^{T} \hat{v}(j\omega)\, e^{j\omega t}\, d\omega\,,$$

with an analogous convention regarding the convergence of the integral, and we use $v = \Theta^{-1}\hat{v}$ to indicate this transformation.

As the notation would indicate, for certain classes of functions these maps are inverses of each other. For our purposes, it suffices to note that for $u \in L_2(-\infty, \infty)$, and $\hat{v} \in \hat{L}_2(j\mathbb{R})$, we have

$$u(t) = \Theta^{-1}(\Theta u)(t) \quad \text{for almost every t,}$$
$$\hat{v}(j\omega) = \Theta(\Theta^{-1}\hat{v})(j\omega) \quad \text{for almost every } \omega.$$

"For almost" means: "For all except for a set of measure zero". Furthermore, we have the following result known as the *Plancherel theorem.*

Theorem B.7 [36] *With the Fourier transform and its inverse defined as above;*

i) The map $\Theta :\ L_2(-\infty, \infty) \to \hat{L}_2(j\mathbb{R})$*, and given any* $u, v \in L_2(-\infty, \infty)$ *the equality;* $\langle u\,, v\rangle_2 = \langle \Theta u\,, \Theta v\rangle_2$ *holds.*

ii) The map $\Theta^{-1} : \hat{L}_2(j\mathbb{R}) \to L_2(-\infty, \infty)$*, and if* $\hat{u}, \hat{v} \in \hat{L}_2(j\mathbb{R})$*, then the equality;* $\langle \hat{u}\,, \hat{v}\rangle_2 = \langle \Theta\hat{u}\,, \Theta\hat{v}\rangle_2$ *is satisfied.*

This theorem says that the Fourier transform is an invertible map between the spaces $L_2(-\infty, \infty)$ and $\hat{L}_2(j\mathbb{R})$, and more importantly, this map preserves the inner product. Hence Fourier transform is a unitary operator which makes the two spaces

isomorphic. In particular, for any $\hat{v} \in \hat{L}_2(j\mathbb{R})$ and $u \in L_2(-\infty, \infty)$, we have: $\|\Theta u\|_2 = \|u\|_2$ and $\|\Theta^{-1}\hat{v}\|_2 = \|\hat{v}\|_2$; namely, norms are preserved under this mapping.

Recall the definition of $L_2[0, \infty)$, the Hilbert space of functions in $L_2(-\infty, \infty)$ which are zero in $L_2(-\infty, 0)$. Given $v \in L_2[0, \infty)$, we define its *Laplace transform* by the integral

$$\hat{v}(s) := \lim_{T \to \infty} \int_0^T v(t)\, e^{-st}\, dt = \int_0^\infty v(t)\, e^{-st}\, dt\,,$$

when the limit exists, and set $\hat{v}(s) = 0$ at the divergent values of s. However, the integral converges absolutely when $\mathrm{Re}(s) > 0$ [41]. We use the notation; $\hat{v} = \Lambda v$ to indicate this transformation. If we evaluate the transform at $s = j\omega$ for $\omega \in \mathbb{R}$, since $v(t) \in L_2[0, \infty)$ we get

$$\hat{v}(jw) = \int_0^\infty v(t)\, e^{-j\omega t}\, dt = \int_{-\infty}^\infty v(t)\, e^{-j\omega t}\, dt\,,$$

which is the Fourier transform of $v(t)$. Notice that the notation $\hat{v}(\cdot)$ refers to the Laplace transform or the Fourier restriction depending on its argument.

Note also that by the Plancherel theorem $\hat{v}(j\omega) \in \hat{L}_2(j\mathbb{R})$ and $\|\hat{v}\|_2 = \|v\|_2$, and this can be proved by the following useful corollary of Plancherel theorem, which is known as *Parseval's theorem*.

Theorem B.8 [44] *Given* $u, v \in L_2[0, \infty)$ *, and* $\hat{u}, \hat{v} \in \hat{L}_2(j\mathbb{R})$ *. Then*:

$$\langle u\,, v\rangle_{L_2} = \int_0^\infty u(t)^T v(t)\, dt = \langle u\,, v\rangle_{\hat{L}_2} = \frac{1}{2\pi} \int_{-\infty}^\infty \hat{u}(j\omega)^* \hat{v}(j\omega)\, d\omega$$

Proof: Since the inverse Fourier transform of $\hat{v}(j\omega)$ is given by

$$v(t) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{v}(j\omega)e^{j\omega t}d\omega\,,$$

we have the following:

$$\begin{aligned}
\langle u\,,v\rangle_{L_2} &= \int_0^{\infty} u(t)^T v(t)\,dt \\
&= \int_0^{\infty} u(t)^T \left(\frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{v}(j\omega)e^{j\omega t}d\omega\right)\,dt \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} \left(\int_0^{\infty} u(t)e^{j\omega t}dt\right)^T \hat{v}(j\omega)\,d\omega \\
&= \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{u}(-j\omega)^T\hat{v}(j\omega)\,d\omega = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{u}(j\omega)^*\hat{v}(j\omega)\,d\omega\,.
\end{aligned}$$

□

Let $S \subset \mathbb{C}$ be an open set, and let $\hat{v}(s)$ be a complex-valued function defined on S as: $\hat{v}(s)\ :\ S \to \mathbb{C}$. Then the function $\hat{v}(s)$ is said to be *analytic* in S if it differentiable at every point z in S, and also at each point in some neighborhood of every point z in S. A matrix-valued function is analytic in S, if every element of the matrix is analytic in S. A well-known property of the analytic functions is the so-called *maximum modulus theorem.*

Theorem B.9 [26] *If $\hat{v}(s)$ is defined and continuous on a closed-bounded set $\mathcal{S}$ and analytic in the interior of $\mathcal{S}$, then $|\hat{v}(s)|$ can not attain the maximum in the interior of $\mathcal{S}$ unless $\hat{v}(s)$ is a constant.*

The theorem implies that $|\hat{v}(s)|$ can only achieve its maximum on the boundary of $\mathcal{S}$. That is; $\max_{s\in\mathcal{S}}|\hat{v}(s)| = \max_{s\in\partial\mathcal{S}}|\hat{v}(s)|$, where $\partial\mathcal{S}$ denotes the boundary of $\mathcal{S}$. Next, we consider some frequently used complex function spaces in control theory.

Definition B.10 *A function $\hat{v} : \overline{\mathbb{C}}^+ \to \mathbb{C}^n$ is in Hardy space H_2 if:*

i) $\hat{v}(s)$ is analytic in the open right half plane $\mathbb{C}^+$,

ii) for almost every real number ω *;* $\lim_{\sigma\to 0^+} \hat{v}(\sigma + j\omega) = \hat{v}(j\omega)$,

iii) $\sup_{\sigma\geq 0} \int_{-\infty}^{\infty} |\hat{v}(\sigma + j\omega)|_2^2 \, d\omega < \infty$.

We remark that, according to maximum modulus theorem, the supremum in *(c)* is always achieved at $\sigma = 0$.

The key result we now state is that Laplace transform of a function $L_2[0, \infty)$ satisfies these properties, and furthermore every function in H_2 can be obtained from this form:

Theorem B.11 [36]

i) If $v \in L_2[0, \infty)$*, then* $\Lambda v \in H_2$.

ii) If $\hat{v} \in H_2$*, then there exists* $v \in L_2[0, \infty)$ *satisfying* $\Lambda v = \hat{v}$.

A useful consequence of Theorem B.11 is the following:

Corollary B.12 *Let* $\hat{u}$ *and* $\hat{v}$ *be the functions in* H_2*, and* $u(j\omega) = v(j\omega)$ *for all* $\omega \in \mathbb{R}$*. Then;* $\hat{u}(s) = \hat{v}(s)$ *for all* $s \in \overline{\mathbb{C}}^+$.

Proof: From Theorem B.11 we find $u(t)$ and $v(t)$ in $L_2[0, \infty)$, such that; $\hat{u} = \Lambda u$, $\hat{v} = \Lambda v$. Then $\hat{u} - \hat{v} = \Lambda(u - v)$. Since $\hat{u}(j\omega) - \hat{v}(j\omega) = 0$, we conclude that the Fourier transform $\Theta(u - v) = 0$. Since Θ is a isomorphism we must have $u(t) - v(t) = 0$ for almost all t. Now the Laplace transform gives $\hat{u}(s) = \hat{v}(s)$ for all $s \in \overline{\mathbb{C}}^+$. □

From the previous corollary we find that there is one-to-one correspondence between function in H_2 and their restrictions to the imaginary axis. We say a function $\hat{v}(j\omega) \in \hat{L}_2(j\mathbb{R})$ is in H_2 is it has an analytic continuation $\hat{v}(s)$ to the right half plane which satisfies the conditions of Definition B.10.

Given this identification, we can endow H_2 with the inner product inherited from $\hat{L}_2(j\mathbb{R})$, defining for $\hat{u}$ and $\hat{v}$ in H_2 as

$$\langle \hat{u}\,,\hat{v}\rangle_{H_2} = \frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{u}(j\omega)^*\hat{v}(j\omega)\,d\omega\,.$$

In this way, Hardy space H_2 is a Hilbert subspace of $\hat{L}_2(j\mathbb{R})$.
Finally note that H_2 norm is defined both for vector functions (signals) and matrix valued functions (systems), which can respectively be denoted as $\|v\|_2$ and $\|G\|_2$.
The corresponding H_2 norms are

$$\|v\|_2 = \|\hat{v}\|_2 := \sqrt{\langle \hat{v}\,,\hat{v}\rangle_{H_2}} = \sqrt{\frac{1}{2\pi}\int_{-\infty}^{\infty} \hat{v}(j\omega)^*\hat{v}(j\omega)\,d\omega}\ ,$$

$$\|G\|_2 = \|\hat{G}\|_2 := \sqrt{\langle \hat{G}\,,\hat{G}\rangle_{H_2}} = \sqrt{\frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{trace}\big(\hat{G}(j\omega)^*\hat{G}(j\omega)\big)\,d\omega}\ .$$

Having explored the structure of L_2 and $\hat{L}_2$ spaces, we now turn to the *operators* introduced previously in the induced matrix norms. Since the spaces L_2 and $\hat{L}_2$ are isomorphic, so must be the $\mathcal{L}(L_2)$ and $\mathcal{L}(\hat{L}_2)$. Thus we can choose to work on operators in whichever domain is the most convenient. However, frequency domain provides a very elegant, exact characterization of an important class of operators on L_2: The operators which are linear and time invariant (LTI).

The space $\hat{L}_\infty(j\omega)$ of matrix valued functions mapping $j\mathbb{R} \to \mathbb{C}^{m\times n}$ is defined as

$$\|\hat{G}\|_\infty := \sup_{\omega\in\mathbb{R}} \overline{\sigma}\big(\hat{G}(j\omega)\big) < \infty\,.$$

The subscript ∞ in this norm corresponds to the essential supremum over frequency, analogously to the time domain L_∞ space introduced before. However, as we are using matrix functions, we take a maximum singular value norm (induced 2-norm) at each frequency. Notice that $\hat{L}_\infty(j\mathbb{R})$ is not related to $L_\infty(-\infty,\infty)$ via the Fourier transform.

Every function $\hat{G} \in \hat{L}_\infty(j\mathbb{R})$ defines a bounded linear multiplicative operator;

$M_{\hat{G}} \,:\, \hat{L}_2(j\mathbb{R}) \to \hat{L}_2(j\mathbb{R})$, via the relationship; $(M_{\hat{G}}\hat{v})(j\omega) = \hat{G}(j\omega)\hat{v}(j\omega)$.
Hence; $\|\hat{G}\|_\infty = \|M_{\hat{G}}\|_{\hat{L}_2 \to \hat{L}_2}$.

$\hat{G}$ refers to an operator on $L_2(-\infty, \infty)$, denoted by G, and defined by; $G = \Theta^{-1} M_{\hat{G}} \Theta$. We call $\hat{G}(j\omega)$ the *frequency response* of G, when G satisfies the last equation which is a *similarity transformation* (see Appendix B.2).

Definition B.13 *A function $\hat{G} : \overline{\mathbb{C}}^+ \to \mathbb{C}^{m\times n}$ is in Hardy space H_∞ if:*

i) $\hat{G}(s)$ *is analytic in the open right half plane* $\mathbb{C}^+$,
ii) for almost every real number ω *;* $\lim_{\sigma\to 0^+} \hat{G}(\sigma + j\omega) = \hat{G}(j\omega)$,
iii) (c) $\sup_{s\in\mathbb{C}^+} \overline{\sigma}(\hat{G}(s)) < \infty$.

As with H_2, it can be shown that H_∞ functions are determined by their values on the imaginary axis. Thus we can regard H_∞ as a subspace of $\hat{L}_\infty(j\mathbb{R})$, with the same norm

$$\|G\|_\infty = \|\hat{G}\|_\infty := \sup_{\omega\in\mathbb{R}} \overline{\sigma}\big(\hat{G}(j\omega)\big) .$$

Note that the above quantity is also coincides with the supremum given in *(c)* of the Definition B.12, which can be induced from the maximum modulus theorem.

In order to gain more insight into the frequency domain spaces that have been introduced, it is useful to discuss which *rational functions* belong to each of them.
A scalar rational function is a ratio of the polynomials

$$\hat{v}(s) = \frac{p(s)}{q(s)}$$

and the concept extends to vector and matrix funtions with components of this form.

We denote by $R\hat{L}_2$, RH_2, $R\hat{L}_\infty$, and RH_∞ for the set of rational functions which belongs respectively to $\hat{L}_2(j\mathbb{R})$, H_2, $\hat{L}_\infty(j\mathbb{R})$, and H_∞. For example, $\hat{v}(s)$ can be in $R\hat{L}_2$

if $\hat{v}(j\omega)$ is in $\hat{L}_2(j\mathbb{R})$.

A vector or matrix rational function is;
(i) in $R\hat{L}_2$ if it is strictly proper and has no poles on the imaginary axis,
(ii) in RH_2 if it is strictly proper and has no poles on the closed right half-plane $\overline{\mathbb{C}}^+$.

The subspaces $R\hat{L}_2$, and RH_2 are themselves inner product spaces, but are not complete.

Similarly, a matrix rational function is
(i) in $R\hat{L}_\infty$ if it is proper and has no poles on the imaginary axis,
(ii) in RH_∞ if it is proper and has no poles on the closed right half-plane $\overline{\mathbb{C}}^+$.

Every function in $R\hat{L}_\infty$ can be expressed in the form of

$$\hat{G}(j\omega) = C(j\omega I - A)^{-1}B + D\,,$$

where A has no eigenvalues on the imaginary axis, and every function in RH_∞ can be expressed in the form

$$\hat{G}(s) = C(sI - A)^{-1}B + D\,, \quad \text{where A is } \textit{Hurwitz}\,.$$

RH_2 and RH_∞ have their respective orthogonal complements denoted as $RH_2^\perp$ and $RH_\infty^\perp$, which are proper and have no poles on the closed left half plane $\overline{\mathbb{C}}^-$.
Hence, every function in $R\hat{L}_2$ can be written as a sum of functions in RH_2 and in $RH_2^\perp$.
Similarly, every function in $R\hat{L}_\infty$ can be written as a sum of functions in RH_∞ and functions in $RH_\infty^\perp$.

B.2. Some Facts from Basic Matrix Theory

We start with the concept of *linear mapping* between linear (vector) spaces. The mapping $A : \mathcal{V} \to \mathcal{W}$ is linear if; $A(\alpha v_1 + \beta v_2) = \alpha A v_1 + \beta A v_2$ for all v_1, v_2 in $\mathcal{V}$, and for all scalars α_1 and α_2. Here $\mathcal{V}$ and $\mathcal{W}$ are linear spaces with the same associated field $\mathbb{F}$. Given bases $\{v_1, ...v_n\}$ and $\{w_1, ...w_m\}$ for $\mathcal{V}$ and $\mathcal{W}$, respectively. We associate scalars a_{ik} with the mapping A, defining them such that they satisfy

$$Av_k = a_{1k}w_1 + a_{2k}w_2 + \cdots + a_{mk}w_m \,,$$

for each $1 \leq k \leq n$. Namely, given any basis vector v_k, the coefficients a_{ik} are the coordinates of Av_k in the chosen basis for $\mathcal{W}$. It turns out that these mn numbers a_{ik} completely specify the linear mapping A.

To express the relationship in a more convenient form, we can write the set of numbers a_{ik} as the $m \times n$ matrix

$$[A] = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} .$$

Then via the standard matrix product we have

$$\begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} .$$

In summary, any linear mapping A between linear vector spaces can be regarded as a matrix $[A]$ mapping $\mathbb{F}^n$ to $\mathbb{F}^m$ via matrix multiplication. Note that the numbers a_{ik} depend intimately on the bases $\{v_1, ...v_n\}$ and $\{w_1, ...w_m\}$. Frequently we use only one basis for $\mathcal{V}$ and one for $\mathcal{W}$ and thus there is no need to distinguish between the map A and the basis dependent matrix $[A]$. Associated with any linear map (or matrix)

$A : \mathcal{V} \to \mathcal{W}$ is its *range*, or *image space*, defined by: $\mathrm{Im}(A) = \mathcal{R}(A) := \{w \in \mathcal{W} :$ there exists $v \in \mathcal{V}$ satisfying $Av = w\}$.
This set contains all the elements of $\mathcal{W}$ which are the image of some region in $\mathcal{V}$.
If $\{v_1, ...v_n\}$ is a basis for $\mathcal{V}$, then; $\mathrm{Im}(A)$ =span$\{Av_1, ..., Av_n\}$, is a subspace of $\mathcal{W}$.
The map A is called *surjective* when $\mathrm{Im}(A) = \mathcal{W}$. The dimension of the image space is called the *rank* of the linear mapping A, and the concept is applied as well to the associated matrix $[A]$. Namely; rank$[A]$ =dim$\big(\mathrm{Im}(A)\big)$. Another important set related to A is its *kernel*, or *null space*, which is defined by
$\mathrm{Ker}(A) = \mathcal{N}(A) := \{v \in \mathcal{V} : \ Av = 0\}$. $\mathrm{Ker}(A)$ is the set of vectors in $\mathcal{V}$ which get mapped by A to the zero element in $\mathcal{W}$, and is easily verified to be a subspace of $\mathcal{V}$. In particular, when $\mathrm{Ker}(A)$ is the zero subspace, there is at most a unique solution to the equation $Av = w$. This means $Av_a = Av_b$ only when $v_a = v_b$. A mapping with this property is called *injective*.

In summary, a solution to the equation $Av = w$ will exist, if and only if, $w \in \mathrm{Im}(A)$, and it will be unique only when $\mathrm{Ker}(A)$ is the zero subspace. The dimensions of the image and kernel of A are linked by the relationship

$$\dim(\mathcal{V}) = \dim\big(\mathrm{Im}(A)\big) + \dim\big(\mathrm{Ker}(A)\big) .$$

A mapping is called *bijective* when it is both injective and surjective. That is for every $w \in \mathcal{W}$ there exists a unique v satisfying $Av = w$. In this case there is a well-defined inverse mapping $A^{-1} : \mathcal{W} \to \mathcal{V}$, such that; $A^{-1}A = I_{\mathcal{V}}$, $AA^{-1} = I_{\mathcal{W}}$.
In the above, I denotes the *identity mapping* in each space, that is the map that leaves elements unchanged. We also use the terms *nonsingular* or *invertible* to describe *bijective mappings*, and apply these terms as well to their associated matrices.

We have already discussed in Appendix B.1 the idea of choosing a basis $\{v_1, ..., v_n\}$ for the linear (vector) space $\mathcal{V}$. We can associate every vector $x \in \mathcal{V}$ with its coordinates $(\alpha_1, \cdots, \alpha_n) \in \mathbb{F}^n$ in this basis, which are the unit scalars satisfying; $x = \alpha_1 v_1 +, ..., \alpha_n v_n$.
This raises the question, suppose we choose another basis $\{u_1, ..., u_n\}$ for $\mathcal{V}$, *how can*

we effectively move between these basis representations ? That is, given $x \in \mathcal{V}$, how are the coordinate vectors x_v and $x_u \in \mathbb{F}^n$ related ?

The answer is as follows. Suppose that each basis vector u_k is expressed in the basis $\{v_1, ..., v_n\}$ by; $u_k = t_{1k}v_1 + \cdots + t_{nk}v_n$. Then the coefficients t_{ik} define the matrix

$$T = \begin{bmatrix} t_{11} & \dots & t_{1n} \\ \vdots & \ddots & \vdots \\ t_{n1} & \dots & t_{nn} \end{bmatrix} .$$

Note that such a matrix is nonsingular, since it represents the identity mapping $I_{\mathcal{V}}$ in the bases $\{v_1, ..., v_n\}$ and $\{u_1, ..., u_n\}$. Then the relationship between the two coordinate vectors is: $Tx_u = x_v$. Now suppose $A : \mathcal{V} \to \mathcal{V}$ and that $A_v : \mathbb{F}^n \to \mathbb{F}^n$ is the representation of A on the basis $\{v_1, ..., v_n\}$, and A_u is the representation of A using the basis $\{u_1, ..., u_n\}$. The question is: *How is A_u related to A_v ?*

To study this, take any $x \in \mathcal{V}$ and let x_v, x_u be its coordinates in the respective bases, and z_v, z_u be the coordinates of Ax. Then; $z_u = T^{-1}z_v = T^{-1}A_v x_v = T^{-1}A_v T x_u$. Since the above identity and $z_u = A_u x_u$ both hold for every x_u, we conclude that

$$A_u = T^{-1}A_v T .$$

The above relationship is called a *similarity transformation.*

Next we examine mappings when viewed with respect to a subspace. Suppose that $\mathcal{S} \subset \mathcal{V}$ is a k-dimensional subspace of $\mathcal{V}$, and that $\{v_1, ..., v_n\}$ is a basis for $\mathcal{V}$ with; $\text{span}\{v_1, ..., v_k\} = \mathcal{S}$. That is the first k vectors of this basis forms a basis for $\mathcal{S}$.
If $E : \mathcal{V} \to \mathbb{F}^n$ is the associated map which maps the basis vectors in $\mathcal{V}$ to the standard basis on $\mathbb{F}^n$, then; $E\mathcal{S} = \mathbb{F}^k \times \{0\} \subset \mathbb{F}^n$. Thus in $\mathbb{F}^n$ we can view $\mathcal{S}$ as the elements of the form $(x\ 0)^T$ where $x \in \mathbb{F}^k$. From the point of view of linear mapping $A : \mathcal{V} \to \mathcal{V}$, this partitioning of $\mathbb{F}^n$ gives a useful decomposition of the

corresponding matrix $[A]$. Namely, we can regard A as; $[A] = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}$, where $A_1 : \mathbb{F}^k \to \mathbb{F}^k,\ A_2 : \mathbb{F}^{n-k} \to \mathbb{F}^k,\ A_3 : \mathbb{F}^k \to \mathbb{F}^{n-k},\ A_4 : \mathbb{F}^{n-k} \to \mathbb{F}^{n-k}$.

We say that a subspace $\mathcal{S} \subset \mathcal{V}$ is $A-$invariant if $A : \mathcal{V} \to \mathcal{V}$ and $A\mathcal{S} \subset \mathcal{S}$. Every map has at least two *invariant subspaces*, the zero subspace and the entire domain $\mathcal{V}$. For subspaces $\mathcal{S}$ of intermediate dimension, the invariance property is expressed most clearly by saying the associated matrix has the form

$$[A] = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}.$$

Here we are assuming, as above, that our basis for $\mathcal{V}$ is obtained by extending a basis for $\mathcal{S}$. Similarly if a matrix has this form, the subspace $\mathbb{F}^k \times \{0\}$ is A-invariant.

Definition B.14 *Given $A \in \mathbb{C}^{n\times n}$. The determinant of A is defined as*

$$\begin{aligned} \det(A) :=& \sum_{i=1}^{n} (-1)^{i+j}\, a_{ij}\ \det(A_{[i,j]}) \ \text{ for any } j = 1 : n \\ =& \sum_{j=1}^{n} (-1)^{i+j}\, a_{ij}\ \det(A_{[i,j]}) \ \text{ for any } i = 1 : n \end{aligned}$$

where $A_{[i,j]}$ denotes the submatrix of A obtained by deleting the ith row and jth column and the determinant of a scalar is defined as its scalar value.

Having introduced the concept of determinant, we can now state an important characteristic of *similar matrices.*

Lemma B.15 [42] *If matrices $A, B \in \mathbb{R}^{n\times n}$ are similar, then they have the same eigenvalues.*

Proof: Let $B = T^{-1}AT$ for some invertible T. Then the eigenvalues of the transformed

matrix B can be found from the equation

$$\det(\lambda I - T^{-1}AT) = \det\big(T^{-1}(\lambda I - A)T\big) = \det(T^{-1})\det(\lambda I - A)\det(T) = \det(\lambda I - A) = 0.$$

Hence the eigenvalues of A and B are equal; i.e., A and B have the same *spectrum.* $\square$

Definition B.16 *Given an invertible $A \in \mathbb{C}^{n\times n}$. The inverse of A can be computed as*

$$A^{-1} = \frac{1}{\det A} C^T, \quad \textit{where } C \textit{ is the cofactor matrix of } A,$$
$$\textit{and its elements are computed as; } c_{ij} = (-1)^{i+j} \det A_{[i,j]}.$$

There are some useful formulas for the inversion of a sum of matrices and a matrix defined by matrix blocks [42]:

(i) $(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$

(ii) Define $\Delta := A - BD^{-1}C$ and $\nabla := D - CA^{-1}B$, if the inverses exist. Then

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} \Delta^{-1} & -\Delta^{-1}BD^{-1} \\ -D^{-1}C\Delta^{-1} & D^{-1} + D^{-1}C\Delta^{-1}BD^{-1} \end{bmatrix}$$

$$= \begin{bmatrix} A^{-1} + A^{-1}B\nabla^{-1}CA^{-1} & -A^{-1}B\nabla^{-1} \\ -\nabla^{-1}CA^{-1} & \nabla^{-1} \end{bmatrix}.$$

Using matrix inverses we can express determinants of block-partitioned matrices in terms of their submatrices.

Theorem B.17 [37] *Suppose that matrices $A \in \mathbb{C}^{n\times n}$, $B \in \mathbb{C}^{n\times m}$, $C \in \mathbb{C}^{m\times n}$, and*

$D \in \mathbb{C}^{m\times m}$ *are given. The following holds where* A *or* D *is invertible*:

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det A \cdot \det(D - CA^{-1}B)$$
$$= \det D \cdot \det(A - BD^{-1}C)\,.$$

Definition B.18 *The trace of* $A \in \mathbb{C}^{n\times n}$*, denoted by* $\mathrm{trace}(A)$*, is the sum of its diagonal elements, i.e.,* $\mathrm{trace}(A) = \sum_{i=1}^{n} a_{ii}$.

Needless to say, trace is a linear operator. We also have the following very useful result regarding the trace of the product of two matrices.

Proposition B.19 *Given* $A \in \mathbb{C}^{m\times n}$ *and* $B \in \mathbb{C}^{n\times m}$; $\mathrm{trace}(AB) = \mathrm{trace}(BA)$.

Proof: $\mathrm{trace}(AB) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{ij}b_{ji}\right) = \sum_{j=1}^{n} (\sum_{i=1}^{m} b_{ji}a_{ij}) = \mathrm{trace}(BA)$. □

Consider the space $\mathbb{C}^{m\times n}$ of complex $m \times n$ matrices of the form

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix}.$$

We define the *transpose* of the above matrix $A \in \mathbb{C}^{m\times n}$ by

$$A^T = \begin{bmatrix} a_{11} & \dots & a_{m1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \dots & a_{mn} \end{bmatrix} \in \mathbb{C}^{n\times m},$$

and its *Hermitian conjugate* or *adjoint* by

$$A^* = \begin{bmatrix} a_{11}^* & \dots & a_{m1}^* \\ \vdots & \ddots & \vdots \\ a_{1n}^* & \dots & a_{mn}^* \end{bmatrix} \in \mathbb{C}^{n\times m}.$$

In both cases the indices have been transposed, but in the latter we also take the complex conjugate of each element. Clearly, both operations coincide if the matrix is real.

The square matrix $A \in \mathbb{C}^{n\times n}$ is *Hermitian* or *self-adjoint* if $A = A^*$. The space of Hermitian matrices is denoted by $\mathbb{H}^n$ and is a real vector space. If a Hermitian matrix $A \in \mathbb{R}^{n\times n}$ it is more specifically referred to as *symmetric*. The set of symmetric matrices is also a real vector space and denoted as $\mathbb{S}^n$. For matrices A and B of compatible dimensions, it can be verified that; $(AB)^* = B^*A^*$.

A square matrix $U \in \mathbb{C}^{n\times n}$ is called *unitary* if it satisfies; $U^*U = I$. From this definition we see that the columns of any unitary matrix forms an orthonormal basis for $\mathbb{C}^n$. Further, since U is square it must be that $U^* = U^{-1}$ and therefore $UU^* = I$. So the columns of U^* also form an orthonormal basis. Unitary matrices are the only matrices that leave the length of every vector unchanged. Now we can state the *spectral theorem* for Hermitian matrices.

Theorem B.20 [38] *Suppose H is a matrix in $\mathbb{H}^n$. Then there exists a unitary matrix U and a real diagonal matrix Λ such that; $H = U\Lambda U^*$.*

Note that since $U^* = U^{-1}$ for a unitary U, the above expression is a *similarity transformation*. Therefore the theorem says that a self-adjoint matrix can be diagonalized by a unitary similarity transformation. Thus the columns of U are the eigenvectors of H.

Definition B.21 (Congruence) *Two matrices $A, B \in \mathbb{C}^{n\times n}$ are said to be congru-*

ent to each other if there exists a nonsingular $T \in \mathbb{C}^{n\times n}$ *such that;* $B = T^*AT$. *The transformation* $A \to T^*AT$ *is called congruence transformation of* A *under* T.

Since all eigenvalues of a Hermitian matrix are real, we can group the ones that are positive, negative, and zero. This brings us to the definition of *inertia* of a Hermitian matrix.

Definition B.22 *Given a matrix* $A = A^* \in \mathbb{C}^{n\times n}$. *The inertia of* A *is the triplet* $\mathrm{in}(A) = (n_+, n_-, n_0)$, *where* n_+, n_-, n_0 *denote the number of positive, negative, and zero eigenvalues (counting multiplicities) of* A *respectively.*

Analogous to similar matrices having the same spectrum, congruent matrices have the same inertia as stated below.

Lemma B.23 [42] *Given two Hermitian matrices* $A, B \in \mathbb{C}^{n\times n}$. *Then*:

$$A \text{ and } B \text{ are congruent} \iff \mathrm{in}(A) = \mathrm{in}(B)$$

Given $Q \in \mathbb{H}^n$, we say it is *positive definite* denoted $Q \succ 0$, if $x^*Qx > 0$ for all nonzero $x \in \mathbb{C}^n$. Similarly Q is *positive semidefinite* denoted $Q \succeq 0$, if the inequality is non-strict. *Negative definiteness* and *negative semidefiniteness* is similarly defined. If a matrix is not positive or negative semidefinite, then it is *indefinite*.
It can be derived from Therom B.15 that for a matrix $Q \in \mathbb{H}^n$, $Q \succ 0$ if and only if eigenvalues of Q are all positive. Notice in particular that a positive definite matrix is always invertible and its inverse is also positive definite. Hence, the number of strictly positive eigenvalues is equal to the rank of the matrix.
An additional useful property of for positive matrices is the existence of a square root. Let $Q = U\Lambda U^* \succeq 0$, i.e., the diagonal elements of Λ are nonnegative. Then we can define $\Lambda^{1/2}$ to be the matrix with diagonal elements $\lambda_k^{1/2}$, and $Q^{1/2} := U\Lambda^{1/2}U^*$. Therefore $Q^{1/2} \succeq 0$ (also $Q^{1/2} \succ 0$ when $Q \succ 0$) and it can be verified that; $Q^{1/2}Q^{1/2} = Q$.

Theorem B.24 [36] (Schur Complement Formula) *Suppose that Q,M, and R are matrices and that M and Q are self-adjoint. Then the following are equivalent:*

i) The matrix inequalities $Q \succ 0$ *and* $M - RQ^{-1}R^* \succ 0$ *both hold.*

ii) The matrix inequality: $\begin{bmatrix} M & R \\ R^* & Q \end{bmatrix} \succ 0$ *is satisfied.*

Proof: The two inequalities listed in (i) are equivalent to the single block inequality

$$\begin{bmatrix} M - RQ^{-1}R^* & 0 \\ 0 & Q \end{bmatrix} \succ 0\,.$$

Now left- and right-multiply this inequality by the nonsingular matrix: $\begin{bmatrix} I & RQ^{-1} \\ 0 & I \end{bmatrix}$ and its adjoint, respectively, to get:

$$\begin{bmatrix} M & R \\ R^* & Q \end{bmatrix} = \begin{bmatrix} I & RQ^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} M - RQ^{-1}R^* & 0 \\ 0 & Q \end{bmatrix} \begin{bmatrix} I & 0 \\ Q^{-1}R^* & I \end{bmatrix} \succ 0\,.$$

Therefore inequality (ii) holds, if and only if, (i) holds. □

If $A \in \mathbb{C}^{n\times n}$, we say that $\lambda \in \mathbb{C}$ is an *eigenvalue* of A if; $Ax = \lambda x$ can be satisfied for some nonzero vector $x \in \mathbb{C}^n$. Such a vector x is called an *eigenvector*. Equivalently, this means that $\mathrm{Ker}(\lambda I - A) \neq 0$ or $\lambda I - A$ is singular. If A is a real matrix, then any nonreal eigenvalues must appear in conjugate pairs. Also, a matrix has the eigenvalue zero, if and only if, it is singular. Associated with every eigenvalue λ_k is the subspace

$$\varepsilon_k = \mathrm{Ker}(\lambda_k I - A)\,,$$

where every nonzero element in ε_k is an eigenvector corresponding to the eigenvalue λ_k.

Suppose that a set of eigenvectors satisfies; $\text{span}\{x_1, ..., x_n\} = \mathbb{C}^n$. Then we define the invertible matrix $X = [x_1 \cdots x_n]$, and find; $AX = [Ax_1 \cdots Ax_n] = [\lambda_1 x_1 \cdots \lambda_n x_n] = X\Lambda$, where $\Lambda = diag(\lambda_i)\ \ i = 1, ..., n$. Thus we have a similarity transformation X such that $X^{-1}AX = \Lambda$ is diagonal, and the matrix A is *diagonalizable*. Notice that a matrix A is diagonalizable, if and only if, $\varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_n = \mathbb{C}^n$ holds.
Hence, matrices with linearly dependent eigenvectors can not be diagonalized.

We define the $n \times n$ matrix N by

$$N = \begin{bmatrix} 0 & 1 & & 0 \\ & & \ddots & \\ & & & 1 \\ 0 & & & 0 \end{bmatrix},$$

where $N = 0$ if the dimension $n = 1$. Such matrices are called *nilpotent* as $N^n = 0$. Using these we define a matrix to be a *Jordan block* if it is of the form

$$J = \lambda I + N = \begin{bmatrix} \lambda & 1 & & 0 \\ & & \ddots & \\ & & & 1 \\ 0 & & & \lambda \end{bmatrix}.$$

A Jordan block has one eigenvalue λ of multiplicity n. However, it has only one linearly independent eigenvector. A key feature of Jordan block is that it has precisely n subspaces which are J-invariant. They are given by; $\mathbb{C}^k \times \{0\}$, for $1 \leq k \leq n$. When $k = 1$, this corresponds exactly to the subspace associated with its eigenvector. We can now state the *Jordan decomposition* theorem.

Theorem B.25 [38] *Suppose $A \in \mathbb{C}^{n \times n}$. Then there exists a nonsingular matrix*

$T \in \mathbb{C}^{n\times n}$, and an integer $1 \leq p \leq n$, such that

$$T^{-1}AT = J = \begin{bmatrix} J_1 & & & 0 \\ & J_2 & & \\ & & \ddots & \\ 0 & & & J_p \end{bmatrix},$$

where the matrices J_k are Jordan blocks.

This theorem states that a matrix can be transformed to one that is block diagonal, where each of the diagonal matrices is a Jordan block. Notice that if a matrix is diagonalizable, each Jordan block J_k will simply be a scalar equal to an eigenvalue of A. The relevance of the Jordan composition is that it provides a *canonical form* to characterize matrix similarity. Namely, *two matrices are similar if and only if they share the same Jordan form.*

Another related feature is that the Jordan form exhibits the structure of *invariant subspaces* of a given matrix. This can be seen by writing the above equation as; $AT = TJ$. Suppose we denote by T_1 the submatrix of T formed by its first n_1 columns, where n_1 is the dimension of the block J_1. Then the first n_1 columns of the preceding equation give; $AT_1 = T_1J_1$, which implies that $\mathcal{S}_1 = \mathrm{Im}(T_1)$ is invariant under A. Furthermore, we can use this formula to study the linear mapping on $\mathcal{S}_1$ obtained by restriction of A. In fact we find that in the basis defined by the columns of T_1, this linear mapping has the associated matrix J_1. The only eigenvalue of A restricted to $\mathcal{S}_1$ is λ_1.

We conclude our discussion with the singular value decomposition (SVD) of a rectangular matrix, which has many applications in control theory.

Theorem B.26 [36] *Suppose $A \in \mathbb{C}^{m\times n}$ and that $p = \min\{m, n\}$. Then there exists unitary matrices $U \in \mathbb{C}^{m\times m}$ and $V \in \mathbb{C}^{n\times n}$ such that; $A = U\Sigma V^*$, where $\Sigma \in \mathbb{R}^{m\times n}$ and it scalar entries satisfy;*

i) the condition $\sigma_{ir} = 0$ *, for* $i \neq r$*,* *(ii) the ordering* $\sigma_{11} \geq \sigma_{22} \geq \cdots \geq \sigma_{pp} \geq 0$ *.*

Proof: Since the result holds for A if and only if it holds for A^*, we assume without loss of generality that $n \geq m$. To start let r be the rank of A^*A, which is Hermitian and therefore by Theorem B.15 we have

$$A^*A = V \begin{bmatrix} \bar{\Sigma}^2 & 0 \\ 0 & 0 \end{bmatrix} V^*, \text{ where } \bar{\Sigma} = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_r \end{bmatrix} \succ 0 \text{, and } V \text{ is unitary.}$$

We also assume that the nonstrict ordering $\sigma_1 \geq \cdots \geq \sigma_r$ holds. Defining $J = \begin{bmatrix} \bar{\Sigma} & 0 \\ 0 & I \end{bmatrix}$, we have; $J^{-1}V^*A^*AVJ^{-1} = (AVJ^{-1})^*(AVJ^{-1}) = \begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$, where I_r denotes the $r \times r$ identity matrix. From the right-hand side we see that the first r columns of AVJ^{-1} form an orthonormal set, and the remaining columns must be zero.

Thus $AVJ^{-1} = [U_1 \ 0]$, where $U_1 \in \mathbb{C}^{m \times r}$. This leads to:

$A = [U_1 \ 0] \begin{bmatrix} \bar{\Sigma} & 0 \\ 0 & I \end{bmatrix} V^* = [U_1 \ U_2] \begin{bmatrix} \bar{\Sigma} & 0 \\ 0 & 0 \end{bmatrix} V^*$, where the right-hand side is valid for any $U_2 \in \mathbb{C}^{m \times (m-r)}$. So choose U_2 such that $[U_1 \ U_2]$ is unitary. □

When $n = m$, the matrix Σ in the SVD is diagonal. When these dimensions are not equal, Σ has the form of either

$$\begin{bmatrix} \sigma_{11} & & & 0 \\ & \ddots & & \\ 0 & & \sigma_{mm} & 0 \end{bmatrix} \text{ when } n > m \text{, or } \begin{bmatrix} \sigma_{11} & & 0 \\ & \ddots & \\ 0 & & \sigma_{nn} \\ 0 & & 0 \end{bmatrix} \text{ when } n < m \text{.}$$

The first p nonnegative scalars σ_{kk} are called the *singular values* of the matrix A, and denoted by the ordered set $\sigma_1, ..., \sigma_p$, where $\sigma_k = \sigma_{kk}$.

As shown in the proof, the decomposition of the theorem immediately gives us that;
$A^*A = V(\Sigma^*\Sigma)V^*$ and $AA^* = U(\Sigma\Sigma^*)U^*$, which are the singular value decompositions of A^*A and AA^*. But since $V^* = V^{-1}$ and $U^* = U^{-1}$, it follows that these are also the diagonalizations of the matrices. Thus $\sigma_1^2 \geq \sigma_2^2 \geq \cdots \geq \sigma_p^2 \geq 0$ are exactly the p largest eigenvalues of A^*A and AA^*. The remaining eigenvalues of either matrix are all necessarily equal to zero. This observation provides a straightforward method to obtain the singular value decomposition of any matrix A, by diagonalizing the Hermitian matrices A^*A and AA^*.

The SVD of a matrix has many useful properties. We use $\overline{\sigma}(A)$ to denote the largest singular value σ_1, which has the property; $\overline{\sigma}(A) = \max\{|Av| : v \in \mathbb{C}^n \text{ and } |v| = 1\}$.
It is the maximum magnification a vector v can experience when acted upon by A.

B.3. Convexity and Semidefinite Programming

We begin by defining the line segment that joins two points in a linear space $\mathcal{V}$. Suppose that v_1 and v_2 are in $\mathcal{V}$, then we define the line segment $L(v_1, v_2)$ between them as the set of points; $L(v_1, v_2) = \{v \in \mathcal{V} : v = \tau v_1 + (1-\tau)v_2 \text{ for } \tau \in [0,1]\}$.

Suppose that $\mathcal{Q}$ is a nonempty subset of the linear space $\mathcal{V}$. Then $\mathcal{Q}$ is defined to be *convex set* if for any $v_1, v_2 \in \mathcal{Q}$ the line segment $L(v_1, v_2)$ is a subset of $\mathcal{Q}$. That is, $\mathcal{Q}$ is convex if it contains all the line segments between its points.

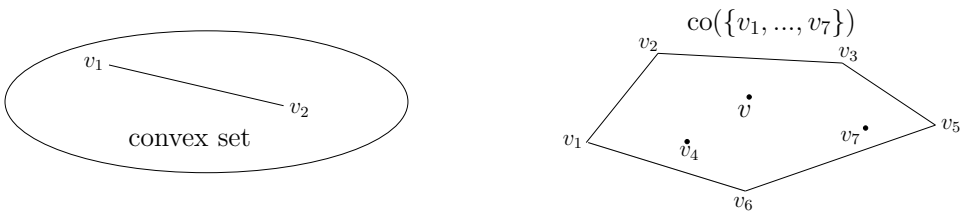

Figure B.1. Convex set and convex hull of finite number of points (polytope) in $\mathbb{R}^2$.

Building on this intuition from $\mathbb{R}^2$, we define the *convex hull* of $\mathcal{Q}$ is defined as the subset

of $\mathcal{V}$ composed of *all* vectors obtained via convex combination from the elements of $\mathcal{Q}$ as

$$\mathrm{co}(\mathcal{Q}) = \{v \in \mathcal{V} \,:\, v = \sum_{i=1} \tau_i v_i \,,\ \text{with } \tau_i \in [0,1], \text{ and } \sum_{i=1} \tau_i = 1 \,,\ i = 1, 2...\} \,.$$

Notice that for two points: $L(v_1\,, v_2) = \mathrm{co}(\{v_1\,, v_2\})$. If the set $\mathcal{Q}$ is convex, then it necessarily contains any convex hull formed from a collection of its points. Note also that, by definition, the intersection of convex sets is always convex. In other words, if $\mathcal{R}$ is convex and $\mathcal{Q} \subset \mathcal{R}$, then necessarily $\mathrm{co}(\mathcal{Q})$ is a subset of $\mathcal{R}$.

Given a linear space $\mathcal{V}$ over $\mathbb{R}$ and m points $v_i \in \mathcal{V}$, $i = 1, ..., m$. A *polytope* of vertices v_i is a set in the form

$$\Pi = \{v \in \mathcal{V} : v = \sum_{i=1}^{m} \tau_i v_i \,,\ \sum_{i=1}^{m} \tau_i = 1 \,,\ \tau_i \geq 0 \,,\ i = 1, ..., m\} \,.$$

We denote the set of the vertices of the polytope by $\Pi^v := \{v_i,\ i = 1, ..., m\}$. Obviously, $co(\mathcal{Q}) \supseteq \mathcal{Q}$ and equality holds if $\mathcal{Q}$ is a polytope. Moreover, notice that *convex hull of a finite number of points is a polytope.* A *hyper-box* $\mathcal{P} \subset \mathbb{R}^m$ is a particular polytope with 2^m vertices.

Lemma B.27 [29] *Let $F : \mathcal{P} \to \mathbb{R}^{n\times n}$, $\rho \to F(\rho)$ be a multi-affine matrix-valued function, where $\mathcal{P} \subset \mathbb{R}^m$ is a hyper-box. Then; $co\big(F(\mathcal{P})\big) = co\big(F(\mathbb{R}^m)\big)$.*

When the dependence is affine (rather than multi-affine), the image (rather than the convex hull of the image) of a matrix function defined over a hyper-box coincides with the convex hull of the images of the vertices of the hyper-box.

Theorem B.28 [29] *Let $F : \mathcal{P} \to \mathbb{R}^{n\times n}$, $\rho \to F(\rho)$ be an affine matrix-valued function, where $\mathcal{P} \subset \mathbb{R}^m$ is a hyper-box. Then; $F(\mathcal{P}) = \mathrm{co}\big(F(\mathcal{P}^v)\big)$.*
Proof: Note that $H \in F(\mathcal{P})$ implies that $H \in \mathrm{co}\big(F(\mathcal{P})\big)$, which, by virtue of Lemma B.27, in turn implies that H also belongs to $\mathrm{co}\big(F(\mathcal{P}^v)\big)$.

Conversely, assume that $H \in \mathrm{co}(F(\mathcal{P}^v))$. Therefore there exists;
$\tau_i \geq 0$, $i = 1, ..., 2^m$, $\sum_{i=1}^{2^m} \tau_i = 1$, such that

$$H = \sum_{i=1}^{2^m} \tau_i F(\rho_i) = F\left(\sum_{i=1}^{2^m} \tau_i \rho_i\right) ,$$

where the last equality is a consequence of the fact that $F(\rho)$ is affine. Since $\mathcal{P}$ is a hyper-box, we have that; $\sum_{i=1}^{2^m} \tau_i \rho_i \in \mathcal{P}$. □

Given a set of symmetric matrices Υ, we write $\Upsilon > 0$ meaning that $M > 0$ for all $M \in \Upsilon$

Lemma B.29 [29] *A set of symmetric matrices $\Upsilon \subset \mathbb{R}^{n\times n}$ is positive definite iff* $\mathrm{co}(\Upsilon)$ *is positive definite.*
Proof: The fact that $\mathrm{co}(\Upsilon) > 0$ implies $\Upsilon > 0$ is obvious. Conversely assume that $\Upsilon > 0$. Then for any matrix $M \in \mathrm{co}(\Upsilon)$ can be written, according to the definition of convex hull, as

$$M = \sum_{i=1}^{h} \tau_i M_i ,$$

where h is some integer, $\sum_{i=1}^{h} \tau_i = 1$, $\tau_i \geq 0$, and $M_i \in \Upsilon$, $i = 1, ..., h$.
Since $M_i > 0$, $i = 1, ..., h$, we conclude that M is positive definite. From the arbitrariness of M, the proof follows. □

From Lemmas B.27 and B.29 we obtain the following fundamental result:

Theorem B.30 [29] *Consider the matrix valued function $F : \mathcal{P} \to \mathbb{R}^{n\times n}$, $\rho \to F(\rho)$, where $\mathcal{P} \subset \mathcal{P}^m$ is a hyper-box and*

$$F(\rho) = \frac{N_F(\rho)}{d_F(\rho)}$$

with N_F and d_F multi-affine functions of ρ and $d_F(\rho) \neq 0$ for all $\rho \in \mathcal{P}$. Then we have

that $F(\mathcal{P}) > (<)0$ *if and only if* $F(\mathcal{P}^v) > (<)0$.

Proof: $F(\mathcal{P}) > 0$ implies $F(\mathcal{P}^v) > 0$ is obvious; conversely, assume that $F(\mathcal{P}^v) > 0$. Since $d_F(\rho)$ can not change sign in $\mathcal{P}$ assume that $d_F(\rho) > 0$ for all $\rho \in \mathcal{P}$. This implies that $d_F(\rho) > 0$ for all $\rho \in \mathcal{P}^v$; since $F(\mathcal{P}^v) > 0$, we have that $N_F(\mathcal{P}^v) > 0$.

We have the following chain of statements

$$\begin{aligned} N_F(\mathcal{P}^v) > 0 &\Rightarrow \mathrm{co}\big(N_F(\mathcal{P}^v)\big) > 0 && \text{by Lemma B.29} \\ &\Rightarrow \mathrm{co}\big(N_F(\mathcal{P})\big) > 0 && \text{by Lemma B.27} \\ &\Rightarrow N_F(\mathcal{P}) > 0 && \text{by Lemma B.29} \end{aligned}$$

The inequality $N_F(\mathcal{P}) > 0$ with the condition $d_F(\rho) > 0$ for all $\rho \in \mathcal{P}$ guarantee that $F(\mathcal{P}) > 0$. The proof is analogous when $d_F(\rho) < 0$ in $\mathcal{P}$. □

A set in an n-dimensional subspace, obtained by shifting an $n-1$ dimensional subspace away from the origin, is called a *hyperplane*. An important property of a hyperplane is that it always breaks up the space into two half-spaces. These half-spaces have the form $\{v : F(v) \leq a\}$ and $\{v : F(v) \geq a\}$, where $F : \mathcal{V} \to \mathbb{R}$ is a linear mapping.

Given two sets $\mathcal{Q}_1$ and $\mathcal{Q}_2$ in $\mathcal{V}$, we say that the hyperplane defined by (F, a) separates the sets if i) $F(v_1) \leq a$, for all $v_1 \in \mathcal{Q}_1$,and ii) $F(v_2) \geq a$, for all $v_2 \in \mathcal{Q}_2$. Further, it is a *strictly separating hyperplane* if (b) is changed to; $F(v_2) \geq a + \epsilon$, for all $v_2 \in \mathcal{Q}_2$, for some fixed $\epsilon > 0$.

We have the following major result, which says that if two sets are convex and disjoint, then there exists a separating hyperplane between them.

Theorem B.31 [45] *Suppose* $\mathcal{Q}_1$ *and* $\mathcal{Q}_2$ *are two nonempty convex subsets of the linear space* $\mathcal{V}$.

i) If the intersection $\mathcal{Q}_1 \cap \mathcal{Q}_2$ *is empty, then there exists a hyperplane which separates*

$\mathcal{Q}_1$ *from* $\mathcal{Q}_2$.

ii) There exists a strictly separating hyperplane between the sets, if and only if, the sets are strictly separated.

A set $\mathcal{Q} \subset \mathcal{V}$ is called a *cone* if it is closed under positive scalar multiplication. That is, if $v \in \mathcal{Q}$ implies $tv \in \mathcal{Q}$ for every $t > 0$. Subspaces are cones, but the latter definition is broader since it includes, for example, the half-line $\mathcal{C}_v = \{\alpha v : \alpha > 0\}$.

Of particular interest are cones which have the convexity property. In fact *convex cones* are precisely those which are closed under addition (i.e., $v_1, v_2 \in \mathcal{C}$ implies $v_1 + v_2 \in \mathcal{C}$). A canonical example of a convex cone is a half-space; $F(v) \geq 0$ that goes through the origin, or an intersection of half spaces of this form. One of the important reasons of our interest in the concept of convexity is the fact that the *linear matrix inequalities*, which are extensively used in control theory [40], define a convex set.

A linear matrix inequality (LMI) in the variable X is an inequality of the form

$$F(X) \prec Q\,, \text{ in which}$$

- the variable X takes values on a real linear space $\mathcal{X}$,
- the matrix Q is in the set of *Hermitian* matrices $\mathbb{H}^n$, i.e., $Q = Q^*$,
- the mapping $F : \mathcal{X} \to \mathbb{H}^n$ is linear.

Thus to determine whether an inequality is an LMI, we simply check whether the above conditions are satisfied.

An important property of LMIs is that a list of them can always be converted to a single LMI. Suppose we have the two LMIs

$$F_1(X_1) \prec Q_1 \text{ and } F_2(X_2) \prec Q_2$$

in the variables X_1 and X_2 in linear spaces $\mathcal{X}_1$ and $\mathcal{X}_2$, respectively. Then this is

equivalent to a single LMI

$$F(X) \prec Q\,, \text{ where } X = (X_1, X_2) \in \mathcal{X}_1 \times \mathcal{X}_2 \text{ and } Q = \begin{bmatrix} Q_1 & 0 \\ 0 & Q_2 \end{bmatrix},$$

and we define the function $F : \mathcal{X}_1 \times \mathcal{X}_2 \to \mathbb{H}^{n_1+n_2}$ by

$$F(X) = \begin{bmatrix} F_1(X_1) & 0 \\ 0 & F_2(X_2) \end{bmatrix}.$$

Most LMIs are not formulated in the standard appearance as shown above, but they can be rewritten in the original form.

Consider the following *Lyapunov inequality* (dual) problem:

$$\mathcal{A}\mathcal{Y} + \mathcal{Y}\mathcal{A}^T \prec 0 \;\text{ with }\; \mathcal{Y} = \mathcal{Y}^T \succ 0\,,$$

which can be rewritten as; $\mathcal{A}\mathcal{Y} + \mathcal{Y}\mathcal{A}^T \prec -Q$ with $\mathcal{Y} = \mathcal{Y}^T \succ 0$, and $Q = Q^* \succ 0$.

Note that $\mathcal{Y}$ enters linearly in both inequalities. We can combine the three conditions as

$$\begin{bmatrix} -\mathcal{Y} & 0 \\ 0 & \mathcal{A}\mathcal{Y} + \mathcal{Y}\mathcal{A}^T \end{bmatrix} \prec \begin{bmatrix} -Q_1 & 0 \\ 0 & -Q_2 \end{bmatrix}.$$

To show how to rewrite this in the standard LMI form, assume $\mathcal{Y} \in \mathbb{R}^{2\times 2}$. Parameterizing $\mathcal{Y}$ as a linear function in X

$$\mathcal{Y}(X) = X_1 \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + X_2 \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + X_3 \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} = X_1\mathcal{Y}_1 + X_2\mathcal{Y}_2 + X_3\mathcal{Y}_3$$

Then, $F(X) = \operatorname{diag}\left[-\mathcal{Y}(X)\,,\, \mathcal{A}\mathcal{Y}(X) + \mathcal{Y}(X)\mathcal{A}^T\right] \prec -Q$, or

$$F(X) = X_1 \begin{bmatrix} -\mathcal{Y}_1 & 0 \\ 0 & \mathcal{A}\mathcal{Y}_1 + \mathcal{Y}_1\mathcal{A}^T \end{bmatrix} + X_2 \begin{bmatrix} -\mathcal{Y}_2 & 0 \\ 0 & \mathcal{A}\mathcal{Y}_2 + \mathcal{Y}_2\mathcal{A}^T \end{bmatrix} + X_3 \begin{bmatrix} -\mathcal{Y}_3 & 0 \\ 0 & \mathcal{A}\mathcal{Y}_3 + \mathcal{Y}_3\mathcal{A}^T \end{bmatrix} \prec \begin{bmatrix} -Q_1 & 0 \\ 0 & -Q_2 \end{bmatrix},$$

which is in the form of the original inequality.

An important point which appears repeatedly in analysis and synthesis formulations, is that conditions that do not look like LMIs at first glance can sometimes be converted to them. The *Schur complement* given in Appendix B.2 is very useful in this regard.

In fact, many control problems can be formulated in terms of finding solutions to LMIs. More generally, we have optimization problems with LMI constraints on the solution space. This class of optimization problem is known as *semidefinite programming*. The simplest semidefinite program is a decision program, in which a search is performed to answer the question: *Does there exist* $X \in \mathcal{X}$, *satisfying* $F(X) \prec Q$ *?* This is known as a *feasibility problem* and asks only whether there is any element which satisfies the LMI. An important property of the solution set is given below:

Proposition B.32 *The set* $\mathcal{C} := \{X \in \mathcal{X}$ *such that* $F(X) \prec Q\}$ *is convex in* $\mathcal{X}$.
Proof: Suppose $X_1, X_2 \in \mathcal{C}$, which means they satisfy $F(X_1) \prec Q$ and $F(X_2) \prec Q$. Consider any point $X_3 = \tau X_1 + (1-\tau)X_2$, for some value $\tau \in [0,1]$. Using linearity of the function F we have: $F(X_3) = \tau F(X_1) + (1-\tau)F(X_2) < \tau Q + (1-\tau)Q = Q$. The equality follows from the fact that the positive definite matrices are convex cone. Therefore, $X_3 \in \mathcal{C}$. $\square$

We now define a more general semidefinite optimization problem:

$$\begin{aligned} &\text{minimize: } c(X); \\ &\text{subject to: } F(X) \leq Q \text{ and } X \in \mathcal{X}, \end{aligned}$$

where $c(X)$ is a linear *functional* (i.e., function of a function) on $\mathcal{X}$. This is referred to as *linear objective problem.* In this formulation we are being informal about the meaning of "minimize". In fact, we want to compute the infimum, which may or may not be achieved at a given X. Alternatively, the constraint could be specified as the strict version $F(X) < Q$, in which case we are always dealing with an infimum as the search will be made in an open set.

Conversely, the feasibility problem can be recast as the problem:

$$\begin{aligned} &\text{find:} \quad J := \inf t\,; \\ &\text{subject to:} \;\; F(X) - tI \leq Q\,. \end{aligned}$$

This is a linear objective problem in the variables $t \in \mathbb{R}$ and $X \in \mathcal{X}$, and the corresponding LMI constraint is automatically (strictly) feasible. Notice that $J < 0$, if and only if, the LMI $F(X) < Q$ is feasible.

Another type of LMI optimization problem is the so-called *generalized eigenvalue minimization problem*, which is

$$\begin{aligned} &\text{minimize: } \gamma; \\ &\quad \text{subject to: } F_0(X) + \gamma F_1(X) < Q_0 + \gamma Q_1\,, \\ &\qquad\qquad\quad F_1(X) < Q_1\,, \\ &\qquad\qquad\quad \text{and } X \in \mathcal{X}\,, \end{aligned}$$

where F_1 and F_2 are linear mappings from $\mathcal{X}$ to $\mathbb{H}^n$, and Q_0, $Q_1 \in \mathbb{H}^n$. Linear objective problems can be reformulated in this format to be solved by iteration, which is also the method used in LMI based H_2 and H_∞ synthesis algorithms.

A pictorial view of the linear objective problem is shown in the Figure B.2 [36], for the case $X = [x_1\ x_2]^T \in \mathbb{R}^2$. The convex set depicted in the figure represents the feasibility set $\overline{\mathcal{C}} = \{X : F(X) \preceq Q\}$ for the linear objective problem. Note that while

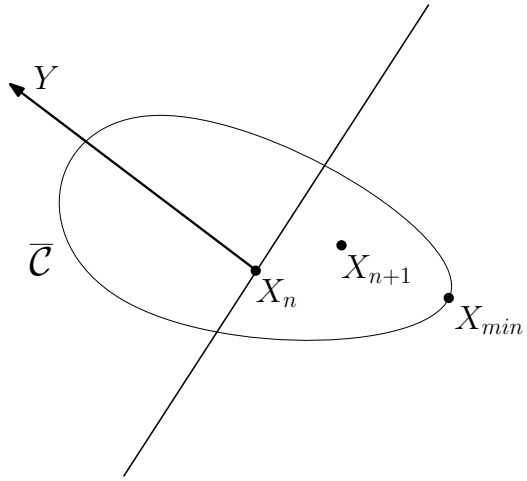

Figure B.2. Illustration of semidefinite programming.

we have drawn a bounded set, this is not necessarily the case.

Since $c(X)$ is a linear functional on $\mathbb{R}^2$, it has the form: $c(X) = Y^*X = y_1x_1 + y_2x_2$, for a fixed vector $Y \in \mathbb{R}^2$. Therefore the point X_{min} that solves the problem is the element $\overline{\mathcal{C}}$ with the most negative projection in the direction of Y, as depicted in Figure (B.2). Also the picture suggests that there are no other local minima for the function in the set; namely, for every point there is so-called *descent direction*. This property, fundamental to *convex optimization* problems, is stated precisely below.

Proposition B.31 [36] *Suppose X_0 is a local minimum of the linear objective problem. That is $c(X_0) \leq c(X)$ for every $X \in \mathcal{N}(X_0) \bigcap \overline{\mathcal{C}}$. Then X_0 is the global minimum of the problem over $\overline{\mathcal{C}}$.*

Proof: Let X_1 be another point in $\overline{\mathcal{C}}$. Since $\overline{\mathcal{C}}$ is convex, it contains the line segment: $L(X_1, X_0) = \{\tau X_1 + (1-\tau)X_0\,,\ \tau \in [0,1]\}$. Also the definition the neighborhood of $\mathcal{N}(X_0)$ will contain the points $\tau X_1 + (1-\tau)X_0$ for sufficiently small τ, say $\tau \in [0, \epsilon)$. Now the function $f\tau = c\big(\tau X_1 + (1-\tau)X_0\big)$ is *affine* in $\tau \in [0,1]$, and $f(0) \leq f(\tau)$ for $\tau \in [0, \epsilon)$ by hypothesis. Then $f(\tau)$ is non-decreasing and $f(0) \leq f(1)$, or equivalently $c(X_0) \leq c(X_1)$. $\square$

The above property generalizes to any convex optimization problem, and plays a strong role in ensuring that these problems can be solved globally, and not just locally, by numerical algorithms. Further discussion on the nature of the semidefinite program-

ming algorithms would take us far away from the field of control theory. It is indeed an interesting feature of this area that theory and computation decouple nicely so that detailed knowledge of the latter is not necessary to appreciate the former. Moreover, good tools for convex optimization are available [49,50,51] for a control engineer who might be quite ignorant on the details of computation. However, we will provide a few remarks aimed at reinforcing the idea that these problems are fundamentally *tractable*.

A first observation is that clearly the minimum, if it exists, must lie on the boundary of the feasible set. Thus one could think of restricting the search to the boundary. This is in fact a popular alternative in *linear programming*, which corresponds to the case where the feasible set is a *polytope*, i.e., the intersection of a finite number of half-spaces. However, in the semidefinite programming problem the boundary is in general quite complicated, and thus methods involving *interior points* are favored.

Refering back to Figure (B.2), suppose that we have a point X_n in our feasible set. An immediate consequence is that we need only keep the set $\{X \in \overline{\mathcal{C}} : c(X) \leq c(X_n)\}$ for our remaining search for the global optimum. This amounts to intersecting $\overline{\mathcal{C}}$ with a half-space. Thus we can progressively shrink the feasibility region to zero, provided that we are able to generate successively a "good" new feasible point X_{n+1}. Many optimization algorithms are based on this principle. One of the simplest is so called *ellipsoid algorithm* that alternates between cutting and overbounding the resulting set by an ellipsoid. Then X_{n+1} would be the center of such an ellipsoid.

More efficient methods for semidefinite programming are based on *barrier functions* to impose the feasibility constraint. The idea is the minimize the function $c(X) + \alpha\phi(X)$ where $\alpha > 0$, and the barrier function $\phi(X)$ is convex and approaches infinity on the boundary of a feasible set.

The barrier function $\phi(X)$ is defined as a mapping from a convex set $\mathcal{C} \in \mathcal{X}$, such that: $\phi : \mathcal{C} \to \mathbb{R}$ is convex if it satisfies; $\phi\big(\tau X_1 + (1-\tau)X_2\big) \leq \tau\phi(X_1) + (1-\tau)\phi(X_2)$. In fact, minimization of such a function is called a *convex optimization* problem. As an important example, the function $\phi(X) = -\log(X)$ is convex in $(0, \infty)$. As must

be clear from the definition, any linear function is convex. Many algorithms use the function $\phi(X) = -\log\big(\det[Q - F(X)]\big)$ as a barrier function [45].

Provided we start from a feasible point, the minimization in $c(X)+\alpha\phi(X)$ can be globally solved by *unconstrained optimization methods* (e.g., *Newton's algorithm*). By successively reducing the weight of the barrier function, an iteration is produced which can be shown to converge to the global minimum. The *computational complexity* of these algorithms has moderate polynomial growth (which is so-called *class* P problems) with problem size, being characterized by the dimensionality of the variable space $\mathcal{X}$ and of the constraint set $\mathbb{H}^n$.

In the eyes of a computer scientist, the word of computation is divided between tractability and intractability, and one deals with easy versus hard problems, *tractable* versus *intractable* problems, and *solvable* versus *unsolvable* problems. Whether a problem is tractable or not is measured by the *efficiency* in computing a solution, that is the demand in computer time and space for solving the problem. Thus a problem may be solvable or *computable* in principle, but may be intractable in a realistic sense that no efficient algorithm may possibly exist, and that solution would require a prohibitive amount of resource unlikely to be found in existing computers. Obviously, for any computation problem, efficiency may vary with individual algorithms one may design, and differs from one instance to another. But such details can neither be fully captured nor are they necessary, and thus it is more meaningful to define efficiency by the intrinsic complexity of a problem. That is, one considers the complexity of the problem in the worst case instance independently of algorithm design, rather than on a case-by-case basis. This would give a fundamental measure of the difficulty, i.e., *computational complexity* in solving a problem, instead of that of the complexity of an algorithm.

When speaking of computational efficiency, it is implicit that certain standard, universal computer is available, which must be intrinsic as well as generic of all computations, and on which algorithms are tested and compared. This universal computer is furnished by the so-called *Turing machines* [54]. A Turing machine is an abstract, theoretical model of computation, consisting of a finite-state processing unit of finite

memory, and a storage unit of unlimited capacity. It operates on a finite set of instructions and attempts to complete a computation, or "halt" in computer science jargon, in finitely many operations. Simply put, the complexity of a decision problem can be measured in terms of the complexity of a deterministic Turing machine that solves it, whereas the complexity of a machine is the cost of computation, such as number of steps, which is taken as a function of the length of input strings to the machine.

For a given machine, let $\sigma(x)$ denote the number of basic operations required for computing a solution from input x to output. Let also n be the size of the input string. Then a Turing machine is said to be a polynomial-time machine if there exists a positive integer q such that for all inputs x, such that: $\sigma(x) = \mathcal{O}(n^q)$ [54].
A decision problem is then said to be in *class* P, or *polynomial-time solvable*, if it is computable on a polynomial time machine. Here a problem is said to be computable if the machine accepts the input string and halts in finitely many operations [54].

The task, however, would seem much simpler if we are just asked to verify a solution, instead of finding it. In other words, we might have been provided with a candidate solution to a decision problem for which we know of no polynomial-time algorithm, and we are given the task to check whether the alleged answer is correct. A machine used as a verifying instead of deciding computer is called a *nondeterministic* Turing machine. The *class* NP is the class of decision problems for which solutions can be solved on a nondeterministic Turing machines. More specifically, a decision problem is in class NP if for each "yes" instance of the problem (the instance that admits a solution) there exists a solution that can be verified by a polynomial-time algorithm, and for each "no" instance of the problem (the instance that does not admit a solution) there exists no solution that can be verified by the algorithm.

A problem $\mathbb{P}_1$ is said to be *polynomial-time reducible*, or polynomially reducible, to a problem $\mathbb{P}_2$ if there exists a deterministic Turing machine on which any instance p_1 in $\mathbb{P}_1$ can be transformed in a polynomial time to an instance p_2 in $\mathbb{P}_2$, so that p_1 admits a solution if and only if p_2 does, and p_1 does not admit a solution if and only if p_2 does not. A problem $\mathbb{P}$ is said to be NP-*hard* if every problem in NP is polynomially

reducible to $\mathbb{P}$, and NP-*complete* if it is NP-hard and $\mathbb{P} \in \mathrm{NP}$ [54].

In other words, a problem $\mathbb{P}_1$ is polynomially reducible to a problem $\mathbb{P}_2$ if it can be "easily" reduced to $\mathbb{P}_2$. It follows that if $\mathbb{P}_2$ is an easy problem, then $\mathbb{P}_1$ must also be easy. Conversely, a "hard" $\mathbb{P}_1$ would mandate a "hard" $\mathbb{P}_2$ [54]. Thus, if we could find a deterministic polynomial time algorithm for one NP-hard problem, then we would have found polynomial-time algorithms for all problem in NP, and if we succeed in finding a polynomial-time algorithm for one NP-complete problem then we would find polynomial-time algorithms for all NP-complete problems, thereby solving all NP-complete problems efficiently.

APPENDIX C: BALANCED MODEL REALIZATIONS

We start with a basic tool for the study of stability, controllability, and observability in linear systems. *Lyapunov equations* come in the two dual forms

$$A^*X + XA + Q = 0\,,$$
$$AX + XA^* + Q = 0\,.$$

In both cases A and Q are given square matrices, and X is the unknown. For concreteness we will focus on the first case, but by the trivial substitution of A by A^*, all the following properties have their dual counterpart. Notice in particular that the Hurwitz property is preserved by the "$*$" operation.

Theorem C.1 [28] *Suppose A and Q are square matrices and that A is Hurwitz. Then*;

$$X = \int_0^\infty e^{A^*\tau} Q e^{A\tau}\, d\tau$$

is the unique solution to the Lyapunov equation $A^*X + XA + Q = 0$.

Proof: Note that the integral converges as A is Hurwitz. It can be verified that the given X is a solution, observing that:

$$A^*X + XA = \int_0^\infty \frac{d}{d\tau}\{e^{A^*\tau} Q e^{A\tau}\}\, d\tau = e^{A^*\tau} Q e^{A\tau}\Big|_0^\infty = -Q$$

It remains to show that it is the unique solution. For this purpose, define the linear mapping $\mathcal{F} : \mathbb{R}^{n\times n} \to \mathbb{R}^{n\times n}$ by $\mathcal{F}(X) := A^*X + XA$, where $X \in \mathbb{R}^{n\times n}$.

For any matrix Q we have verified that the equation $\mathcal{F}(X) = -Q$ has a solution. This means that the dimension of the image of $\mathcal{F}$ is n^2. But the dimension of the domain of $\mathcal{F}$ is also n^2 and therefore $\mathrm{Ker}\mathcal{F} = 0$. This means that $\mathcal{F}(X) = -Q$ has a unique solution for each Q. $\square$

In terms of computation, Lyapunov equation is linear in the matrix variable X, and therefore it is in essence no more than a system of n^2 entries of X, which can be solved by standard techniques. In particular when A is Hurwitz we are guaranteed to have a unique solution, hence direct computation of the integral in Theorem C.1 is never required. We will focus on the case where $Q = Q^*$, for which one always seeks solutions $X = X^*$ to the Lyapunov equations.

Proposition C.2 [28] *Suppose $Q \succ 0$. Then A is Hurwitz, if and only if, there exists a solution $X \succ 0$ to the Lyapunov equation $A^*X + XA + Q = 0$*
Proof: We first establish the "only if" direction. Since $\mathcal{A}$ is Hurwitz we already know that the unique solution is

$$X = \int_0^\infty e^{A^*\tau} Q e^{A\tau}\, d\tau\,.$$

Since $Q \succ 0$, it is straightforward to verify that $X \succ 0$. For the converse, suppose $X \succ 0$ solves the equation. If λ is an eigenvalue of A, with eigenvector $v \neq 0$, we have

$$0 = v^*(A^*X + XA + Q)v = \lambda^* v^* X v + \lambda v^* X v + v^* Q v\,.$$

Since $v^*Xv > 0$, we have

$$2\mathrm{Re}(\lambda) = -\frac{v^*Qv}{v^*Xv} < 0\,, \quad \text{so } A \text{ is Hurwitz.} \qquad \square$$

The preceding result remains true if the hypothesis $Q \succ 0$ is weakened to $Q = C^*C \succeq 0$, with (C, A) observable. In that case, Lyapunov solution is denoted by Y_o, satisfying

$$A^*Y_o + Y_oA + C^*C = 0\,,$$

and is called the *observability gramian* of (C, A).

Proposition C.3 [28] *Suppose that A is Hurwitz, and X_0 satisfies $A^*X_0+X_0A+Q=0$, where Q is a symmetric matrix. If X satisfies $A^*X+XA+Q \preceq 0$, then $X \succ X_0$.*

Proof: Subtracting the equality from the inequality leads

$$\begin{aligned} &A^*X+XA+Q\ -(A^*X_0+X_0A+Q) \preceq 0 \\ &A^*(X-X_0)+(X-X_0)A \preceq 0 \end{aligned}$$

Since A is Hurwitz, this is equivalent to $X-X_0 \succeq 0 \Rightarrow X \succeq X_0$. □

An autonomous and stable LTI system is given by

$$\begin{aligned} &\dot{x}(t)=Ax(t)\,, \quad x(0)=x_0 \in \mathbb{C}^n \\ &y(t)=Cx(t)\,, \end{aligned}$$

where A is a Hurwitz matrix. The solution to this system is $Ce^{At}x_0$ for $t \geq 0$.
We define the *observability operator* $\Psi_o : \mathbb{C}^n \to L_2[0,\infty)$ by

$$x_0 \to \begin{cases} Ce^{At}x_0\,, & \text{for } t \geq 0\,, \\ 0\,, & \text{otherwise}\,. \end{cases}$$

Note that as A is Hurwitz, $y \in L_2[0,\infty)$ and Ψ_o is a bounded operator.
The energy of $y=\Psi_o x_0$, is given by; $\|y\|_2 = \big(\langle \Psi_o x_0\,, \Psi_o x_0\rangle\big)^{1/2} = \big(\langle x_0\,, \Psi_o^*\Psi_o x_0\rangle\big)^{1/2}$, where $\Psi_o^* : L_2[0,\infty) \to \mathbb{C}^n$ is the adjoint of Ψ_o. For each $z \in L_2[0,\infty)$ and $x_0 \in \mathbb{C}^n$

$$\begin{aligned} \langle \Psi_o^* z\,, x_0\rangle_{\mathbb{C}^n} &= \langle z\,, \Psi_o x_0\rangle_{L_2} = \int_0^\infty z(\tau)^* Ce^{A\tau}x_0\,d\tau \\ &= \left(\int_0^\infty e^{A^*\tau}C^*z(\tau)\,d\tau\right)^* x_0 = \left\langle \int_0^\infty e^{A^*\tau}C^*z(\tau)\,d\tau, x_0\right\rangle_{\mathbb{C}^n}. \end{aligned}$$

Thus we see that Ψ_o^* is given by

$$\Psi_o^* z = \int_0^\infty e^{A^*\tau}C^*z(\tau)\,d\tau\,, \quad \text{for } z \in L_2[0,\infty)\,.$$

The 2-norm of $y = \Psi_o x_0$ is given by $\left(\langle x_0\,,\Psi_o^*\Psi_o x_0\rangle\right)^{1/2}$ and we have:

$$\begin{aligned}\Psi_o^*(\Psi_o x_0) &= \int_0^\infty e^{A^*\tau}C^*Ce^{A\tau}x_0\,d\tau \\ &= \left(\int_0^\infty e^{A^*\tau}C^*Ce^{A\tau}\,d\tau\right)x_0 = (\Psi_o^*\Psi_o)x_0\,.\end{aligned}$$

Therefore, the operator $\Psi_o^*\Psi_o$ is given by the matrix

$$Y_o := \Psi_o^*\Psi_o = \int_0^\infty e^{A^*\tau}C^*Ce^{A\tau}\,d\tau\,,$$

which is the *observability gramian* of (C, A). In fact, using Theorem C.1 it is clear that Y_o satisfies the Lyapunov equation. This gramian is positive semidefinite ($Y_o \geq 0$), and it is *positive definite* ($Y_o > 0$) when (C, A) is *observable.*

However, the previous calculation shows that the gramian carries quantitative information about observability; this is in addition to simply indicating whether the matrix pair is observable. We have shown that the energy of the output $y = \Psi_o x_0$, starting from an initial condition $x_0 \in \mathbb{C}^n$, is given by; $\|y\|_2 = (x_0^* Y_o x_0)^{1/2}$. In particular, if we only consider states that satisfy $|x_0| = 1$, then clearly some of these states will yield higher output norms $\|y\|_2$ than others. *Hence, the observability gramian measures "how observable" a given initial condition is.*

This idea can be described geometrically in a very intuitive way. Consider the vector $Y_o^{1/2}x_0$, where x_0 has unit length. This vector has length equal to the 2-norm of y, and as we will see, contains directional information as well. Defining the set

$$\varepsilon_0 := \{Y_o^{1/2}x_0 : \, x_0 \in \mathbb{C}^n \text{ and } |x_0| = 1\}\,,$$

which is the collection of all such vectors.

Since Y_o is positive semidefinite, this set is an ellipsoid and is depicted two dimensionally in Figure C.1.

This shape is called the *observability ellipsoid*, and tells us the output norm associated with a particular direction in state-space. Let $\eta_1 \geq \eta_2 \geq \cdots \geq \eta_n \geq 0$ be the

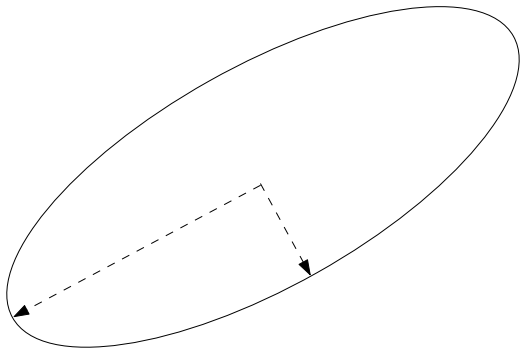

Figure C.1. Observability ellipsoid.

eigenvalues of $Y_o^{1/2}$ and $v_1, ..., v_n$ their respective unit-norm eigenvectors. Then the v_k give the directions of the principle axes of the ellipsoid, and the η_k the length in each axis. If $\eta_k = 0$ for some k, the corresponding v_k produces no output energy and is therefore unobservable. In fact the span of all vectors v_k with $\eta_k = 0$ is precisely the unobservable subspace $\mathcal{N}_{CA}$. If η_k, η_l are both nonzero but, for example $\eta_k \gg \eta_l$, then the output energy resulting from initial state v_l is much smaller than that observed when the initial condition is v_k. Thus both states are observable, but intuitively state v_k is "more observable" than state v_l. Hence, observability gramian provides us with a way to assess the relative observability of various directions in the state-space. When a state corresponds to a large eigenvalue η_k we say it is *strongly observable*. If, $\eta_k > 0$ is small compared with other eigenvalues we say that the state is *weakly observable*.

Given matrices (A, B), we define the *controllability* operator $\Psi_c : L_2(-\infty, 0] \to \mathbb{C}^n$ by

$$u \to \int_{-\infty}^{0} e^{A\tau} Bu(\tau)\, d\tau\,.$$

Notice that this can be thought of as the response of the system described by:

$$\dot{x}(t) = Ax(t) + Bu(t)\,, \quad x(-\infty) = 0$$

to an input function $u \in L_2(-\infty, 0]$, where the output is the vector $x(0)$.
To see the significance of this operator, we ask the question: *Given the final condition $x(0) \in \mathbb{C}^n$ with unit norm, $|x_0| = 1$, what $u \in L_2(-\infty, 0]$ solves $\Psi_c u = x_0$ with the*

smallest norm $\|u\|_2$ *?*

Clearly, if the matrix pair (A, B) is not controllable it may not be possible even to satisfy $\Psi_c u = x_0$, so for the next result we focus on the case where the pair is controllable.

Proposition C.4 [36] *Suppose that* (A, B) *is controllable. Then;*

i) the matrix $\Psi_c \Psi_c^* =: X_c$ *is nonsingular,*

ii) for any $x_0 \in \mathbb{C}^n$ *the input* $u_{opt} = \Psi_c^* X_c^{-1} x_0$ *is the element of maximum norm in the set* $\{u \in L_2(-\infty, 0],\ \Psi_c u = x_0\}$ *.*

The proposition says that if we want to reach a state x_0, then $u_{opt} = \Psi_c^* X_c^{-1} x_0$ is the most economical input in terms of energy. This input energy is given by

$$\begin{aligned} \|u_{opt}\|_2^2 &= \langle \Psi_c^* X_c^{-1} x_0 \,, \Psi_c^* X_c^{-1} x_0 \rangle \\ &= \langle X_c^{-1} x_0 \,, \Psi_c \Psi_c^* X_c^{-1} x_0 \rangle \\ &= x_0^* X_c^{-1} x_0 \,. \end{aligned}$$

We now provide a geometric interpretation of the controllability gramian, similar to that for Y_o. The key question is; *what are all the final states* $x_0 = \Psi_c u$ *that can result from an input* $u \in L_2(-\infty, 0]$ *of unit norm ?* The answer is given in the following proposition, which holds even if the pair (A, B) is not controllable.

Proposition C.5 [36] *The following sets are equal:*

(i) $\{\Psi_c u : u \in L_2(-\infty, 0]$ *and* $\|u\|_2 \leq 1\}$ *, (ii)* $\{X_c^{1/2} x_c : x_c \in \mathbb{C}^n$ *and* $|x_c| \leq 1\}$ *.*

This result says that all the states reachable with u satisfying $\|u\|_2 \leq 1$ are given by; $X_c^{1/2} x_c$, where $|x_c| \leq 1$. Notice the norm squared of any such state is given by;

$x_c^* X_c x_c$. We define the *controllability ellipsoid* by

$$\varepsilon_c = \{X_c^{1/2} x_c : x_c \in \mathbb{C}^n \text{ and } |x_c| = 1\} .$$

This gives us the boundary of the set given in the proposition.

Let $\eta_1 \geq \eta_2 \geq \cdots \geq \eta_n \geq 0$, be the eigenvalues of $X_c^{1/2}$ and $v_1, ..., v_n$ their corresponding orthonormal eigenvectors. Then the principle axes of the ellipsoid ε_c are given by; $\eta_k v_k$. The unit vectors $v_1, ..., v_n$ give a basis for the state-space and the values $\eta_1, ..., \eta_n$ tells us that given an input $\|u\|_2 = 1$, the largest we can make a state in the direction v_k is η_k. Thus we conclude that if $\eta_k = 0$, then v_k is an unreachable state. Similarly, if $\eta_k \gg \eta_l$ then the state direction v_k is "more" controllable than direction v_l. As with observability, this gives rise to the terms *strong* and *weak controllability.*

We have explored the geometric ways, in terms of the observability and controllability gramians introduced, to assess which directions in the state-space are strongly and weakly controllable and observable. We now consider our usual system given by

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \qquad \text{for } t \geq 0 . \end{aligned}$$

Here the direct feed-through term given by the matrix D is not included, since it makes no difference to our investigation. We are interested in finding a natural basis for this state-space that gives some idea as to which states dominate the system behavior. So far, we have concluded that states corresponding to small eigenvectors of the observability gramian are not very observable. However this does not mean that such states do not contribute to the input-output behavior of the above system, since such states may be very controllable. This phenomenon is easily appreciated by looking at Figure C.2, which shows both the observability and controllability ellipsoids in two dimensions on the same plot. As the picture shows, these two ellipsoids need not be aligned, and therefore we can have a situation as drawn here, where the major and minor axes of

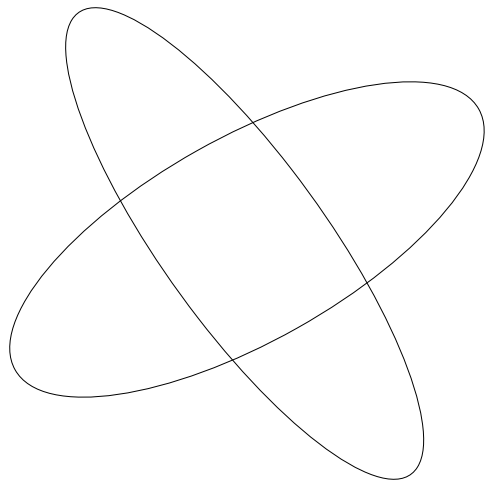

Figure C.2. Unbalanced system ellipsoids

the ellipsoids are nearly opposite. Therefore we reason that the most intuition would be gained about the system if the controllability and observability ellipsoids were exactly aligned. This raises the question of whether or not it is possible to arrange such alignment by means of a state transformation.

A change of basis to the state-space of the above system yields the familiar transformed realization

$$\tilde{A} = TAT^{-1}, \qquad \tilde{B} = TB\,, \quad \text{and} \quad \tilde{C} = CT^{-1},$$

where T is a state transformation.
The controllability gramian associated with this new realization is

$$\begin{aligned}\tilde{X}_c &= \int_0^\infty e^{\tilde{A}\tau}\tilde{B}\tilde{B}^* e^{\tilde{A}^*\tau} d\tau \\ &= \int_0^\infty Te^{A\tau}T^{-1}TBB^*T^*(T^*)^{-1}e^{A^*\tau}T^* d\tau \\ &= TX_cT^*.\end{aligned}$$

Similarly the observability gramian in the transformed basis is

$$\tilde{Y}_o = (T^*)^{-1}Y_oT^{-1}.$$

The above are called *congruence transformations*, which typically arise from quadratic forms under a change of basis. The following result concerns *simultaneous diagonalization* of positive definite matrices by this kind of transformation.

Proposition C.6 [36] *Given positive definite matrices X and Y, there exists a nonsingular matrix T such that*

$$TXT^* = (T^*)^{-1}YT^{-1} = \Sigma\,,$$

where Σ *is a diagonal, positive definite matrix.*

Proof: Perform a singular value decomposition of the matrix $X^{1/2}YX^{1/2}$ to get; $X^{1/2}YX^{1/2} = U\Sigma^2U^*$, where U is unitary and Σ is diagonal, positive definite. Therefore we get; $\Sigma^{-1/2}U^*X^{1/2}YX^{1/2}U\Sigma^{-1/2} = \Sigma$. Now set $T^{-1} = X^{1/2}U\Sigma^{-1/2}$ and the above states that $(T^{-1})^*YT^{-1} = \Sigma$ and also $TXT^* = (\Sigma^{1/2}U^*X^{-1/2})X(X^{-1/2}U\Sigma^{1/2}) = \Sigma$. $\square$

Applying this proposition to the gramians, the following conclusion can be drawn.

Corollary C.7 *Suppose (A,B,C) is a controllable and observable realization. Then there exists a state transformation T such that the equivalent realization* $(\tilde{A}, \tilde{B}, \tilde{C}) = (TAT^{-1}, TB, CT^{-1})$ *satisfies;* $\tilde{X}_c = \tilde{Y}_o = \Sigma$, *with* $\Sigma > 0$ *diagonal.*

A state-space realization such that the controllability and observability gramians are equal and diagonal is called a *balanced realization.* The controllability and observability ellipsoids for the system are exactly aligned when a system is balanced. Thus the states which are the least controllable are also the least observable. An extension to the Proposition C.5 is stated below, which covers the general case where the original realization is not necessarily controllable or observable.

Proposition C.8 [36] *Given positive semidefinite matrices X and Y. There exists a nonsingular matrix T such that*

$$(a)\ TXT^* = \begin{bmatrix} \Sigma_1 & & & \\ & \Sigma_2 & & \\ & & 0 & \\ & & & 0 \end{bmatrix}, \qquad (b)\ (T^*)^{-1}YT^{-1} = \begin{bmatrix} \Sigma_1 & & & \\ & 0 & & \\ & & \Sigma_3 & \\ & & & 0 \end{bmatrix},$$

where the matrices Σ_k are diagonal and positive definite.

When applied to the gramians X_c and Y_o, we find that if the system is either uncontrollable or unobservable, then each of the Σ_k blocks of $\tilde{Y}_o$ and $\tilde{X}_c$ have the following interpretation:

i) Σ_1 captures the controllable and observable states,
ii) Σ_2 captures the controllable and unobservable states,
iii) Σ_3 captures the uncontrollable and observable states.

Under such a transformation the state matrix $\tilde{A}$ has the form

$$\tilde{A} = \begin{bmatrix} \tilde{A}_1 & 0 & \tilde{A}_6 & 0 \\ \tilde{A}_2 & \tilde{A}_3 & \tilde{A}_4 & \tilde{A}_5 \\ 0 & 0 & \tilde{A}_7 & 0 \\ 0 & 0 & \tilde{A}_8 & \tilde{A}_9 \end{bmatrix},$$

which is so-called *Kalman decomposition* [28]. This can be verified from the various invariance properties of the controllability and observability subspaces.

APPENDIX D: PROOFS

D.1. Proof of Theorem 3.3

(i) $\varrho(M) \leq \mu_{\mathbf{\Delta}}(M)$:

$$\begin{aligned}\varrho(M) &= \sup\{|\lambda| : \det(\lambda I - M) = 0\} = \frac{1}{\inf\{\frac{1}{|\lambda|} : \det(\lambda I - M) = 0\}} \\ &= \frac{1}{\inf\{|k| : \det(\frac{1}{k} I - M) = 0\}} = \frac{1}{\inf\{|k| : \det(I - kM) = 0\}} \\ &= \frac{1}{\inf\{\overline{\sigma}(kI) : \det(I - MkI) = 0\}} \leq \frac{1}{\inf_{\Delta}\{\overline{\sigma}(\Delta) : \det(I - M\Delta) = 0\}}\end{aligned}$$

since $kI \in \Delta$.

(ii) $\mu_{\mathbf{\Delta}}(M) \leq \overline{\sigma}(M)$:

First note that: $\underline{\sigma}(I - M\Delta) \geq \overline{\sigma}(I) - \overline{\sigma}(M\Delta)$ (see Appendix B.2). However;

$$\begin{aligned}&\overline{\sigma}(I) - \overline{\sigma}(M\Delta) = 1 - \sup_{x \neq 0} \frac{\|M\Delta x\|}{\|x\|} \\ &= 1 - \sup_{x\,,\,\Delta x \neq 0} \left(\frac{\|M\Delta x\|}{\|\Delta x\|} \cdot \frac{\|\Delta x\|}{\|x\|} \right) \geq 1 - \sup_{\Delta x \neq 0} \frac{\|M\Delta x\|}{\|\Delta x\|} \cdot \sup_{x \neq 0} \frac{\|\Delta x\|}{\|x\|} = 1 - \overline{\sigma}(M)\overline{\sigma}(\Delta)\end{aligned}$$

Hence; $\underline{\sigma}(I - M\Delta) \geq 1 - \overline{\sigma}(M)\overline{\sigma}(\Delta)$. Since $\det(I - M\Delta) = 0$, if and only if, $\underline{\sigma}(I - M\Delta)$. Then for any $\Delta \in \{\det(I - M\Delta) = 0\}$, we have; $1 - \overline{\sigma}(M)\overline{\sigma}(\Delta) < 0 \Rightarrow \frac{1}{\overline{\sigma}(\Delta)} \leq \overline{\sigma}(M)$. This is valid for any Δ which makes $\det(I - M\Delta) = 0$. Therefore;

$$\frac{1}{\inf_{\Delta}\{\overline{\sigma}(\Delta) : \det(I - M\Delta) = 0\}} = \mu_{\mathbf{\Delta}}(M) \leq \overline{\sigma}(M) \qquad \square$$

D.2. Proof of Lemma 3.4

(i) $\varrho(M) = \mu_{\mathbf{\Delta}}(M)$:

Obvious from the proof of $\varrho(M) \leq \mu_{\mathbf{\Delta}}(M)$ for the general case.
(ii) $\mu_{\mathbf{\Delta}}(M) = \overline{\sigma}(M)$:

We have already proved that $\mu_{\mathbf{\Delta}}(M) \leq \overline{\sigma}(M)$ for any Δ. We need to prove the opposite inequality $\mu_{\mathbf{\Delta}}(M) \geq \overline{\sigma}(M)$ for $\Delta = \mathbb{C}^{n\times n}$. Constructing a $\Delta^* \in \mathbb{C}^{n\times n}$ for which; $\det(I - M\Delta^*) = 0$, and $\frac{1}{\overline{\sigma}(\Delta^*)} = \overline{\sigma}(M)$ would lead;

$$\frac{1}{\inf_{\Delta \in \mathbb{C}^{n\times n}}\{\overline{\sigma}(\Delta) : \det(I - M\Delta) = 0\}} \geq \mu_{\mathbb{C}^{n\times n}}(M) \qquad \square$$

D.3. Proof of Theorem 3.5

For robust stability and performance the requirements are (according to Figure 3.4):
(i) $\mu_{\mathbf{\Delta}}\big(M_{qp}(j\omega)\big) < 1 \quad \forall\omega \in \mathbb{R}_e$, and (ii) $\|N\|_\infty = \|F_u(M,\Delta)\|_\infty < \gamma \quad \forall\Delta \in \mathcal{B}_{\mathbf{\Delta}}$.

(ii) is equivalent to; $\mu_{\mathbf{\Delta}f}\big(F_u(M,\Delta)\big) < \gamma \quad \Leftrightarrow \quad \mu_{\mathbf{\Delta}f}\big(\frac{1}{\gamma}F_u(M,\Delta)\big) < 1 \quad \forall\omega \in \mathbb{R}_e$, where $\Delta f = \mathbb{C}^{n\times n}$.
Considering the last μ-condition:

$$\det\left(I - \frac{1}{\gamma}F_u(M,\Delta)\Delta_f\right) = 0 \;\Leftrightarrow\; \det\left(I - \big(\frac{1}{\gamma}M_{zw} + \frac{1}{\gamma}M_{zp}\Delta(I - M_{qp}\Delta)^{-1}M_{qw}\big)\Delta_f\right) = 0$$

with $\Delta_f \in \mathcal{B}_{\mathbf{\Delta}f}$. Note that the last determinant condition can be written as:

$$\det\left(I - \frac{1}{\gamma}M_{zw}\Delta_f - \frac{1}{\gamma}M_{zp}\Delta(I - M_{qp}\Delta)^{-1}M_{qw}\Delta_f\right) = 0$$

If $\det(I - M_{qp}\Delta) \neq 0 \ \forall \Delta \in \mathcal{B}_{\mathbf{\Delta}}$ (which is the requirement for robust stability), and using the identity; $\det \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \det(A)\det(D - CA^{-1}B)$, the last determinant condition becomes:

$$\det \begin{pmatrix} I - M_{qp}\Delta & -M_{qw}\Delta_f \\ -\frac{1}{\gamma}M_{zp}\Delta & I - \frac{1}{\gamma}M_{zw}\Delta_f \end{pmatrix} = \det \left(\begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix} - \begin{pmatrix} M_{qp} & M_{qw} \\ M_{zp}/\gamma & M_{zw}/\gamma \end{pmatrix} \begin{pmatrix} \Delta & 0 \\ 0 & \Delta_f \end{pmatrix} \right) = 0$$

which means that $\det(I - \tilde{M}\Delta_e) = 0$, where $\Delta_e := \begin{pmatrix} \Delta & 0 \\ 0 & \Delta_f \end{pmatrix}$. $\square$

D.4. Proof of Corollary 3.7

To show the "only if" part; by the hypothesis, the controllability gramian X_c satisfies; $\mathrm{trace}(\mathcal{C}X_c\mathcal{C}^*) < \beta^2$, and $\mathcal{A}$ is Hurwitz. Now the expression

$$\mathcal{X} = \int_0^\infty e^{\mathcal{A}t}(\mathcal{B}\mathcal{B}^* + \epsilon I)e^{\mathcal{A}^*t}\, dt\,,$$

is continuous in ϵ, and equal to X_c when $\epsilon = 0$ (see Appendix C). Therefore, for some $\epsilon > 0$, we have $\mathrm{trace}(\mathcal{C}\mathcal{X}\mathcal{C}^*) < \beta^2$. This matrix $\mathcal{X}$ clearly satisfies the Lyapunov equation $\mathcal{A}\mathcal{X} + \mathcal{X}\mathcal{A}^* + \mathcal{B}\mathcal{B}^* + \epsilon I = 0$, which means that the inequality of the claim is met.

The "if" direction follows the same idea. By assumption there exists $\mathcal{X}$ which is a positive definite solution to $\mathcal{A}\mathcal{X} + \mathcal{X}\mathcal{A}^* + \mathcal{B}\mathcal{B}^* \prec 0$, and satisfies $\mathrm{trace}(\mathcal{C}\mathcal{X}\mathcal{C}^*) < \beta^2$. Thus A is necessarily Hurwitz, and $\mathcal{X}$ is a generalized controllability gramian. As shown in Proposition C.3 in Appendix C, $\mathcal{X} \succeq X_c$ where X_c is the controllability gramian. This implies that; $\mathrm{trace}(\mathcal{C}\mathcal{X}\mathcal{C}^*) \geq \mathrm{trace}(\mathcal{C}X_c\mathcal{C}^*)$ and thus $\|M\|_2 < \beta$. $\square$

D.5. Proof of Theorem 3.9

The closed-loop system has a H_2 norm less than β, if and only if, there exists an $\mathcal{X} = \mathcal{X}^T \succ 0$ such that $\mathcal{A}\mathcal{X} + \mathcal{X}\mathcal{A}^T + \mathcal{B}\mathcal{B}^T \prec 0$, and $\mathrm{trace}(\mathcal{C}\mathcal{X}\mathcal{C}^T) < \beta^2$ (see Theorem

B. in Appendix B.2).

The trace inequality is equivalent to the existence of a $Z \succ 0$ such that: $\mathcal{C}\mathcal{X}\mathcal{C}^T \prec Z$ and $\text{trace}(Z) < \beta^2$. Equivalently, using the Schur complement formula and a congruence transformation; $\begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}^T \begin{bmatrix} * \end{bmatrix} \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}$ (see Appendix B.2), trace inequality entails to:

$$\begin{bmatrix} \mathcal{X}^{-1} & \mathcal{C}^T \\ \mathcal{C} & Z \end{bmatrix} \succ 0 \quad \text{and} \quad \text{trace}(Z) < \beta^2 \,.$$

The first inequality $\mathcal{A}\mathcal{X} + \mathcal{X}\mathcal{A}^T + \mathcal{B}\mathcal{B}^T \prec 0$ can also be written as (by the Schur complement):

$$\begin{bmatrix} \mathcal{A}^T\mathcal{X}^{-1} + \mathcal{X}^{-1}\mathcal{A} & \mathcal{X}^{-1}\mathcal{B} \\ \mathcal{B}^T\mathcal{X}^{-1} & -I \end{bmatrix} \prec 0$$

Let $\mathcal{Y} := \mathcal{X}^{-1}$. Then, we obtain the LMIs:

$$\begin{bmatrix} \mathcal{A}^T\mathcal{Y} + \mathcal{Y}\mathcal{A} & \mathcal{Y}\mathcal{B} \\ \mathcal{B}^T\mathcal{Y} & -I \end{bmatrix} \prec 0 \,, \qquad \begin{bmatrix} \mathcal{Y} & \mathcal{C}^T \\ \mathcal{C} & Z \end{bmatrix} \succ 0 \,, \quad \text{and} \quad \text{trace}(Z) < \beta^2 \,.$$

Without loss of generality, one can assume $\mathcal{Y} = \begin{bmatrix} Y & -Y \\ -Y & S^{-1} + Y \end{bmatrix}$.

If we apply the congruence transformation; $\mathcal{T}^T(*)\mathcal{T}$ to both of the linear matrix inequalities, with;

$\mathcal{T} := \begin{bmatrix} X & I & 0 \\ S & 0 & 0 \\ 0 & 0 & I \end{bmatrix}$, where $X := S + Y^{-1}$, and defining;

$$
\begin{aligned}
E &:= C_K S\,, \\
F &:= -Y B_K\,, \\
G &:= Y(A - B_K C_2)X + Y(B_2 C_K - A_K)S\,,
\end{aligned}
$$

we can obtain the LMI conditions of Theorem 3.4. in terms of the variables X, Y, E, F, and G. $\square$

D.6. Proof of Theorem 3.10

We can rewrite the condition $\|M\|_\infty < \gamma$ as follows:

$$
\begin{aligned}
\|M\|_\infty = \sup_{\omega \in \mathbb{R}} \overline{\sigma}\big(M(j\omega)\big) < \gamma \iff & \sup_{\omega \in \mathbb{R}} \lambda_{max}\big(M(j\omega)^* M(j\omega)\big) < \gamma^2 \\
\iff & \lambda_{max}\big(M(j\omega)^* M(j\omega)\big) < \gamma^2 \quad \forall \omega \in \mathbb{R}_e \\
\iff & M(j\omega)^* M(j\omega) \prec \gamma^2 I \quad \forall \omega \in \mathbb{R}_e \\
\iff & \gamma^{-1} M(j\omega)^* M(j\omega) - \gamma I \prec 0 \quad \forall \omega \in \mathbb{R}_e \\
\iff & \begin{bmatrix} M(j\omega)^* & I \end{bmatrix} \begin{bmatrix} \gamma^{-1} I & 0 \\ 0 & -\gamma I \end{bmatrix} \begin{bmatrix} M(j\omega) \\ I \end{bmatrix} \prec 0 \quad \forall \omega \in \mathbb{R}_e \\
\iff & \begin{bmatrix} M(j\omega) \\ I \end{bmatrix}^* \begin{bmatrix} \gamma^{-1} I & 0 \\ 0 & -\gamma I \end{bmatrix} \begin{bmatrix} M(j\omega) \\ I \end{bmatrix} \prec 0 \quad \forall \omega \in \mathbb{R}_e
\end{aligned}
$$

Note that we can write;

$$
\begin{bmatrix} M(s) \\ I \end{bmatrix} = \left[\begin{array}{c} \left[\begin{array}{c|c} \mathcal{A} & \mathcal{B} \\ \hline \mathcal{C} & \mathcal{D} \end{array}\right] \\ I \end{array}\right] = \left[\begin{array}{c|c} \mathcal{A} & \mathcal{B} \\ \hline \mathcal{C} & \mathcal{D} \\ 0 & I \end{array}\right] =: \mathbf{M}(s)
$$

Let $\Gamma := \begin{bmatrix} \gamma^{-1}I & 0 \\ 0 & -\gamma I \end{bmatrix}$. Then $\|M\|_\infty < \gamma$, if and only if:

$$\mathbf{M}(j\omega)^*\Gamma\mathbf{M}(j\omega) \prec 0 \quad \forall\omega \in \mathbb{R}_e\,.$$

Applying *KYP Lemma* (Theorem 3.8), with

$$\left[\begin{array}{c|c} \mathbf{A} & \mathbf{B} \\ \hline \mathbf{C} & \mathbf{D} \end{array}\right] = \left[\begin{array}{c|c} \mathcal{A} & \mathcal{B} \\ \hline \mathcal{C} & \mathcal{D} \\ 0 & I \end{array}\right],$$

we obtain:

$$\begin{bmatrix} \mathcal{A}^T\mathcal{Y} + \mathcal{Y}\mathcal{A} + \gamma^{-1}\mathcal{C}^T\mathcal{C} & \mathcal{Y}\mathcal{B} + \gamma^{-1}\mathcal{C}^T\mathcal{D} \\ \mathcal{B}^T\mathcal{Y} + \gamma^{-1}\mathcal{D}^T\mathcal{C} & -\gamma I + \gamma^{-1}\mathcal{D}^T\mathcal{D} \end{bmatrix} \prec 0\,, \quad \text{with} \quad \mathcal{Y} = \mathcal{Y}^T \succ 0\,.$$

Unfortunately, this inequality is not linear in γ, and therefore we can not minimize γ using this inequality to compute $\|M\|_\infty$ with arbitrary accuracy. However, rewriting the inequality in a more explicit form;

$$\begin{bmatrix} \mathcal{A}^T\mathcal{Y} + \mathcal{Y}\mathcal{A} & \mathcal{Y}\mathcal{B} \\ \mathcal{B}^T\mathcal{Y} & -\gamma I \end{bmatrix} + \gamma^{-1}\begin{bmatrix} \mathcal{C}^T \\ \mathcal{D}^T \end{bmatrix}\begin{bmatrix} \mathcal{C} & \mathcal{D} \end{bmatrix} \prec 0\,,$$

and using the Schur complement formula (see Appendix B.2), leads to the LMI in Theorem 3.10. □

D.7. Proof of Lemma 3.11

Applying Theorem 3.10, the closed-loop system M is stable with an H_∞ performance γ if and only if there exists a $\mathcal{Y} = \mathcal{Y}^T \succ 0$ such that:

$$\begin{bmatrix} \mathcal{A}^T\mathcal{Y} + \mathcal{Y}\mathcal{A} & \mathcal{Y}\mathcal{B} & \mathcal{C}^T \\ \mathcal{B}^T\mathcal{Y} & -\gamma I & \mathcal{D}^T \\ \mathcal{C} & \mathcal{D} & -\gamma I \end{bmatrix} \prec 0\,.$$

We can assume, without loss of generality, that; $\mathcal{Y} = \begin{bmatrix} Y & -Y \\ -Y & Y + S^{-1} \end{bmatrix}$, for some $S = S^T \succ 0$.

If we apply the congruence transformations (see Appendix B.2):

$$\begin{bmatrix} \mathcal{T} & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}^T \begin{bmatrix} \mathcal{A}^T\mathcal{Y} + \mathcal{Y}\mathcal{A} & \mathcal{Y}\mathcal{B} & \mathcal{C}^T \\ \mathcal{B}^T\mathcal{Y} & -\gamma I & \mathcal{D}^T \\ \mathcal{C} & \mathcal{D} & -\gamma I \end{bmatrix} \begin{bmatrix} \mathcal{T} & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \prec 0 \quad \text{and} \quad \mathcal{T}^T\mathcal{Y}\mathcal{T} \succ 0\,,$$

where; $\mathcal{T} := \begin{bmatrix} X & I \\ S & 0 \end{bmatrix}$, $\quad X := S + Y^{-1}$.

Carrying out the multiplications and defining:

$$\begin{aligned} E &:= D_K C_2 X + C_K S\,, \\ F &:= Y(B_2 D_K - B_K)\,, \\ G &:= D_K\,, \\ H &:= Y(A + B_2 D_K C_2 - B_K C_2)X + Y(B_2 C_K - A_K)S\,, \end{aligned}$$

we can obtain the LMI formulations of Lemma 3.11 in terms of the variables X, Y, E, F, G, and H. $\square$

D.8. Proof of Theorem 3.22

Choose a Lyapunov function of the form $V(t,x) = x^T I x$.
Then, $\dot{V}(t,x) = x^T[\mathcal{A}(t) + \mathcal{A}(t)^T]x \leq \lambda_{max}(t)\, x^T x = \lambda_{max}(t) V(t,x) \leq -\alpha\, V(t,x)\,,$
as $[\mathcal{A}(t) + \mathcal{A}(t)^T]$ is always symmetric (see Appendix B.2).
Hence; $V(t,x) \leq V(0,x)\, e^{-\alpha t} \;\Rightarrow\; \lim_{t\to\infty} V(t,x) = 0 \;\Rightarrow\; \lim_{t\to\infty} \|x\|_2 = 0\,.$ □

D.9. Proof of Theorem 3.24

Let $\rho(t)$ be a *piecewise continuous* function mapping $\mathbb{R}^+ \to \mathcal{P}$.
Then, $\tilde{\mathcal{A}}(t) := \mathcal{A}\big(\rho(t)\big)$ is a also piecewise continuous mapping $\mathbb{R}^+ \to \mathbb{R}^{n\times n}$.
Therefore, for the given $\rho(t)$, system admits a unique solution. □

D.10. Proof of Lemma 3.25

Pre- and post-multiply (3.43) by $\mathcal{X}^{-1}$, then let $\mathcal{X}^{-1} = \mathcal{Y}$. □

D.11. Proof of Theorem 3.29

Let $\rho(t)$ be any parameter realization satisfying the hypothesis of the theorem. Then $\tilde{\mathcal{A}}(t) := \mathcal{A}\big(\rho(t)\big) \in (\mathbb{R}^+, \mathbb{R}^{n\times n})$. Note that from Theorem B.30 in Appendix B.3, inequality (3.52) is equivalent to:

$$\mathcal{A}^T(\rho)\mathcal{X}(\rho) + \mathcal{X}(\rho)\mathcal{A}(\rho) + \sum_{i=1}^{m} \frac{\partial \mathcal{X}(\rho)}{\partial \rho_i}\nu_i \prec 0$$

for all $\rho \in \mathcal{P}$, and for all $\nu \in \dot{\mathcal{P}}$.
Since $\mathcal{P}$ is compact, the inequality above implies there exists a positive definite and continuous matrix-valued function $\tilde{\mathcal{X}}(t) := \mathcal{X}\big(\rho(t)\big) \in (\mathbb{R}^+, \mathbb{R}^{n\times n})$ such that:

$$\dot{\tilde{\mathcal{X}}} + \tilde{A}^T\tilde{\mathcal{X}} + \tilde{\mathcal{X}}\tilde{A} \prec 0$$

The proof follows from Corollary 3.19 and the arbitrariness of the function $\rho(t)$. $\square$

D.12. Proof of Lemma 3.33

If (3.62) has a solution then (3.63) just follow from (3.62) by canceling the first third block row/column. Suppose that (3.63) holds that implies $Q_2 \prec 0$. The central trick is to cancel in (3.62) the block Q_{21} by a congruence transformation. To be specific (3.62) is equivalent to

$$\begin{bmatrix} I & -Q_{21}^T Q_2^{-1} & 0 \\ 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} (3.62) \begin{bmatrix} I & 0 & 0 \\ -Q_2^{-1} Q_{21} & I & 0 \\ 0 & 0 & I \end{bmatrix} \prec 0$$

which rewrites to

$$\begin{bmatrix} Q_1 - Q_{21}^T Q_2^{-1} Q_{21} & 0 & Q_{31}^T - Q_{21}^T Q_2^{-1} Q_{32}^T + X^T \\ 0 & Q_2 & Q_{32}^T \\ Q_{31} - Q_{32} Q_2^{-1} Q_{21} + X & Q_{32} & Q_3 \end{bmatrix} \prec 0 \,.$$

X defined in (3.64) is a solution since (3.63) implies $Q_1 - Q_{21}^T Q_2^{-1} Q_{21}$. $\square$

D.13. Proof of Theorem 3.34

Proof of necessity: Suppose for some controller of size $k \leq n$ there exists a smooth time matrix function $\mathcal{X}$ with (3.61). We fix $\mathcal{X}$ and view the BRL inequality in (3.61) as an inequality in A_K. To be specific define $\mathcal{J} := [0 \; I]$ to get

$$\mathcal{A} = \begin{bmatrix} A + B_2 D_K C & B_2 C_K \\ B_K C & 0 \end{bmatrix} + \begin{bmatrix} 0 \\ I \end{bmatrix} A_K \begin{bmatrix} 0 & I \end{bmatrix} = \hat{\mathcal{A}} + \mathcal{J}^T A_K \mathcal{J} \,.$$

Then the BRL inequality reads as

$$\begin{bmatrix} \dot{\mathcal{X}} + \hat{\mathcal{A}}^T\mathcal{X} + \mathcal{X}\hat{\mathcal{A}} + \mathcal{J}^T A_K^T \mathcal{J}\mathcal{X} + \mathcal{X}\mathcal{J}^T A_K \mathcal{J} & \mathcal{X}\mathcal{B} & \mathcal{C}_1^T \\ \mathcal{B}^T\mathcal{X} & -\gamma I & \mathcal{D}_1^T \\ \mathcal{C}_1 & \mathcal{D}_1 & -\gamma I \end{bmatrix} \prec 0\,.$$

The matrix $\mathfrak{J} := \begin{bmatrix} I \\ 0 \end{bmatrix}$ clearly satisfies $\mathfrak{J}\mathcal{J} = 0$. Hence the inequality above implies

$$\begin{bmatrix} \mathfrak{J}^T[\dot{\mathcal{X}} + \hat{\mathcal{A}}^T\mathcal{X} + \mathcal{X}\hat{\mathcal{A}}]\mathfrak{J} & \mathfrak{J}^T\mathcal{X}\mathcal{B} & \mathfrak{J}^T\mathcal{C}_1^T \\ \mathcal{B}^T\mathcal{X}\mathfrak{J} & -\gamma I & \mathcal{D}_1^T \\ \mathcal{C}_1\mathfrak{J} & \mathcal{D}_1 & -\gamma I \end{bmatrix} \prec 0\,.$$

Similarly with $\mathcal{Y} := \mathcal{X}^{-1}$, we infer $(\mathcal{J}\mathcal{X})\mathcal{Y}\mathfrak{J} = 0$ and thus

$$\begin{bmatrix} \mathfrak{J}^T[-\dot{\mathcal{Y}} + \hat{\mathcal{A}}\mathcal{Y} + \mathcal{Y}\hat{\mathcal{A}}^T]\mathfrak{J} & \mathfrak{J}^T\mathcal{B} & \mathfrak{J}^T\mathcal{Y}\mathcal{C}_1^T \\ \mathcal{B}^T\mathfrak{J} & -\gamma I & \mathcal{D}_1^T \\ \mathcal{C}_1\mathcal{Y}\mathfrak{J} & \mathcal{D}_1 & -\gamma I \end{bmatrix} \prec 0\,.$$

Let us partition $\mathcal{X}, \mathcal{Y}$ according to $\mathcal{A}$ into n and k rows/columns as

$$\mathcal{X} = \begin{bmatrix} X & U \\ U^T & \tilde{X} \end{bmatrix}, \qquad \mathcal{Y} = \begin{bmatrix} Y & W \\ W^T & \tilde{Y} \end{bmatrix},$$

and let us recall that (see Appendix B.2)

$$\mathcal{Y} = \begin{bmatrix} [X - U\tilde{X}^{-1}U^T]^{-1} & -[X - U\tilde{X}^{-1}U^T]^{-1}U\tilde{X}^{-1} \\ -\tilde{X}^{-1}U^T[X - U\tilde{X}^{-1}U^T]^{-1} & [\tilde{X} - U^TX^{-1}U]^{-1} \end{bmatrix}.$$

We can assume without loss of generality that U (of dimension $n \times k$) satisfies $U^TU \succ 0$. If not true, we can perturb $\mathcal{X}$ suitably without violating (3.61). The formula above for $\mathcal{Y}$ reveals $Y^{-1} = X - U\tilde{X}^{-1}U^T$. If *Cholesky factorizing* $\tilde{X} = V^TV$ such that

W is smooth, bounded, and a has a bounded inverse, we arrive at the condition of $X - Y^{-1} = ZZ^T,\ \ Z^T Z \succ 0$ of Theorem 3.34 for $Z := UV^{-1}$.

A simple computation shows, with the identities

$$E = GCY + C_K W^T \quad \text{and} \quad F = XB_2G + UB_K\,,$$

that the left-hand sides of the above inequalities with X and Y are identical with (3.65) and (3.66) respectively. We remark that we define E and F via these equations for providing necessity and we view them as equations in B_K and C_K for constructing a controller in the sufficiency proof.

Under the hypothesis of $X - Y^{-1} = ZZ^T,\ \ Z^T Z \succ 0$, it is not difficult to see that the equations above are solvable as equations in B_K and C_K if and only if the rank conditions in Theorem 3.34 are true. We clarify this for the first equation: It is solvable if and only if $\ker(W^T) \subset \ker(E - GCY)$. The (1,1) block of $\mathcal{X}\mathcal{Y} = I$ implies $XY + UW^T = I$ which reveals $\ker(I - XY) = \ker(W^T)$ since U has full column rank. Hence, solvability is equivalent to $\ker(I - XY) \subset \ker(E - GCY)$. By $X - Y^{-1} = ZZ^T,\ \ Z^T Z \succ 0$, $(I - XY)$ has rank k and, therefore, this inclusion is equivalent to first rank condition in Theorem 3.34 since

$$\begin{bmatrix} E & GC \\ Y & I \\ I & X \end{bmatrix} \begin{bmatrix} I & 0 \\ -Y & I \end{bmatrix} = \begin{bmatrix} E - GCY & GC \\ 0 & I \\ I - XY & X \end{bmatrix}.$$

Finally, there is $\epsilon > 0$ with $\mathcal{C}_2 \mathcal{X}^{-1} \mathcal{C}_2^T \preceq (\beta - \epsilon) I$, which is equivalent to

$$\begin{bmatrix} (\beta - \epsilon) I & \mathcal{C}_2 \\ \mathcal{C}_2^T & \mathcal{X} \end{bmatrix} \succeq 0\,.$$

With the $(n+k)\times 2n$ function $\mathcal{Z} := \begin{bmatrix} Y & I \\ W^T & 0 \end{bmatrix}$, the inequality above implies

$$\begin{bmatrix} (\beta-\epsilon)I & \mathcal{C}_2\mathcal{Z} \\ \mathcal{Z}^T\mathcal{C}_2^T & \mathcal{Z}^T\mathcal{X}\mathcal{Z} \end{bmatrix} \succeq 0\,.$$

Recalling the definition of E and computing $\mathcal{Z}^T\mathcal{X}\mathcal{Z} = \mathcal{Z}^T \begin{bmatrix} I & X \\ 0 & U^T \end{bmatrix} = \begin{bmatrix} Y & I \\ I & X \end{bmatrix}$, which leads to (3.67). This proves necessity.

Proof of sufficiency: Define $U := Z$, $\tilde{X} = I$, and $\mathcal{X} := \begin{bmatrix} X & U \\ U^T & I \end{bmatrix}$.
Due to $X - UU^T \succ 0$, $\mathcal{X}$ is smooth, bounded, and strictly positive, and the same holds for $\mathcal{Y} := \mathcal{X}^{-1}$. Again because of $X - Y^{-1} = ZZ^T$, $Z^TZ \succ 0$ and the definition of $\mathcal{Y}$, $\mathcal{Y}$ has Y as its left-upper block. Using the partitions of $\mathcal{X}$ and $\mathcal{Y}$ as defined above, we have $U^TU \succ 0$ by hypothesis, and since $W = YU$, we have $W^TW \succ 0$ as a consequence. Defining the time functions

$$C_K := (E - GCY)W(W^TW)^{-1} \quad \text{and} \quad B_K := (U^TU)^{-1}U^T(F - XB_2G)\,,$$

and recalling that U and V have full column rank, these are the unique solutions of

$$\begin{bmatrix} E & GC \\ Y & I \\ I & X \end{bmatrix} \begin{bmatrix} I & 0 \\ -Y & I \end{bmatrix} = \begin{bmatrix} E - GCY & GC \\ 0 & I \\ I - XY & X \end{bmatrix}.$$

if they are solvable at all. With B_K, C_K, G given so far, we can define $\hat{\mathcal{A}}, \mathcal{B}, \mathcal{C}_1, \mathcal{C}_2, \mathcal{D}_1, \mathcal{D}_2$. If we introduce $\mathcal{Z}$ as before, we infer that (3.67) is the same as $\begin{bmatrix} (\beta-\epsilon)I & \mathcal{C}_2\mathcal{Z} \\ \mathcal{Z}^T\mathcal{C}_2^T & \mathcal{Z}^T\mathcal{X}\mathcal{Z} \end{bmatrix} \succeq 0\,.$
Since $\mathcal{Z}$ has a full column rank, we can get back to $\begin{bmatrix} (\beta-\epsilon)I & \mathcal{C}_2 \\ \mathcal{C}_2^T & \mathcal{X} \end{bmatrix} \succeq 0$, which leads to the third inequality in (3.61). Moreover, (3.65) and (3.66) are equivalent to DLMIs with $\mathcal{X}$ and $\mathcal{Y}$ in the sufficiency proof respectively. Hence it remains to find a time

function A_K which satisfies

$$\begin{bmatrix} \dot{\mathcal{X}} + \hat{\mathcal{A}}^T\mathcal{X} + \mathcal{X}\hat{\mathcal{A}} + \mathcal{J}^T A_K^T \mathcal{J}\mathcal{X} + \mathcal{X}\mathcal{J}^T A_K \mathcal{J} & \mathcal{X}\mathcal{B} & \mathcal{C}_1^T \\ \mathcal{B}^T\mathcal{X} & -\gamma I & \mathcal{D}_1^T \\ \mathcal{C}_1 & \mathcal{D}_1 & -\gamma I \end{bmatrix} \prec 0\,,$$

and, therefore, leads to the BRL inequality in (3.61). We use $\mathcal{X}\mathcal{J}^T A_K \mathcal{J} = \begin{bmatrix} Z \\ I_k \end{bmatrix} A_K \begin{bmatrix} 0 & I_k \end{bmatrix}$, where we display the size of the identity blocks by using index.

Due to $Z^T Z \succ 0$ there exists a smooth time function Z_e such that $[Z\ Z_e]$ has a smooth bounded inverse. With the first k rows S_1 and the last $n-k$ rows S_2 of this inverse, we conclude that

$$\begin{bmatrix} S_1 \\ S_2 \end{bmatrix} \text{ has a bounded inverse and } \begin{bmatrix} S_1 \\ S_2 \end{bmatrix} Z = \begin{bmatrix} I_k \\ 0 \end{bmatrix},$$

with

$$\mathcal{S} := \begin{bmatrix} S_1 & 0 \\ S_2 & 0 \\ -S_1 & I_k \end{bmatrix} \text{ we get } \mathcal{S}\begin{bmatrix} Z \\ I_k \end{bmatrix} = \begin{bmatrix} I_k \\ 0 \\ 0 \end{bmatrix}, \text{ and } \mathcal{S}\begin{bmatrix} 0 \\ I_k \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ I_k \end{bmatrix}.$$

Note that both $\mathcal{S}$ and $\mathcal{S}^{-1}$ are smooth time functions. We can left-multiply the first row of the DLMI above with $\mathcal{S}$ and right-multiply the first column with $\mathcal{S}^T$. After this *congruence transformation* the same DLMI becomes equivalent to

$$\begin{bmatrix} Q_1 & Q_{12} + A_K & Q_3 \\ Q_{21} + A_K^T & Q_2 & Q_{23} \\ Q_{31} & Q_{32} & Q_3 \end{bmatrix}$$

for some computable Q. Note that the dual DLMIs in the necessity part of the proof with $\mathcal{X}$ and $\mathcal{Y}$ are equivalent to

$$\begin{bmatrix} Q_1 & Q_{13} \\ Q_{31} & Q_3 \end{bmatrix} \prec 0 \quad \text{and} \quad \begin{bmatrix} Q_2 & Q_{23} \\ Q_{32} & Q_3 \end{bmatrix} \prec 0\,.$$

Lemma 3.33 then reveals that $A_K := Q_{13}Q_3^{-1}Q_{32} - Q_{12}$ is indeed a time matrix function which leads to the DLMI;

$$\begin{bmatrix} \dot{\mathcal{X}} + \hat{\mathcal{A}}^T\mathcal{X} + \mathcal{X}\hat{\mathcal{A}} + \mathcal{J}^T A_K^T \mathcal{J}\mathcal{X} + \mathcal{X}\mathcal{J}^T A_K \mathcal{J} & \mathcal{X}\mathcal{B} & \mathcal{C}_1^T \\ \mathcal{B}^T\mathcal{X} & -\gamma I & \mathcal{D}_1^T \\ \mathcal{C}_1 & \mathcal{D}_1 & -\gamma I \end{bmatrix} \prec 0\,.$$

This completes the proof and the controller construction. $\square$

APPENDIX E: SIMULATION ENVIRONMENT

Structure of the simulation model used in SIMULINK® for H_∞ control are shown in the Figure E.1

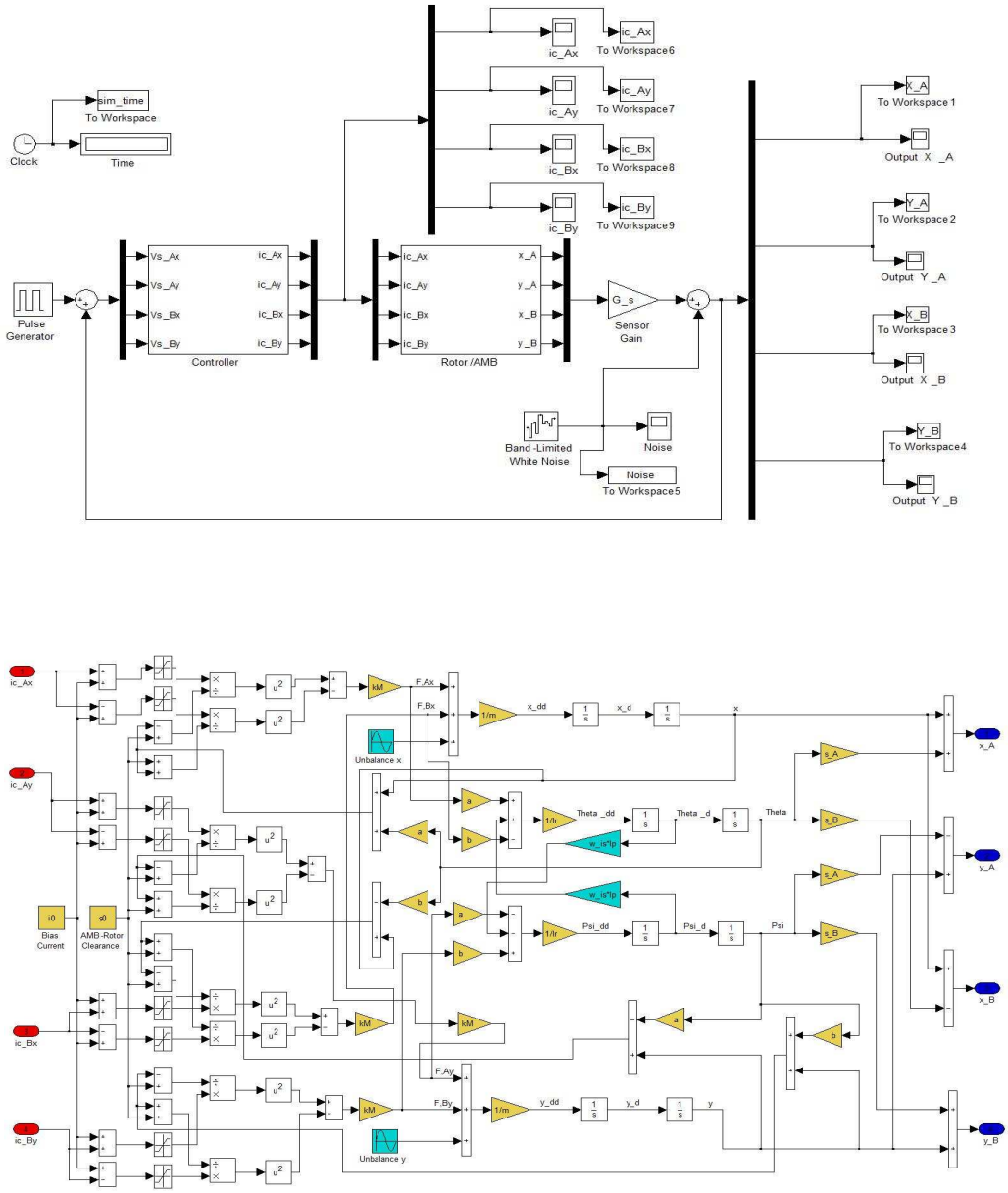

Figure E.1. Closed-loop system and rotor/AMB model with H_∞ control.

Simulations for the LPV system are made using the LFR Toolbox from ONERA [52-53]. Structure of the model is shown in the Figure E.2

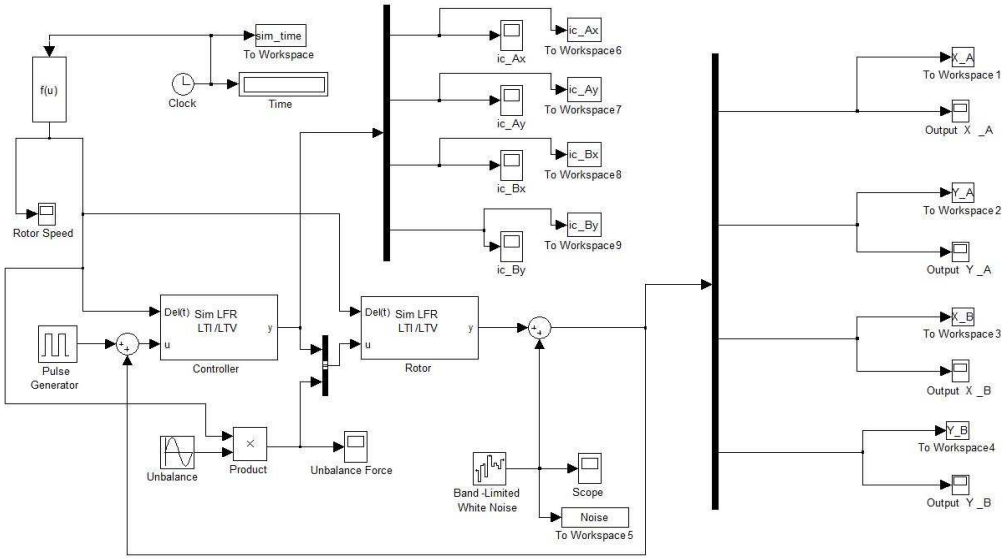

Figure E.2. Closed-loop parameter dependent system with LPV control.

REFERENCES

1. Genta G. *Dynamics of Rotating Systems.* Springer Mech. Eng'g Series, (2005).

2. Yamamoto T., Y. Ishida *Linear and Nonlinear Rotordynamics.* Wiley Series on Nonlinear Science, (2001).

3. Shabana A.A. *Vibration of Discrete and Continuos Systems.* Springer Mechanical Engineering Series, (1997).

4. Chiba A. *Magnetic Bearings and Bearingless Drives.* Elsevier-Newness, (2005).

5. Apkarian P., R.J. Adams *Advanced gain scheduling techniques for uncertain systems.* IEEE Transactions on Control Systems Technology, 6(1), pp 21-32, (1997).

6. Tsiotras P., S. Mason *Self-scheduled H_∞ controllers for magnetic bearings.* International Mech. Eng'g Congress and Exhibition, Atlanta, GA, pp. 151-158, (1996).

7. Burrows C.R., N. Sahinkaya, A. Traxler, and G. Schweitzer *Design and application of a magnetic bearing for vibration control and stabilization of a flexible rotor.* 1st International Symposium on Magnetic Bearings, pp 159-168, Springer-Verlag, (1988).

8. Zhang X. *Active Regel und Kompensationsstratigien für magnetgelagerte Mehrefreiheitsgrad Rotoren.* Ph.D. Thesis, Techniche Universität Darmstadt, (2002).

9. Herzog R., P. Bühler, C. Gähler, and R. Larsonneur *Unbalance compensation using generalized notch filters in the multivariable feedback of magnetic bearings.* IEEE Transactions on Control Systems Technology, Vol 4, no.5, pp 580-586, (1996).

10. Glover K., J.C. Doyle *State space formulae for all stabilizing controllers that satisfy an H_∞ norm bound and relation to risk sensitivity.* Systems and Control Letters, Vol 11, pp 167-172, (1988).

11. Glover K., J.C. Doyle, P.P. Khargonekar, and B.A. Francis *State space solutions to standard H_2 and H_∞ control problems.* IEEE Transactions on Automatic Control, Vol 34, no.8, pp 831-847, (1989).

12. Gahinet P., P. Apkarian *A linear matrix inequality approach to H_∞ control.* International Journal of Robust and Nonlinear Control, Vol 4, pp 421-448, (1994).

13. Gahinet P. *Explicit controller formulas for LMI-based H_∞ synthesis.* Automatica, Vol 32, no.7, pp 1007-1014, (1996).

14. Scherer C., P. Gahinet, and M. Chilali *Multi-objective output-feedback control via LMI optimization.* IEEE Transactions on Automatic Control, Vol 42, no.7, pp 896-911, (1997).

15. Packard A. *Gain Scheduling via linear fractional transformations.* Systems and Control Letters, Vol 22, no.2, pp 79-92, (1994).

16. Becker G., A. Packard *Robust performance of linear parametrically varying systems using parametrically-dependent linear feedback.* Systems and Control Letters, Vol 23, no.3, pp 205-214, (1994).

17. Apkarian P., P. Gahinet *A convex characterization of gain-scheduled H_∞ controllers.* IEEE Transactions on Automatic Control, Vol 40, no.5, pp 853-864, (1995).

18. Gahinet P., P. Apkarian, G. Becker *Self-scheduled H_∞ control of linear parameter-varying systems: a design example controller formulas for LMI-based H_∞ synthesis.* Automatica, Vol 31, no.9, pp 1251-1261, (1995).

19. Sivrioğlu S., K. Nanomi *LMI approach to gain scheduled H_∞ beyond PID control for gyroscopic rotor-magnetic bearing system.* Proceedings of the 35th Conference on Decision and Control, pp 3694-3699, (1996).

20. Tsiotras P., C. Knospe *Reducing conservatism for gain scheduled H_∞ controllers for AMBs.* Proceedings of MAG'97, pp 290-299, (1997).

21. Scherer C. *Mixed H_2/H_∞ Control.* In: A. Isidori, Ed., *Trends in Control, A European Perspective*, Springer-Verlag, pp 173-216, (1995).

22. Tsiotras P., M. Arcak *Low-bias control of AMB subject to voltage saturation: State-feedback and observer designs.* IEEE Transactions on Control Systems Technology, Vol 13, no. 2, pp 262-273, (2005).

23. Sivrioğlu S. *Adaptive control of nonlinear zero-bias current magnetic bearing system.* Nonlinear Dynamics, Vol 48, pp 175-184, (2007).

24. Husain A.R., M.N. Ahmad, A.H.M. Yatim *Asymptotic stabilization of an active meagnetic bearing system using LMI-based sliding mode control.* Proceedings of World Academy of Science, Eng'g and Technology, Vol 27, ISSN 1307-6884, (2008).

25. Lösch F. *Identification and Automated Controller Design for Active Magnetic Bearing Systems.* Ph.D. Thesis, ETH Zurich, (2002).

26. Zhou G., J.C. Doyle, K. Glover *Robust and Optimal Control.* Prentice-Hall, (1996).

27. Zhou G., J.C. Doyle *Essentials of Robust Control.* Prentice-Hall, (1998).

28. Antsaklis P.J., A.N. Michel *Linear Systems.* Birkhäuser Boston, (2006).

29. Amato F. *Robust Control of Linear Systems Subject to Time-Varying Parameters.* Springer, Lecture Notes in Control and Information Sciences , (2006).

30. Shamma J., M. Athans *Guaranteed properties of gain scheduled control for linear parameter-varying systems.* Automatica, Vol 27, no.3, pp 559-564, (1991).

31. Zhou K., P.P. Khargonekar *An algebraic Riccati equation approach to H_∞ optimization.* System and Control Letters, Vol 11, no.2, pp 85-91, (1988).

32. Skogestad S., I. Postlethwaite *Multivariable Feedback Control.* John Wiley and Sons Ltd., (2005).

33. Schweitzer G., E. Maslen (Editors) *Magnetic Bearings.* Springer-Verlag Berlin Heidelberg, (2009).

34. Li G., E. Maslen, P. Allaire *A note on ISO AMB-Rotor System Stability Margin* . Proceedings of the Tenth International Symposium on Magnetic Bearings, Martigny, Switzerland, (2006).

35. Herzog H., P. Bühler, C. Gähler, R. Larsonneur *Unbalance compensation using generalized notch filters in the multivariable feedback of magnetic bearings.* IEEE Transactions on Control Systems Technology, 4(5):580-586, (2006).

36. Dullerud G.E., F. Paganini *A Course on Robust Control Theory.* Springer Science+Business Media Inc., (2000).

37. Horn R.A., F. Johnson *Matrix Analysis.* Cambridge University Press, (1985).

38. Strang G. *Linear Algebra and Its Applications.* Thompson Brooks/Cole Corp., (2006).

39. Kreyzig E. *Introductory Functional Analysis with Applications.* John Wiley and Sons Inc., (1987).

40. Boyd S., L. El Ghaoui, E. Feron, V. Balakrishnan *Linear Matrix Inequalities in System and Control Theory.* SIAM Studies in Applied Mathematics, Vol.15, (1994).

41. Rudin W. *Real and Complex Analysis.* Mc-Graw Hill, (1986).

42. Bernstein D.S. *Matrix Mathematics.* Princeton University Press, (2009).

43. Köse İ.E., F. Jabbari *Control of LPV Systems with Partly Measured Parameters.* IEEE Transections on Automatic Control, Vol.44, no.3, pp 658-663, (1999).

44. Luenberger D.G. *Optimization by Vector Space Methods.* John-Wiley and Sons Ltd., (1969).

45. Boyd S., L. Vandenberghe *Convex Optimization*. Cambridge University Press, (2004).

46. Megretski A., A. Rantzer *System analysis via integral quadratic constraints*. IEEE Transactions on Automatic Control, Vol.42, no.6, pp 819-830, (1997).

47. Scherer C. *Robust mixed control and linear parameter varying control with full block scalings*, in Niculescu S.I. and El Ghaoi (editors) *Advances in Linear Matrix Inequality Methods in Control*, pp 187-207, (1999).

48. MATLAB® *The Language of Technical Computing*. The MathWorks Inc., *http://www.mathworks.com*.

49. Balas G., R. Chiang, A. Packard, M. Sofonov *User's Guide for Robust Control Toolbox for* MATLAB, Ver.3. The MathWorks Inc., (2006).

50. Löfberg J. YALMIP: *A toolbox for modeling and optimization in* MATLAB. Proceedings of the CACSD Conference. Taipei, Taiwan, (2004). Available from *http://control.ee.ethz.ch/joloef/yalmip.php*.

51. Sturm J.F. *Using* SEDUMI 1.02: *A* MATLAB *Toolbox for Optimization Over Symmetric Cones*. Optimization Methods and Software, 11(12), pp 625-653, (1999).

52. Magni J.F. *User Manual for the Linear Fractional Representation Toolbox for* MATLAB, Ver.2. ONERA System Control and Flight Dynamics Department. Available from *http://www.onera.fr/dcsd/*, (2006).

53. Biannic J.M., C. Döll SIMULINK® *Handling of LFR Objects*. Free Web publication avialable from *http://www.cert.fr/dcsd/idco/perso/Biannic/*, (2006).

54. Blum L., F. Cucker, M. Shub, S. Smale *Complexity and Real Computation*. Springer-Verlag, Berlin, Germany, (1997).