
Unsolved Problems in Mathematical Systems and Control Theory

Edited by
Vincent D. Blondel
Alexandre Megretski

PRINCETON UNIVERSITY PRESS
PRINCETON AND OXFORD

Copyright © 2004 by Princeton University Press

Published by Princeton University Press, 41 William Street, Princeton, New Jersey 08540, USA

In the United Kingdom: Princeton University Press, 3 Market Place, Woodstock, Oxfordshire OX20 1SY, UK

All rights reserved

Library of Congress Cataloging-in-Publication Data

Unsolved problems in mathematical systems and control theory

Edited by Vincent D. Blondel, Alexandre Megretski. p. cm.

Includes bibliographical references.

ISBN 0-691-11748-9 (cl : alk. paper)

1. System analysis. 2. Control theory. I. Blondel, Vincent. II. Megretski, Alexandre.

QA402.U535 2004 2003064802

003—dc22

The publisher would like to acknowledge the editors of this volume for providing the camera-ready copy from which this book was printed.

Printed in the United States of America

10 9 8 7 6 5 4 3 2 1

I have yet to see any problem, however complicated, which, when you looked at it in the right way, did not become still more complicated.

Poul Anderson

Contents

Preface	xiii
Associate Editors	xv
Website	xvii
 PART 1. LINEAR SYSTEMS	 1
Problem 1.1. Stability and composition of transfer functions <i>Guillermo Fernández-Anaya, Juan Carlos Martínez-García</i>	3
Problem 1.2. The realization problem for Herglotz-Nevanlinna functions <i>Seppo Hassi, Henk de Snoo, Eduard Tsekanovskii</i>	8
Problem 1.3. Does any analytic contractive operator function on the polydisk have a dissipative scattering nD realization? <i>Dmitry S. Kalyuzhnyi-Verbovetzky</i>	14
Problem 1.4. Partial disturbance decoupling with stability <i>Juan Carlos Martínez-García, Michel Malabre, Vladimír Kučera</i>	18
Problem 1.5. Is Monopoli's model reference adaptive controller correct? <i>A. S. Morse</i>	22
Problem 1.6. Model reduction of delay systems <i>Jonathan R. Partington</i>	29
Problem 1.7. Schur extremal problems <i>Lev Sakhnovich</i>	33
Problem 1.8. The elusive iff test for time-controllability of behaviors <i>Amol J. Sasane</i>	36

Problem 1.9. A Farkas lemma for behavioral inequalities <i>A.A. (Tonny) ten Dam, J.W. (Hans) Nieuwenhuis</i>	40
Problem 1.10. Regular feedback implementability of linear differential behaviors <i>H. L. Trentelman</i>	44
Problem 1.11. Riccati stability <i>Erik I. Verriest</i>	49
Problem 1.12. State and first order representations <i>Jan C. Willems</i>	54
Problem 1.13. Projection of state space realizations <i>Antoine Vandendorpe, Paul Van Dooren</i>	58
PART 2. STOCHASTIC SYSTEMS	65
Problem 2.1. On error of estimation and minimum of cost for wide band noise driven systems <i>Agamirza E. Bashirov</i>	67
Problem 2.2. On the stability of random matrices <i>Giuseppe C. Calafiore, Fabrizio Dabbene</i>	71
Problem 2.3. Aspects of Fisher geometry for stochastic linear systems <i>Bernard Hanzon, Ralf Peeters</i>	76
Problem 2.4. On the convergence of normal forms for analytic control systems <i>Wei Kang, Arthur J. Krener</i>	82
PART 3. NONLINEAR SYSTEMS	87
Problem 3.1. Minimum time control of the Kepler equation <i>Jean-Baptiste Caillaud, Joseph Gergaud, Joseph Noailles</i>	89
Problem 3.2. Linearization of linearly controllable systems <i>R. Devanathan</i>	93
Problem 3.3. Bases for Lie algebras and a continuous CBH formula <i>Matthias Kowski</i>	97

CONTENTS	ix
Problem 3.4. An extended gradient conjecture <i>Luis Carlos Martins Jr., Geraldo Nunes Silva</i>	103
Problem 3.5. Optimal transaction costs from a Stackelberg perspective <i>Geert Jan Olsder</i>	107
Problem 3.6. Does cheap control solve a singular nonlinear quadratic problem? <i>Yuri V. Orlov</i>	111
Problem 3.7. Delta-Sigma modulator synthesis <i>Anders Rantzer</i>	114
Problem 3.8. Determining of various asymptotics of solutions of nonlinear time-optimal problems via right ideals in the moment algebra <i>G. M. Sklyar, S. Yu. Ignatovich</i>	117
Problem 3.9. Dynamics of principal and minor component flows <i>U. Helmke, S. Yoshizawa, R. Evans, J.H. Manton, and I.M.Y. Mareels</i>	122
PART 4. DISCRETE EVENT, HYBRID SYSTEMS	129
Problem 4.1. \mathcal{L}_2 -induced gains of switched linear systems <i>João P. Hespanha</i>	131
Problem 4.2. The state partitioning problem of quantized systems <i>Jan Lunze</i>	134
Problem 4.3. Feedback control in flowshops <i>S.P. Sethi and Q. Zhang</i>	140
Problem 4.4. Decentralized control with communication between controllers <i>Jan H. van Schuppen</i>	144
PART 5. DISTRIBUTED PARAMETER SYSTEMS	151
Problem 5.1. Infinite dimensional backstepping for nonlinear parabolic PDEs <i>Andras Balogh, Miroslav Krstic</i>	153
Problem 5.2. The dynamical Lamé system with boundary control: on the structure of reachable sets <i>M.I. Belishev</i>	160

Problem 5.3. Null-controllability of the heat equation in unbounded domains <i>Sorin Micu, Enrique Zuazua</i>	163
Problem 5.4. Is the conservative wave equation regular? <i>George Weiss</i>	169
Problem 5.5. Exact controllability of the semilinear wave equation <i>Xu Zhang, Enrique Zuazua</i>	173
Problem 5.6. Some control problems in electromagnetics and fluid dynamics <i>Lorella Fatone, Maria Cristina Recchioni, Francesco Zirilli</i>	179
PART 6. STABILITY, STABILIZATION	187
Problem 6.1. Copositive Lyapunov functions <i>M. K. Çamlibel, J. M. Schumacher</i>	189
Problem 6.2. The strong stabilization problem for linear time-varying systems <i>Avraham Feintuch</i>	194
Problem 6.3. Robustness of transient behavior <i>Diederich Hinrichsen, Elmar Plischke, Fabian Wirth</i>	197
Problem 6.4. Lie algebras and stability of switched nonlinear systems <i>Daniel Liberzon</i>	203
Problem 6.5. Robust stability test for interval fractional order linear systems <i>Ivo Petráš, YangQuan Chen, Blas M. Vinagre</i>	208
Problem 6.6. Delay-independent and delay-dependent Aizerman problem <i>Vladimir Răşvan</i>	212
Problem 6.7. Open problems in control of linear discrete multidimensional systems <i>Li Xu, Zhiping Lin, Jiang-Qian Ying, Osami Saito, Yoshihisa Anazawa</i>	221
Problem 6.8. An open problem in adaptative nonlinear control theory <i>Leonid S. Zhitckij</i>	229
Problem 6.9. Generalized Lyapunov theory and its omega-transformable regions <i>Sheng-Guo Wang</i>	233

CONTENTS	xi
Problem 6.10. Smooth Lyapunov characterization of measurement to error stability <i>Brian P. Ingalls, Eduardo D. Sontag</i>	239
PART 7. CONTROLLABILITY, OBSERVABILITY	245
Problem 7.1. Time for local controllability of a 1-D tank containing a fluid modeled by the shallow water equations <i>Jean-Michel Coron</i>	247
Problem 7.2. A Hautus test for infinite-dimensional systems <i>Birgit Jacob, Hans Zwart</i>	251
Problem 7.3. Three problems in the field of observability <i>Philippe Jouan</i>	256
Problem 7.4. Control of the KdV equation <i>Lionel Rosier</i>	260
PART 8. ROBUSTNESS, ROBUST CONTROL	265
Problem 8.1. H_∞ -norm approximation <i>A.C. Antoulas, A. Astolfi</i>	267
Problem 8.2. Noniterative computation of optimal value in H_∞ control <i>Ben M. Chen</i>	271
Problem 8.3. Determining the least upper bound on the achievable delay margin <i>Daniel E. Davison, Daniel E. Miller</i>	276
Problem 8.4. Stable controller coefficient perturbation in floating point implementation <i>Jun Wu, Sheng Chen</i>	280
PART 9. IDENTIFICATION, SIGNAL PROCESSING	285
Problem 9.1. A conjecture on Lyapunov equations and principal angles in subspace identification <i>Katrien De Cock, Bart De Moor</i>	287

Problem 9.2. Stability of a nonlinear adaptive system for filtering and parameter estimation <i>Masoud Karimi-Ghartemani, Alireza K. Ziarani</i>	293
PART 10. ALGORITHMS, COMPUTATION	297
Problem 10.1. Root-clustering for multivariate polynomials and robust stability analysis <i>Pierre-Alexandre Bliman</i>	299
Problem 10.2. When is a pair of matrices stable? <i>Vincent D. Blondel, Jacques Theys, John N. Tsitsiklis</i>	304
Problem 10.3. Freeness of multiplicative matrix semigroups <i>Vincent D. Blondel, Julien Cassaigne, Juhani Karhumäki</i>	309
Problem 10.4. Vector-valued quadratic forms in control theory <i>Francesco Bullo, Jorge Cortés, Andrew D. Lewis, Sonia Martínez</i>	315
Problem 10.5. Nilpotent bases of distributions <i>Henry G. Hermes, Matthias Kowski</i>	321
Problem 10.6. What is the characteristic polynomial of a signal flow graph? <i>Andrew D. Lewis</i>	326
Problem 10.7. Open problems in randomized μ analysis <i>Onur Toker</i>	330

Preface

Five years ago, a first volume of open problems in Mathematical Systems and Control Theory appeared.¹ Some of the 53 problems that were published in this volume attracted considerable attention in the research community.

The book in front of you contains a new collection of 63 open problems. The contents of both volumes show the evolution of the field in the half decade since the publication of the first volume. One noticeable feature is the shift toward a wider class of questions and more emphasis on issues driven by physical modeling.

Early versions of some of the problems in this book have been presented at the Open Problem sessions of the Oberwolfach Tagung on Regelungstheorie, on February 27, 2002, and of the Conference on Mathematical Theory of Networks and Systems (MTNS) in Notre Dame, Indiana, on August 12, 2002. The editors thank the organizers of these meetings for their willingness to provide the problems this welcome exposure.

Since the appearance of the first volume, open problems have continued to meet with large interest in the mathematical community. Undoubtedly, the most spectacular event in this arena was the announcement by the Clay Mathematics Institute² of the *Millennium Prize Problems* whose solution will be rewarded by one million U.S. dollars each. Modesty and modesty of means have prevented the editors of the present volume from offering similar rewards toward the solution of the problems in this book. However, we trust that, notwithstanding this absence of a financial incentive, the intellectual challenge will stimulate many readers to attack the problems.

The editors thank in the first place the researchers who have submitted the problems. We are also very thankful to the Princeton University Press, and in particular Vickie Kearn, for their willingness to publish this volume. The full text of the problems, together with comments, additions, and solutions, will be posted on the book website at Princeton University Press (link available from <http://pup.princeton.edu/math/>) and on <http://www.inma.ucl.ac.be/~blondel/op/>. Readers are encouraged to submit contributions by following the instructions given on these websites.

The editors,

Louvain-la-Neuve, March 15, 2003.

¹Vincent D. Blondel, Eduardo D. Sontag, M. Vidyasagar, and Jan C. Willems, *Open Problems in Mathematical Systems and Control Theory*, Springer Verlag, 1998.

²See <http://www.claymath.org>.

Associate Editors

Roger Brockett, Harvard University, USA

Jean-Michel Coron, University of Paris (Orsay), France

Roland Hildebrand, University of Louvain (Louvain-la-Neuve), Belgium

Miroslav Krstic, University of California (San Diego), USA

Anders Rantzer, Lund Institute of Technology, Sweden

Joachim Rosenthal, University of Notre Dame, USA

Eduardo Sontag, Rutgers University, USA

M. Vidyasagar, Tata Consultancy Services, India

Jan Willems, University of Leuven, Belgium

Website

The full text of the problems presented in this book, together with comments, additions and solutions, are freely available in electronic format from the book website at Princeton University Press:

`http://pup.princeton.edu/math/`

and from an editor website:

`http://www.inma.ucl.ac.be/~blondel/op/`

Readers are encouraged to submit contributions by following the instructions given on these websites.

PART 1

Linear Systems

Problem 1.1

Stability and composition of transfer functions

G. Fernández-Anaya

Departamento de Ciencias Básicas

Universidad Iberoamericana

Lomas de Santa Fe

01210 México D.F.

México

guillermo.fernandez@uia.mx

J. C. Martínez-García

Departamento de Control Automático

CINVESTAV-IPN

A.P. 14-740

07300 México D.F.

México

martinez@ctrl.cinvestav.mx

1 INTRODUCTION

As far as the frequency-described continuous linear time-invariant systems are concerned, the study of control-oriented properties (like stability) resulting from the substitution of the complex Laplace variable s by rational transfer functions have been little studied by the Automatic Control community. However, some interesting results have recently been published:

Concerning the study of the so-called uniform systems, i.e., **LTI** systems consisting of identical components and amplifiers, it was established in [8] a general criterion for robust stability for rational functions of the form $D(f(s))$, where $D(s)$ is a polynomial and $f(s)$ is a rational transfer function. By applying such a criterium, it gave a generalization of the celebrated Kharitonov's theorem [7], as well as some robust stability criteria under H_∞ -uncertainty. The results given in [8] are based on the so-called *H-domains*.¹ As far as robust stability of polynomial families is concerned, some Kharito-

¹The *H-domain* of a function $f(s)$ is defined to be the set of points h on the complex plane for which the function $f(s) - h$ has no zeros on the open right-half complex plane.

nov's like results [7] are given in [9] (for a particular class of polynomials), when interpreting substitutions as nonlinearly correlated perturbations on the coefficients.

More recently, in [1], some results for proper and stable real rational SISO functions and coprime factorizations were proved, by making substitutions with $\alpha(s) = (as + b) / (cs + d)$, where a, b, c , and d are strictly positive real numbers, and with $ad - bc \neq 0$. But these results are limited to the bilinear transforms, which are very restricted.

In [4] is studied the preservation of properties linked to control problems (like weighted nominal performance and robust stability) for Single-Input Single-Output systems, when performing the substitution of the Laplace variable (in transfer functions associated to the control problems) by strictly positive real functions of zero relative degree. Some results concerning the preservation of control-oriented properties in Multi-Input Multi-Output systems are given in [5], while [6] deals with the preservation of solvability conditions in algebraic Riccati equations linked to robust control problems.

Following our interest in substitutions we propose in section 22.2 three interesting problems. The motivations concerning the proposed problems are presented in section 22.3.

2 DESCRIPTION OF THE PROBLEMS

In this section we propose three closely related problems. The first one concerns the characterization of a transfer function as a composition of transfer functions. The second problem is a modified version of the first problem: the characterization of a transfer function as the result of substituting the Laplace variable in a transfer function by a strictly positive real transfer function of zero relative degree. The third problem is in fact a conjecture concerning the preservation of stability property in a given polynomial resulting from the substitution of the coefficients in the given polynomial by a polynomial with non-negative coefficients evaluated in the substituted coefficients.

Problem 1: *Let a Single Input Single Output (SISO) transfer function $G(s)$ be given. Find transfer functions $G_0(s)$ and $H(s)$ such that:*

1. $G(s) = G_0(H(s))$;
2. $H(s)$ preserves proper stable transfer functions under substitution of the variable s by $H(s)$, and:
3. The degree of the denominator of $H(s)$ is the maximum with the properties 1 and 2.

Problem 2: Let a SISO transfer function $G(s)$ be given. Find a transfer function $G_0(s)$ and a Strictly Positive Real transfer function of zero relative degree (SPR0), say $H(s)$, such that:

1. $G(s) = G_0(H(s))$ and:
2. The degree of the denominator of $H(s)$ is the maximum with the property 1.

Problem 3: (Conjecture) Given any stable polynomial:

$$a_n s^n + a_{n-1} s^{n-1} + \cdots + a_1 s + a_0$$

and given any polynomial $q(s)$ with non-negative coefficients, then the polynomial:

$$q(a_n) s^n + q(a_{n-1}) s^{n-1} + \cdots + q(a_1) s + q(a_0)$$

is stable (see [3]).

3 MOTIVATIONS

Consider the closed-loop control scheme:

$$y(s) = G(s)u(s) + d(s), \quad u(s) = K(s)(r(s) - y(s)),$$

where: $P(s)$ denotes the SISO plant; $K(s)$ denotes a *stabilizing controller*; $u(s)$ denotes the control input; $y(s)$ denotes the control input; $d(s)$ denotes the disturbance and $r(s)$ denotes the reference input. We shall denote the closed-loop transfer function from $r(s)$ to $y(s)$ as $\mathcal{F}_r(G(s), K(s))$ and the closed-loop transfer function from $d(s)$ to $y(s)$ as $\mathcal{F}_d(G(s), K(s))$.

- Consider the closed-loop system $\mathcal{F}_r(G(s), K(s))$, and suppose that the plant $G(s)$ results from a particular substitution of the s Laplace variable in a transfer function $G_0(s)$ by a transfer function $H(s)$, i.e., $G(s) = G_0(H(s))$. It has been proved that a controller $K_0(s)$ which stabilizes the closed-loop system $\mathcal{F}_r(G_0(s), K_0(s))$ is such that $K_0(H(s))$ stabilizes $\mathcal{F}_r(G(s), K_0(H(s)))$ (see [2] and [8]). Thus, the simplification of procedures for the synthesis of stabilizing controllers (profiting from transfer function compositions) justifies problem 1.
- As far as problem 2 is concerned, consider the synthesis of a controller $K(s)$ stabilizing the closed-loop transfer function $\mathcal{F}_d(G(s), K(s))$, and such that $\|\mathcal{F}_d(G(s), K(s))\|_\infty < \gamma$, for a fixed given $\gamma > 0$. If we known that $G(s) = G_0(H(s))$, being $H(s)$ a SPR0 transfer function, the solution of problem 2 would arise to the following procedure:

1. Find a controller $K_0(s)$ which stabilizes the closed-loop transfer function $\mathcal{F}_d(G_0(s), K_0(s))$ and such that:

$$\|\mathcal{F}_d(G_0(s), K_0(s))\|_\infty < \gamma.$$

2. The composed controller $K(s) = K_0(H(s))$ stabilizes the closed-loop system $\mathcal{F}_d(G(s), K(s))$ and:

$$\|\mathcal{F}_d(G(s), K(s))\|_\infty < \gamma$$

(see [2], [4], and [5]).

It is clear that condition 3 in the first problem, or condition 2 in the second problem, can be relaxed to the following condition: *the degree of the denominator of $H(s)$ is as high as be possible with the appropriate conditions.* With this new condition, the open problems are a bit less difficult.

- Finally, problem 3 can be interpreted in terms of robustness under positive polynomial perturbations in the coefficients of a stable transfer function.

BIBLIOGRAPHY

- [1] G. Fernández, S. Muñoz, R. A. Sánchez, and W. W. Mayol, "Simultaneous stabilization using evolutionary strategies," *Int. J. Contr.*, vol. 68, no. 6, pp. 1417-1435, 1997.
- [2] G. Fernández, "Preservation of SPR functions and stabilization by substitutions in SISO plants," *IEEE Transaction on Automatic Control*, vol. 44, no. 11, pp. 2171-2174, 1999.
- [3] G. Fernández and J. Alvarez, "On the preservation of stability in families of polynomials via substitutions," *Int. J. of Robust and Nonlinear Control*, vol. 10, no. 8, pp. 671-685, 2000.
- [4] G. Fernández, J. C. Martínez-García, and V. Kučera, " H_∞ -Robustness Properties Preservation in SISO Systems when applying SPR Substitutions," Submitted to the *International Journal of Automatic Control*.
- [5] G. Fernández and J. C. Martínez-García, "MIMO Systems Properties Preservation under SPR Substitutions," *International Symposium on the Mathematical Theory of Networks and Systems (MTNS'2002)*, University of Notre Dame, USA, August 12-16, 2002.
- [6] G. Fernández, J. C. Martínez-García, and D. Aguilar-George, "Preservation of solvability conditions in Riccati equations when applying SPR0 substitutions," submitted to *IEEE Transactions on Automatic Control*, 2002.
- [7] V. L. Kharitonov, "Asymptotic stability of families of systems of linear differential equations," *Differential'nye Uravneniya*, vol. 14, pp. 2086-2088, 1978.

- [8] B. T. Polyak and Ya. Z. Tsyppkin, "Stability and robust stability of uniform systems, " *Automation and Remote Contr.*, vol. 57, pp. 1606-1617, 1996.
- [9] L. Wang, "Robust stability of a class of polynomial families under non-linearly correlated perturbations," *System and Control Letters*, vol. 30, pp. 25-30, 1997.

Problem 1.2

The realization problem for Herglotz-Nevanlinna functions

Seppo Hassi

Department of Mathematics and Statistics
University of Vaasa
P.O. Box 700, 65101 Vaasa
Finland
`sha@uwasa.fi`

Henk de Snoo

Department of Mathematics
University of Groningen
P.O. Box 800, 9700 AV Groningen
Nederland
`desnoo@math.rug.nl`

Eduard Tsekanovskii

Department of Mathematics
Niagara University, NY 14109
USA
`tsekanov@niagara.edu`

1 MOTIVATION AND HISTORY OF THE PROBLEM

Roughly speaking, realization theory concerns itself with identifying a given holomorphic function as the transfer function of a system or as its linear fractional transformation. Linear, conservative, time-invariant systems whose main operator is *bounded* have been investigated thoroughly. However, many realizations in different areas of mathematics including system theory, electrical engineering, and scattering theory involve *unbounded* main operators, and a complete theory is still lacking. The aim of the present proposal is to outline the necessary steps needed to obtain a general realization theory along the lines of M. S. Brodskii and M. S. Livšic [8], [9], [16], who have

considered systems with a bounded main operator.

An operator-valued function $V(z)$ acting on a Hilbert space \mathfrak{E} belongs to the Herglotz-Nevanlinna class \mathbf{N} , if outside \mathbb{R} it is holomorphic, symmetric, i.e., $V(z)^* = V(\bar{z})$, and satisfies $(\operatorname{Im} z)(\operatorname{Im} V(z)) \geq 0$. Here and in the following it is assumed that the Hilbert space \mathfrak{E} is finite-dimensional. Each Herglotz-Nevanlinna function $V(z)$ has an integral representation of the form

$$V(z) = Q + Lz + \int_{\mathbb{R}} \left(\frac{1}{t-z} - \frac{t}{1+t^2} \right) d\Sigma(t), \quad (1)$$

where $Q = Q^*$, $L \geq 0$, and $\Sigma(t)$ is a nondecreasing matrix-function on \mathbb{R} with $\int_{\mathbb{R}} d\Sigma(t)/(t^2 + 1) < \infty$. Conversely, each function of the form (1) belongs to the class \mathbf{N} . Of special importance (cf. [15]) are the class \mathbf{S} of Stieltjes functions

$$V(z) = \gamma + \int_0^\infty \frac{d\Sigma(t)}{t-z}, \quad (2)$$

where $\gamma \geq 0$ and $\int_0^\infty d\Sigma(t)/(t+1) < \infty$, and the class \mathbf{S}^{-1} of inverse Stieltjes functions

$$V(z) = \alpha + \beta z + \int_0^\infty \left(\frac{1}{t-z} - \frac{1}{t} \right) d\Sigma(t), \quad (3)$$

where $\alpha \leq 0$, $\beta \geq 0$, and $\int_0^\infty d\Sigma(t)/(t^2 + 1) < \infty$.

2 SPECIAL REALIZATION PROBLEMS

One way to characterize Herglotz-Nevanlinna functions is to identify them as (linear fractional transformations of) transfer functions:

$$V(z) = i[W(z) + I]^{-1}[W(z) - I]J, \quad (4)$$

where $J = J^* = J^{-1}$ and $W(z)$ is the transfer function of some generalized linear, stationary, conservative dynamical system (cf. [1], [3]). The approach based on the use of Brodskii-Livšic operator colligations Θ yields to a simultaneous representation of the functions $W(z)$ and $V(z)$ in the form

$$W_\Theta(z) = I - 2iK^*(T - zI)^{-1}KJ, \quad (5)$$

$$V_\Theta(z) = K^*(T_R - zI)^{-1}K, \quad (6)$$

where T_R stands for the real part of T . The definitions and main results associated with Brodskii-Livšic type operator colligations in realization of Herglotz-Nevanlinna functions are as follows, cf. [8], [9], [16].

Let $T \in [\mathfrak{H}]$, i.e., T is a bounded linear mapping in a Hilbert space \mathfrak{H} , and assume that $\operatorname{Im} T = (T - T^*)/2i$ of T is represented as $\operatorname{Im} T = KJK^*$, where $K \in [\mathfrak{E}, \mathfrak{H}]$, and $J \in [\mathfrak{E}]$ is self-adjoint and unitary. Then the array

$$\Theta = \begin{pmatrix} T & K & J \\ \mathfrak{H} & & \mathfrak{E} \end{pmatrix} \quad (7)$$

defines a Brodskii-Livšic operator colligation, and the function $W_\Theta(z)$ given by (5) is the transfer function of Θ . In the case of the directing operator $J = I$ the system (7) is called a scattering system, in which case the main operator T of the system Θ is dissipative: $\text{Im } T \geq 0$. In system theory $W_\Theta(z)$ is interpreted as the transfer function of the conservative system (i.e., $\text{Im } T = KJK^*$) of the form $(T - zI)x = KJ\varphi_-$ and $\varphi_+ = \varphi_- - 2iK^*x$, where $\varphi_- \in \mathfrak{E}$ is an input vector, $\varphi_+ \in \mathfrak{E}$ is an output vector, and x is a state space vector in \mathfrak{H} , so that $\varphi_+ = W_\Theta(z)\varphi_-$. The system is said to be minimal if the main operator T of Θ is completely non self-adjoint (i.e., there are no nontrivial invariant subspaces on which T induces self-adjoint operators), cf. [8], [16]. A classical result due to Brodskii and Livšic [9] states that the compactly supported Herglotz-Nevanlinna functions of the form $\int_a^b d\Sigma(t)/(t - z)$ correspond to minimal systems Θ of the form (7) via (4) with $W(z) = W_\Theta(z)$ given by (5) and $V(z) = V_\Theta(z)$ given by (6).

Next consider a linear, stationary, conservative dynamical system Θ of the form

$$\Theta = \begin{pmatrix} \mathbb{A} & K & J \\ \mathfrak{H}_+ \subset \mathfrak{H} \subset \mathfrak{H}_- & \mathfrak{E} \end{pmatrix}. \quad (8)$$

Here $\mathbb{A} \in [\mathfrak{H}_+, \mathfrak{H}_-]$, where $\mathfrak{H}_+ \subset \mathfrak{H} \subset \mathfrak{H}_-$ is a rigged Hilbert space, $\mathbb{A} \supset T \supset A$, $\mathbb{A}^* \supset T^* \supset A$, A is a Hermitian operator in \mathfrak{H} , T is a non-Hermitian operator in \mathfrak{H} , $K \in [\mathfrak{E}, \mathfrak{H}_-]$, $J = J^* = J^{-1}$, and $\text{Im } \mathbb{A} = KJK^*$. In this case Θ is said to be a Brodskii-Livšic rigged operator colligation. The transfer function of Θ in (8) and its linear fractional transform are given by

$$W_\Theta(z) = I - 2iK^*(\mathbb{A} - zI)^{-1}KJ, \quad V_\Theta(z) = K^*(\mathbb{A}_R - zI)^{-1}K. \quad (9)$$

The functions $V(z)$ in (1) which can be realized in the form (4), (9) with a transfer function of a system Θ as in (8) have been characterized in [2], [5], [6], [7], [18]. For the significance of rigged Hilbert spaces in system theory, see [14], [16]. Systems (7) and (8) naturally appear in electrical engineering and scattering theory [16].

3 GENERAL REALIZATION PROBLEMS

In the particular case of Stieltjes functions or of inverse Stieltjes functions general realization results along the lines of [5], [6], [7] remain to be worked out in detail, cf. [4], [10].

The systems (7) and (8) are not general enough for the realization of general Herglotz-Nevanlinna functions in (1) without any conditions on $Q = Q^*$ and $L \geq 0$. However, a generalization of the Brodskii-Livšic operator colligation (7) leads to analogous realization results for Herglotz-Nevanlinna functions $V(z)$ of the form (1) whose spectral function is compactly supported: such functions $V(z)$ admit a realization via (4) with

$$\begin{aligned} W(z) &= W_\Theta(z) = I - 2iK^*(M - zF)^{-1}KJ, \\ V(z) &= W_\Theta(z) = K^*(M_R - zF)^{-1}K, \end{aligned} \quad (10)$$

where $M = M_R + iKJK^*$, $M_R \in [\mathfrak{H}]$ is the real part of M , F is a finite-dimensional orthogonal projector, and Θ is a generalized Brodskii-Livšic operator colligation of the form

$$\Theta = \begin{pmatrix} M & F & K & J \\ \mathfrak{H} & & & \mathfrak{E} \end{pmatrix}, \quad (11)$$

see [11], [12], [13]. The basic open problems are:

Determine the class of linear, conservative, time-invariant dynamical systems (new type of operator colligations) such that an arbitrary matrix-valued Herglotz-Nevanlinna function $V(z)$ acting on \mathfrak{E} can be realized as a linear fractional transformation (4) of the matrix-valued transfer function $W_\Theta(z)$ of some minimal system Θ from this class.

Find criteria for a given matrix-valued Stieltjes or inverse Stieltjes function acting on \mathfrak{E} to be realized as a linear fractional transformation of the matrix-valued transfer function of a minimal Brodskii-Livšic type system Θ in (8) with: (i) an accretive operator \mathbb{A} , (ii) an α -sectorial operator \mathbb{A} , or (iii) an extremal operator \mathbb{A} (accretive but not α -sectorial).

The same problem for the (compactly supported) matrix-valued Stieltjes or inverse Stieltjes functions and the generalized Brodskii-Livšic systems of the form (11) with the main operator M and the finite-dimensional orthogonal projector F .

There is a close connection to the so-called regular impedance conservative systems (where the coefficient of the derivative is invertible) that were recently considered in [17] (see also [19]). It is shown that any function $D(s)$ with non-negative real part in the open right half-plane and for which $D(s)/s \rightarrow 0$ as $s \rightarrow \infty$ has a realization with such an impedance conservative system.

BIBLIOGRAPHY

- [1] D. Alpay, A. Dijksma, J. Rovnyak, and H.S.V. de Snoo, "Schur functions, operator colligations, and reproducing kernel Pontryagin spaces," *Oper. Theory Adv. Appl.*, 96, Birkhäuser Verlag, Basel, 1997.
- [2] Yu. M. Arlinskii, "On the inverse problem of the theory of characteristic functions of unbounded operator colligations", *Dopovidi Akad. Nauk Ukrain. RSR*, 2 (1976), 105–109 (Russian).
- [3] D. Z. Arov, "Passive linear steady-state dynamical systems," *Sibirsk. Mat. Zh.*, 20, no. 2, (1979), 211–228, 457 (Russian) [English transl.: *Siberian Math. J.*, 20 no. 2, (1979) 149–162].

- [4] S. V. Belyi, S. Hassi, H. S. V. de Snoo, and E. R. Tsekanovskii, "On the realization of inverse Stieltjes functions," *Proceedings of the 15th International Symposium on Mathematical Theory of Networks and Systems*, Editors D. Gillian and J. Rosenthal, University of Notre Dame, South Bend, Indiana, USA, 2002, http://www.nd.edu/~mtns/papers/20160_6.pdf
- [5] S. V. Belyi and E. R. Tsekanovskii, "Realization and factorization problems for J -contractive operator-valued functions in half-plane and systems with unbounded operators," *Systems and Networks: Mathematical Theory and Applications*, Akademie Verlag, 2 (1994), 621–624.
- [6] S. V. Belyi and E. R. Tsekanovskii, "Realization theorems for operator-valued R -functions," *Oper. Theory Adv. Appl.*, 98 (1997), 55–91.
- [7] S. V. Belyi and E. R. Tsekanovskii, "On classes of realizable operator-valued R -functions," *Oper. Theory Adv. Appl.*, 115 (2000), 85–112.
- [8] M. S. Brodskii, "Triangular and Jordan representations of linear operators," Moscow, Nauka, 1969 (Russian) [English trans.: Vol. 32 of *Transl. Math. Monographs*, Amer. Math. Soc., 1971].
- [9] M. S. Brodskii and M. S. Livšic, "Spectral analysis of non-selfadjoint operators and intermediate systems," *Uspekhi Mat. Nauk*, 13 no. 1, 79, (1958), 3–85 (Russian) [English trans.: *Amer. Math. Soc. Transl.*, (2) 13 (1960), 265–346].
- [10] I. Dovshenko and E. R. Tsekanovskii, "Classes of Stieltjes operator-functions and their conservative realizations," *Dokl. Akad. Nauk SSSR*, 311 no. 1 (1990), 18–22.
- [11] S. Hassi, H. S. V. de Snoo, and E. R. Tsekanovskii, "An addendum to the multiplication and factorization theorems of Brodskii-Livšic-Potapov," *Appl. Anal.*, 77 (2001), 125–133.
- [12] S. Hassi, H. S. V. de Snoo, and E. R. Tsekanovskii, "On commutative and noncommutative representations of matrix-valued Herglotz-Nevanlinna functions," *Appl. Anal.*, 77 (2001), 135–147.
- [13] S. Hassi, H. S. V. de Snoo, and E. R. Tsekanovskii, "Realizations of Herglotz-Nevanlinna functions via F -systems," *Oper. Theory: Adv. Appl.*, 132 (2002), 183–198.
- [14] J.W. Helton, "Systems with infinite-dimensional state space: the Hilbert space approach," *Proc. IEEE*, 64 (1976), no. 1, 145–160.
- [15] I. S. Kač and M. G. Kreĭn, "The R -functions: Analytic functions mapping the upper half-plane into itself," Supplement I to the Russian edition of F. V. Atkinson, *Discrete and Continuous Boundary Problems*, Moscow, 1974 [English trans.: *Amer. Math. Soc. Trans.*, (2) 103 (1974), 1–18].

- [16] M. S. Livšic, “Operators, Oscillations, Waves,” Moscow, Nauka, 1966 (Russian) [English trans.: Vol. 34 of *Trans. Math. Monographs*, Amer. Math. Soc., 1973].
- [17] O. J. Staffans, “Passive and conservative infinite-dimensional impedance and scattering systems (from a personal point of view),” *Proceedings of the 15th International Symposium on Mathematical Theory of Networks and Systems*, Ed., D. Gillian and J. Rosenthal, University of Notre Dame, South Bend, Indiana, USA, 2002, Plenary talk, <http://www.nd.edu/~mtns>
- [18] E. R. Tsekanovskii and Yu. L. Shmul’yan, “The theory of biextensions of operators in rigged Hilbert spaces: Unbounded operator colligations and characteristic functions,” *Uspekhi Mat. Nauk*, 32 (1977), 69–124 (Russian) [English transl.: *Russian Math. Surv.*, 32 (1977), 73–131].
- [19] G. Weiss, “Transfer functions of regular linear systems. Part I: characterizations of regularity”, *Trans. Amer. Math. Soc.*, 342 (1994), 827–854.

Problem 1.3

Does any analytic contractive operator function on the polydisk have a dissipative scattering nD realization?

Dmitry S. Kalyuzhniy-Verbovetzky

Department of Mathematics
The Weizmann Institute of Science
Rehovot 76100
Israel
dmitryk@wisdom.weizmann.ac.il

1 DESCRIPTION OF THE PROBLEM

Let $\mathcal{X}, \mathcal{U}, \mathcal{Y}$ be finite-dimensional or infinite-dimensional separable Hilbert spaces. Consider nD linear systems of the form

$$\alpha : \begin{cases} x(t) &= \sum_{k=1}^n (A_k x(t - e_k) + B_k u(t - e_k)), \\ y(t) &= \sum_{k=1}^n (C_k x(t - e_k) + D_k u(t - e_k)), \end{cases} \quad (t \in \mathbb{Z}^n : \sum_{k=1}^n t_k > 0) \quad (1)$$

where $e_k := (0, \dots, 0, 1, 0, \dots, 0) \in \mathbb{Z}^n$ (here unit is on the k -th place), for all $t \in \mathbb{Z}^n$ such that $\sum_{k=1}^n t_k \geq 0$ one has $x(t) \in \mathcal{X}$ (the state space), $u(t) \in \mathcal{U}$ (the input space), $y(t) \in \mathcal{Y}$ (the output space), A_k, B_k, C_k, D_k are bounded linear operators, i.e., $A_k \in L(\mathcal{X}), B_k \in L(\mathcal{U}, \mathcal{X}), C_k \in L(\mathcal{X}, \mathcal{Y}), D_k \in L(\mathcal{U}, \mathcal{Y})$ for all $k \in \{1, \dots, n\}$. We use the notation $\alpha = (n; \mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ for such a system (here $\mathbf{A} := (A_1, \dots, A_n)$, etc.). For $\mathbf{T} \in L(\mathcal{H}_1, \mathcal{H}_2)^n$ and $z \in \mathbb{C}^n$ denote $z\mathbf{T} := \sum_{k=1}^n z_k T_k$. Then the *transfer function* of α is

$$\theta_\alpha(z) = z\mathbf{D} + z\mathbf{C}(I_{\mathcal{X}} - z\mathbf{A})^{-1}z\mathbf{B}.$$

Clearly, θ_α is analytic in some neighbourhood of $z = 0$ in \mathbb{C}^n . Let

$$G_k := \begin{pmatrix} A_k & B_k \\ C_k & D_k \end{pmatrix} \in \mathcal{L}(\mathcal{X} \oplus \mathcal{U}, \mathcal{X} \oplus \mathcal{Y}), \quad k = 1, \dots, n.$$

We call $\alpha = (n; \mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ a *dissipative scattering nD system* (see [5, 6]) if for any $\zeta \in \mathbb{T}^n$ (the unit torus) $\zeta\mathbf{G}$ is a contractive operator, i.e.,

$\|\zeta \mathbf{G}\| \leq 1$. It is known [5] that the transfer function of a dissipative scattering nD system $\alpha = (n; \mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ belongs to the subclass $B_n^0(\mathcal{U}, \mathcal{Y})$ of the class $B_n(\mathcal{U}, \mathcal{Y})$ of all analytic contractive $L(\mathcal{U}, \mathcal{Y})$ -valued functions on the open unit polydisk \mathbb{D}^n , which is segregated by the condition of vanishing of its functions at $z = 0$. The question whether the converse is true was implicitly asked in [5] and still has not been answered. Thus, we pose the following problem.

Problem: Either prove that an arbitrary $\theta \in B_n^0(\mathcal{U}, \mathcal{Y})$ can be realized as the transfer function of a dissipative scattering nD system of the form (1) with the input space \mathcal{U} and the output space \mathcal{Y} , or give an example of a function $\theta \in B_n^0(\mathcal{U}, \mathcal{Y})$ (for some $n \in \mathbb{N}$, and some finite-dimensional or infinite-dimensional separable Hilbert spaces \mathcal{U}, \mathcal{Y}) that has no such a realization.

2 MOTIVATION AND HISTORY OF THE PROBLEM

For $n = 1$ the theory of dissipative (or passive, in other terminology) scattering linear systems is well developed (see, e.g., [2, 3]) and related to various problems of physics (in particular, scattering theory), stochastic processes, control theory, operator theory, and 1D complex analysis. It is well known (essentially, due to [8]) that the class of transfer functions of dissipative scattering 1D systems of the form (1) with the input space \mathcal{U} and the output space \mathcal{Y} coincides with $B_1^0(\mathcal{U}, \mathcal{Y})$. Moreover, this class of transfer functions remains the same when one is restricted within the important special case of conservative scattering 1D systems, for which the system block matrix G is unitary, i.e., $G^*G = I_{\mathcal{X} \oplus \mathcal{U}}$, $GG^* = I_{\mathcal{X} \oplus \mathcal{Y}}$. Let us note that in the case $n = 1$ a system (1) can be rewritten in an equivalent form (without a unit delay in output signal y) that is the standard form of a linear system, then a transfer function does not necessarily vanish at $z = 0$, and the class of transfer functions turns into the Schur class $S(\mathcal{U}, \mathcal{Y}) = B_1(\mathcal{U}, \mathcal{Y})$. The classes $B_1^0(\mathcal{U}, \mathcal{Y})$ and $B_1(\mathcal{U}, \mathcal{Y})$ are canonically isomorphic due to the relation $B_1^0(\mathcal{U}, \mathcal{Y}) = zB_1(\mathcal{U}, \mathcal{Y})$.

In [1] an important subclass $S_n(\mathcal{U}, \mathcal{Y})$ in $B_n(\mathcal{U}, \mathcal{Y})$ was introduced. This subclass consists of analytic $L(\mathcal{U}, \mathcal{Y})$ -valued functions on \mathbb{D}^n , say, $\theta(z) = \sum_{t \in \mathbb{Z}_+^n} \theta_t z^t$ (here $\mathbb{Z}_+^n = \{t \in \mathbb{Z}^n : t_k \geq 0, k = 1, \dots, n\}$, $z^t := \prod_{k=1}^n z_k^{t_k}$ for $z \in \mathbb{D}^n, t \in \mathbb{Z}_+^n$) such that for any n -tuple $\mathbf{T} = (T_1, \dots, T_n)$ of commuting contractions on some common separable Hilbert space \mathcal{H} and any positive $r < 1$ one has $\|\theta(r\mathbf{T})\| \leq 1$, where $\theta(r\mathbf{T}) = \sum_{t \in \mathbb{Z}_+^n} \theta_t \otimes (r\mathbf{T})^t \in L(\mathcal{U} \otimes \mathcal{H}, \mathcal{Y} \otimes \mathcal{H})$, and $(r\mathbf{T})^t := \prod_{k=1}^n (rT_k)^{t_k}$. For $n = 1$ and $n = 2$ one has $S_n(\mathcal{U}, \mathcal{Y}) = B_n(\mathcal{U}, \mathcal{Y})$. However, for any $n > 2$ and any non-zero spaces \mathcal{U} and \mathcal{Y} the class $S_n(\mathcal{U}, \mathcal{Y})$ is a proper subclass of $B_n(\mathcal{U}, \mathcal{Y})$. J. Agler in [1] constructed a representation of an arbitrary function from $S_n(\mathcal{U}, \mathcal{Y})$, which in a system-theoretical language was interpreted in [4] as follows: $S_n(\mathcal{U}, \mathcal{Y})$

coincides with the class of transfer functions of nD systems of Roesser type with the input space \mathcal{U} and the output space \mathcal{Y} , and certain conservativity condition imposed. The analogous result is valid for conservative systems of the form (1). A system $\alpha = (n; \mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ is called a *conservative scattering nD system* if for any $\zeta \in \mathbb{T}^n$ the operator $\zeta \mathbf{G}$ is unitary. Clearly, a conservative scattering system is a special case of a dissipative one. By [5], the class of transfer functions of conservative scattering nD systems coincides with the subclass $S_n^0(\mathcal{U}, \mathcal{Y})$ in $S_n(\mathcal{U}, \mathcal{Y})$, which is segregated from the latter by the condition of vanishing of its functions at $z = 0$. Since for $n = 1$ and $n = 2$ one has $S_n^0(\mathcal{U}, \mathcal{Y}) = B_n^0(\mathcal{U}, \mathcal{Y})$, this gives the whole class of transfer functions of dissipative scattering nD systems of the form (1), and the solution to the problem formulated above for these two cases.

In [6] the dilation theory for nD systems of the form (1) was developed. It was proven that $\alpha = (n; \mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}; \mathcal{X}, \mathcal{U}, \mathcal{Y})$ has a conservative dilation if and only if the corresponding linear function $L_{\mathbf{G}}(z) := z\mathbf{G}$ belongs to $S_n^0(\mathcal{X} \oplus \mathcal{U}, \mathcal{X} \oplus \mathcal{Y})$. Systems that satisfy this criterion are called n -dissipative scattering ones. In the cases $n = 1$ and $n = 2$ the subclass of n -dissipative scattering systems coincides with the whole class of dissipative ones, and in the case $n > 2$ this subclass is proper. Since transfer functions of a system and of its dilation coincide, the class of transfer functions of n -dissipative scattering systems with the input space \mathcal{U} and the output space \mathcal{Y} is $S_n^0(\mathcal{U}, \mathcal{Y})$. According to [7], for any $n > 2$ there exist $p \in \mathbb{N}, m \in \mathbb{N}$, operators $D_k \in L(\mathbb{C}^p)$ and commuting contractions $T_k \in L(\mathbb{C}^m)$, $k = 1, \dots, n$, such that

$$\max_{\zeta \in \mathbb{T}^n} \left\| \sum_{k=1}^n \zeta_k D_k \right\| = 1 < \left\| \sum_{k=1}^n T_k \otimes D_k \right\|.$$

The system $\alpha = (n; 0, 0, 0, \mathbf{D}; \{0\}, \mathbb{C}^p, \mathbb{C}^p)$ is a dissipative scattering one, however not, n -dissipative. Its transfer function $\theta_{\alpha}(z) = L_{\mathbf{G}}(z) = z\mathbf{D} \in B_n^0(\mathbb{C}^p, \mathbb{C}^p) \setminus S_n^0(\mathbb{C}^p, \mathbb{C}^p)$.

Since for functions in $B_n^0(\mathcal{U}, \mathcal{Y}) \setminus S_n^0(\mathcal{U}, \mathcal{Y})$ the realization technique elaborated in [1] and developed in [4] and [5] is not applicable, our problem is of current interest.

BIBLIOGRAPHY

- [1] J. Agler, "On the representation of certain holomorphic functions defined on a polydisc," *Topics in Operator Theory: Ernst D. Hellinger Memorial Volume* (L. de Branges, I. Gohberg, and J. Rovnyak, Eds.), Oper. Theory Adv. Appl. 48, pp. 47-66 (1990).
- [2] D. Z. Arov, "Passive linear steady-state dynamic systems," *Sibirsk. Math. Zh.* 20 (2), 211-228 (1979), (Russian).
- [3] J. A. Ball and N. Cohen, "De Branges-Rovnyak operator models and systems theory: A survey," *Topics in Matrix and Operator Theory* (H.

- Bart, I. Gohberg, and M.A. Kaashoek, eds.), *Oper. Theory Adv. Appl.*, 50, pp. 93-136 (1991).
- [4] J. A. Ball and T. Trent, "Unitary colligations, reproducing kernel hilbert spaces, and Nevanlinna-Pick interpolation in several variables," *J. Funct. Anal.* 157, pp. 1-61 (1998).
- [5] D. S. Kalyuzhniy, "Multiparametric dissipative linear stationary dynamical scattering systems: Discrete case," *J. Operator Theory*, 43 (2), pp. 427-460 (2000).
- [6] D. S. Kalyuzhniy, "Multiparametric dissipative linear stationary dynamical scattering systems: Discrete case, II: Existence of conservative dilations," *Integr. Eq. Oper. Th.*, 36 (1), pp. 107-120 (2000).
- [7] D. S. Kalyuzhniy, "On the von Neumann inequality for linear matrix functions of several variables," *Mat. Zametki* 64 (2), pp. 218-223 (1998), (Russian); translated in *Math. Notes* 64 (2), pp. 186-189 (1998).
- [8] B. Sz.-Nagy and C. Foias, *Harmonic Analysis of Operators on Hilbert Spaces*, North Holland, Amsterdam, 1970.

Problem 1.4

Partial disturbance decoupling with stability

J. C. Martínez-García

Programa de Investigación en Matemáticas Aplicadas y Computación
Instituto Mexicano del Petróleo
Eje Central Lázaro Cárdenas No. 152
Col San Bartolo Atepehuacan, 07730 México D.F.,
México
`martinez@ctrl.cinvestav.mx`

M. Malabre

Institut de Recherche en Communications et Cybernétique de Nantes
CNRS- & (Ecole Centrale-Université-Ecole des Mines) de Nantes
1 rue de la Noë, F-44321 Nantes Cedex 03,
France
`Michel.Malabre@irccyn.ec-nantes.fr`

V. Kučera

Faculty of Electrical Engineering
Czech Technical University in Prague
Technická 2, 16627 Prague 6,
Czech Republic
`kucera@fel.vcut.cz`

1 DESCRIPTION OF THE PROBLEM

Consider a linear time-invariant system (A, B, C, E) described by:

$$\begin{cases} \sigma x(t) = Ax(t) + Bu(t) + Ed(t), \\ z(t) = Cx(t), \end{cases} \quad (1)$$

where σ denotes either the derivation or the shift operator, depending on the continuous-time or discrete-time context; $x(t) \in \mathcal{X} \simeq \mathbb{R}^n$ denotes the state; $u(t) \in \mathcal{U} \simeq \mathbb{R}^m$ denotes the control input; $z(t) \in \mathcal{Z} \simeq \mathbb{R}^m$ denotes the output, and $d(t) \in \mathcal{D} \simeq \mathbb{R}^p$ denotes the disturbance. $A : \mathcal{X} \rightarrow \mathcal{X}$, $B : \mathcal{U} \rightarrow \mathcal{X}$, $C : \mathcal{X} \rightarrow \mathcal{Z}$, and $E : \mathcal{D} \rightarrow \mathcal{X}$ denote linear maps represented by real constant matrices.

Let a system (A, B, C, E) and an integer $k \geq 1$ be given. Find necessary and sufficient conditions for the existence of a static state feedback control law $u(t) = Fx(t) + Gd(t)$, where $F: \mathcal{X} \rightarrow \mathcal{U}$ and $G: \mathcal{D} \rightarrow \mathcal{U}$ are linear maps such as zeroing the first k Markov parameters of T_{zd} , the transfer function between the disturbance and the controlled output, while insuring internal stability, i.e.:

- $C(A + BF)^l(BG + E) \equiv 0$, for $i \in \{0, 1, \dots, k - 1\}$, and
- $\sigma(A + BF) \subseteq C_g$,

where $\sigma(A + BF)$ stands for the spectrum of $A + BF$ and C_g stands for the (good) stable part of the complex plane, e.g., the open left-half complex plane (continuous-time case) or the open unit disk (discrete-time case)

2 MOTIVATION

The literature contains a lot of contributions related to disturbance rejection or attenuation. The early attempts were devoted to canceling the effect of the disturbance on the controlled output, i.e., insuring $T_{zd} \equiv 0$. This problem is usually referred to as the disturbance decoupling problem with internal stability, noted as DDPS (see [11], [1]).

The solvability conditions for DDPS can be expressed as matching of infinite and unstable (invariant) zeros of certain systems (see, for instance, [8]), namely those of (A, B, C) , i.e., (1) with $d(t) \equiv 0$, and those of $(A, \begin{bmatrix} B & E \end{bmatrix}, C)$, i.e., (1) with $d(t)$ considered as a control input. However, the rigid solvability conditions for DDPS are hardly met in practical cases. This is why alternative design procedures have been considered, such as almost disturbance decoupling (see [10]) and optimal disturbance attenuation, i.e., minimization of a norm of T_{zd} (see, for instance, [12]).

The partial version of the problem, as defined in Section 1, offers another alternative from the rigid design of DDPS. The partial disturbance decoupling problem (PDDP) amounts to zeroing the first, say k , Markov parameters of T_{zd} . It was initially introduced in [2] and later revisited in [5], without stability, [6, 7] with dynamic state feedback and stability, [4] with static state feedback and stability (sufficient solvability conditions for the single-input single-output case), [3] with dynamic measurement feedback, stability, and H_∞ -norm bound. When no stability constraint is imposed, solvability conditions of PDDP involve only a subset of the infinite structure of (A, B, C) and $(A, \begin{bmatrix} B & E \end{bmatrix}, C)$, namely the orders which are less than or equal to $k - 1$ (see details in [5]). For PDDPS (i.e., PDDP with internal stability), the role played by the finite invariant zeros must be clarified to obtain the necessary and sufficient conditions that we are looking for, and solve the open problem.

Several extensions of this problem are also important:

- solve PDDPS while reducing the H_∞ -norm of T_{zd} ;
- consider static measurement feedback in place of static state feedback.

BIBLIOGRAPHY

- [1] G. Basile and G. Marro, *Controlled and Conditioned Invariants in Linear System Theory*, Prentice-Hall, 1992.
- [2] E. Emre and L. M. Silverman, "Partial model matching of linear systems," *IEEE Trans. Automat. Contr.*, vol. AC-25, no. 2, pp. 280-281, 1980.
- [3] V. Eldem, H. Özbay, H. Selbuz, and K. Özcaldiran, "Partial disturbance rejection with internal stability and H_∞ norm bound," *SIAM Journal on Control and Optimization*, vol. 36, no. 1, pp. 180-192, 1998.
- [4] F. N. Koumboulis and V. Kučera, "Partial model matching via static feedback (The multivariable case)," *IEEE Trans. Automat. Contr.*, vol. AC-44, no. 2, pp. 386-392, 1999.
- [5] M. Malabre and J. C. Martínez-García, "The partial disturbance rejection or partial model matching: Geometric and structural solutions," *IEEE Trans. Automat. Contr.*, vol. AC-40, no. 2, pp. 356-360, 1995.
- [6] V. Kučera, J. C. Martínez-García, and M. Malabre, "Partial model matching: Parametrization of solutions," *Automatica*, vol. 33, no. 5, pp. 975-977, 1997.
- [7] J. C. Martínez-García, M. Malabre, and V. Kučera, "The partial model matching problem with stability," *Systems and Control Letters*, no. 24, pp. 61-74, 1994.
- [8] J. C. Martínez-García, M. Malabre, J.-M. Dion, and C. Commault, "Condensed structural solutions to the disturbance rejection and decoupling problems with stability," *International Journal of Control*, vol. 72, No. 15, pp. 1392-1401, 1999.
- [9] A. Saberi, P. Sannuti, A. A. Stoorvogel, and B. M. Chen, *H_2 Optimal Control*, Prentice-Hall, 1995.
- [10] J. C. Willems, "Almost invariant subspaces: An approach to high gain feedback design - part I: Almost controlled invariant subspaces," *IEEE Trans. Automat. Contr.*, vol. AC-26, no.1, pp. 235-252, 1981.
- [11] M. M. Wonham, *Linear Multivariable Control: A Geometric Approach*, 3rd ed., Springer Verlag, New York, 1985.

- [12] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*, Upper Saddle River, NJ: Prentice-Hall, Inc., Simon & Schuster, 1995.

Problem 1.5

Is Monopoli's model reference adaptive controller correct?

A. S. Morse¹

Center for Computational Vision and Control
Department of Electrical Engineering
Yale University, New Haven, CT 06520
USA

1 INTRODUCTION

In 1974 R. V. Monopoli published a paper [1] in which he posed the now classical model reference adaptive control problem, proposed a solution and presented arguments intended to establish the solution's correctness. Subsequent research [2] revealed a flaw in his proof, which placed in doubt the correctness of the solution he proposed. Although provably correct solutions to the model reference adaptive control problem now exist (see [3] and the references therein), the problem of deciding whether or not Monopoli's original proposed solution is in fact correct remains unsolved. The aim of this note is to review the formulation of the classical model reference adaptive control problem, to describe Monopoli's proposed solution, and to outline what's known at present about its correctness.

2 THE CLASSICAL MODEL REFERENCE ADAPTIVE CONTROL PROBLEM

The *classical model reference adaptive control problem* is to develop a dynamical controller capable of causing the output y of an imprecisely modeled SISO process \mathbb{P} to approach and track the output y_{ref} of a prespecified reference model \mathbb{M}_{ref} with input r . The underlying assumption is that the process model is known only to the extent that it is one of the members of a pre-specified class \mathcal{M} . In the classical problem \mathcal{M} is taken to be the set of

¹This research was supported by DARPA under its SEC program and by the NSF.

all SISO controllable, observable linear systems with strictly proper transfer functions of the form $g \frac{\beta(s)}{\alpha(s)}$ where g is a nonzero constant called the *high frequency gain* and $\alpha(s)$ and $\beta(s)$ are monic, coprime polynomials. All g have the same sign and each transfer function is *minimum phase* (i.e., each $\beta(s)$ is stable). All transfer functions are required to have the same relative degree \bar{n} (i.e., $\deg \alpha(s) - \deg \beta(s) = \bar{n}$.) and each must have a McMillan degree not exceeding some prespecified integer n (i.e., $\deg \alpha(s) \leq n$). In the sequel we are going to discuss a simplified version of the problem in which all $g = 1$ and the reference model transfer function is of the form $\frac{1}{(s+\lambda)^{\bar{n}}}$ where λ is a positive number. Thus \mathbb{M}_{ref} is a system of the form

$$\dot{y}_{\text{ref}} = -\lambda y_{\text{ref}} + \bar{c}x_{\text{ref}} + \bar{d}r \quad \dot{x}_{\text{ref}} = \bar{A}x_{\text{ref}} + \bar{b}r \quad (1)$$

where $\{\bar{A}, \bar{b}, \bar{c}, \bar{d}\}$ is a controllable, observable realization of $\frac{1}{(s+\lambda)^{(\bar{n}-1)}}$.

3 MONOPOLI'S PROPOSED SOLUTION

Monopoli's proposed solution is based on a special representation of \mathbb{P} that involves picking any n -dimensional, single-input, controllable pair (A, b) with A stable. It is possible to prove [1, 4] that the assumption that the process \mathbb{P} admits a model in \mathcal{M} , implies the existence of a vector $p^* \in \mathbb{R}^{2n}$ and initial conditions $z(0)$ and $\bar{x}(0)$, such that u and y exactly satisfy

$$\begin{aligned} \dot{z} &= \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} z + \begin{bmatrix} b \\ 0 \end{bmatrix} y + \begin{bmatrix} 0 \\ b \end{bmatrix} u \\ \dot{\bar{x}} &= \bar{A}\bar{x} + \bar{b}(u - z'p^*) \\ \dot{y} &= -\lambda y + \bar{c}\bar{x} + \bar{d}(u - z'p^*) \end{aligned}$$

Monopoli combined this model with that of \mathbb{M}_{ref} to obtain the *direct control model reference parameterization*

$$\dot{z} = \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} z + \begin{bmatrix} b \\ 0 \end{bmatrix} y + \begin{bmatrix} 0 \\ b \end{bmatrix} u \quad (2)$$

$$\dot{\bar{x}} = \bar{A}\bar{x} + \bar{b}(u - z'p^* - r) \quad (3)$$

$$\dot{\mathbf{e}}_{\mathbf{T}} = -\lambda \mathbf{e}_{\mathbf{T}} + \bar{c}\bar{x} + \bar{d}(u - z'p^* - r) \quad (4)$$

Here $\mathbf{e}_{\mathbf{T}}$ is the *tracking error*

$$\mathbf{e}_{\mathbf{T}} \triangleq y - y_{\text{ref}} \quad (5)$$

and $x \triangleq \bar{x} - x_{\text{ref}}$. Note that it is possible to generate an asymptotically correct estimate \hat{z} of z using a copy of (2) with \hat{z} replacing z . To keep the exposition simple, we are going to ignore the exponentially decaying estimation error $\hat{z} - z$ and assume that z can be measured directly.

To solve the MRAC problem, Monopoli proposed a control law of the form

$$u = z'\hat{p} + r \quad (6)$$

where \hat{p} is a suitably defined estimate of p^* . Motivation for this particular choice stems from the fact that if one knew p^* and were thus able to use the control $u = z'p^* + r$ instead of (6), then this would cause \mathbf{e}_T to tend to zero exponentially fast and tracking would therefore be achieved.

Monopoli proposed to generate \hat{p} using two subsystems that we will refer to here as a “multi-estimator” and a “tuner” respectively. A *multi-estimator* $\mathbb{E}(\hat{p})$ is a parameter-varying linear system with parameter \hat{p} , whose inputs are u , y , and r and whose output is an estimate \hat{e} of \mathbf{e}_T that would be asymptotically correct were \hat{p} held fixed at p^* . It turns out that there are two different but very similar types of multi-estimators that have the requisite properties. While Monopoli focused on just one, we will describe both since each is relevant to the present discussion. Both multi-estimators contain (2) as a subsystem.

Version 1

There are two versions of the adaptive controller that are relevant to the problem at hand. In this section we describe the multi-estimator and tuner that, together with reference model (1) and control law (6), comprise the first version.

Multi-Estimator 1

The form of the first multi-estimator $\mathbb{E}_1(\hat{p})$ is suggested by the readily verifiable fact that if H_1 and w_1 are $\bar{n} \times 2n$ and $\bar{n} \times 1$ signal matrices generated by the equations

$$\dot{H}_1 = \bar{A}H_1 + \bar{b}z' \quad \text{and} \quad \dot{w}_1 = \bar{A}w_1 + \bar{b}(u - r) \quad (7)$$

respectively, then $w_1 - H_1p^*$ is a solution to (3). In other words $x = w_1 - H_1p^* + \epsilon$ where ϵ is an initial condition dependent time function decaying to zero as fast as $e^{\bar{A}t}$. Again, for simplicity, we shall ignore ϵ . This means that (4) can be re-written as

$$\dot{\mathbf{e}}_T = -\lambda \mathbf{e}_T - (\bar{c}H_1 + \bar{d}z')p^* + \bar{c}w_1 + \bar{d}(u - r)$$

Thus a natural way to generate an estimate \hat{e}_1 of \mathbf{e}_T is by means of the equation

$$\dot{\hat{e}}_1 = -\lambda \hat{e}_1 - (\bar{c}H_1 + \bar{d}z')\hat{p} + \bar{c}w_1 + \bar{d}(u - r) \quad (8)$$

From this it clearly follows that the multi-estimator $\mathbb{E}_1(\hat{p})$ defined by (2), (7) and (8) has the required property of delivering an asymptotically correct estimate \hat{e}_1 of \mathbf{e}_T if \hat{p} is fixed at p^* .

Tuner 1

From (8) and the differential equation for \mathbf{e}_T directly above it, it can be seen that the estimation error²

$$e_1 \triangleq \hat{e}_1 - \mathbf{e}_T \quad (9)$$

satisfies the error equation

$$\dot{e}_1 = -\lambda e_1 + \phi_1'(\hat{p} - p^*) \quad (10)$$

where

$$\phi_1' = -(\bar{c}H_1 + \bar{d}z') \quad (11)$$

Prompted by this, Monopoli proposed to tune \hat{p}_1 using the pseudo-gradient tuner

$$\dot{\hat{p}}_1 = -\phi_1 e_1 \quad (12)$$

The motivation for considering this particular tuning law will become clear shortly, if it is not already.

What is known about Version 1?

The overall model reference adaptive controller proposed by Monopoli thus consists of the reference model (1), the control law (6), the multi-estimator (2), (7), (8), the output estimation error (9) and the tuner (11), (12). *The open problem is to prove that this controller either solves the model reference adaptive control problem or that it does not.*

Much is known that is relevant to the problem. In the first place, note that (1), (2) together with (5) - (11) define a parameter varying linear system $\Sigma_1(\hat{p})$ with input r , state $(y_{\text{ref}}, x_{\text{ref}}, z, H_1, w_1, \hat{e}_1, e_1)$ and output e_1 . The consequence of the assumption that every system in \mathcal{M} is minimum phase is that $\Sigma_1(\hat{p})$ is detectable through e_1 for every fixed value of \hat{p} [5]. Meanwhile the form of (10) enables one to show by direct calculation, that the rate of change of the partial Lyapunov function $V \triangleq e_1^2 + \|\hat{p} - p^*\|^2$ along a solution to (12) and the equations defining $\Sigma_1(\hat{p})$, satisfies

$$\dot{V} = -2\lambda e_1^2 \leq 0 \quad (13)$$

From this it is evident that V is a bounded monotone nonincreasing function and consequently that e_1 and \hat{p} are bounded wherever they exist. Using and the fact that $\Sigma_1(\hat{p})$ is a *linear* parameter-varying system, it can be concluded that solutions exist globally and that e_1 and \hat{p} are bounded on $[0, \infty)$. By integrating (13) it can also be concluded that e_1 has a finite $\mathcal{L}^2[0, \infty)$ -norm and that $\|e_1\|^2 + \|\hat{p} - p^*\|^2$ tends to a finite limit as $t \rightarrow \infty$. Were it possible to deduce from these properties that \hat{p} tended to a limit \bar{p} , then it would be possible to establish correctness of the overall adaptive controller using the detectability of $\Sigma_1(\bar{p})$.

²Monopoli called e_1 an *augmented error*.

There are two very special cases for which correctness has been established. The first is when the process models in \mathcal{M} all have relative degree 1; that is when $\bar{n} = 1$. See the references cited in [3] for more on this special case. The second special case is when p^* is taken to be of the form q^*k where k is a known vector and q^* is a scalar; in this case $\hat{p} \triangleq \hat{q}k$ where \hat{q} is a scalar parameter tuned by the equation $\dot{\hat{q}} = -k'\phi_1 e_1$ [6].

Version 2

In the sequel we describe the multi-estimator and tuner that, together with reference model (1) and control law (6), comprise the second version of them adaptive controller relevant to the problem at hand.

Multi-Estimator 2

The second multi-estimator $\mathbb{E}_2(\hat{p})$, which is relevant to the problem under consideration, is similar to $\mathbb{E}_1(\hat{p})$ but has the slight advantage of leading to a tuner that is somewhat easier to analyze. To describe $\mathbb{E}_2(\hat{p})$, we need first to define matrices

$$\bar{A}_2 \triangleq \begin{bmatrix} \bar{A} & 0 \\ \bar{c} & -\lambda \end{bmatrix} \quad \text{and} \quad \bar{b}_2 \triangleq \begin{bmatrix} \bar{b} \\ \bar{d} \end{bmatrix}$$

The form of $\mathbb{E}_2(\hat{p})$ is motivated by the readily verifiable fact that if H_2 and w_2 are $(\bar{n}+1) \times 2n$ and $(\bar{n}+1) \times 1$ signal matrices generated by the equations

$$\dot{H}_2 = \bar{A}_2 H_2 + \bar{b}_2 z' \quad \text{and} \quad \dot{w}_2 = \bar{A}_2 w_2 + \bar{b}_2 (u - r) \quad (14)$$

then $w_2 - H_2 p^*$ is a solution to (3) - (4). In other words, $\begin{bmatrix} x' & \mathbf{e}_T \end{bmatrix}' = w_2 - H_2 p^* + \epsilon$ where ϵ is an initial condition dependent time function decaying to zero as fast as $e^{\bar{A}_2 t}$. Again, for simplicity, we shall ignore ϵ . This means that

$$\mathbf{e}_T = \bar{c}_2 w_2 - \bar{c}_2 H_2 p^*$$

where $\bar{c}_2 = \begin{bmatrix} 0 & \cdots & 0 & 1 \end{bmatrix}$. Thus, in this case, a natural way to generate an estimate \hat{e}_2 of \mathbf{e}_T is by means of the equation

$$\hat{e}_2 = \bar{c}_2 w_2 - \bar{c}_2 H_2 \hat{p} \quad (15)$$

It is clear that the multi-estimator $\mathbb{E}_2(\hat{p})$ defined by (2), (14) and (15) has the required property of delivering an asymptotically correct estimate \hat{e}_2 of \mathbf{e}_T if \hat{p} is fixed at p^* .

Tuner 2

Note that in this case the estimation error

$$e_2 \triangleq \hat{e}_2 - \mathbf{e}_T \quad (16)$$

satisfies the error equation

$$e_2 = \phi_2'(\hat{p}_2 - p^*) \quad (17)$$

where

$$\phi_2' = -\bar{c}_2 H_2 \quad (18)$$

Equation (17) suggests that one consider a pseudo-gradient tuner of the form

$$\dot{\hat{p}} = -\phi_2 e_2 \quad (19)$$

What is Known about Version 2?

The overall model reference adaptive controller in this case thus consists of the reference model (1), the control law (6), the multi-estimator (2), (14), (15), the output estimation error (16) and the tuner (18), (19). *The open problem is here to prove that this version of the controller either solves the model reference adaptive control problem or that it does not.*

Much is known about the problem. In the first place, (1), (2) together with (5), (6) (14) - (18) define a parameter varying linear system $\Sigma_2(\hat{p})$ with input r , state $(y_{\text{ref}}, x_{\text{ref}}, z, H_2, w_2)$ and output e_2 . The consequence of the assumption that every system in \mathcal{M} is minimum phase is that this $\Sigma_2(\hat{p})$ is detectable through e_2 for every fixed value of \hat{p} [5]. Meanwhile the form of (17) enables one to show by direct calculation that the rate of change of the partial Lyapunov function $V \triangleq \|\hat{p} - p^*\|^2$ along a solution to (19) and the equations defining $\Sigma_2(\hat{p})$, satisfies

$$\dot{V} = -2\lambda e_2^2 \leq 0 \quad (20)$$

It is evident that V is a bounded monotone nonincreasing function and consequently that \hat{p} is bounded wherever they exist. From this and the fact that $\Sigma_2(\hat{p})$ is a linear parameter-varying system, it can be concluded that solutions exist globally and that \hat{p} is bounded on $[0, \infty)$. By integrating (20) it can also be concluded that e_2 has a finite $\mathcal{L}^2[0, \infty)$ -norm and that $\|\hat{p} - p^*\|^2$ tends to a finite limit as $t \rightarrow \infty$. Were it possible to deduce from these properties that \hat{p} tended to a limit \bar{p} , then it would establish correctness using the detectability of $\Sigma_2(\bar{p})$.

There is one very special cases for which correctness has been established [6]. This is when p^* is taken to be of the form q^*k where k is a known vector and q^* is a scalar; in this case $\hat{p} \triangleq \hat{q}k$ where \hat{q} is a scalar parameter tuned by the equation $\dot{\hat{q}} = -k'\phi_2 e_2$. The underlying reason why things go through is because in this special case, the fact that $\|\hat{p} - p^*\|^2$ and consequently $\|\hat{q} - q^*\|$ tend to a finite limits, means that \hat{q} tends to a finite limit as well.

4 THE ESSENCE OF THE PROBLEM

In this section we transcribe a stripped down version of the problem that retains all the essential feature that need to be overcome in order to decide

whether or not Monopoli's controller is correct. We do this only for version 2 of the problem and only for the case when $r = 0$ and $\bar{n} = 1$. Thus, in this case, we can take $A_2 = -\lambda$ and $\bar{b}_2 = 1$. Assuming the reference model is initialized at 0, dropping the subscript 2 throughout, and writing ϕ' for $-H$, the system to be analyzed reduces to

$$\dot{z} = \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} z + \begin{bmatrix} b \\ 0 \end{bmatrix} (w + \phi' p^*) + \begin{bmatrix} 0 \\ b \end{bmatrix} \hat{p}' z \quad (21)$$

$$\dot{\phi} = -\lambda\phi - z \quad (22)$$

$$\dot{w} = -\lambda w + \hat{p}' z \quad (23)$$

$$e = \phi'(\hat{p} - p^*) \quad (24)$$

$$\dot{\hat{p}} = -\phi e \quad (25)$$

To recap, p^* is unknown and constant but is such that the linear parameter-varying system $\Sigma(\hat{p})$ defined by (21) to (24) is detectable through e for each fixed value of \hat{p} . Solutions to the system (21) - (25) exist globally. The parameter vector \hat{p} and integral square of e are bounded on $[0, \infty)$ and $\|\hat{p} - p^*\|$ tends to a finite limit as $t \rightarrow \infty$. *The open problem here is to show for every initialization of (21)-(25), that the state of $\Sigma(\hat{p})$ tends to 0 or that it does not.*

BIBLIOGRAPHY

- [1] R. V. Monopoli, "Model reference adaptive control with an augmented error," *IEEE Transactions on Automatic Control*, pp. 474-484, October 1974.
- [2] A. Feuer, B. R. Barmish, and A. S. Morse, "An unstable system associated with model reference adaptive control," *IEEE Transactions on Automatic Control*, 23:499-500, 1978.
- [3] A. S. Morse, "Overcoming the obstacle of high relative degree," *European Journal of Control*, 2(1):29-35, 1996.
- [4] K. J. Åström and B. Wittenmark, "On self-tuning regulators," *Automatica*, 9:185-199, 1973.
- [5] A. S. Morse, "Towards a unified theory of parameter adaptive control - Part 2: Certainty equivalence and implicit tuning," *IEEE Transactions on Automatic Control*, 37(1):15-29, January 1992.
- [6] A. Feuer, *Adaptive Control of Single Input Single Output Linear Systems*, Ph.D. thesis, Yale University, 1978.

Problem 1.6

Model reduction of delay systems

Jonathan R. Partington

School of Mathematics

University of Leeds

Leeds, LS2 9JT

U.K.

J.R.Partington@leeds.ac.uk

1 DESCRIPTION OF THE PROBLEM

Our concern here is with stable single input single output delay systems, and we shall restrict to the case when the system has a transfer function of the form $G(s) = e^{-sT}R(s)$, with $T > 0$ and R rational, stable, and strictly proper, thus bounded and analytic on the right half plane \mathbb{C}_+ . It is a fundamental problem in robust control design to approximate such systems by finite-dimensional systems. Thus, for a fixed natural number n , we wish to find a rational approximant $G_n(s)$ of degree at most n in order to make small the approximation error $\|G - G_n\|$, where $\|\cdot\|$ denotes an appropriate norm. See [9] for some recent work on this subject.

Commonly used norms on a linear time-invariant system with impulse response $g \in L^1(0, \infty)$ and transfer function $G \in H^\infty(\mathbb{C}_+)$ are the H^∞ norm $\|G\|_\infty = \sup_{\operatorname{Re} s > 0} |G(s)|$, the L^p norms $\|g\|_p = (\int_0^\infty |g(t)|^p dt)^{1/p}$ ($1 \leq p < \infty$), and the Hankel norm $\|\Gamma\|$, where $\Gamma : L^2(0, \infty) \rightarrow L^2(0, \infty)$ is the Hankel operator defined by

$$(\Gamma u)(t) = \int_0^\infty g(t + \tau)u(\tau) d\tau.$$

These norms are related by

$$\|\Gamma\| \leq \|G\|_\infty \leq \|g\|_1 \leq 2n\|\Gamma\|,$$

where the last inequality holds for systems of degree at most n .

Two particular approximation techniques for finite-dimensional systems are well-established in the literature [14], and they can also be used for some infinite-dimensional systems [5]:

- Truncated balanced realizations, or, equivalently, output normal realizations [11, 13, 5];
- Optimal Hankel-norm approximants [1, 4, 5].

As we explain in the next section, these techniques are known to produce H^∞ -convergent sequences of approximants for many classes of delay systems (systems of nuclear type). We are thus led to pose the following question: *Do the sequences of reduced order models produced by truncated balanced realizations and optimal Hankel-norm approximations converge for all stable delay systems?*

2 MOTIVATION AND HISTORY OF THE PROBLEM

Balanced realizations were introduced in [11], and many properties of truncations of such realizations were given in [13]. An H^∞ error bound for the reduced-order system produced by truncating a balanced realization was given for finite-dimensional systems in [3, 4], and extended to infinite-dimensional systems in [5]. This commonly used bound is expressed in terms of the sequence $(\sigma_k)_{k=1}^\infty$ of singular values of the Hankel operator Γ corresponding to the original system G ; in our case Γ is compact, and so $\sigma_k \rightarrow 0$. Provided that $g \in L^1 \cap L^2$ and Γ is nuclear (i.e., $\sum_{k=1}^\infty \sigma_k < \infty$) with distinct singular values, then the inequality

$$\|G - G_n^b\|_\infty \leq 2(\sigma_{n+1} + \sigma_{n+2} + \dots)$$

holds for the degree- n balanced truncation G_n^b of G . The elementary lower bound $\|G - G_n\| \geq \sigma_{n+1}$ holds for any degree- n approximation to G .

Another numerically convenient approximation method is the optimal Hankel-norm technique [1, 4, 5], which involves finding a best rank- n Hankel approximation Γ_n^H to Γ , in the Hankel norm, so that $\|\Gamma - \Gamma_n^H\| = \sigma_{n+1}$. In this case the bound

$$\|G - G_n^H - D_0\|_\infty \leq \sigma_{n+1} + \sigma_{n+2} + \dots$$

is available for the corresponding transfer function G_n^H with a suitable constant D_0 . Again, we require the nuclearity of Γ for this to be meaningful.

3 AVAILABLE RESULTS

In the case of a delay system $G(s) = e^{-sT}R(s)$ as specified above, it is known that the Hankel singular values σ_k are asymptotic to $A \left(\frac{T}{\pi k} \right)^r$, where r is

the relative degree of R and $|s^r R(s)|$ tends to the finite nonzero limit A as $|s| \rightarrow \infty$. Hence Γ is nuclear if and only if the relative degree of R is at least 2. (Equivalently, if and only if g is continuous.) We refer to [6, 7] for these and more precise results.

Even for a very simple non-nuclear system such as $G(s) = \frac{e^{-sT}}{s+1}$, for which $k\sigma_k \rightarrow T/\pi$, no theoretical upper bound is known for the H^∞ errors in the rational approximants produced by truncated balanced realizations and optimal Hankel-norm approximation, although numerical evidence suggests that they should still tend to zero.

A related question is to find the best error bounds in L^1 approximation of a delay system. For example, a smoothing technique gives an L^1 approximation error $O\left(\frac{\ln n}{n}\right)$ for systems of relative degree $r = 1$ (see [8]), and it is possible that the optimal Hankel norm might yield a similar rate of convergence. (A lower bound of C/n for some constant $C > 0$ follows easily from the above discussion.)

One approach that may be useful in these analyses is to exploit Bonsall's theorem that a Hankel integral operator Γ is bounded if and only if it is uniformly bounded on the set of all normalized L^2 functions whose Laplace transforms are rational of degree one [2, 12]. An explicit constant in Bonsall's theorem is not known, and would be of great interest in its own right.

Another approach which may be relevant is that of Megretski [10], who introduces maximal real part norms. Their interest stems from the inequality $\|G\|_\infty \geq \|\operatorname{Re} G\|_\infty \geq \|\Gamma\|/2$.

BIBLIOGRAPHY

- [1] V. M. Adamjan, D. Z. Arov, and M. G. Kreĭn, "Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur–Takagi problem," *Math. USSR Sbornik*, 15:31–73, 1971.
- [2] F. F. Bonsall, "Boundedness of Hankel matrices", *J. London Math. Soc.* (2), 29(2):289–300, 1984.
- [3] D. Enns, *Model Reduction for Control System Design*, Ph.D. dissertation, Stanford University, 1984.
- [4] K. Glover, "All optimal Hankel-norm approximations of linear multivariable systems and their L^∞ -error bounds, *Internat. J. Control*, 39(6):1115–1193, 1984.

- [5] K. Glover, R. F. Curtain, and J. R. Partington, "Realisation and approximation of linear infinite-dimensional systems with error bounds," *SIAM J. Control Optim.*, 26(4):863–898, 1988.
- [6] K. Glover, J. Lam, and J. R. Partington, "Rational approximation of a class of infinite-dimensional systems. I. Singular values of Hankel operators," *Math. Control Signals Systems*, 3(4):325–344, 1990.
- [7] K. Glover, J. Lam, and J. R. Partington, "Rational approximation of a class of infinite-dimensional systems. II. Optimal convergence rates of L_∞ approximants," *Math. Control Signals Systems*, 4(3):233–246, 1991.
- [8] K. Glover and J. R. Partington, "Bounds on the achievable accuracy in model reduction," In: *Modelling, Robustness and Sensitivity Reduction in Control Systems* (Groningen, 1986), pp. 95–118. Springer, Berlin, 1987.
- [9] P. M. Mäkilä and J. R. Partington, "Shift operator induced approximations of delay systems," *SIAM J. Control Optim.*, 37(6):1897–1912, 1999.
- [10] A. Megretski, "Model order reduction using maximal real part norms," Presented at *CDC 2000*, Sydney, 2000.
<http://web.mit.edu/ameg/www/images/lund.ps>.
- [11] B. C. Moore, "Principal component analysis in linear systems: controllability, observability, and model reduction," *IEEE Trans. Automat. Control*, 26(1):17–32, 1981.
- [12] J. R. Partington and G. Weiss, "Admissible observation operators for the right-shift semigroup," *Math. Control Signals Systems*, 13(3):179–192, 2000.
- [13] L. Pernebo and L. M. Silverman, "Model reduction via balanced state space representations," *IEEE Trans. Automat. Control*, 27(2):382–387, 1982.
- [14] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*, Upper Saddle River, NJ: Prentice Hall 1996.

Problem 1.7

Schur extremal problems

Lev Sakhnovich

Courant Institute of Mathematical Science

New York, NY 11223

USA

Lev.Sakhnovich@verizon.net

1 DESCRIPTION OF THE PROBLEM

In this paper we consider the well-known Schur problem the solution of which satisfy in addition the extremal condition

$$w^*(z)w(z) \leq \rho_{min}^2, \quad |z| < 1, \quad (1)$$

where $w(z)$ and ρ_{min} are $m \times m$ matrices and $\rho_{min} > 0$. Here the matrix ρ_{min} is defined by a certain minimal-rank condition (see Definition 1). We remark that the extremal Schur problem is a particular case. The general case is considered in book [1] and paper [2]. Our approach to the extremal problems does not coincide with the superoptimal approach [3],[4]. In paper [2] we compare our approach to the extremal problems with the superoptimal approach. Interpolation has found great applications in control theory [5],[6].

Schur Extremal Problem: The $m \times m$ matrices a_0, a_1, \dots, a_n are given. Describe the set of $m \times m$ matrix functions $w(z)$ holomorphic in the circle $|z| < 1$ and satisfying the relation

$$w(z) = a_0 + a_1 z + \dots + a_n z^n + \dots \quad (2)$$

and inequality (1.1).

A necessary condition of the solvability of the Schur extremal problem is the inequality

$$R_{min}^2 - S \geq 0, \quad (3)$$

where the $(n+1)m \times (n+1)m$ matrices S and R_{min} are defined by the relations

$$S = C_n C_n^*, \quad R_{min} = \text{diag}[\rho_{min}, \rho_{min}, \dots, \rho_{min}], \quad (4)$$

$$C_n = \begin{pmatrix} a_0 & 0 & \dots & 0 \\ a_1 & a_0 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ a_n & a_{n-1} & \dots & a_0 \end{pmatrix}. \quad (5)$$

Definition 1: We shall call the matrix $\rho = \rho_{min} > 0$ minimal if the following two requirements are fulfilled:

1. The inequality

$$R_{min}^2 - S \geq 0 \quad (6)$$

holds.

2. If the $m \times m$ matrix $\rho > 0$ is such that

$$R^2 - S \geq 0, \quad (7)$$

then

$$\text{rank}(R_{min}^2 - S) \leq \text{rank}(R^2 - S), \quad (8)$$

where $R = \text{diag}[\rho, \rho, \dots, \rho]$.

Remark 1: The existence of ρ_{min} follows directly from definition 1.

Question 1: Is ρ_{min} unique?

Remark 2: If $m = 1$ then ρ_{min} is unique and $\rho_{min}^2 = \lambda_{max}$, where λ_{max} is the largest eigenvalue of the matrix S .

Remark 3: Under some assumptions the uniqueness of ρ_{min} is proved in the case $m > 1, n = 1$ (see [2],[7]).

If ρ_{min} is known then the corresponding $w_{min}(\xi)$ is a rational matrix function. This generalizes the well-known fact for the scalar case (see [7]).

Question 2: How to find ρ_{min} ?

In order to describe some results in this direction we write the matrix $S = C_n C_n^*$ in the following block form

$$\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix}, \quad (9)$$

where S_{22} is an $m \times m$ matrix.

Proposition 1: [1] If $\rho = q > 0$ satisfies inequality (1.7) and the relation

$$q^2 = S_{22} + S_{12}^*(Q^2 - S_{11})^{-1}S_{12}, \quad (10)$$

where $Q = \text{diag}[q, q, \dots, q]$, then $\rho_{min} = q$.

We shall apply the method of successive approximation when studying equation (1.10). We put $q_0^2 = S_{22}$, $q_{k+1}^2 = S_{22} + S_{12}^*(Q_k^2 - S_{11})^{-1}S_{12}$, where $k \geq 0$, $Q_k = \text{diag}[q_k, q_k, \dots, q_k]$. We suppose that

$$Q_0^2 - S_{11} > 0. \quad (11)$$

Theorem 1: [1] The sequence $q_0^2, q_2^2, q_4^2, \dots$ monotonically increases and has the limit m_1 . The sequence $q_1^2, q_3^2, q_5^2, \dots$ monotonically decreases and has the limit m_2 . The inequality $m_1 \leq m_2$ holds. If $m_1 = m_2$ then $\rho_{min}^2 = q^2$.

Question 3: Suppose relation (1.11) holds. Is there a case when $m_1 \neq m_2$?

The answer is “no” if $n = 1$ (see [2],[8]).

Remark 4: In book [1] we give an example in which ρ_{min} is constructed in explicit form.

BIBLIOGRAPHY

- [1] Lev Sakhnovich. *Interpolation Theory and Applications*, Kluwer Acad.Publ. (1997).
- [2] J. Helton and L. Sakhnovich, *Extremal Problems of Interpolation Theory* (forthcoming).
- [3] N. J. Young, "The Nevanlinna-Pick Problem for matrix-valued functions," *J. Operator Theory*, 15, pp. 239–265, (1986).
- [4] V. V. Peller and N. J. Young, "Superoptimal analytic approximations of matrix functions," *Journal of Functional Analysis*, 120, pp. 300–343, (1994).
- [5] M. Green and D. Limebeer, *Linear Robust Control*, Prentice-Hall, (1995).
- [6] "Lecture Notes in Control and Information," *Sciences*, Springer Verlag, 135 (1989).
- [7] N. I. Akhiezer, "On a Minimum in Function Theory," *Operator Theory, Adv. and Appl.* 95, pp. 19–35, (1997).
- [8] A. Ferrante and B. C. Levy, "Hermitian Solutions of the Equations $X = Q + NX^{-1}N^*$," *Linear Algebra and Applications* 247, pp. 359–373, (1996).

Problem 1.8

The elusive iff test for time-controllability of behaviours

Amol Sasane

Faculty of Mathematical Sciences

University of Twente

7500 AE, Enschede

The Netherlands

A.J.Sasane@math.utwente.nl

1 DESCRIPTION OF THE PROBLEM

Problem: Let $R \in \mathbb{C}[\eta_1, \dots, \eta_m, \xi]^{\mathfrak{g} \times \mathfrak{w}}$ and let \mathfrak{B} be the behavior given by the kernel representation corresponding to R . Find an algebraic test on R characterizing the time-controllability of \mathfrak{B} .

In the above, we assume \mathfrak{B} to comprise of only smooth trajectories, that is,

$$\mathfrak{B} = \{w \in \mathcal{C}^\infty(\mathbb{R}^{m+1}, \mathbb{C}^{\mathfrak{w}}) \mid D_R w = 0\},$$

where $D_R : \mathcal{C}^\infty(\mathbb{R}^{m+1}, \mathbb{C}^{\mathfrak{w}}) \rightarrow \mathcal{C}^\infty(\mathbb{R}^{m+1}, \mathbb{C}^{\mathfrak{g}})$ is the differential map that acts as follows: if $R = [r_{ij}]_{\mathfrak{g} \times \mathfrak{w}}$, then

$$D_R \begin{bmatrix} w_1 \\ \vdots \\ w_{\mathfrak{w}} \end{bmatrix} = \begin{bmatrix} \sum_{k=1}^{\mathfrak{w}} r_{1k} \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_m}, \frac{\partial}{\partial t} \right) w_k \\ \vdots \\ \sum_{k=1}^{\mathfrak{w}} r_{\mathfrak{g}k} \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_m}, \frac{\partial}{\partial t} \right) w_k \end{bmatrix}.$$

Time-controllability is a property of the behavior, defined as follows. The behavior \mathfrak{B} is said to be *time-controllable* if for any w_1 and w_2 in \mathfrak{B} , there exists a $w \in \mathfrak{B}$ and a $\tau \geq 0$ such that

$$w(\bullet, t) = \begin{cases} w_1(\bullet, t) & \text{for all } t \leq 0 \\ w_2(\bullet, t - \tau) & \text{for all } t \geq \tau \end{cases}.$$

2 MOTIVATION AND HISTORY OF THE PROBLEM

The behavioral theory for systems described by a set of linear constant coefficient partial differential equations has been a challenging and fruitful area of research for quite some time (see, for instance, Pillai and Shankar [5], Oberst [3] and Wood et al. [4]). An excellent elementary introduction to the behavioral theory in the 1–D case (corresponding to systems described by a set of linear constant coefficient ordinary differential equations) can be found in Polderman and Willems [6].

In [5], [3] and [4], the behaviours arising from systems of partial differential equations are studied in a general setting in which the time-axis does not play a distinguished role in the formulation of the definitions pertinent to control theory. Since in the study of systems with “dynamics,” it is useful to give special importance to time in defining system theoretic concepts, recent attempts have been made in this direction (see, for example, Cotroneo and Sasane [2], Sasane et al. [7], and Çamlıbel and Sasane [1]). The formulation of definitions with special emphasis on the time-axis is straightforward, since they can be seen quite easily as extensions of the pertinent definitions in the 1–D case. However, the algebraic characterization of the properties of the behavior, such as time-controllability, turn out to be quite involved.

Although the traditional treatment of distributed parameter systems (in which one views them as an ordinary differential equation with an infinite-dimensional Hilbert space as the state-space) is quite successful, the study of the present problem will have its advantages, since it would give a test that is *algebraic* in nature (and hence computationally easy) for a property of the sets of trajectories, namely time-controllability. Another motivation for considering this problem is that the problem of patching up of solutions of partial differential equations is also an interesting question from a purely mathematical point of view.

3 AVAILABLE RESULTS

In the 1–D case, it is well-known (see, for example, theorem 5.2.5 on page 154 of [6]) that time-controllability is equivalent with the following condition: There exists a $\mathbf{r}_0 \in \mathbb{N} \cup \{0\}$ such that for all $\lambda \in \mathbb{C}$, $\text{rank}(R(\lambda)) = \mathbf{r}_0$. This condition is in turn equivalent with the torsion freeness of the $\mathbb{C}[\xi]$ -module $\mathbb{C}[\xi]^w / \mathbb{C}[\xi]^g R$.

Let us consider the following statements

- A1. The $\mathbb{C}(\eta_1, \dots, \eta_m)[\xi]$ -module $\mathbb{C}(\eta_1, \dots, \eta_m)[\xi]^w / \mathbb{C}(\eta_1, \dots, \eta_m)[\xi]^g R$ is torsion free.
- A2. There exists a $\chi \in \mathbb{C}[\eta_1, \dots, \eta_m, \xi]^w \setminus \mathbb{C}(\eta_1, \dots, \eta_m)[\xi]^g R$ and there exists a nonzero $p \in \mathbb{C}[\eta_1, \dots, \eta_m, \xi]$ such that $p \cdot \chi \in \mathbb{C}(\eta_1, \dots, \eta_m)[\xi]^g R$, and

$\deg(p) = \deg(j(p))$, where j denotes the homomorphism
 $p(\xi, \eta_1, \dots, \eta_m) \mapsto p(\xi, 0, \dots, 0) : \mathbb{C}[\xi, \eta_1, \dots, \eta_m] \rightarrow \mathbb{C}[\xi]$.

In [2], [7] and [1], the following implications were proved:

$$\begin{array}{ccc} \mathfrak{B} & \text{is time-controllable} & \\ \Downarrow \nexists & & \Uparrow \\ \neg A2 & \begin{array}{c} \Leftarrow \\ \neq \end{array} & A1 \end{array}$$

Although it is tempting to conjecture that the condition $A1$ might be the iff test for time-controllability, the diffusion equation reveals the precariousness of hazarding such a guess. In [1] it was shown that the diffusion equation is time-controllable with respect to¹ the space \mathcal{W} defined below. Before defining the set \mathcal{W} , we recall the definition of the (small) Gevrey class of order 2, denoted by $\gamma^{(2)}(\mathbb{R})$: $\gamma^{(2)}(\mathbb{R})$ is the set of all $\varphi \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{C})$ such that for every compact set K and every $\epsilon > 0$ there exists a constant C_ϵ such that for every $\mathbf{k} \in \mathbb{N}$, $|\varphi^{(\mathbf{k})}(t)| \leq C_\epsilon \epsilon^{\mathbf{k}} (\mathbf{k}!)^2$ for all $t \in K$. \mathcal{W} is then defined to be the set of all $w \in \mathfrak{B}$ such that $w(0, \bullet) \in \gamma^{(2)}(\mathbb{R})$. Furthermore, it was also shown in [1], that the control could then be implemented by the two *point control* input functions acting at the point $x = 0$: $u_1(t) = w(0, t)$ and $u_2(t) = \frac{\partial}{\partial x} w(0, t)$ for all $t \in \mathbb{R}$. The subset \mathcal{W} of $\mathcal{C}^\infty(\mathbb{R}^2, \mathbb{C})$ functions comprises a large class of solutions of the diffusion equation. In fact, an interesting open problem is the problem of constructing a trajectory in the behavior that is not in the class \mathcal{W} . Also whether the whole behavior (and not just trajectories in \mathcal{W}) of the diffusion equation is time-controllable or not is an open question. The answers to these questions would either strengthen or discard the conjecture that the behavior corresponding to $p \in \mathbb{C}[\eta_1, \dots, \eta_m, \xi]$ is time-controllable iff $p \in \mathbb{C}[\eta_1, \dots, \eta_m]$, which would eventually help in settling the question of the equivalence of $A1$ and time-controllability.

BIBLIOGRAPHY

- [1] M. K. Çamlıbel and A. J. Sasane, "Approximate time-controllability versus time-controllability," submitted to *the 15th MTNS*, U.S.A., June 2002.
- [2] T. Cotroneo and A. J. Sasane, "Conditions for time-controllability of behaviours," *International Journal of Control*, 75, pp. 61-67 (2002).
- [3] U. Oberst, "Multidimensional constant linear systems," *Acta Appl. Math.*, 20, pp. 1-175 (1990).

¹That is, for any two trajectories in $\mathcal{W} \cap \mathfrak{B}$, there exists a concatenating trajectory in $\mathcal{W} \cap \mathfrak{B}$.

- [4] D. H. Owens, E. Rogers and J. Wood, "Controllable and autonomous n -D linear systems," *Multidimensional Systems and Signal Processing*, 10, pp. 33-69 (1999).
- [5] H. K. Pillai and S. Shankar, "A Behavioural Approach to the Control of Distributed Systems," *SIAM Journal on Control and Optimization*, 37, pp. 388-408 (1998).
- [6] J. W. Polderman and J. C. Willems, *Introduction to Mathematical Systems Theory*, Springer-Verlag, 1998.
- [7] A. J. Sasane, E. G. F. Thomas and J. C. Willems, "Time-autonomy versus time-controllability," accepted for publication in *The Systems and Control Letters*, 2001.

Problem 1.9

A Farkas lemma for behavioral inequalities

A. A. (Tonny) ten Dam

Information, Communication and Technology Division
National Aerospace Laboratory NLR
P.O. Box 90502 , 1006 BM Amsterdam
The Netherlands
`tendam@nlr.nl`

J.W. (Hans) Nieuwenhuis

Faculty of Economics
University of Groningen
Postbus 800, 9700 AV Groningen
The Netherlands
`j.w.nieuwenhuis@eco.rug.nl`

1 DESCRIPTION OF THE PROBLEM

Within the systems and control community there has always been an interest in minimality issues. In this chapter we conjecture a Farkas Lemma for behavioral inequalities that, when true, will allow to study minimality and elimination issues for behavioral systems described by inequalities.

Let $\mathbb{R}^{n \times m}[s, s^{-1}]$ denote the $(n \times m)$ polynomial matrices with real coefficients and positive and negative powers in the indeterminate s . Let $\mathbb{R}_+^{n \times m}[s, s^{-1}]$ denote the set of matrices in $\mathbb{R}^{n \times m}[s, s^{-1}]$ with non-negative coefficients only. In this chapter we consider discrete-time systems with time-axis \mathbb{Z} . Let σ denote the (backward) shift operator, and let $R(\sigma, \sigma^{-1})$ denote polynomial operators in the shift.

Of interest is the relation between two polynomial matrices $R(s, s^{-1})$ and $R'(s, s^{-1})$ when they satisfy

$$R(\sigma, \sigma^{-1})w \geq 0 \Rightarrow R'(\sigma, \sigma^{-1})w \geq 0. \quad (1)$$

Based on the static case, one may expect that such a relation should be the extension of Farkas's lemma to the behavioral case. This leads to the raison d'être of this chapter.

Conjecture: Let $R \in \mathbb{R}^{g \times q}[s, s^{-1}]$ and $R' \in \mathbb{R}^{g' \times q}[s, s^{-1}]$. Then we have $\{R(\sigma, \sigma^{-1})w \geq 0 \Rightarrow R'(\sigma, \sigma^{-1})w \geq 0\}$ if and only if there exists a polynomial matrix $H \in \mathbb{R}_+^{g' \times g}[s, s^{-1}]$ such that $R'(s, s^{-1}) = H(s, s^{-1})R(s, s^{-1})$.

In order to prove this conjecture, one could try to extend the original proof given by Farkas in [4]. However, this proof explicitly uses the fact that every scalar that is unequal to zero is invertible. Such a general statement does not hold for elements of $\mathbb{R}^{q \times q}[s, s^{-1}]$. The most promising approach for the dynamic case seems to be the use of mathematical tools such as the separation theorem of Hahn-Banach (see, for instance, [5]). The basic mathematical preliminaries read as follows.

Denote $\mathbb{E} := (\mathbb{R}^q)^\mathbb{Z}$ with the topology of point-wise convergence. The dual of \mathbb{E} , denoted by \mathbb{E}^* , consists of all \mathbb{R}^q -valued sequences that have compact support. Let $R \in \mathbb{R}^{g \times q}[s, s^{-1}]$. Let $\mathfrak{B} = \{w \in \mathbb{E}^q | R(\sigma, \sigma^{-1})w \geq 0\}$. The polar cone of \mathfrak{B} , denoted by $\mathfrak{B}^\#$, is given by $\{w \in \mathbb{E}^* | \forall w \in \mathfrak{B} : \sum_{t \in \mathbb{Z}} w^*(t)w(t) \geq 0\}$. We would like to establish that $\mathfrak{B}^\# = \{w \in \mathbb{E} | \exists \alpha \in \mathbb{E}^*, \alpha \geq 0 \text{ such that } w^* = R^T(\sigma^{-1}, \sigma)\alpha\}$, but we have so far not been able to prove or disprove these statements. These statements, together with the fact that $\{\mathfrak{B}_1 \subseteq \mathfrak{B}_2\}$ implies $\{\mathfrak{B}_2^\# \subseteq \mathfrak{B}_1^\#\}$, are believed to be useful in a proof of the conjecture.

2 MOTIVATION AND HISTORY OF THE PROBLEM

In the early nineties the first author started to investigate minimality issues for so-called behavior inequality systems, e.g., systems whose behavior \mathfrak{B} allows a description $\mathfrak{B} = \{w \in \mathbb{R}^q | R(\sigma, \sigma^{-1})w \geq 0\}$. Examples can be found in [2].

The first publication that we are aware of that deals with this class of systems is [1]. And the conjecture mentioned above can already be found in that paper. As the problem proved hard to solve, a number of investigations were carried out in the context of linear static inequalities, where the problem of minimal representations of systems containing both equalities and inequalities was solved [2]. The conjecture, however, withstood our efforts, and it became a part of the Ph.D. thesis of the first author [2]. As the study is placed in the context of behaviors, the Farkas lemma for behavioral inequalities is also discussed in the Willem's Festschrift [3] (chapter 16).

Until the Farkas lemma for behavioral inequalities has been proven, issues like minimal representations, elimination of latent variables etcetera cannot be solved in their full generality. It is our belief that the Farkas lemma for behavior inequalities, as conjectured here, will be a cornerstone for further investigations in a theory for behavioral inequalities.

3 AVAILABLE RESULTS

For the static case, the conjecture is nothing else than the famous Farkas lemma for linear inequalities. For the dynamic case, the conjecture holds true for a special case.

Proposition: Let $R \in \mathbb{R}^{g \times q}[s, s^{-1}]$ be a full-row rank polynomial matrix. Let $R' \in \mathbb{R}^{g' \times q}[s, s^{-1}]$. Then: $\{R(\sigma, \sigma^{-1})w \geq 0 \Rightarrow R'(\sigma, \sigma^{-1})w \geq 0\}$ if and only if there exists a unique polynomial matrix $H \in \mathbb{R}_+^{g' \times g}[s, s^{-1}]$ such that $R'(s, s^{-1}) = H(s, s^{-1})R(s, s^{-1})$.

The proof of this proposition can be found in [2] (proposition 4.5.12).

4 A RELATED CONJECTURE

It is of interest to present a related conjecture, whose resolution is closely linked to the Farkas lemma for behavioral inequalities.

Recall from [6] that a matrix $U \in \mathbb{R}^{g \times g}[s, s^{-1}]$ is said to be unimodular if it has an inverse $U^{-1} \in \mathbb{R}^{g \times g}[s, s^{-1}]$. We will call a matrix $H \in \mathbb{R}_+^{g \times g}[s, s^{-1}]$ posimodular if it is unimodular and $H^{-1} \in \mathbb{R}_+^{g \times g}[s, s^{-1}]$. Omitting the formal definitions, we will call a representation minimal if the number of equations used to describe the behavior is minimal.

Conjecture: Let $\{w \in (\mathbb{R}^q)^{\mathbb{Z}} | R_1(\sigma, \sigma^{-1})w = 0 \text{ and } R_2(\sigma, \sigma^{-1})w \geq 0\}$ and $\{w \in (\mathbb{R}^q)^{\mathbb{Z}} | R'_1(\sigma, \sigma^{-1})w = 0 \text{ and } R'_2(\sigma, \sigma^{-1})w \geq 0\}$ both be two minimal representations. They represent the same behavior if and only if there are polynomial matrices $U(s, s^{-1})$, $H(s, s^{-1})$ and $S(s, s^{-1})$ such that

$$\begin{bmatrix} R'_1(s, s^{-1}) \\ R'_2(s, s^{-1}) \end{bmatrix} = \begin{bmatrix} U(s, s^{-1}) & 0 \\ S(s, s^{-1}) & H(s, s^{-1}) \end{bmatrix} \begin{bmatrix} R_1(s, s^{-1}) \\ R_2(s, s^{-1}) \end{bmatrix} \quad (2)$$

with U unimodular, H posimodular and no conditions on S .

We remark that this conjecture holds true for static inequalities and for that case is given as proposition 3.4.5 in [2].

BIBLIOGRAPHY

- [1] A. A. ten Dam, "Representations of dynamical systems described by behavioural inequalities," In: *Proceedings European Control Conference ECC'93*, vol. 3, pp. 1780-1783, June 28-July 1, Groningen, The Netherlands, 1993.
- [2] A. A. ten Dam, *Unilaterally Constrained Dynamical Systems*, Ph.D. Dissertation, University Groningen, The Netherlands, 1997. URL: <http://www.ub.rug.nl/eldoc/dis/science/a.a.ten.dam/>

- [3] A. A. ten Dam and J. W. Nieuwenhuis, "On behavioural inequalities," In: *The Mathematics from Systems and Control: From Intelligent Control to Behavioral Systems*, Groningen, 1999, pp. 165-176.
- [4] J. Farkas, "Die algebraische Grundlage der Anwendungen des Mechanischen Princips von Fourier," *Mathematische und naturwissenschaftliche Berichte aus Ungarn*, 16, pp. 154-157, 1899. (Translation of: Gy. Farkas, "A Fourier-féle mechanikai elv alkalmazásának algebrai alapja," *Matematikai és Természettudományi Értesítő*, 16, pp. 361-364, 1898.)
- [5] W. Rudin, *Functional Analysis*, Tata McGraw-Hill Publishing Company Ltd., 1973.
- [6] J.C. Willems, "Paradigms and puzzles in the theory of dynamical systems," *IEEE Transactions on Automatic Control*, vol. 36, no. 3, pp. 259-294, 1991.

Problem 1.10

Regular feedback implementability of linear differential behaviors

H. L. Trentelman

Mathematics Institute

University of Groningen

P.O. Box 800, 9700 AV Groningen

The Netherlands

H.L.Trentelman@math.rug.nl

1 INTRODUCTION

In this short paper, we want to discuss an open problem that appears in the context of interconnection of systems in a behavioral framework. Given a system behavior, playing the role of plant to be controlled, the problem is to characterize all system behaviors that can be achieved by interconnecting the plant behavior with a controller behavior, where the interconnection should be a regular feedback interconnection.

More specifically, we will deal with linear time-invariant differential systems, i.e., dynamical systems Σ given as a triple $\{\mathbb{R}, \mathbb{R}^w, \mathcal{B}\}$, where \mathbb{R} is the time-axis, and where \mathcal{B} , called the *behavior* of the system Σ , is equal to the set of all solutions $w : \mathbb{R} \rightarrow \mathbb{R}^w$ of a set of higher order, linear, constant coefficient, differential equations. More precisely,

$$\mathcal{B} = \{w \in \mathcal{C}^\infty(\mathbb{R}, \mathbb{R}^w) \mid R(\frac{d}{dt})w = 0\},$$

for some polynomial matrix $R \in \mathbb{R}^{\bullet \times w}[\xi]$. The set of all such systems Σ is denoted by \mathcal{L}^w . Often, we simply refer to a system by talking about its behavior, and we write $\mathcal{B} \in \mathcal{L}^w$ instead of $\Sigma \in \mathcal{L}^w$. Behaviors $\mathcal{B} \in \mathcal{L}^w$ can hence be described by differential equations of the form $R(\frac{d}{dt})w = 0$, typically with the number of rows of R strictly less than its number of columns. Mathematically, $R(\frac{d}{dt})w = 0$ is then an *under-determined* system of equations. This results in the fact that some of the components of $w = (w_1, w_2, \dots, w_w)$ are unconstrained. This number of unconstrained components is an integer “invariant” associated with \mathcal{B} , and is called the *input cardinality* of \mathcal{B} , denoted by $\mathfrak{m}(\mathcal{B})$, its number of free, “input,” variables. The remaining number of

variables, $\mathfrak{w} - \mathfrak{m}(\mathcal{B})$, is called the *output cardinality* of \mathcal{B} and is denoted by $\mathfrak{p}(\mathcal{B})$. Finally, a third integer invariant associated with a system behavior $\mathcal{B} \in \mathcal{L}^{\mathfrak{w}}$ is its McMillan degree. It can be shown that (modulo permutation of the components of the external variable w) any $\mathcal{B} \in \mathcal{L}^{\mathfrak{w}}$ can be represented by a state space representation of the form $\frac{d}{dt}x = Ax + Bu$, $y = Cx + Du$, $w = (u, y)$. Here, A, B, C , and D are constant matrices with real components. The minimal number of components of the state variable x needed in such an input/state/output representation of \mathcal{B} is called the *McMillan degree of \mathcal{B}* , and is denoted by $\mathfrak{n}(\mathcal{B})$.

Suppose now $\Sigma_1 = \{\mathbb{R}, \mathbb{R}^{\mathfrak{w}_1} \times \mathbb{R}^{\mathfrak{w}_2}, \mathcal{B}_1\} \in \mathcal{L}^{\mathfrak{w}_1 + \mathfrak{w}_2}$ and $\Sigma_2 = \{\mathbb{R}, \mathbb{R}^{\mathfrak{w}_2} \times \mathbb{R}^{\mathfrak{w}_3}, \mathcal{B}_2\} \in \mathcal{L}^{\mathfrak{w}_2 + \mathfrak{w}_3}$ are linear differential systems with common factor $\mathbb{R}^{\mathfrak{w}_2}$ in the signal space. The manifest variable of Σ_1 is (w_1, w_2) and that of Σ_2 is (w_2, w_3) . The variable w_2 is shared by the systems, and it is through this variable, called the *interconnection variable*, that we can interconnect the systems. We define the *interconnection of Σ_1 and Σ_2 through w_2* as the system

$$\Sigma_1 \wedge_{w_2} \Sigma_2 := \{\mathbb{R}, \mathbb{R}^{\mathfrak{w}_1} \times \mathbb{R}^{\mathfrak{w}_2} \times \mathbb{R}^{\mathfrak{w}_3}, \mathcal{B}_1 \wedge_{w_2} \mathcal{B}_2\},$$

with interconnection behavior

$$\mathcal{B}_1 \wedge_{w_2} \mathcal{B}_2 := \{(w_1, w_2, w_3) \mid (w_1, w_2) \in \mathcal{B}_1 \text{ and } (w_2, w_3) \in \mathcal{B}_2\}.$$

The interconnection $\Sigma_1 \wedge_{w_2} \Sigma_2$ is called a *regular interconnection* if the output cardinalities of Σ_1 and Σ_2 add up to that of $\Sigma_1 \wedge_{w_2} \Sigma_2$:

$$\mathfrak{p}(\mathcal{B}_1 \wedge_{w_2} \mathcal{B}_2) = \mathfrak{p}(\mathcal{B}_1) + \mathfrak{p}(\mathcal{B}_2).$$

It is called a *regular feedback interconnection* if, in addition, the sum of the McMillan degrees of \mathcal{B}_1 and \mathcal{B}_2 is equal to the McMillan degree of the interconnection:

$$\mathfrak{n}(\mathcal{B}_1 \wedge_{w_2} \mathcal{B}_2) = \mathfrak{n}(\mathcal{B}_1) + \mathfrak{n}(\mathcal{B}_2).$$

It can be proven that the interconnection of Σ_1 and Σ_2 is a regular feedback interconnection if, possibly after permutation of components within w_1 , w_2 and w_3 , there exists a component-wise partition of w_2 into $w_2 = (u, y_1, y_2)$, of w_1 into $w_1 = (v_1, z_1)$, and of w_3 into $w_3 = (v_2, z_2)$ such that the following four conditions hold:

1. in the system Σ_1 , (v_1, y_2, u) is input and (z_1, y_1) is output, and the transfer matrix from (v_1, y_2, u) to (z_1, y_1) is proper.
2. in the system Σ_2 , (v_2, y_1, u) is input and (z_2, y_2) is output, and the transfer matrix from (v_2, y_1, u) to (z_2, y_2) is proper.
3. in the system $\Sigma_1 \wedge_{w_2} \Sigma_2$, (v_1, v_2, u) is input and (z_1, z_2, y_1, y_2) is output, and the transfer matrix from (v_1, v_2, u) to (z_1, z_2, y_1, y_2) is proper.
4. if we introduce new (“perturbation signals”) e_1 and e_2 and, instead of y_1 and y_2 we apply inputs $y_1 + e_2$ and $y_2 + e_1$ to Σ_2 and Σ_1 respectively, then the transfer matrix from (v_1, v_2, u, e_1, e_2) to (z_1, z_2, y_1, y_2) is proper.

The first three of these conditions state that, in the interconnection of Σ_1 and Σ_2 , along the terminals of the interconnected system one can identify a signal flow that is compatible with the signal flow diagram of a feedback configuration with proper transfer matrices. The fourth condition states that this feedback interconnection is “well-posed.” The equivalence of the property of being a regular feedback interconnection with these four conditions was studied for the “full interconnection case” in [8] and [2].

2 STATEMENT OF THE PROBLEM

Suppose $\mathcal{P}_{\text{full}} \in \mathcal{L}^{\mathfrak{w}+c}$ is a system (the plant) with two types of external variables, namely c and w . The first of these, c , is the interconnection variable through which it can be interconnected to a second system $\mathcal{C} \in \mathcal{L}^c$ (the controller) with external variable c . The external variable c is shared by $\mathcal{P}_{\text{full}}$ and \mathcal{C} . The remaining variable w is the variable through which $\mathcal{P}_{\text{full}}$ interacts with the rest of its environment. After interconnecting plant and controller through the shared variable c , we obtain the *full controlled behavior* $\mathcal{P}_{\text{full}} \wedge_c \mathcal{C} \in \mathcal{L}^{\mathfrak{w}+c}$. The *manifest controlled behavior* $\mathcal{K} \in \mathcal{L}^{\mathfrak{w}}$ is obtained by projecting all trajectories $(w, c) \in \mathcal{P}_{\text{full}} \wedge_c \mathcal{C}$ on their first coordinate:

$$\mathcal{K} := \{w \mid \text{there exists } c \text{ such that } (w, c) \in \mathcal{P}_{\text{full}} \wedge_c \mathcal{C}\}. \quad (1)$$

If this holds, then we say that \mathcal{C} *implements* \mathcal{K} . If, for a given $\mathcal{K} \in \mathcal{L}^{\mathfrak{w}}$ there exists $\mathcal{C} \in \mathcal{L}^c$ such that \mathcal{C} implements \mathcal{K} , then we call \mathcal{K} *implementable*. If, in addition, the interconnection of $\mathcal{P}_{\text{full}}$ and \mathcal{C} is regular, we call \mathcal{K} *regularly implementable*. Finally, if the interconnection of $\mathcal{P}_{\text{full}}$ and \mathcal{C} is a regular feedback interconnection, we call \mathcal{K} *implementable by regular feedback*.

This now brings us to the statement of our problem: the problem is to characterize, for a given $\mathcal{P}_{\text{full}} \in \mathcal{L}^{\mathfrak{w}+c}$, the set of all behaviors $\mathcal{K} \in \mathcal{L}^{\mathfrak{w}}$ that are implementable by regular feedback. In other words:

Problem statement: Let $\mathcal{P}_{\text{full}} \in \mathcal{L}^{\mathfrak{w}+c}$ be given. Let $\mathcal{K} \in \mathcal{L}^{\mathfrak{w}}$. Find necessary and sufficient conditions on \mathcal{K} under which there exists $\mathcal{C} \in \mathcal{L}^c$ such that

1. \mathcal{C} implements \mathcal{K} [meaning that (1) holds],
2. $\mathbf{p}(\mathcal{P}_{\text{full}} \wedge_c \mathcal{C}) = \mathbf{p}(\mathcal{P}_{\text{full}}) + \mathbf{p}(\mathcal{C})$,
3. $\mathbf{n}(\mathcal{P}_{\text{full}} \wedge_c \mathcal{C}) = \mathbf{n}(\mathcal{P}_{\text{full}}) + \mathbf{n}(\mathcal{C})$.

Effectively, a characterization of all such behaviors $\mathcal{K} \in \mathcal{L}^{\mathfrak{w}}$ gives a characterization of the “limits of performance” of the given plant under regular feedback control.

3 BACKGROUND

Our open problem is to find conditions for a given $\mathcal{K} \in \mathcal{L}^w$ to be implementable by regular feedback. An obvious necessary condition for this is that \mathcal{K} is implementable, i.e., it can be achieved by interconnecting the plant with a controller by (just any) interconnection through the interconnection variable c . Necessary and sufficient conditions for implementability have been obtained in [7]. These conditions are formulated in terms of two behaviors derived from the full plant behavior $\mathcal{P}_{\text{full}}$:

$$\mathcal{P} := \{w \mid \text{there exists } c \text{ such that } (w, c) \in \mathcal{P}_{\text{full}}\}$$

and

$$\mathcal{N} := \{w \mid (w, 0) \in \mathcal{P}_{\text{full}}\}.$$

\mathcal{P} and \mathcal{N} are both in \mathcal{L}^w , and are called the *manifest plant behavior* and *hidden behavior* associated with the full plant behavior $\mathcal{P}_{\text{full}}$, respectively. In [7] it has been shown that $\mathcal{K} \in \mathcal{L}^w$ is implementable if and only if

$$\mathcal{N} \subseteq \mathcal{K} \subseteq \mathcal{P}, \quad (2)$$

i.e., \mathcal{K} contains \mathcal{N} , and is contained in \mathcal{P} . This elegant characterization of the set of implementable behaviors still holds true if, instead of (ordinary) linear differential system behaviors, we deal with nD linear system behaviors, which are system behaviors that can be represented by partial differential equations of the form

$$R\left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n}\right)w(x_1, x_2, \dots, x_n) = 0,$$

with $R(\xi_1, \xi_2, \dots, \xi_n)$ a polynomial matrix in n indeterminates. Recently, in [6] a variation of condition (2) was shown to be sufficient for implementability of system behaviors in a more general (including nonlinear) context.

For a system behavior $\mathcal{K} \in \mathcal{L}^w$ to be implementable by regular feedback, another necessary condition is of course that \mathcal{K} is regularly implementable, i.e., it can be achieved by interconnecting the plant with a controller by regular interconnection through the interconnection variable c . Also for regular implementability necessary and sufficient conditions can already be found in the literature. In [1] it has been shown that a given $\mathcal{K} \in \mathcal{L}^w$ is regularly implementable if and only if, in addition to condition (2), the following condition holds:

$$\mathcal{K} + \mathcal{P}_{\text{cont}} = \mathcal{P}. \quad (3)$$

Condition (3) states that the sum of \mathcal{K} and the controllable part of \mathcal{P} is equal to \mathcal{P} . The controllable part $\mathcal{P}_{\text{cont}}$ of the behavior \mathcal{P} is defined as the largest controllable subbehavior of \mathcal{P} , which is the unique behavior $\mathcal{P}_{\text{cont}}$ with the properties that 1.) $\mathcal{P}_{\text{cont}} \subseteq \mathcal{P}$, and 2.) \mathcal{P}' controllable and $\mathcal{P}' \subseteq \mathcal{P}$ implies $\mathcal{P}' \subseteq \mathcal{P}_{\text{cont}}$. Clearly, if the manifest plant behavior \mathcal{P} is controllable, then $\mathcal{P} = \mathcal{P}_{\text{cont}}$, so condition (3) automatically holds. In this case, implementability and regular implementability are equivalent properties. For the special

case $\mathcal{N} = 0$ (which is equivalent to the “full interconnection case”), conditions (2) and (3) for regular implementability in the context of nD system behaviors can also be found in [4]. In the same context, results on regular implementability can also be found in [9].

We finally note that, again for the full interconnection case, the open problem stated in this paper has recently been studied in [3], using a somewhat different notion of linear system behavior, in discrete time. Up to now, however, the general problem has remained unsolved.

BIBLIOGRAPHY

- [1] M. N. Belur and H. L. Trentelman, “Stabilization, pole placement and regular implementability,” *IEEE Transactions on Automatic Control*, May 2002.
- [2] M. Kuijper, “Why do stabilizing controllers stabilize ?” *Automatica*, vol. 31, pp. 621-625, 1995.
- [3] V. Lomadze, *On interconnections and control*, manuscript, 2001.
- [4] P. Rocha and J. Wood, “Trajectory control and interconnection of nD systems,” *SIAM Journal on Contr. and Opt.*, vol. 40, no 1, pp. 107-134, 2001.
- [5] J. W. Polderman and J. C. Willems, *Introduction to Mathematical Systems Theory: A Behavioral Approach*, Springer Verlag, 1997.
- [6] A. J. van der Schaft, *Achievable behavior of general systems*, manuscript, submitted for publication, 2002.
- [7] J. C. Willems and H. L. Trentelman, “Synthesis of dissipative systems using quadratic differential forms, Part 1,” *IEEE Transactions on Automatic Control*, vol. 47, no. 1, pp. 53-69, 2002.
- [8] J. C. Willems, “On Interconnections, Control and Feedback,” *IEEE Transactions on Automatic Control*, vol. 42, pp. 326-337, 1997.
- [9] E. Zerz and V. Lomadze, “A constructive solution to interconnection and decomposition problems with multidimensional behaviors,” *SIAM Journal on Contr. and Opt.*, vol. 40, no 4, pp. 1072-1086, 2001.

Problem 1.11

Riccati stability

Erik I. Verriest¹

School of Electrical and Computer Engineering

Georgia Institute of Technology

Atlanta, GA,30332-0250

USA

erik.verriest@ee.gatech.edu

1 DESCRIPTION OF THE PROBLEM

Given two $n \times n$ real matrices, A and B , consider the matrix Riccati equation

$$A'P + PA + Q + PBQ^{-1}B'P + R = 0. \quad (1)$$

Can one *characterize* the pairs (A, B) for which the above equation has a solution for positive definite symmetric matrices P , Q , and R ?

In [8] a pair (A, B) was defined to be *Riccati stable* if a triple of positive definite matrices P, Q, R exists such that (1) holds.

The problem may be stated equivalently as an LMI:

Can one characterize all pairs (A, B) without invoking additional matrices, for which there exist positive definite matrices P and Q such that

$$\begin{bmatrix} A'P + PA + Q & PB \\ B'P & -Q \end{bmatrix} < 0. \quad (2)$$

2 MOTIVATION AND HISTORY OF THE PROBLEM

Equation (1) plays an important role in the stability analysis of linear time-invariant delay-differential systems. It is known [9] that the autonomous system

$$\dot{x}(t) = Ax(t) + Bx(t - \tau) \quad (3)$$

¹Support by the NSF-CNRS collaborative grant INT-9818312 is gratefully acknowledged.

is asymptotically stable, for all values of $\tau \geq 0$, if the pair (A, B) is Riccati stable. Note that since (3) has to be stable for $\tau = 0$ and $\tau \rightarrow \infty$, the matrices $A + B$ and A have to be Hurwitz stable, i.e., has its spectrum in the open left half plane. Recall also that a matrix C is Schur-Cohn stable, if its spectrum lies in the open unit disk.

If $B = 0$, thus reducing the problem to a finite dimensional time-invariant system, the Riccati equation reduces to the ubiquitous Lyapunov equation,

$$A'P + PA + S = 0, \quad (4)$$

where we have set $Q + R = S$. It is well known that a positive definite pair (P, S) exists if and only if A is Hurwitz. This condition is necessary and sufficient.

The above result and its equivalent LMI formulation, initiated a whole set of extensions: for multiple delays, distributed delays, time-variant systems (with time-variant delays) [3, 5]. In addition all the above variants can further be extended to include parameter variations (robust stability) and noise (stochastic stability). Also other types of functional differential equations (scale delay) lead to such conditions [7]. The main idea in deriving these results is the use of the Lyapunov-Krasovskii theory with appropriate Lyapunov *functionals*. The equation (1) appears also in H_∞ control theory and in game theory.

3 AVAILABLE RESULTS

In [8], where Riccati-stability was called “d-stability,” referring to “delay,” the following connections with spectral properties of A and B were obtained

Theorem 1: *If there exists a triple of symmetric positive definite matrices P, Q , and R , satisfying (1), then A is Hurwitz and $A^{-1}B$ is Schur-Cohn.*

There is no complete converse of this theorem, however, two partial converses are easily proven:

Theorem 2: *If the matrix product $A^{-1}B$ is Schur-Cohn, then there exists an orthogonal matrix Θ such that ΘA is Hurwitz, and the pair $(\Theta A, \Theta B)$ is Riccati-stable.*

Theorem 3: *If the matrix A is Hurwitz, then there exists a matrix B such that $A^{-1}B$ is Schur-Cohn and (A, B) is Riccati-stable.*

In addition the following scaling properties are shown in [8].

Lemma 1: *If (A, B) is Riccati-stable, then $(\alpha A, \alpha B)$ is Riccati-stable for all $\alpha > 0$.*

Lemma 2: *If (A, B) is Riccati-stable, then (SAS^{-1}, SBS^{-1}) is Riccati-stable, for all nonsingular S .*

Lemma 3: *If (A, B) is Riccati-stable, and B has full rank, then (A', B') is*

Riccati-stable.

The full rank condition on B can be relaxed. Lemma 3 is a *duality* result.

In [8] a detailed construction was given for a subset of Riccati-stable pairs for the case $n = 2$. It leads to an (over-)parameterization, but the construction readily extends to arbitrary dimensions, by using

Theorem 4: *Assume that the pairs $\{(A_i, B_i) \mid i = 1 \dots N\}$ are Riccati-stable for the same P -matrix. i.e., there exist $Q_i, R_i > 0$, $i = 1 \dots N$ such that*

$$A_i'P + PA_i + Q_i + PB_iQ_i^{-1}B_i'P + R_i = 0.$$

Then all pairs in the positive cone generated by the above pairs are Riccati-stable. i.e., $\forall \alpha_i \geq 0$, but not all zero, the pair $(\sum_i \alpha_i A_i, \sum_i \alpha_i B_i)$ is Riccati-stable.

The invariance of Riccati-stability under similarity (lemma 2) ensures that if (A, B) is Riccati-stable, one can transform the system to one for which the new P matrix, i.e., $S^{-T}PS^{-1}$ is the identity. Thus motivated, we provide a simplified form:

Given B , denote by \mathcal{A}_B the set of matrices A for which (A, B) is Riccati stable, with $P = I$, i.e.,

$$\mathcal{A}_B = \{A \mid \exists Q = Q' > 0, \text{ s.t. } A + A' + Q + BQ^{-1}B' < 0\}.$$

Hence, a *necessary* condition for $A \in \mathcal{A}_B$ is that its symmetric part A_s satisfies

$$A_s < -\frac{1}{2}(Q + BQ^{-1}B'),$$

for *some* $Q > 0$. If for each B the set \mathcal{A}_B can be determined, the proposed problem will be solved. The following special case is proven:

Theorem 5: *If B is in the real-diagonal form,*

$$\mathbf{B} = \text{Blockdiag} \{\Lambda_+, 0, -\Lambda_-, B_1, \dots, B_c\}$$

where $\Lambda_+ = \text{diag}\{\lambda_1, \dots, \lambda_p\}$ are the positive real eigenvalues, $-\Lambda_- = \text{diag}\{-\lambda_{p+1}, \dots, -\lambda_{p+m}\}$ are the negative real ones, and the B_k 's are 2×2 blocks $B_k = \begin{bmatrix} \sigma_k & \omega_k \\ -\omega_k & \sigma_k \end{bmatrix}$, associated with the complex eigenvalues $\sigma_k \pm i\omega_k$, then the set \mathcal{A}_B is characterized by the set of all matrices, A , whose symmetric part satisfies $A_s < -2 \text{Blockdiag} \{\Lambda_+, 0, \Lambda_-, |\sigma_1|I_2, \dots, |\sigma_c|I_2\}$.

Proof: In this block diagonal form, it is clear that it suffices to choose the same blockdiagonal structure for Q and the problem decouples. For real eigenvalues in the sets Λ_+ and $-\Lambda_-$, observe that $q + \frac{\lambda_k^2}{q} \geq 2|\lambda_k|$ and equality is obtained for $q = |\lambda_k|$. Likewise for a zero eigenvalue, the corresponding q may be taken infinitesimally small. For complex conjugate eigenvalue pairs, observe that

$$\begin{bmatrix} q_1 & q \\ q & q_2 \end{bmatrix} + \begin{bmatrix} \sigma_k & \omega_k \\ -\omega_k & \sigma_k \end{bmatrix} \begin{bmatrix} q_1 & q \\ q & q_2 \end{bmatrix}^{-1} \begin{bmatrix} \sigma_k & \omega_k \\ -\omega_k & \sigma_k \end{bmatrix} \geq 2|\sigma_k| \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Equality is achieved with $\begin{bmatrix} q_1 & q \\ q & q_2 \end{bmatrix} = \begin{bmatrix} |\sigma| & \omega \\ \omega & |\sigma| \end{bmatrix}$, if $|\sigma| \geq |\omega|$ and, switching to polar form, with $q_1 = \frac{\rho}{|\cos \phi|}(1 + \cos \chi)$, $q_2 = \frac{\rho}{|\cos \phi|}(1 - \cos \chi)$, and $q = \rho \sqrt{\tan^2 \phi - \frac{\cos^2 \chi}{\cos^2 \phi}}$, where ρ and ϕ are respectively the modulus and the argument of the complex eigenvalue $\sigma + i\omega$, and χ arbitrary with $|\cos \chi| < |\sin \phi|$ if $|\sigma| < |\omega|$. In the latter case, the solution was obtained by direct optimization of the minimal eigenvalue of the matrix $Q + B_k Q^{-1} B'_k$ over all positive definite matrices Q . Hence $Q + BQ^{-1}B' \geq 0$ if \mathbf{B} is singular, and $Q + BQ^{-1}B' \geq 2zI$, where $z = \min(\{|\lambda_k|; k = 1 \dots p + m\} \cup \{|\sigma_\ell|, \ell = 1, \dots, c\})$ if \mathbf{B} has full rank, from which the theorem follows. \square

Equations related to (1) are also discussed in [1,2,4,6].

BIBLIOGRAPHY

- [1] W. N. Anderson, Jr., T. D. Morley and G. E. Trapp, "Positive solutions to $X = A - B * X^{-1} B^*$," *LAA*, 134, pp. 53-62, 1990.
- [2] R. Datko, "Solutions of the operator equation $A^*K + KA + KKK = -W$," In: *Semigroups of Operators: Theory and Applications*, A.V. Balakrishnan, Ed., Birkhäuser, 2000.
- [3] L. Dugard and E. I. Verriest, *Stability and Control of Time-Delay Systems*, Springer-Verlag, LNCIS, vol. 228, 1998.
- [4] J. Engwerda, "On the existence of a positive definite solution to the matrix equation $X + A^T X^{-1} A = I$," *LAA*, 194, pp. 91-108, 1993.
- [5] V. B. Kolmanovskii, S.-I. Niculescu and K. Gu, "Delay effects on stability: A survey," *Proc. 38th IEEE Conf. Dec. Control*, Phoenix, AZ, pp. 1993-1998, 1999).
- [6] A. C. M. Ran and M. C. B. Reurings, "On the nonlinear matrix equation $X + A^* \mathcal{F}(X) A = Q$: solutions and perturbation theory," *Linear Algebra and its Applications*, vol. 346, pp. 15-26, 2002.
- [7] E.I. Verriest, "Robust stability, adjoints, and LQ control of scale-delay systems," *Proc. 38-th Conf. Dec. Control*, Phoenix, AZ, pp. 209-214, 1999.
- [8] E.I. Verriest, "Robust stability and stabilization: from linear to nonlinear," *Proceedings of the 2nd IFAC Workshop on Linear Time Delay Systems*, Ancona, Italy, pp. 184-195, September 2000.

- [9] E.I. Verriest and A.F. Ivanov, "Robust stabilization of systems with delayed feedback," *Proceedings of the Second International Symposium on Implicit and Robust Systems*, Warszawa, Poland, pp. 190-193, July 1991.

Problem 1.12

State and first order representations

Jan C. Willems¹

Department of Electrical Engineering - SCD (SISTA)

University of Leuven

Kasteelpark Arenberg 10

B-3001 Leuven-Heverlee

Belgium

Jan.Willems@esat.kuleuven.ac.be

1 DESCRIPTION OF THE PROBLEM

We conjecture that the solution set of a system of linear constant coefficient PDEs is Markovian if and only if it is the solution set of a system of first order PDEs. An analogous conjecture regarding state systems is also made.

Notation

First, we introduce our notation for the solution sets of linear PDEs in the \mathbf{n} real independent variables $x = (x_1, \dots, x_n)$. Let $\mathfrak{D}'_{\mathbf{n}}$ denote, as usual, the set of real distributions on \mathbb{R}^n , and $\mathfrak{L}_{\mathbf{n}}^{\mathbf{w}}$ the linear subspaces of $(\mathfrak{D}'_{\mathbf{n}})^{\mathbf{w}}$ consisting of the solutions of a system of linear constant coefficient PDEs in the \mathbf{w} real-valued dependent variables $w = \text{col}(w_1, \dots, w_{\mathbf{w}})$. More precisely, each element $\mathfrak{B} \in \mathfrak{L}_{\mathbf{n}}^{\mathbf{w}}$ is defined by a polynomial matrix $R \in \mathbb{R}^{\bullet \times \mathbf{w}}[\xi_1, \xi_2, \dots, \xi_n]$, with \mathbf{w} columns, but any number of rows, such that

$$\mathfrak{B} = \{w \in (\mathfrak{D}'_{\mathbf{n}})^{\mathbf{w}} \mid R(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n})w = 0\}.$$

We refer to elements of $\mathfrak{L}_{\mathbf{n}}^{\mathbf{w}}$ as *linear differential \mathbf{n} -D systems*. The above PDE is called a *kernel representation* of $\mathfrak{B} \in \mathfrak{L}_{\mathbf{n}}^{\mathbf{w}}$. Note that each $\mathfrak{B} \in \mathfrak{L}_{\mathbf{n}}^{\mathbf{w}}$ has many kernel representations. For an in-depth study of $\mathfrak{L}_{\mathbf{n}}^{\mathbf{w}}$, see [1] and [2].

¹This research is supported by the Belgian Federal Government under the DWTC program Interuniversity Attraction Poles, Phase V, 2002 - 2006, Dynamical Systems and Control: Computation, Identification and Modelling.

Next, we introduce a class of special three-way partitions of \mathbb{R}^n . Denote by \mathfrak{P} the following set of partitions of \mathbb{R}^n :

$$[(S_-, S_0, S_+) \in \mathfrak{P}] :\Leftrightarrow [(S_-, S_0, S_+ \text{ are disjoint subsets of } \mathbb{R}^n) \\ \wedge (S_- \cup S_0 \cup S_+ = \mathbb{R}^n) \wedge (S_- \text{ and } S_+ \text{ are open, and } S_0 \text{ is closed})].$$

Finally, we define concatenation of maps on \mathbb{R}^n . Let $f_-, f_+ : \mathbb{R}^n \rightarrow \mathfrak{F}$, and let $\pi = (S_-, S_0, S_+) \in \mathfrak{P}$. Define the map $f_- \wedge_\pi f_+ : \mathbb{R}^n \rightarrow \mathfrak{F}$, called the *concatenation* of (f_-, f_+) along π , by

$$(f_- \wedge_\pi f_+)(x) := \begin{cases} f_-(x) & \text{for } x \in S_- \\ f_+(x) & \text{for } x \in S_0 \cup S_+ \end{cases}$$

Markovian systems

Define $\mathfrak{B} \in \mathfrak{L}_n^w$ to be *Markovian* $:\Leftrightarrow$

$$[(w_-, w_+ \in \mathfrak{B} \cap \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R}^w)) \wedge (\pi = (S_-, S_0, S_+) \in \mathfrak{P}) \\ \wedge (w_-|_{S_0} = w_+|_{S_0})] \Rightarrow [(w_- \wedge_\pi w_+ \in \mathfrak{B})].$$

Think of S_- as the “*past*”, S_0 as the “*present*”, and S_+ as the “*future*.” Markovian means that if two solutions of the PDE agree on the present, then their pasts and futures are compatible, in the sense that the past (and present) of one, concatenated with the (present and) future of the other, is also a solution. In the language of probability: the past and the future are independent given the present.

We come to our first conjecture:

$$\mathfrak{B} \in \mathfrak{L}_n^w \text{ is Markovian} \\ \text{if and only if} \\ \text{it has a kernel representation that is first order.}$$

Thus, it is conjectured that a Markovian system admits a kernel representation of the form

$$R_0 w + R_1 \frac{\partial}{\partial x_1} w + R_2 \frac{\partial}{\partial x_2} w + \cdots R_n \frac{\partial}{\partial x_n} w = 0.$$

Oberst [2] has proven that there is a one-to-one relation between \mathfrak{L}_n^w and the submodules of

$\mathbb{R}^w[\xi_1, \xi_2, \dots, \xi_n]$, each $\mathfrak{B} \in \mathfrak{L}_n^w$ being identifiable with its set of *annihilators*

$$\mathfrak{N}_{\mathfrak{B}} := \{n \in \mathbb{R}^w[\xi_1, \xi_2, \dots, \xi_n] \mid n^\top \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_n} \right) \mathfrak{B} = 0\}.$$

Markovianity is hence conjectured to correspond exactly to those $\mathfrak{B} \in \mathfrak{L}_n^w$ for which the submodule $\mathfrak{N}_{\mathfrak{B}}$ has a set of first order generators.

State systems

In this section we consider systems with two kind of variables: \mathbf{w} real-valued *manifest* variables, $w = \text{col}(w_1, \dots, w_{\mathbf{w}})$, and \mathbf{z} real-valued *state* variables, $z = \text{col}(z_1, \dots, z_{\mathbf{z}})$. Their joint behavior is again assumed to be the solution set of a system of linear constant coefficient PDEs. Thus we consider behaviors in $\mathfrak{L}_{\mathbf{n}}^{\mathbf{w}+\mathbf{z}}$, whence each element $\mathfrak{B} \in \mathfrak{L}_{\mathbf{n}}^{\mathbf{w}+\mathbf{z}}$ is described in terms of two polynomial matrices $(R, M) \in \mathbb{R}^{\bullet \times (\mathbf{w}+\mathbf{z})}[\xi_1, \xi_2, \dots, \xi_{\mathbf{n}}]$ by

$$\mathfrak{B} = \{(w, z) \in (\mathfrak{D}'_{\mathbf{n}})^{\mathbf{w}} \times (\mathfrak{D}'_{\mathbf{n}})^{\mathbf{z}} \mid R(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_{\mathbf{n}}})w + M(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \dots, \frac{\partial}{\partial x_{\mathbf{n}}})z = 0\}.$$

Define $\mathfrak{B} \in \mathfrak{L}_{\mathbf{n}}^{\mathbf{w}+\mathbf{z}}$ to be a *state system* with state $z : \Leftrightarrow$

$$[((w_-, z_-), (w_+, z_+)) \in \mathfrak{B} \cap \mathfrak{C}^{\infty}(\mathbb{R}^{\mathbf{n}}, \mathbb{R}^{\mathbf{w}+\mathbf{z}})) \wedge (\pi = (S_-, S_0, S_+) \in \mathfrak{P}) \\ \wedge (z_-|_{S_0} = z_+|_{S_0})] \Rightarrow [(w_-, z_-) \wedge_{\pi} (w_+, z_+) \in \mathfrak{B}].$$

Think again of S_- as the “*past*”, S_0 as the “*present*”, S_+ as the “*future*”. State means that if the state components of two solutions agree on the present, then their pasts and futures are compatible, in the sense that the past of one solution (involving both the manifest and the state variables), concatenated with the present and future of the other solution, is also a solution. In the language of probability: the present state “*splits*” the past and the present plus future of the manifest and the state trajectory combined.

We come to our second conjecture:

$$\begin{aligned} &\mathfrak{B} \in \mathfrak{L}_{\mathbf{n}}^{\mathbf{w}+\mathbf{z}} \text{ is a state system} \\ &\quad \text{if and only if} \\ &\quad \text{it has a kernel representation} \\ &\quad \text{that is first order in the state variables } \mathbf{z} \\ &\quad \text{and zero-th order in the manifest variables } \mathbf{w}. \end{aligned}$$

I.e., it is conjectured that a state system admits a kernel representation of the form

$$R_0 w + M_0 z + M_1 \frac{\partial}{\partial x_1} z + M_2 \frac{\partial}{\partial x_2} z + \dots M_{\mathbf{n}} \frac{\partial}{\partial x_{\mathbf{n}}} z = 0.$$

2 MOTIVATION AND HISTORY OF THE PROBLEM

These open problems aim at understanding state and state construction for \mathbf{n} -D systems. Maxwell’s equations constitute an example of a Markovian system. The diffusion equation and the wave equation are non-examples.

3 AVAILABLE RESULTS

It is straightforward to prove the “if”-part of both conjectures. The conjectures are true for $n = 1$, i.e., in the ODE case, see [3].

BIBLIOGRAPHY

- [1] H. K. Pillai and S. Shankar, “A behavioral approach to control of distributed systems,” *SIAM Journal on Control and Optimization*, vol. 37, pp. 388-408, 1999.
- [2] U. Oberst, “Multidimensional constant linear systems,” *Acta Applicandae Mathematicae*, vol. 20, pp. 1-175, 1990.
- [3] P. Rapisarda and J. C. Willems, “State maps for linear systems,” *SIAM Journal on Control and Optimization*, vol. 35, pp. 1053-1091, 1997.

Problem 1.13

Projection of state space realizations

Antoine Vandendorpe and Paul Van Dooren

Department of Mathematical Engineering

Université catholique de Louvain

B-1348 Louvain-la-Neuve

Belgium

1 DESCRIPTION OF THE PROBLEM

We consider two $m \times p$ strictly proper transfer functions

$$T(s) = C(sI_n - A)^{-1}B, \quad \hat{T}(s) = \hat{C}(sI_k - \hat{A})^{-1}\hat{B}, \quad (1)$$

of respective Mc Millan degrees n and $k < n$. We want to characterize the set of projecting matrices $Z, V \in \mathbb{C}^{n \times k}$ such that

$$\hat{C} = CV, \quad \hat{A} = Z^T AV, \quad \hat{B} = Z^T B, \quad Z^T V = I_k. \quad (2)$$

Given only $T(s)$, we are interested in characterizing the set of all transfer functions $\hat{T}(s)$ that can be obtained via the projection equations (1,2). Here is our first conjecture.

Conjecture 1. Any minimal state space realization of $\hat{T}(s)$ can be obtained by a projection from any minimal state space realization of $T(s)$ if

$$\frac{m+p}{2} \leq n-k. \quad (3)$$

In the case that condition (3) is not satisfied, we give a second, more detailed conjecture in section 3 in terms of the zero structure of $T(s) - \hat{T}(s)$. A justification of conjecture 1 is that it actually holds for SISO systems. Indeed, the following result was shown in [4]:

Theorem. Let $T(s) = C(sI_n - A)^{-1}B$ and $\hat{T}(s) = \hat{C}(sI_k - \hat{A})^{-1}\hat{B}$ be arbitrary strictly proper SISO transfer functions of McMillan degrees n and $k < n$, respectively. Then any minimal state space realization of $\hat{T}(s)$ can be constructed via projection of any minimal state space realization of $T(s)$ using equations (2).

2 MOTIVATION AND HISTORY OF THE PROBLEM

Equation (2) arises naturally in the general framework of model reduction of large scale linear systems [1]. In this context we are given a transfer function $T(s)$ of McMillan degree n , which we want to approximate by a transfer function $\hat{T}(s)$ of smaller McMillan degree k , in order to solve a simpler analysis or design problem.

Classical model reduction techniques include *modal approximation* (where the dominant poles of the original transfer function are copied in the reduced order transfer function), *balanced truncation* and *optimal Hankel norm approximation* (related to the controllability and observability Grammians of the transfer function [10]). These methods either provide a global error bound between the original and reduced-order system and/or guarantee stability of the reduced order system. Unfortunately, their exact calculation involves $O(n^3)$ floating point operations even for systems with sparse model matrices $\{A, B, C\}$, which becomes untractable for a very large state dimension n .

Only the image of the projecting matrices Z and V are important since choosing other bases satisfying the bi-orthogonality condition (2) amounts to a state-space transformation of the realization of $\hat{T}(s)$.

A more recent approach involves *generalized Krylov spaces* ([3]) which are defined as the images of the *generalized Krylov matrices*

$$[(\sigma I_n - A)^{-1}B \cdots (\sigma I_n - A)^{-k}B] X, \quad X = \begin{bmatrix} x_0 & & \\ \vdots & \ddots & \\ x_{k-1} & \cdots & x_0 \end{bmatrix} \quad (4)$$

and

$$[(\gamma I_n - A^T)^{-1}C^T \cdots (\gamma I_n - A^T)^{-\ell}C^T] Y, \quad Y = \begin{bmatrix} y_0 & & \\ \vdots & \ddots & \\ y_{\ell-1} & \cdots & y_0 \end{bmatrix}. \quad (5)$$

These are related to the respective right and left *tangential interpolation* conditions

$$[T(s) - \hat{T}(s)] x(s) = O(s - \sigma)^k, \quad x(s) \doteq \sum_{i=0}^{k-1} x_i (s - \sigma)^i \quad (6)$$

and

$$[T(s) - \hat{T}(s)]^T y(s) = O(s - \gamma)^\ell, \quad y(s) \doteq \sum_{i=0}^{\ell-1} y_i (s - \gamma)^i. \quad (7)$$

In the most general form, one imposes such conditions in several points σ_i and γ_j as well as bi-tangential conditions (see [2], [5] for more details). The calculation of Krylov spaces and the solution of the corresponding tangential interpolation problem typically exploits the sparsity or the structure of the

model matrices (A, B, C) of the original system and are therefore efficient for large scale dynamical systems with such structure. Their drawbacks are that the resulting reduced order systems have no guaranteed error bound and that stability is not necessarily preserved.

The conjecture -and open problem- is that these methods are in fact quite universal (i.e., they contain the classical methods as special cases) and can be formulated in terms of Sylvester equations and generalized eigenvalue problems. Tangential interpolation would then be a unifying procedure to construct reduced-order transfer functions in which only the interpolation points and tangential conditions need to be specified.

3 OUR CONJECTURE

The error transfer function $E(s) \doteq T(s) - \hat{T}(s)$ is realized by the following pencil:

$$M - Ns \doteq \left[\begin{array}{cc|c} A & 0 & B \\ 0 & \hat{A} & \hat{B} \\ \hline C & -\hat{C} & 0 \end{array} \right] - s \left[\begin{array}{cc|c} I_n & & \\ & I_k & \\ \hline & & 0 \end{array} \right]. \quad (8)$$

The transmission zeros of the system matrix (i.e., the system zeros of its minimal part) can be chosen as interpolation points between $T(s)$ and $\hat{T}(s)$ since the normal rank of $E(s)$ drops below its normal rank. Therefore one can impose interpolation conditions of the type (6,7) for appropriate choices of $x(s)$ and $y(s)$ and generalized eigenvalues σ and γ of (8).

Our conjecture tries to give necessary and sufficient conditions for this in terms of the system zero matrix.

Conjecture 2. A minimal state space realization of the strictly proper transfer function $\hat{T}(s)$ of McMillan degree k can be obtained by projection from a minimal state space realization of the strictly proper transfer function $T(s)$ of McMillan degree $n > k$ if and only if there exist two regular pencils, $M_r - sN_r$ and $M_l - sN_l$ such that the matrices $L, \hat{L}, R, \hat{R}, Q_l$ and Q_r of the following equations

$$\left[\begin{array}{ccc} A - sI_n & 0 & B \\ 0 & \hat{A} - sI_k & \hat{B} \\ C & -\hat{C} & 0 \end{array} \right] \left[\begin{array}{c} RN_r \\ \hat{R}N_r \\ Q_r \end{array} \right] = \left[\begin{array}{c} R \\ \hat{R} \\ 0 \end{array} \right] (M_r - sN_r), \quad (9)$$

$$\left[\begin{array}{ccc} A^T - sI_n & 0 & C^T \\ 0 & \hat{A}^T - sI_k & -\hat{C}^T \\ B^T & \hat{B}^T & 0 \end{array} \right] \left[\begin{array}{c} LN_l \\ -\hat{L}N_l \\ Q_l \end{array} \right] = \left[\begin{array}{c} L \\ -\hat{L} \\ 0 \end{array} \right] (M_l - sN_l), \quad (10)$$

satisfy the following conditions :

$$1. \begin{bmatrix} N_l^T L^T & -N_l^T \hat{L}^T & Q_l^T \end{bmatrix} (M - Ns) \begin{bmatrix} RN_r \\ \hat{R}N_r \\ Q_r \end{bmatrix} = 0,$$

$$2. \dim \left(\text{Im}(\hat{R}N_r) \right) = \dim \left(\text{Im}(\hat{L}N_l) \right) = k.$$

Moreover, such matrices always exist provided $2k \leq 2n - m - p$.

The conditions given by our conjecture are at least sufficient. Indeed, from equations (10), and (9) and the regularity assumption of $M_r - sN_r$ and $M_l - sN_l$, it follows that

$$CRN_r = \hat{C}\hat{R}N_r \quad , \quad N_l^T L^T B = N_l^T \hat{L}^T \hat{B}. \quad (11)$$

Then, from condition 1,

$$N_l^T L^T RN_r = N_l^T \hat{L}^T \hat{R}N_r \quad , \quad N_l^T L^T ARN_r = N_l^T \hat{L}^T \hat{A}\hat{R}N_r. \quad (12)$$

Finally, conditions 1 and 2 imply that the matrices $\hat{R}N_r$ and $\hat{L}N_l$ are right invertible. Defining $Z, V \in \mathbb{C}^{n \times k}$ by

$$Z = LN_l(\hat{L}N_l)^{-r}, \quad V = RN_r(\hat{R}N_r)^{-r}, \quad (13)$$

we can easily verify equations (1) and (2). Another justification is that (by looking carefully at the proof of theorem 1) Conjecture 3 is true for the SISO case.

We now present the link with the Krylov techniques. Equations (9) and (10) give us the following Sylvester equations:

$$ARN_r - RM_r + BQ_r = 0 \quad , \quad A^T LN_l - LM_l + C^T Q_l = 0. \quad (14)$$

These Sylvester equations correspond to generalized left and right eigenspaces of the system zero matrix (8). More precisely, $\text{Im}(RN_r)$ and $\text{Im}(LN_l)$ can be expressed as *generalized Krylov spaces* of the form (4) and (5). The choice of matrices M_l, N_l, M_r, N_r, Q_l , and Q_r correspond respectively to left and right tangential interpolation conditions at the eigenvalues σ_i of $(M_r - sN_r)$ and γ_j of $(M_l - sN_l)$, that are satisfied between $T(s)$ and $\hat{T}(s)$ (see [5]). These eigenspaces correspond to disjoint parts of the spectrum of $M - Ns$ such that the product $N_l^T L^T RN_r = N_l^T \hat{L}^T \hat{R}N_r$ is invertible (see [5] for more details).

In other words, our conjecture is that any projected reduced-order transfer function can be obtained by imposing some interpolation conditions or some modal approximation conditions with respect to the original transfer function. Moreover, a solution always exists provided $2k \leq 2n - m - p$ (i.e., for all $\hat{T}(s)$ of sufficiently small degree k). If this turns out to be true, we could hope to find the interpolation conditions that yield, e.g., the optimal Hankel norm or optimal H_∞ norm reduced-order models using cheap interpolation techniques.

4 AVAILABLE RESULTS

Independently, Halevi recently proved in [6] new results concerning the general framework of model order reduction via projection. The unknowns

Z and V have $2nk$ parameters (or degrees of freedom), while (2) imposes $(2k + m + p)k$ constraints. He shows that the case $k = n - \frac{m+p}{2}$ corresponds to a finite number of solutions. Moreover, for the particular case $m = p$ and $k = n - m$, he shows that any pair of projecting matrices Z, V satisfying (2) can be seen as generalized eigenspaces of a certain matrix pencil. The matrix pencil used by Halevi can be linked to the system zero matrix of the error transfer function defined in equation (8).

Matrices Z and V satisfying (2) are also the k trailing rows of S^{-1} , respectively columns of S which transform the system (A, B, C) to the system $(S^{-1}AS, S^{-1}B, CS)$:

$$\left[\begin{array}{c|c} S^{-1}AS - sI_n & S^{-1}B \\ \hline CS & 0 \end{array} \right] = \left[\begin{array}{cc|c} * & * & * \\ * & \hat{A} - sI_k & \hat{B} \\ * & \hat{C} & 0 \end{array} \right]. \quad (15)$$

The existence of projecting matrices Z, V satisfying (1 and 2) is therefore related to the above submatrix problem. A square matrix \hat{A} is said to be *embedded* in a square matrix A when there exists a change of coordinates S such that $\hat{A} - sI_k$ is a submatrix of $S^{-1}(A - sI_n)S$. Necessary and sufficient conditions for the embedding of such monic pencils are given in [9], [8].

As for monic pencils, we say that the pencil $\hat{M} - \hat{N}s$ is embedded in the pencil $M - Ns$ when there exist invertible matrices Le, Ri such that $\hat{M} - \hat{N}s$ is a sub-matrix of $Le(M - Ns)Ri$. Finding necessary and sufficient conditions for the embedding of such general pencils is still an open problem [7]. Nevertheless, one obtains from [9], [8], [7] *necessary* conditions on $(\hat{C}, \hat{A}, \hat{B})$ and (C, A, B) for $\left[\begin{array}{c|c} \hat{A} - sI_k & \hat{B} \\ \hline \hat{C} & 0 \end{array} \right]$ to be embedded in $\left[\begin{array}{c|c} A - sI_n & B \\ \hline C & 0 \end{array} \right]$. These obviously give *necessary* conditions for the existence of projecting matrices Z, V satisfying (1 and 2). We hope to be able to shed new light on the necessary and sufficient conditions for the embedding problem via the connections developed in this paper.

BIBLIOGRAPHY

- [1] A. C. Antoulas, "Lectures on the approximation of large-scale dynamical systems," *SIAM Book Series: Advances in Design and Control*, 2002.
- [2] J. A. Ball, I. Gohberg and L. Rodman, *Interpolation of rational matrix functions*, Birkhäuser Verlag, Basel, 1990.
- [3] K. Gallivan, A. Vandendorpe, and P. Van Dooren, "Model reduction via tangential interpolation," *MTNS 2002 (15th Symp. on the Mathematical Theory of Networks and Systems)*, University of Notre Dame, South Bend, Indiana, USA, August 2002.

- [4] K. Gallivan, A. Vandendorpe, and P. Van Dooren, "Model Reduction via truncation: an interpolation point of view," *Linear Algebra Appl.*, submitted.
- [5] K. Gallivan, A. Vandendorpe, and P. Van Dooren, "Model reduction of MIMO systems via tangential interpolation," *Internal report, Université catholique de Louvain*, 2002. Available at <http://www.auto.ucl.ac.be/~vandendorpe/>.
- [6] Y. Halevi, "On model order reduction via projection," *15th IFAC World Congress on automatic control*, Barcelona, July 2002.
- [7] J. J. Loiseau, S. Mondié, I. Zaballa and P. Zagalak, "Assigning the Kronecker invariants of a matrix pencil by row or column completions," *Linear Algebra Appl.*, 278, pp. 327-336, 1998.
- [8] E. Marques de Sá, "Imbedding conditions for λ -matrices," *Linear Algebra Appl.*, 24, pp. 33-50, 1979.
- [9] R. C. Thompson, "Interlacing Inequalities for Invariant Factors," *Linear Algebra Appl.*, 24, pp. 1-31, 1979.
- [10] K. Zhou, J. C. Doyle, and K. Glover, *Robust and optimal control*, Upper Saddle River, Prentice Hall, Inc, N.J.: 1996.

PART 2

Stochastic Systems

Problem 2.1

On error of estimation and minimum of cost for wide band noise driven systems

Agamirza E. Bashirov
Department of Mathematics
Eastern Mediterranean University
Mersin 10
Turkey
`agamirza.bashirov@emu.edu.tr`

1 DESCRIPTION OF THE PROBLEM

The suggested open problem concerns the error of estimation and the minimum of the cost in the filtering and optimal control problems for a partially observable linear system corrupted by wide band noise processes.

Recent results allow to construct a wide band noise process in a certain integral form on the basis of its autocovariance function and design the optimal filter and the optimal control for a partially observable linear system corrupted by such wide band noise processes. Moreover, explicit formulae for the error of estimation and for the minimum of the cost have been obtained. But, the information about wide band noise contained in its autocovariance function is incomplete. Hence, every autocovariance function generates infinitely many wide band noise processes represented in the integral form. Consequently, the error of estimation and the minimum of the cost mentioned above are for a sample wide band noise process corresponding to the given autocovariance function.

The following problem arises: given an autocovariance function, what are the least upper and greatest lower bounds of the respective error of estimation and the respective minimum of the cost? What are the distributions of the error of estimation and the minimum of the cost? What are the parameters of the wide band noise process producing the average error and the average minimum of the cost?

2 MOTIVATION AND HISTORY OF THE PROBLEM

Modern stochastic optimal control and filtering theories use white noise driven systems. Results such as the separation principle and the Kalman-Bucy filtering are based on the white noise model. In fact, white noise, being a mathematical idealization, gives only an approximate description of real noise. In some fields the parameters of real noise are near to the parameters of white noise and, so, the mathematical methods of control and filtering for white noise driven systems can be satisfactorily applied to them. But in many fields white noise is a crude approximation to real noise. Consequently, the theoretical optimal controls and the theoretical optimal filters for white noise driven systems become not optimal and, indeed, might be quite far from being optimal. It becomes important to develop the control and estimation theories for the systems driven by noise models that describe real noise more adequately. Such a noise model is the wide band noise model.

The importance of wide band noise processes was mentioned by Fleming and Rishel [1]. An approach to wide band noise based on approximations by white noise was used in Kushner [2]. Another approach to wide band noise based on representation in a certain integral form was suggested in [3] and its applications to space engineering and gravimetry was discussed in [4, 5]. Filtering, smoothing, and prediction results for wide band noise driven linear systems are obtained in [3, 6]. The proofs in [3, 6] are given through the duality principle and, technically, they are routine, making further developments in the theory difficult. A more handle technique based on the reduction of a wide band noise driven system to a white noise driven system was developed in [7, 8, 9]. This technique allows to find the explicit formulae for the optimal filter and for the optimal control, as well as for the error of estimation and for the minimum of the cost in the filtering and optimal control problems for a wide band noise driven linear system. In particular the open problem described here was originally formulated in [9]. A complete discussion of the subject can be found in the recent book [10].

3 AVAILABLE RESULTS AND DISCUSSION

The random process φ with the property $\text{cov}(\varphi(t+s), \varphi(t)) = \lambda(t, s)$ if $0 \leq s < \varepsilon$ and $\text{cov}(\varphi(t+s), \varphi(t)) = 0$ if $s \geq \varepsilon$, where $\varepsilon > 0$ is a small value and λ is a nonzero function, is called a *wide band noise process* and it is said to be stationary (in wide sense) if the function λ (called the *autocovariance function* of φ) depends only on s (see Fleming and Rishel [8]).

Starting from the autocovariance function λ , one can construct the respective wide band noise process φ in the integral form

$$\varphi(t) = \int_{-\min(t, \varepsilon)}^0 \phi(\theta) w(t + \theta) d\theta, \quad t \geq 0, \quad (1)$$

where w is a white noise process with $\text{cov}(w(t), w(s)) = \delta(t - s)$, δ is the Dirac's delta-function, $\varepsilon > 0$ and ϕ is a solution of the equation

$$\int_{-\varepsilon}^{-s} \phi(\theta)\phi(\theta + s) d\theta = \lambda(s), \quad 0 \leq s \leq \varepsilon. \quad (2)$$

The solution φ of (2) is called a *relaxing function*. Since in (2) ϕ has only one variable the process φ from (1) is stationary in wide sense (except small time interval $[0, \varepsilon]$). The following theorem from [8, 9] is crucial for the proposed problem.

Theorem: Let $\varepsilon > 0$ and let λ be a continuous real-valued function on $[0, \varepsilon]$. Define the function λ_0 as the even extension of λ to the real line vanishing outside of $[-\varepsilon, \varepsilon]$ and assume that λ_0 is a positive definite function with $\mathcal{F}(\lambda_0)^{1/2} \in L_2(-\infty, \infty)$ where $\mathcal{F}(\lambda_0)$ is the Fourier transformation of λ_0 . Then there exists an infinite number of solutions of the equation (2) in the space $L_2(-\varepsilon, 0)$ if λ is a nonzero function a.e. on $[-\varepsilon, 0]$.

The nonuniqueness of the solution of equation (2) demonstrates that the covariance function λ does not provide complete information about the respective wide band noise process φ . Therefore, for given λ , a sample solution ϕ of (2) generates the random process φ in the form (1) that could be considered as a less or more adequate model of real noise. In order to make a reasonable decision about the relaxing function, one of the ways is studying the distributions of the error of estimation and the minimum of the cost in filtering and control problems, finding the average error and the average minimum and identifying the relaxing function $\bar{\phi}$ producing these average values. For this, the explicit formulae from [7, 8, 9] (they are not given here because of the length) can be used to investigate the problem analytically or numerically. Also, the proof of the theorem from [8, 9] can be useful for construction different solutions of equation (2).

Finally, note that in a partially observable system both the state (signal) and the observations may be disturbed by wide band noise processes. Hence, the suggested problem concerns both these cases and their combination as well.

BIBLIOGRAPHY

- [1] W. H. Fleming and R. W. Rishel, *Deterministic and Stochastic Optimal Control*, New York, Springer Verlag, 1975, p. 126.
- [2] H. J. Kushner, *Weak Convergence Methods and Singularly Perturbed Stochastic Control and Filtering Problems*, Boston, Birkhäuser, 1990.
- [3] A. E. Bashirov, "On linear filtering under dependent wide band noises", *Stochastics*, 23, pp. 413-437, 1988.

- [4] A. E. Bashirov, L. V. Eppelbaum and L. R. Mishne, "Improving Eötvös corrections by wide band noise Kalman filtering," *Geophys. J. Int.*, 108, pp. 193-127, 1992.
- [5] A. E. Bashirov, "Control and filtering for wide band noise driven linear systems", *Journal on Guidance, Control and Dynamics*, 16, pp. 983-985, 1993.
- [6] A. E. Bashirov, H. Etikan and N. Şemi, "Filtering, smoothing and prediction of wide band noise driven systems", *J. Franklin Inst., Eng. Appl. Math.*, 334B, pp. 667-683, 1997.
- [7] A. E. Bashirov, "On linear systems disturbed by wide band noise", *Proceedings of the 14th International Conference on Mathematical Theory of Networks and Systems*, Perpignan, France, June 19-23, 7 p., 2000.
- [8] A. E. Bashirov, "Control and filtering of linear systems driven by wide band noise," *1st IFAC Symposium on Systems Structure and Control*, Prague, Czech Republic, August 29-31, 6 p., 2001.
- [9] A. E. Bashirov and S. Uğural, "Analyzing wide band noise with application to control and filtering", *IEEE Trans. Automatic Control*, 47, pp. 323-327, 2002.
- [10] A. E. Bashirov, *Partially Observable Linear Systems Under Dependent Noises*, Basel, Birkhäuser, 2003.

Problem 2.2

On the stability of random matrices

Giuseppe Calafiore*, Fabrizio Dabbene**

*Dip. di Automatica e Informatica

**IEIT-CNR

Politecnico di Torino

C.so Duca degli Abruzzi, 24 Torino,

Italy

{giuseppe.calafiore, fabrizio.dabbene}@polito.it

1 INTRODUCTION AND MOTIVATION

In the theory of linear systems, the problem of assessing whether the omogeneous system $\dot{x} = Ax$, $A \in \mathbb{R}^{n,n}$ is asymptotically stable is a well understood (and fundamental) one. Of course, the system (and we shall say also the matrix A) is stable if and only if $\text{Re}\lambda_i < 0$, $i = 1, \dots, n$, being λ_i the eigenvalues of A .

Evolving from this basic notion, much research effort has been devoted in recent years to the study of *robust* stability of a system. Without entering in the details of more than thirty years of fruitful research, we could condense the essence of the robust stability problem as follows: given a bounded set Δ and a stable matrix $A \in \mathbb{R}^{n,n}$, state whether $A_\Delta = A + \Delta$ is stable for all $\Delta \in \Delta$. Since the above deterministic problem may be computationally hard in some cases, a recent line of study proposes to introduce a probability distribution over Δ , and then to assess the *probability* of stability of the *random matrix* $A + \Delta$. Actually, in the probabilistic approach to robust stability, this probability is *not* analytically computed but only *estimated* by means of randomized algorithms, which makes the problem feasible from a computational point of view, see, for instance, [3] and the references therein.

Leaving apart the randomized approach, which circumvents the problem of analytical computations, there is a clear disparity between the abundance of results available for the deterministic problem (both positive and negative results, in the form of computational “hardness,” [2]) and their deficiency in the probabilistic one. In this latter case, almost no analytical result is known among control researchers.

The objective of this note is to encourage research on random matrices in the control community. The one who adventures in this field will encounter unexpected and exciting connections among different fields of science and beautiful branches of mathematics.

In the next section, we resume some of the known results on random matrices, and state a simple new (to the best of our knowledge) closed form result on the probability of stability of a certain class of random matrices. Then, in section 3 we propose three open problems related to the analytical assessment of the probability of stability of random matrices. The problems are presented in what we believe is their order of difficulty.

2 AVAILABLE RESULTS

Notation : A real random matrix \mathbf{X} is a matrix whose elements are real random variables. The probability density (pdf) of \mathbf{X} , $f_{\mathbf{X}}(X)$ is defined as the joint pdf of its elements. The notation $\mathbf{X} \sim \mathbf{Y}$ means that \mathbf{X}, \mathbf{Y} are random quantities with the same pdf. The Gaussian density with mean μ and variance σ^2 is denoted as $N(\mu, \sigma^2)$. For a matrix X , $\rho(X)$ denotes the spectral radius, and $\|\mathbf{X}\|$ the Frobenius norm. The multivariate Gamma function is defined as $\Gamma_n(x) = \pi^{n(n-1)/4} \prod_{i=1}^n \Gamma(x - (i-1)/2)$, where $\Gamma(\cdot)$ is the standard Gamma function.

In this note, we consider the class of random matrices (a class of random matrices is often called an “ensemble” in the physics literature) whose density is invariant under orthogonal similarity. For a random matrix \mathbf{X} in this class, we have that $\mathbf{X} \sim U\mathbf{X}U^T$, for any fixed orthogonal matrix U . For *symmetric* orthogonal invariant random matrices, it can be proved that the pdf of \mathbf{X} is a function of only its eigenvalues $\Lambda \doteq \text{diag}(\lambda_1, \dots, \lambda_n)$, i.e.,

$$f_{\mathbf{X}}(X) = g_{\mathbf{X}}(\Lambda). \quad (1)$$

Orthogonal invariant random matrices may seem specialized, but we provide below some notable examples:

1. G_n : Gaussian matrices. It is the class of $n \times n$ real random matrices with independent identically distributed (iid) elements drawn from $N(0, 1)$.
2. W_n : Wishart matrices. Symmetric $n \times n$ random matrices of the form $\mathbf{X}\mathbf{X}^T$, where \mathbf{X} is G_n .
3. GOE_n : Gaussian Orthogonal Ensemble. Symmetric $n \times n$ random matrices of the form $(\mathbf{X} + \mathbf{X}^T)/2$, where \mathbf{X} is G_n .
4. S_n : Symmetric orthogonal invariant ensemble. Generic symmetric $n \times n$ random matrices whose density satisfies (1). W_n and GOE_n are special cases of these.

5. US_n^ρ : Symmetric $n \times n$ random matrices from S_n , which are uniform over the set $\{X \in \mathbb{R}^{n,n} : \rho(X) \leq 1\}$.
6. US_n^F : Symmetric $n \times n$ random matrices from S_n , which are uniform over the set $\{X \in \mathbb{R}^{n,n} : \|X\| \leq 1\}$.

Whishart matrices have a long history, and are well studied in the statistics literature, see [1] for an early reference. The Gaussian Orthogonal Ensemble is a fundamental model used to study the theory of energy levels in nuclear physics, and it has been originally introduced by Wigner [9, 8]. A thorough account of its statistical properties is presented in [7].

A fundamental result for the S_n ensemble is that the joint pdf of the eigenvalues of random matrices belonging to S_n is known analytically. In particular, if $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ are the eigenvalues of a random matrix \mathbf{X} belonging to S_n , then their pdf $f_\Lambda(\Lambda)$ is

$$f_\Lambda(\Lambda) = \frac{\pi^{n^2/2}}{\Gamma_n(n/2)} g_{\mathbf{X}}(\Lambda) \prod_{1 \leq i < j \leq n} (\lambda_i - \lambda_j). \quad (2)$$

This result can be deduced from [7], and it is also presented in [4]. For some of the ensembles listed above, this specializes to:

$$W_n : \frac{\pi^{n^2}}{\Gamma_n^2(n/2)} \exp(-\frac{1}{2} \sum_i \lambda_i) \prod_i \lambda_i^{-1/2} \prod_{1 \leq i < j \leq n} (\lambda_i - \lambda_j) \quad (3)$$

$$\text{GOE}_n : \frac{1}{2^{n/2} \prod_i \Gamma(i/2)} \exp(-\frac{1}{2} \sum_i \lambda_i^2) \prod_{1 \leq i < j \leq n} (\lambda_i - \lambda_j) \quad (4)$$

$$\text{US}_n^\rho : K_u \prod_{1 \leq i < j \leq n} (\lambda_i - \lambda_j), \quad 1 \geq \lambda_1 \geq \dots \geq \lambda_n \geq -1. \quad (5)$$

The normalization constant K_u in the last expression can be determined in closed form solving a Legendre integral, see eq. (17.6.3) of [7]

$$K_u = n! 2^{\frac{n}{2}(n+1)} \prod_{j=0}^{n-1} \frac{\Gamma(3/2 + j/2) \Gamma^2(1 + j/2)}{\Gamma(3/2) \Gamma((n+j+3)/2)}. \quad (6)$$

Clearly, knowing the joint density of the eigenvalues is a key step in the direction of computing the probability of stability of a random matrix. We remark that the above results all refer to the symmetric case, which has the advantage of having all *real* eigenvalues. Very little is known for instance about the distribution of the eigenvalues of generic Gaussian matrices G_n . By consequence, to the best of our knowledge, nothing is known about the probability of stability of Gaussian random matrices (i.e., matrices drawn using Matlab `randn` command). Famous asymptotic results (i.e., for $n \rightarrow \infty$) go under the name of “circular laws” and are presented in [6]. An exact formula for the distribution of the *real* eigenvalues may be found in [5]. We show below a (seemingly new) result regarding the probability of stability for the US_n^ρ ensemble.

2.1 Probability of stability for the US_n^ρ ensemble

Given an $n \times n$ real random matrix \mathbf{X} , let $f_\Lambda(\Lambda)$ be the marginal density of the eigenvalues of \mathbf{X} . The *probability of stability* of \mathbf{X} is defined as

$$P \doteq \int \cdots \int_{\text{Re}\Lambda < 0} f_\Lambda(\Lambda) d\Lambda. \quad (7)$$

We now compute this probability for matrices in the US_n^ρ ensemble, whose pdf is given in (5). To this end, we first remove the ordering of the eigenvalues, and therefore divide by $n!$ the pdf (5). Then, the probability of stability is

$$P_{US} = \frac{K_u}{n!} \int_{-1}^0 \cdots \int_{-1}^0 \prod_{1 \leq i < j \leq n} |\lambda_i - \lambda_j| d\lambda_1 \cdots d\lambda_n. \quad (8)$$

This multiple integral is a Selberg type integral whose solution is reported for instance in [7], p. 339. The above probability results to be

$$P_{US} = 2^{-\frac{1}{2}n(n+1)}.$$

3 OPEN PROBLEMS

The probability of stability can be computed also for the GOE_n ensemble and the US_n^F ensemble, using a technique of integration over alternate variables. We pose this as the first open problem (of medium difficulty):

Problem 1: *Determine the probability of stability for the GOE_n and the US_n^F ensembles.*

A much harder problem would be to determine an analytic expression for the density of the eigenvalues (which are now both real and complex) of Gaussian matrices G_n , and then integrate it to obtain the probability of stability for the G_n ensemble:

Problem 2: *Determine the probability of stability for the G_n ensemble.*

A numerical estimate of the probability of (Hurwitz) stability for G_n matrices is reported in table 2.2.1, as a function of dimension n .

n	1	2	3	4	5	6
Prob.	0.500	0.250	0.104	0.037	0.011	0.003

Table 2.2.1 Estimated probability of stability for G_n matrices.

As the reader may have noticed, all the problems treated so far relate to random matrices with zero mean. From the point of view of robustness analysis it would be much more interesting to consider the case of *biased* random matrices. This motivates our last (and most difficult) open problem:

Problem 3: Let $A \in \mathbb{R}^{n,n}$ be a given stable matrix. Determine the probability of stability of the random matrix $A + \mathbf{X}$, where \mathbf{X} belongs to one of the ensembles listed in section 2.

ACKNOWLEDGEMENTS

The authors wish to thank Professor Madan Lal Mehta for reading and giving his precious advice on this manuscript.

BIBLIOGRAPHY

- [1] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, John Wiley & Sons, New York, 1958.
- [2] V. D. Blondel and J. N. Tsitsiklis, "A survey of computational complexity results in systems and control," *Automatica*, 36:1249–1274, 2000.
- [3] G. Calafiore, F. Dabbene, and R. Tempo, "Randomized algorithms for probabilistic robustness with real and complex structured uncertainty," *IEEE Trans. Aut. Control*, 45(12):2218–2235, December 2000.
- [4] A. Edelman, *Eigenvalues and Condition Numbers of Random Matrices*, Ph.D. thesis, Massachusetts Institute of Technology, Boston, 1989.
- [5] A. Edelman, "How many eigenvalues of a random matrix are real?", *J. Amer. Math. Soc.*, 7:247–267, 1994.
- [6] V. L. Girko, *Theory of Random Determinants*, Kluwer, Boston, 1990.
- [7] M. L. Mehta, *Random Matrices*, Academic Press, Boston, 1991.
- [8] E. P. Wigner, "Distribution laws for the roots of a random Hermitian matrix," In C. E. Porter, ed., *Statistical Theories of Spectra: Fluctuations*. Academic, New York, 1965.
- [9] E. P. Wigner, "Statistical properties of real symmetric matrices with many dimensions," In C. E. Porter, ed., *Statistical Theories of Spectra: Fluctuations*. Academic, New York, 1965.

Problem 2.3

Aspects of Fisher geometry for stochastic linear systems

Bernard Hanzon

Institute for Econometrics, Operations Research and System Theory
Technical University of Vienna
Argentinierstr. 8/119
A-1040 Vienna
Austria
`bhanzon@eos.tuwien.ac.at`

Ralf Peeters

Department of Mathematics
Maastricht University
P.O. Box 616, 6200 MD Maastricht
The Netherlands
`ralf.peeters@math.unimaas.nl`

1 DESCRIPTION OF THE PROBLEM

Consider the space S of stable minimum phase systems in discrete-time, of order (McMillan degree) n , having m inputs and m outputs, driven by a stationary Gaussian white noise (innovations) process of zero mean and covariance Ω . This space is often considered, for instance in system identification, to characterize stochastic processes by means of linear time-invariant dynamical systems (see [8, 18]). The space S is well known to exhibit a differentiable manifold structure (cf. [5]), which can be endowed with a notion of distance between systems, for instance by means of a Riemannian metric, in various meaningful ways.

One particular Riemannian metric of interest on S is provided by the so-called *Fisher metric*. Here the Riemannian metric tensor is defined in terms of local coordinates (i.e., in terms of an actual parametrization at hand) by the Fisher information matrix associated with a given system. The open question raised in this paper reads as follows:

Does there exist a uniform upper bound on the distance induced by the Fisher metric for a fixed $\Omega > 0$, between any two systems in S ?

In case the answer is affirmative, a natural follow-up question from the differential geometric point of view would be whether it is possible to construct a finite atlas of charts for the manifold S , such that the charts as subsets of Euclidean space are bounded (i.e., contained in an open ball in Euclidean space), while the distortion of each chart remains finite.

2 MOTIVATION AND BACKGROUND OF THE PROBLEM

An important and well-studied problem in linear systems identification is the construction of parametrizations for various classes of linear systems. In the literature a great number of parametrizations for linear systems have been proposed and used. From the geometric point of view the question arises whether one can qualify various parametrizations as good or bad. A parametrization is a way to (locally) describe a geometric object. Intuitively, a parametrization is better the more it reflects the (local) structure of the geometric object. An important consideration in this respect is the *scale* of the parametrization, or rather the *spectrum of scales*, see [4]. To explain this, consider the tangent space of a differential manifold of systems, such as S . The differentiable manifold can be supplied with a Riemannian geometry, for example, by smoothly embedding the differentiable manifold in an appropriate Hilbert space: then the tangent spaces to the manifold are linear subspaces of the Hilbert space, which induces an inner product on each of the tangent spaces and a Riemannian metric structure on the manifold. If such a Riemannian metric is defined, then any sufficiently smooth parametrization will have an associated Riemannian metric tensor. In local coordinates (i.e., in terms of the parameters used) it is represented by a symmetric, positive definite matrix at each point. The eigenvalues of this matrix reflect the local scales of the parametrization: the scale of any infinitesimal movement starting from a given point, will vary between the largest and the smallest eigenvalue of the Riemannian metric tensor at the point involved. Over a set of points the scale will clearly vary between the largest eigenvalue to be found in the spectra of the corresponding set of Riemannian metric matrices and the smallest eigenvalue to be found in that same set of spectra. Following Milnor (see [12]), who considered the question of finding good charts for the earth, we define the distortion of a parametrization, which we will call the *Milnor distortion*, as the quotient of the largest scale and the smallest scale of the parametrization.

Note that this concept of Milnor distortion is a generalization of the concept of the condition number of a matrix. However it is (in general) *not* the maximum of the condition numbers of the set of Riemannian metric matrices.

Indeed, the largest eigenvalue and the smallest eigenvalue that enter into the definition of the Milnor distortion do not have to correspond to the Riemannian metric tensor at the same point.

If one has an atlas of overlapping charts, one can calculate the Milnor distortion in each of the charts and consider the largest distortion in any of the charts of the atlas. One could now be tempted to define this number as the distortion of the atlas and look for atlases with relatively small distortion. However, in this case, the problem shows up that it is always possible to take a large number of small charts, each one displaying very little distortion (i.e., distortion close to one), while such an atlas may still not be desirable as it may require a huge number of charts. The difficulty here is to trade off the number of charts in an atlas against the Milnor distortion in each of those charts. At this point, we have no clear natural candidate for this trade-off. But at least for atlases with an equal *finite* number of charts the concept of maximal Milnor distortion could be used to compare the atlases.

3 AVAILABLE RESULTS

In trying to apply these ideas to the question of parametrization of linear systems, the problem arises that many parametrizations turn out to have in fact an infinite Milnor distortion. Consider for example the case of real SISO discrete-time strictly proper stable systems of order one. (See also [9] and [13, section 4.7].) This set can be described by two real parameters, e.g., by writing the associated transfer function into the form $h(z) = b/(z - a)$. Here, the parameter a denotes the pole of the system and the parameter b is associated with the gain. The Riemannian metric tensor induced by the H_2 norm of this parametrization can be computed as $\begin{pmatrix} b^2(1+a^2)/(1-a^2)^3 & ab/(1-a^2)^2 \\ ab/(1-a^2)^2 & 1/(1-a^2) \end{pmatrix}$, see [9]. Therefore it tends to infinity when a approaches the stability boundary $|a| = 1$, whence the Milnor distortion of this parametrization becomes infinity. In this example the geometry is that of a flat double infinite-sheeted Riemann surface. Locally, it is isometric with Euclidean space and therefore one can construct charts that have the identity matrix as their Riemannian metric tensor (see [13]). However, in this case, this means that close to the stability boundary the distances between points become arbitrarily large. Therefore, although it is possible to construct charts with optimal Milnor distortion, this can only be done at the price of having to work with infinitely large (i.e., unbounded) charts. If one wants to work with charts in which the distances remain bounded then one will need infinitely many of them on such occasions.

In the case of stochastic Gaussian time-invariant linear dynamical systems without observed inputs, the class of stable minimum-phase systems plays an important role. For such stochastic systems the (asymptotic) Fisher information matrix is well-defined. This matrix is dependent on the parametrization

used and admits the interpretation of a Riemannian metric tensor (see [15]). There is an extensive literature on the computation of Fisher information, especially for AR and ARMA systems. See, e.g., [6, 7, 11]. Much of this interest derives from the many applications in practical settings: it can be used to establish local parameter identifiability, it is used for parameter estimation in the method of scoring, and it is also known to determine the local convergence properties of the popular Gauss-Newton method for least-squares identification of linear systems based on the maximum likelihood principle (see [10]).

In the case of stable AR systems, the Fisher metric tensor can, for instance, be calculated using the parametrization with Schur parameters. From the formulas in [14] it follows that the Fisher information for scalar AR systems of order one driven by zero mean Gaussian white noise of unit variance is equal to $1/(1 - \gamma_1^2)$. Here γ_1 is required to range between -1 and 1 (to impose stability) and to be nonzero (to impose minimality). Although this again implies an infinite Milnor distortion, the situation here is structurally different from the situation in the previous case: the length of the curve of systems obtained by letting γ_1 range from 0 to 1 is finite! Indeed, the (Fisher) length of this curve is computed as $\int_0^1 \frac{1}{\sqrt{1-\gamma_1^2}} d\gamma_1 = \pi/2$.

Let the *inner geometry* of a connected Riemannian manifold of systems be defined by the shortest path distance: $d(\Sigma_1, \Sigma_2)$ is the Riemannian length of the shortest curve connecting the two systems Σ_1 and Σ_2 . Then, in this simple case, the Fisher geometry has the property that the corresponding inner geometry has a uniform upper bound. Therefore, this example provides an instance of a *subset* of the manifold S for which the answer to the question raised is affirmative.

As a matter of fact, if one now reparametrizes the set of systems as in [17] by θ defined through $\gamma_1 = \sin(\theta)$, then the resulting Fisher information quantity becomes equal to 1 everywhere. Thus, it is bounded and the Milnor distortion of this reparametrization is finite. But at the same time the parameter chart itself remains bounded! Hence, also the “follow-up question” of the previous section is answered affirmative here.

If one considers SISO stable minimum-phase systems of order 1 , it can be shown likewise that also here the Fisher distance between two systems is uniformly bounded and that a finite atlas with bounded charts and finite Milnor distortion can be designed. Whether this also occurs for larger state-space dimensions is still unknown (to the best of the authors’ knowledge) and this is precisely the open problem presented above.

To conclude, we note that the role played by the covariance matrix Ω of the driving white noise is rather limited. It is well known that if the system equations and the covariance matrix are parametrized independently of each other, then the Fisher information matrix attains a block-diagonal structure (see, e.g., [18, Ch. 7]). The covariance matrix Ω then appears as a weighting matrix for the block of the Fisher information matrix associated with the

parameters involved in the system equations. Therefore, if Ω is known, or rather if an upper bound on Ω is known (which is likely to be the case in any practical situation!), its role with respect to the questions raised can be largely disregarded. This allows to restrict attention to the situation where Ω is fixed to the identity matrix I_m .

BIBLIOGRAPHY

- [1] S.-I. Amari, *Differential-Geometrical Methods in Statistics*, Lecture Notes in Statistics 28, Springer Verlag, Berlin, 1985.
- [2] S.-I. Amari, "Differential geometry of a parametric family of invertible linear systems Riemannian metric, dual affine connections, and divergence," *Mathematical Systems Theory*, 20, 53–82, 1987
- [3] C. Atkinson and A. F. S. Mitchell, "Rao's distance measure. *Sankhyā: The Indian Journal of Statistics*, Series A, 43(3), 345–365, 1981.
- [4] R. W. Brockett and P. S. Krishnaprasad "A scaling theory for linear systems," *IEEE Trans. Aut. Contr.*, AC-25, 197–206, 1980.
- [5] J. M. C. Clark, "The consistent selection of parametrizations in system identification," *Proc. Joint Automatic Control Conference*, 576–580. Purdue University, Lafayette, Indiana, 1976.
- [6] B. Friedlander, "On the Computation of the Cramer-Rao Bound for ARMA Parameter Estimation," *IEEE Transactions on Acoustics, Speech and Signal Processing*, ASSP-32 (4), 721–727.
- [7] E. J. Godolphin and J. M. Unwin, "Evaluation of the covariance matrix for the maximum likelihood estimator of a Gaussian autoregressive-moving average process," *Biometrika*, 70 (1), 279–284, 1983.
- [8] E. J. Hannan and M. Deistler, *The Statistical Theory of Linear Systems*. John Wiley & Sons, New York, 1988.
- [9] B. Hanzon, *Identifiability, Recursive Identification and Spaces of Linear Dynamical Systems*, CWI Tracts 63 and 64, Centrum voor Wiskunde en Informatica (CWI), Amsterdam, 1989.
- [10] B. Hanzon and R. L. M. Peeters, "On the Riemannian Interpretation of the Gauss-Newton Algorithm," In: M. Kárný and K. Warwick (eds.), *Mutual Impact of Computing Power and Control Theory*, 111–121. Plenum Press, New York, 1993.
- [11] A. Klein and G. Mélard, "On Algorithms for Computing the Covariance Matrix of Estimates in Autoregressive Moving Average Processes," *Computational Statistics Quarterly*, 1, 1–9, 1989.

- [12] J. Milnor, "A problem in cartography," *American Math. Monthly*, 76, 1101–1112, 1969.
- [13] R. L. M. Peeters, *System Identification Based on Riemannian Geometry: Theory and Algorithms*. Tinbergen Institute Research Series, vol. 64, Thesis Publishers, Amsterdam, 1994.
- [14] R. L. M. Peeters and B. Hanzon, "Symbolic computation of Fisher information matrices for parametrized state-space systems," *Automatica*, 35, 1059–1071, 1999.
- [15] C. R. Rao, "Information and accuracy attainable in the estimation of statistical parameters," *Bull. Calcutta Math. Soc.*, 37, 81–91, 1945.
- [16] N. Ravishanker, E. L. Melnick and C.-L. Tsai, "Differential geometry of ARMA models," *Journal of Time Series Analysis*, 11, 259–274, .
- [17] A. L. Rijkeboer, "Fisher optimal approximation of an AR(n)-process by an AR(n-1)-process," In: J. W. Nieuwenhuis, C. Praagman and H. L. Trentelman eds., *Proceedings of the 2nd European Control Conference ECC '93*, 1225–1229, Groningen, 1993.
- [18] T. Söderström and P. Stoica, *System Identification*, Prentice-Hall, New York, 1989.

Problem 2.4

On the convergence of normal forms for analytic control systems

Wei Kang

Department of Mathematics,
Naval Postgraduate School
Monterey, CA 93943
USA
wkang@nps.navy.mil

Arthur J. Krener

Department of Mathematics,
University of California
Davis, CA 95616
USA
ajkrener@ucdavis.edu

1 BACKGROUND

A fruitful technique for the local analysis of a dynamical system consists of using a local change of coordinates to transform the system to a simpler form, which is called a normal form. The qualitative behavior of the original system is equivalent to that of its normal form which may be easier to analyze. A bifurcation of a parameterized dynamics occurs when a change in the parameter leads to a change in its qualitative properties. Therefore, normal forms are useful in analyzing when and how a bifurcation will occur. In his dissertation, Poincaré studied the problem of linearizing a dynamical system around an equilibrium point, linear dynamics being the simplest normal form. Poincaré's idea is to simplify the linear part of a system first, using a linear change of coordinates. Then the quadratic terms in the system are simplified, using a quadratic change of coordinates, then the cubic terms, and so on. For systems that are not linearizable, the Poincaré-Dulac theorem provides the normal form.

Given a C^∞ dynamical system in its Taylor expansion around $x = 0$,

$$\dot{x} = f(x) = Fx + f^{[2]}(x) + f^{[3]}(x) + \cdots \quad (1)$$

where $x \in \mathbb{R}^n$, F is a diagonal matrix with eigenvalues $\lambda = (\lambda_1, \dots, \lambda_n)$, and $f^{[d]}(x)$ is a vector field of homogeneous polynomial of degree d . The dots $+\dots$ represent the rest of the formal power series expansion of f . Let \mathbf{e}_k be the k -th unit vector in \mathbb{R}^n . Let $m = (m_1, \dots, m_n)$ be a vector of non-negative integers. In the following, we define $|x|$ and x^m by $|m| = \sum |m_i|$ and $x^m = x_1^{m_1} x_2^{m_2} \dots x_n^{m_n}$. A nonlinear term $x^m \mathbf{e}_k$ is said to be resonant if $m \cdot \lambda = \lambda_k$ for some nonzero vector of non-negative integers m and some $1 \leq k \leq n$.

Definition 1 The eigenvalues of F are in the Poincaré domain if their convex hull does not contain zero, otherwise they are in the Siegel domain.

Definition 2: The eigenvalues of F are of type (C, ν) for some $C > 0, \nu > 0$ if

$$|m \cdot \lambda - \lambda_k| \geq \frac{C}{|m|^\nu}$$

For eigenvalues in the Poincaré domain, there are at most a finite number of resonances. For eigenvalues of type (C, ν) , there are no resonances and as $|m| \rightarrow \infty$ the rate at which resonances are approached is controlled.

A formal change of coordinates is a formal power series

$$z = Tx + \theta^{[2]}(x) + \theta^{[3]}(x) + \dots \quad (2)$$

where T is invertible. If $T = I$, then it is called a near identity change of coordinates. If the power series converges locally, then it defines a real analytic change of coordinates.

Theorem 1: (Poincaré-Dulac) *If the system (1) is C^∞ then there exists a formal change of coordinates (2) transforming it to*

$$\dot{z} = Az + w(z)$$

where A is in Jordan form and $w(z)$ consists solely of resonant terms. (If some of the eigenvalues of F are complex then the change of coordinates will also be complex). In this normal form, $w(z)$ need not be unique.

If the system (1) is real analytic and its eigenvalues lie in the Poincaré domain (2), then $w(z)$ is a polynomial vector field and the change of coordinates (2) is real analytic.

Theorem 2: (Siegel) *If the system (1) is real analytic and its eigenvalues are of type (C, ν) for some $C > 0, \nu > 0$, then $w(z) = 0$ and the change of coordinates (2) is real analytic.*

As is pointed out in [1], even in cases where the formal series are divergent, the method of normal forms turns out to be a powerful device in the study of nonlinear dynamical systems. A few low degree terms in the normal form often give significant information on the local behavior of the dynamics.

2 THE OPEN PROBLEM

In [3], [4], [5], [10], and [8], Poincaré's idea is applied to nonlinear control systems. A normal form is derived for nonlinear control systems under change of state coordinates and invertible state feedback. Consider a C^∞ control system

$$\dot{x} = f(x, u) = Fx + Gu + f^{[2]}(x, u) + f^{[3]}(x, u) + \cdots \quad (3)$$

where $x \in \mathbb{R}^n$ is the state variable, $u \in \mathbb{R}$ is a control input. We only discuss scalar input systems, but the problem can be generalized to vector input systems. Such a system is called *linearly controllable* at the origin if the linearization (F, G) is controllable.

In contrast with Poincaré's theory, a homogeneous transformation for (3) consists of both change of coordinates and invertible state feedback,

$$z = x + \theta^{[d]}(x), \quad v = u + \kappa^{[d]}(x, u) \quad (4)$$

where $\theta^{[d]}(x)$ represents a vector field whose components are homogeneous polynomials of degree d . Similarly, $\kappa^{[d]}(x)$ is a polynomial of degree d . A formal transformation is defined by

$$z = Tx + \sum_{d=2}^{\infty} \theta^{[d]}(x), \quad v = Ku + \sum_{d=2}^{\infty} \kappa^{[d]}(x, u) \quad (5)$$

where T and K are invertible. If T and K are identity matrices then this is called a near identity transformation.

The following theorem for the normal form of control systems is a slight generalization of that proved in [3], see also [8] and [10].

Theorem 3: *Suppose (F, G) in (3) is a controllable pair. Under a suitable transformation (5), (3) can be transformed into the following normal form*

$$\begin{aligned} \dot{z}_i &= z_{i+1} + \sum_{j=i+2}^{n+1} p_{i,j}(\bar{z}_j) z_j^2 & 1 \leq i \leq n-1 \\ \dot{z}_n &= v \end{aligned} \quad (6)$$

where $z_{n+1} = v$, $\bar{z}_j = (z_1, z_2, \dots, z_j)$, and $p_{i,j}(\bar{z}_j)$ is a formal series of \bar{z}_j .

Once again, the convergence of the formal series $p_{i,j}$ in (6) is not guaranteed, hence nothing is known about the convergence of the normal form.

Open Problem (The Convergence of Normal Form): *Suppose the controlled vector field $f(x, u)$ in (3) is real analytic and F, G is a controllable pair. Find verifiable necessary and sufficient conditions for the existence of a real analytic transformation (5) that transforms the system to the normal form (6).*

Normal forms of control systems have proven to be a powerful tool in the analysis of local bifurcations and local qualitative performance of control systems. A convergent normal form will make it possible to study a control system over the entire region in which the normal form converges. Global or semi-global results on control systems and feedback design can be proved by studying analytic normal forms.

3 RELATED RESULTS

The convergence of the Poincaré normal form was an active research topic in dynamical systems. According to Poincaré's Theorem and Siegel's theorem, the location of eigenvalues determines the convergence. If the eigenvalues are located in the Poincaré domain with no resonances, or if the eigenvalues are located in the Siegel domain and are of type (C, ν) , then the analytic vector field that defines the system is biholomorphically equivalent to a linear vector field. Clearly, the normal form converges because it has only a linear part. The Poincaré-Dulac theorem deals with a more complicated case. It states that if the eigenvalues of an analytic vector field belong to the Poincaré domain, then the field is biholomorphically equivalent to a polynomial vector field. Therefore, the Poincaré normal form has only finite many terms, and hence is convergent.

As for control systems, it is proved in [5] that if an analytic control system is linearizable by a formal transformation, then it is linearizable by an analytic transformation. It is also proved in [5] that a class of three-dimensional analytic control systems, which are not necessarily linearizable, can be transformed to their normal forms by analytic transformations. No other results on the convergence of control system normal forms are known to us.

The convergence problem for control systems has a fundamental difference from the convergence results of Poincaré-Dulac. For the latter, the location of the eigenvalues are crucial and the eigenvalues are invariant under change of coordinates. However, the eigenvalues of a control system can be changed by linear state feedback. It is unknown what intrinsic factor in control systems determines the convergence of their normal form or if the normal form is always convergent.

The convergence of normal forms is an important problem to be addressed. Applications of normal forms for control systems are proved to be successful. In [6] the normal forms are used to classify the bifurcation of equilibrium sets and controllability for uncontrollable systems. In [7] the control of bifurcations using state feedback is introduced based on normal forms. For discrete-time systems, normal form and the stabilization of Naimark-Sacker bifurcation are addressed in [2]. In [10] a complete characterization for the symmetry of nonlinear systems is found for linearly controllable systems.

In addition to linearly controllable systems, the normal form theory has been generalized to larger family of control systems. Normal forms for systems with uncontrollable linearization are derived in several papers ([6], [7], [8], and [10]). Normal forms of discrete-time systems can be found in [9] and [2]. The convergence of these normal forms is also an open problem.

BIBLIOGRAPHY

- [1] V. I. Arnold, *Geometrical Method in the Theory of Ordinary Differential Equations*, Springer-Verlag, New York, Berlin, 1988.
- [2] B. Hamzi, J.-P. Barbot, S. Monaco, and D. Normand-Cyrot, "Nonlinear discrete-time control of systems with a Naimark-Sacker bifurcation," *Systems and Control Letters*, 44, 4, pp. 245-258, 2001.
- [3] W. Kang, *Extended controller normal form, invariants and dynamical feedback linearization of nonlinear control systems*, Ph.D. dissertation, University of California at Davis, 1991.
- [4] W. Kang and A. J. Krener, "Extended quadratic controller normal form and dynamic feedback linearization of nonlinear systems," *SIAM J. Control and Optimization*, 30 (1992), 1319-1337.
- [5] W. Kang, "Extended controller form and invariants of nonlinear control systems with a single input," *J. of Mathematical Systems, Estimation and Control*, 6 (1996), 27-51.
- [6] W. Kang, "Bifurcation and normal form of nonlinear control systems - Part I and II, *SIAM J. Control and Optimization*, 36 (1998), 193-212 and 213-232.
- [7] W. Kang, "Bifurcation control via state feedback for systems with a single uncontrollable mode," *SIAM J. Control and Optimization*, 38, (2000), 1428-1452.
- [8] A. J. Krener, W. Kang, and D. E. Chang, "Control bifurcations," *IEEE Transactions on Automatic Control*, forthcoming.
- [9] A. J. Krener and L. Li, "Normal forms and bifurcations of discrete time nonlinear control systems," *Proc. of 5th IFAC NOLCOS Symposium*, Saint-Petersburg, 2001.
- [10] I. Tall and W. Respondek, "Feedback classification of nonlinear single-input controls systems with controllable linearization: Normal forms, canonical forms, and invariants," preprint.

PART 3

Nonlinear Systems

Problem 3.1

Minimum time control of the Kepler equation

Jean-Baptiste Caillau, Joseph Gergaud,
and Joseph Noailles

ENSEEIH-IRIT, UMR CNRS 5505

2 rue Camichel

F-31071 Toulouse

France

{caillau, gergaud, jnoaille}@enseeiht.fr

1 DESCRIPTION OF THE PROBLEM

We consider the controlled Kepler equation in three dimensions

$$\ddot{r} = -k \frac{r}{|r|^3} + \gamma \quad (1)$$

where $r = (r_1, r_2, r_3)$ is the position vector—the double dot denoting the second order time derivative—, k a strictly positive constant, $|\cdot|$ the Euclidean norm in \mathbf{R}^3 , and where $\gamma = (\gamma_1, \gamma_2, \gamma_3)$ is the control. The minimum time problem is then stated as follows: find a positive time T and a measurable function γ defined on $[0, T]$ such that (1) holds almost everywhere on $[0, T]$ and:

$$\begin{aligned} T &\rightarrow \min \\ r(0) &= r^0, \quad \dot{r}(0) = \dot{r}^0 \end{aligned} \quad (2)$$

$$h(r(T), \dot{r}(T)) = 0 \quad (3)$$

$$|\gamma| \leq \Gamma. \quad (4)$$

In (2), r^0 and \dot{r}^0 are the known initial position and speed with:

$$\frac{|\dot{r}^0|^2}{2} - \frac{k}{|r^0|} < 0$$

in order that the uncontrolled initial motion be periodic [1]. In (3) h is a fixed submersion of \mathbf{R}^6 onto \mathbf{R}^l , $l \leq 6$, defining a non-trivial endpoint condition. The constraint (4) on the Euclidean norm of the control, with Γ a strictly positive constant, means that almost everywhere on $[0, T]$

$$\gamma_1^2 + \gamma_2^2 + \gamma_3^2 \leq \Gamma^2.$$

Our first concern is uniqueness (see §3 about existence):

Question 1. *Is the optimal control unique?*

The second point is about regularity, namely:

Question 2. *Are there continuous optimal controls?*

Denoting by $T(\Gamma)$ the value function that assigns to any strictly positive Γ (parameter involved in (4)) the associated minimum time, our third and last question is:

Question 3. *Does the product $T(\Gamma) \cdot \Gamma$ have a limit when Γ tends to zero?*

2 MOTIVATION

This problem originates in the computation of optimal orbit transfers in celestial mechanics for satellites with very low thrust engines [5]. Since the 1990s, low electro-ionic propulsion is been considered as an alternative to strong chemical propulsion, but the lower the thrust, the longer the transfer time, hence the idea of minimizing the final time. In this context, γ is the ratio u/m of the engine thrust by the mass of the satellite, and one has moreover to take into account the mass variation due to fuel consumption:

$$\dot{m} = -\beta|u|.$$

Typical boundary conditions in this case consist in inserting the spacecraft on a high geostationary orbit, and the terminal condition is defined by:

$$|r(T)| \text{ and } |\dot{r}(T)| \text{ fixed, } r(T) \cdot \dot{r}(T) = 0, \quad r(T) \times \dot{r}(T) \times \vec{k} = 0$$

where \vec{k} is the normal vector to the equatorial plane.

In contrast with the impulsional manoeuvres performed using the strong classic chemical propulsion, the gradual control by a low thrust engine is sometimes referred to as “continuous?” Thus, question 2 could be rephrased according to:

Are “continuous” optimal controls continuous?

Besides, this question is also relevant in practice since continuity of controls is the basic assumption required by most numerical methods [2]. In the same respect, question 3 is the key to get accurate estimates of the unknown transfer time, needed to ensure convergence of the numerical computation.

3 RELATED RESULTS

The existence of controls achieving the minimum time transfer comes from the controllability of the system (the associated Lie algebra has maximal rank and the drift is periodic, see [7]) and from the convexity properties of the dynamics by Filippov theorem [4]. Regarding regularity, it is proven in [3] (whose results extend straightforwardly to three dimensions) that any time minimal control of (1) has at most finitely many discontinuity points. More precisely, using Pontryagin Maximum Principle [4, 6], one gets that any discontinuity point \bar{t} is a switching point in the sense that the control is instantaneously rotated of an angle π :

$$\gamma(\bar{t}+) = -\gamma(\bar{t}-).$$

Furthermore, bounds are given in [3], not on the total number of switchings but for those located at special points of the osculating ellipse: there cannot be consecutive switchings at perigee or apogee. Since the numerics suggest that the possible discontinuities are exactly located at the perigee, a conjecture would be:

There is at most one switching point and this point is located at the perigee.

Finally, as for question 3, the value function $T(\Gamma)$ is obviously decreasing and is proven to be right-continuous in [2]. Besides, the product $T(\Gamma) \cdot \Gamma$ turns to be nearly constant numerically so that the conjecture would be to answer positively:

There is a positive constant c such that $T(\Gamma) \cdot \Gamma$ tends to c when Γ tends to zero.

BIBLIOGRAPHY

- [1] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, Springer-Verlag, 1978.
- [2] J. B. Caillau, J. Gergaud and J. Noailles, “3D geosynchronous transfer of a satellite: Continuation on the thrust,” *JOTA*, 118:3, 2003 (Forthcoming).
- [3] J.B. Caillau and J. Noailles, “Coplanar control of a satellite around the earth,” *ESAIM COCV*, vol. 6, pp. 239-258, 2001.
- [4] L. Cesari, *Optimization Theory and Applications*, Springer-Verlag, 1983.
- [5] J. Noailles and C. T. Le, “Contrôle en temps minimal et transfert à faible poussée,” *Équations aux dérivées partielles et applications*, Articles in Honour of J. L. Lions for his 70th Birthday, pp. 705-724, Gauthiers-Villars, 1998.

- [6] H. J. Sussmann, "Geometry and optimal control," *Mathematical Control Theory*, Dedicated to R. W. Brockett on his 60th Birthday, J. Baillieul and J. C. Willems eds., Springer-Verlag, 1998.
- [7] V. Jurdjevic, *Geometric Control Theory* Cambridge University Press, 1997.

Problem 3.2

Linearization of linearly controllable systems

R. Devanathan

School of Electrical and Electronic Engineering (Block S1)
Nanyang Technological University
Singapore 639798
Republic of Singapore
edevan@ntu.edu.sg

1 DESCRIPTION OF THE PROBLEM

We consider a class of systems of the form

$$\dot{\xi} = f(\xi) + g(\xi)\zeta \quad (1)$$

where ξ is an n -tuple vector and $f(\xi)$ and $g(\xi)$ are vector fields, i.e., n -tuple vectors whose elements are, in general, functions of ξ . For simplicity, we assume a scalar input ζ . We require that the system (1) be linearly controllable [1], i.e., the pair (F, G) is controllable where $F = \frac{\partial f}{\partial \xi}(0)$ and $G = g(0)$ at the assumed equilibrium point at the origin.

The power series expansion of (1) about the origin can be written, with an appropriate change of variable and input, as

$$\dot{x} = Fx + G\phi + O_1(x)^{(2)} + \gamma_1(x, \phi)^{(1)} \quad (2)$$

where, without loss of generality, F and G can be in Brunovsky form [2], superscript (2) corresponds to terms in x of degree greater than one, superscript (1) corresponds to terms in x of degree greater than or equal to one and x and ϕ are the transformed state and input variables respectively. We introduce state feedback as in [3]

$$\phi = -Kx + u$$

where

$$K = [k_n, k_{n-1}, \dots, k_2, k_1]^t$$

Equation (2) then becomes

$$\dot{x} = Ax + Gu + O(x)^{(2)} + \gamma(x, u)^{(1)} \quad (3)$$

where

$$A = F - GK \quad (4)$$

We can choose the eigenvalues of matrix A in (4), without loss of generality, to be real, distinct, and nonresonant by a proper choice of the matrix K [3]. The nonresonant property, meaning that no integral relation exists among the eigenvalues of matrix A , ensures that (3) can be linearized up to an arbitrary order.

Put (3) into the form

$$\dot{x} = Ax + Gu + f_2(x) + f_3(x) + \cdots + g_1(x)u + g_2(x)u + \cdots \quad (5)$$

where $f_m(x)$ and $g_{m-1}(x)$ correspond to vector-valued polynomials containing terms of the form

$$x^m = x_1^{m_1} x_2^{m_2} \cdots x_n^{m_n}, \quad m_i \in (0, 1, 2, \dots, n), \quad i = 1, 2, \dots, n;$$

$$\sum_{i=1}^n m_i = m, \quad m \geq 2.$$

Consider a near identity (normalizing) transformation as in

$$x = y + h(y) \quad (6)$$

and a change of input as in

$$v = (1 + \beta(x))u + \alpha(x), \quad 1 + \beta(x) \neq 0 \quad (7)$$

where

$$h(y) = h_2(y) + h_3(y) + \cdots \quad (8)$$

$$\alpha(x) = \alpha_2(x) + \alpha_3(x) + \cdots \quad (9)$$

$$\beta(x) = \beta_1(x) + \beta_2(x) + \cdots \quad (10)$$

The problem is to find a solution for $h_m(\cdot)$, $\alpha_m(\cdot)$ and $\beta_{m-1}(\cdot)$, $m \geq 2$ such that the nonlinear terms upto an arbitrary order, viz., " $f_m(\cdot)$ " and " $g_{m-1}(\cdot)u$ " can be removed from (5) by the application of the transformations (6) and (7) to it.

2 MOTIVATION AND HISTORY OF THE PROBLEM

Linearization of a nonlinear dynamic system of the form (1), but without the control input, was originally investigated by Poincaré [4] [5]. It was shown that, around an equilibrium point, a near identity (normalizing) transformation takes it to its normal form where only the residual nonlinearities, that cannot be removed by the transformation, remain. The dynamic system is said to be resonant in the order of these residual nonlinearities.

The solution for the normalizing transformation is in the form of an infinite series as in (8) whose convergence has been proved under certain assumptions [6] [7]. Irrespective of the convergence of the infinite series, the transformation is of interest because, one can remove up to an arbitrary order of nonlinearities (as long as they are nonresonant) through such a transformation, thus providing an approximate linearization of the dynamic system.

3 AVAILABLE RESULTS

Our problem is an analog of Poincaré's problem with the control input provided. Krener et al. [8] have considered a nonlinear system of the form (1) and showed that a generalized form of the homological equation can be formulated in this case. Devanathan [3] has shown that, for a linearly controllable system, the system matrix can be made nonresonant through an appropriate choice of state feedback. This concept is further exploited in [9] to find a solution to the second order linearization. An analogous solution to the case of an arbitrary order linearization, however, is still open.

By application of (6) and (7), one can write

$$f_2(x) + f_3(x) + f_4(x) + \cdots = f'_2(y) + f'_3(y) + f'_4(y) + \cdots \quad (11)$$

$$g_1(x)u + g_2(x)u + g_3(x)u + \cdots = g'_1(y)u + g'_2(y)u + g'_3(y)u + \cdots \quad (12)$$

$$\alpha(x) = \alpha'_2(y) + \alpha'_3(y) + \alpha'_4(y) + \cdots \quad (13)$$

$$\beta(x) = \beta'_1(y) + \beta'_2(y) + \beta'_3(y) + \cdots \quad (14)$$

for some appropriate polynomials $f'_m(\cdot), g'_{m-1}(\cdot), \alpha'_m(\cdot)$ and $\beta'_{m-1}(\cdot), m \geq 2$. Substituting (6) and (7) into (5) and using (11)-(14), consider the polynomials of the form y^m and $y^{m-1}u, m = 2, 3, \dots$. Then the terms " $f_m(x)$ " and " $g_{m-1}(x)u$ " can be removed from (5) progressively, $m = 2, 3$, etc. provided the following generalized homological equations are satisfied [3].

$$\frac{\partial h_m(y)}{\partial y}(Ay) - Ah_m(y) + G\alpha'_m(y) = f''_m(y), m \geq 2 \quad (15)$$

$$\frac{\partial h_m(y)}{\partial y}(Gu) + G\beta'_{m-1}(y)u = g''_{m-1}(y)u, \forall u, m \geq 2 \quad (16)$$

where $f''_2(y) = f'_2(y) = f_2(y)$ and $f''_m(y)$ is expressed in terms of $f'_{m-i}(y), i = 0, 1, 2, \dots, (m-2)$ and $h_{m-j}(y), j = 1, 2, \dots, (m-2), m > 2$. Also, $g''_1(y) = g'_1(y) = g_1(y)$ and $g''_m(y)$ is expressed in terms of $g'_{m-i}(y), i = 0, 1, 2, \dots, (m-1)$ and $h_{m-j}(y), j = 0, 1, 2, \dots, (m-2), m \geq 2$.

Assuming that $h_{m-j}(y), \alpha_{m-j}(y), \beta_{m-j-1}(y), j = 1, 2, \dots, (m-2), m > 2$, are known, $f''_m(y)$ and $g''_{m-1}(y)$ can be computed. Without loss of generality, one can assume matrix A to be diagonal and $G = [1, 1, \dots, 1, 1]^t$ by applying a change of coordinate to (5) involving Vandermonde matrix [10]. One can then proceed to solve (15) for $h_m(y)$ in terms of $\alpha_m(y)$ and substitute the same into (16) to set up a linear system of equations in the unknown coefficients of polynomials $\alpha_m(y)$ and $\beta_{m-1}(y)$.

For $m = 2$, it has been shown in [9] that the corresponding system of linear equations can be reduced to a system of $\left(\frac{n(n-1)}{2}\right)$ equations in n variables whose rank is $(n-1)$. It is conjectured that a similar reduction of the linear system of equations, in the arbitrary order case, should also be possible.

Formulation of the properties and solution, if it exists, of the linear system of equations involving the coefficients of the polynomials $\alpha_m(\cdot), \beta_{m-1}(\cdot)$ and $h_m(\cdot), m > 2$ will constitute the solution to the open problem.

BIBLIOGRAPHY

- [1] A. J. Krener and W. Kang, "Extended normal forms of quadratic systems," *Proc. 29th Conf. Decision and Control*, pp. 2091-2095, 1990.
- [2] P. Brunovsky, "A Classification of linear controllable systems," *Kybernetika cislo*, vol. 3, pp. 173-188, 1970.
- [3] R. Devanathan, "Linearization condition through state feedback," *IEEE Transactions on Automatic Control*, vol. 46, no. 8, pp. 1257-1260, 2001.
- [4] V. I. Arnold, *Geometrical Methods in the Theory of Ordinary Differential Equations*, Springer-Verlag, New York, pp. 177-188, 1983.
- [5] J. Guckenheimer and P. Holmes, *Nonlinear Oscillations, Dynamical Systems and Bifurcation of Vector Fields*, Springer-Verlag, New York, 1983.
- [6] A. D. Bruno, *Local Methods in Nonlinear Differential Equations*, Springer-Verlag, Berlin, 1989.
- [7] Giampaolo Cicogna, "On the convergence of normalizing transformations in the presence of symmetries," Dipartimento di Fisica, Università di Pisa, P.zza Torricelli 2, I-56126, Pisa, Italy.
- [8] A. J. Krener, S. Karahan, and M. Hubbard, "Approximate normal forms of nonlinear systems," *Proc. 27th Conf. Decision and Control*, pp. 1223-1229, 1988.
- [9] R. Devanathan, "Solution of second order linearization," *Fifteenth International Symposium on Mathematical Theory of Networks and Systems*, University of Notre Dame, Notre Dame, IN, USA (2002).
- [10] B. C. Kuo, *Automatic Control Systems*, Fourth Ed., Prentice-Hall, pp. 227-240, 1982.

Problem 3.3

Bases for Lie algebras and a continuous CBH formula

Matthias Kawski¹

Department of Mathematics and Statistics
Arizona State University
Tempe, AZ 85287-1804
USA
kawski@asu.edu

1 DESCRIPTION OF THE PROBLEM

Many time-varying linear systems $\dot{x} = F(t, x)$ naturally split into time-invariant geometric components and time-dependent parameters. A special case are nonlinear control systems that are affine in the control u , and specified by analytic vector fields on a manifold M^n

$$\dot{x} = f_0(x) + \sum_{k=1}^m u_k f_k(x). \quad (1)$$

It is natural to search for solution formulas for $x(t) = x(t, u)$ that separate the time-dependent contributions of the controls u from the invariant, geometric role of the vector fields f_k . Ideally, one may be able to a priori obtain closed-form expressions for the flows of certain vector fields. The quadratures of the control might be done in real-time, or the integrals of the controls may be considered new variables for theoretical purposes such as path-planning or tracking.

To make this scheme work, one needs *simple* formulas for assembling these pieces to obtain the solution curve $x(t, u)$. Such formulas are nontrivial since in general the vector fields f_k do not commute: already in the case of linear systems, $\exp(sA) \cdot \exp(tB) \neq \exp(sA + tB)$ (for matrices A and B). Thus the desired formulas not only involve the flows of the system vector fields f_i but also the flows of their iterated commutators $[f_i, f_j]$, $[[f_i, f_j], f_k]$, and so on.

Using Hall-Viennot bases \mathcal{H} for the free Lie algebra generated by m indeterminates X_1, \dots, X_m , Sussmann [22] gave a general formula as a directed

¹Supported in part by NSF-grant DMS 00-72369

infinite product of exponentials

$$x(T, u) = \prod_{H \in \mathcal{H}}^{\rightarrow} \exp(\xi_H(T, u) \cdot f_H). \quad (2)$$

Here the vector field f_H is the image of the formal bracket H under the canonical Lie algebra homomorphism that maps X_i to f_i . Using the chronological product $(U * V)(t) = \int_0^T U(s)V'(s)ds$, the iterated integrals ξ_H are defined recursively by $\xi_{X_k}(T, u) = \int_0^T u_k(t)dt$ and

$$\xi_{HK} = \xi_H * \xi_K \quad (3)$$

if H, K, HK are Hall words and the left factor of K is not equal to H [9, 22]. (In the case of repeated left factors, the formula contains an additional factorial.)

An alternative to such infinite exponential product (in Lie group language, “coordinates of the 2nd kind”) is a single exponential of an infinite Lie series (“coordinates of the 1st kind”).

$$x(T, u) = \exp\left(\sum_{B \in \mathcal{B}} \zeta_B(T, u) \cdot f_B\right) \quad (4)$$

It is straightforward to obtain explicit formulas for ζ_B for some spanning sets \mathcal{B} of the free Lie algebra [22], but much preferable are series that use *bases* \mathcal{B} , and which, in addition, yield as simple formulas for ζ_B as (3) does for ξ_H .

Problem 1: Construct bases $\mathcal{B} = \{B_k : k \geq 0\}$ for the *free* Lie algebra on a finite number of generators X_1, \dots, X_m such that the corresponding iterated integral functionals ζ_B defined by (4) have *simple formulas* (similar to (3)), suitable for control applications (both analysis and design).

The formulae (4) and (2) arise from the “free control system” on the free associative algebra on m generators. Its universality means that its solutions map to solutions of specific systems (1) on M^n via the evaluation homomorphism $X_i \mapsto f_i$. However, the resulting formulas contain many redundant terms since the vector fields f_B are not linearly independent.

Problem 2: Provide an algorithm that generates for any finite collection of analytic vector fields $\mathcal{F} = \{f_1, \dots, f_m\}$ on M^n a basis for $L(f_1, \dots, f_m)$ together with effective formulas for associated iterated integral functionals.

Without loss of generality, one may assume that the collection \mathcal{F} satisfies the Lie algebra rank condition, i.e., $L(f_1, \dots, f_m)(p) = T_p M$ at a specified initial point p . It makes sense to first consider special classes of systems \mathcal{F} , e.g., which are such that $L(f_1, \dots, f_m)$ is finite, nilpotent, solvable, etc. The words *simple* and *effective* are not used in a technical sense in problems 1 and 2 (as in formal studies of computational complexity) but instead refer to comparison with the elegant formula (3), which has proven convenient for theoretical studies, numerical computation, and practical implementations.

2 MOTIVATION AND HISTORY OF THE PROBLEM

Series expansions of solution to differential equations have a long history. Elementary Picard iteration of the universal control system $\dot{S} = \sum_{i=1}^m X_i u_i$ on the free associative algebra over (X_1, \dots, X_m) yields the Chen Fliess series [5, 11, 21]. Other major tools are Volterra series, and the Magnus expansion [14], which groups the terms in a different way than the Fliess series. The main drawback of the Fliess series is that (unlike its exponential product expansion (2)) no finite truncation is the exact solution of any approximating system. A key innovation is the chronological calculus of 1970s Agrachev and Gamkrelidze [1]. However, it is generally not formulated using explicit bases.

The series and product expansions have manifold uses in control beyond simple computation of integral curves and analysis of reachable sets (which includes controllability and optimality). These include state-space realizations of systems given in input-output operator form [8, 20], output tracking, and path-planning. For the latter, express the target or reference trajectory in terms of the ξ or ζ , now considered as coordinates of a suitably lifted system (e.g., free nilpotent) and invert the restriction of the map $u \mapsto \{\xi_B : B \in \mathcal{B}_N\}$ or $u \mapsto \{\zeta_B : B \in \mathcal{B}_N\}$ (for some finite subbasis \mathcal{B}_N) to a finitely parameterized family of controls u , e.g., piecewise polynomial [7] or trigonometric polynomial [12, 17].

The Campbell-Baker-Hausdorff formula [18] is a classic tool to combine products of exponentials into a single exponential $e^a e^b = e^{H(a,b)}$ where $H(a,b) = a + b + \frac{1}{2}[a, b] + \frac{1}{12}[a, [a, b]] - \frac{1}{12}[b, [a, b]] + \dots$. It has been extensively used for designing piecewise constant control variations that generate high order tangent vectors to reachable sets, e.g., for deriving conditions for optimality. However, repeated use of the formula quickly leads to unwieldy expressions. The expansion (2) is the natural continuous analogue of the CBH formula, and the problem is to find the most useful form.

The uses of these expansions (2) and (4) extend far beyond control, as they apply to any dynamical systems that split into different interacting components. In particular, closely related techniques have recently found much attention in numerical analysis. This started with a search for Runge-Kutta-like integration schemes such that the approximate solutions inherently satisfy algebraic constraints (e.g., conservation laws) imposed on the dynamical system [3]. Much effort has been devoted to optimize such schemes, in particular minimizing the number of costly function evaluations [16]. For a recent survey see, [6]. Clearly, the form (4) is most attractive as it requires the evaluation of only a single (computationally costly) exponential.

The general area of noncommuting formal power series admits both dynamical systems/analytic and purely algebraic/combinatorial approaches. Algebraically, underlying the expansions (2) and (4) is the Chen series [2], which is well-known to be an exponential Lie series, compare [18], thus guarantee-

ing the existence of the alternative expansions

$$\sum_{w \in Z^*} w \otimes w \stackrel{!}{=} \exp \left(\sum_{B \in \mathcal{B}} \zeta_B \otimes B \right) \stackrel{!}{=} \prod_{B \in \mathcal{B}}^{\rightarrow} \exp (\zeta_B \otimes B) \quad (5)$$

The first bases for free Lie algebras build on Hall's work in the 1930s on commutator groups. While several other bases (Lyndon, Shirsov) have been proposed in the sequel, Viennot [23] showed that they are all special cases of generalized Hall bases. Underlying their construction is a unique factorization principle, which in turn is closely related to Poincar-Birckhoff-Witt bases (of the universal enveloping algebra of a Lie algebra) and Lazard elimination. Formulas for the dual PBW bases ζ_B have been given by Schützenberger, Sussmann [22], Grossman, and Melancon and Reutenauer [15]. For an introductory survey, see [11], while [15] elucidates the underlying Hopf algebra structure, and [18] is the principal technical reference for combinatorics of free Lie algebras.

3 AVAILABLE RELATED RESULTS

The direct expansion of the logarithm into a formal power series may be simplified using symmetrization [18, 22], but this still does not yield well-defined “coordinates” with respect to a basis.

Explicit but quite *unattractive* formulas for the first 14 coefficients ζ_H in the case of $m = 2$ and a Hall-basis are calculated in [10]. This calculation can be automated in a computer algebra system for terms of considerably higher order, but no apparent algebraic structure is discernible. These results suffice for some numerical purposes, but they do not provide much structural insight.

Several new algebraic structures introduced in [19] lead to systematic formulas for ζ_B using spanning sets \mathcal{B} that are smaller than those in [22], but are not bases. These formulas can be refined to apply to Hall-bases, but at the cost of losing their simple structure. Further recent insights into the underlying algebraic structures may be found in [4, 13].

The introductory survey [11] lays out in elementary terms the close connections between Lazard elimination, Hall-sets, chronological products, and the particularly attractive formula (3). These intimate connections suggest that to obtain similarly attractive expressions for ζ_B one may have to start from the very beginning by building bases for free Lie algebras that do not rely on recursive use of Lazard elimination. While it is desirable that any such new bases still restrict to bases of the homogeneous subspaces of the free Lie algebra, we suggest consider balancing the simplicity of the basis for the Lie algebra and structural simplicity of the formulas for the dual objects ζ_B . In particular, consider bases whose elements are not necessarily Lie monomials but possibly nontrivial linear combinations of iterated Lie brackets of the generators.

BIBLIOGRAPHY

- [1] A. Agrachev and R. Gamkrelidze, “Chronological algebras and nonstationary vector fields,” *Journal Soviet Math.*, 17, pp. 1650–1675, 1979.
- [2] K. T. Chen, “Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula,” *Annals of Mathematics*, 65, pp. 163–178, 1957.
- [3] P. Crouch and R. Grossman, “The explicit computation of integration algorithms and first integrals for ordinary differential equations with polynomial coefficients using trees,” *Proc. Int. Symposium on Symbolic and Algebraic Computation*, pp. 89–94, ACM Press, 1992.
- [4] A. Dzhumadil’daev, “Non-associative algebras without unit,” *Comm. Algebra*, 2002.
- [5] M. Fliess, “Fonctionelles causales nonlinéaires et indéterminées noncommutatives,” *Bull. Soc. Math. France*, 109, pp. 3–40, 1981.
- [6] A. Iserles, “Expansions that grow on trees,” *Notices AMS*, 49, pp. 430–440, 2002.
- [7] G. Jacob, “Motion planning by piecewise constant or polynomial inputs,” *Proc. NOLCOS*, Bordeaux, 1992.
- [8] B. Jakubczyk, “Convergence of power series along vector fields and their commutators; a Cartan-Kähler type theorem,” *Ann. Polon. Math.*, 74, pp. 117–132, 2000.
- [9] M. Kawski and H. J. Sussmann “Noncommutative power series and formal Lie-algebraic techniques in nonlinear control theory,” In: *Operators, Systems, and Linear Algebra*, U. Helmke, D. Prätzel-Wolters and E. Zerz, eds. Teubner, pp. 111–128, 1997.
- [10] M. Kawski, “Calculating the logarithm of the Chen Fliess series,” *Proc. MTNS*, Perpignan, 2000.
- [11] M. Kawski, “The combinatorics of nonlinear controllability and noncommuting flows,” *Lecture Notes series of the Abdus Salam ICTP*, 2001.
- [12] G. Lafferriere and H. Sussmann, “Motion planning for controllable systems without drift,” *Proc. Int. Conf. Robot. Automat.*, pp. 1148–1153, 1991.
- [13] J.-L. Loday and T. Pirashvili, “Universal enveloping algebras of Leibniz algebras and (co)homology,” *Math. Annalen*, 196, pp. 139–158, 1993.
- [14] W. Magnus, “On the exponential solution of differential equations for a linear operator,” *Comm. Pure Appl. Math.*, VII, pp. 649–673, 1954.

- [15] G. Melançon and C. Reutenauer, "Lyndon words, free algebras and shuffles," *Canadian J. Math.*, XLI, pp. 577–591, 1989.
- [16] H. Munthe-Kaas and B. Owren, "Computations in a free Lie algebra," *Royal Soc. London Philos. Trans. Ser. A Math. Phys. Eng. Sci.*, 357, pp. 957–982, 1999.
- [17] R. Murray and S. Sastry, "Nonholonomic path planning: steering with sinusoids," *IEEE Trans. Aut. Control*, 38, pp. 700–716, 1993.
- [18] C. Reutenauer, *Free Lie algebras*, Oxford: Clarendon Press, 1993.
- [19] E. Rocha, "On computataion of the logarithm of the Chen-Fliess series for nonlinear systems," *Proc. NCN*, Sheffield, 2001.
- [20] E. Sontag and Y. Wang, "Orders of input/output differential equations and state space dimensions," *SIAM J. Control and Optimization*, 33, pp. 1102–1126, 1995.
- [21] H. Sussmann, "Lie brackets and local controllability: A sufficient condition for scalar-input systems," *SIAM J. Cntrl. & Opt.*, 21, pp. 686–713, 1983.
- [22] H. Sussmann, "A product expansion of the Chen series," *Theory and Applications of Nonlinear Control Systems*, C. I. Byrnes and A. Lindquist, eds., Elsevier, North-Holland, pp. 323–335, 1986.
- [23] G. Viennot, "Algèbres de Lie Libres et Monoïdes Libres," *Lecture Notes Math.*, 692, Springer, Berlin, 1978.

Problem 3.4

An extended gradient conjecture

Luiz Carlos Martins Jr.

Universidade Paulista - UNIP

15091-450, S. J. do Rio Preto, SP

Brazil

lcmartinsjr@bol.com.br

Geraldo Nunes Silva

Departamento de Computacao e Estatistica

Universidade Estadual Paulista - UNESP

15054-000, S. J. do Rio Preto, SP

Brazil

gsilva@dcce.ibilce.unesp.br

1 DESCRIPTION OF THE PROBLEM

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}$ be a locally Lipschitz function, i.e., for all $x \in \mathbb{R}^n$ there is $\epsilon > 0$ and a constant K depending on ϵ such that

$$|f(x_1) - f(x_2)| \leq K\|x_1 - x_2\|, \quad \forall x_1, x_2 \in x + \epsilon B.$$

Here B denotes the open unit ball of \mathbb{R}^n .

Let $v \in \mathbb{R}^n$. The *generalized directional derivative* of f at x , in the direction v , denoted by $f^0(x; v)$, is defined as follows:

$$f^0(x; v) = \limsup_{\substack{y \rightarrow x \\ s \rightarrow 0^+}} \frac{f(y + sv) - f(y)}{s}.$$

Here $y \in \mathbb{R}^n$, $s \in (0, +\infty)$. The *generalized gradient* of f at x , denoted by $\partial f(x)$, is the subset of \mathbb{R}^n given by

$$\{\xi \in \mathbb{R}^n : f^0(x; v) \geq \langle \xi, v \rangle, \quad \forall v \in \mathbb{R}^n\}.$$

For the properties and basic calculus of the generalized gradient, standard references are [1] and [2].

The problem we propose here is regarding the following differential inclusion

$$\dot{x}(t) \in \partial f(x(t)) \text{ a.e. } t \in [0, \beta), \quad (1)$$

where β is a positive scalar. A solution of (1) is an absolutely continuous function $x : [0, \beta) \rightarrow \mathbb{R}^n$ that, together with \dot{x} , its derivative with respect to t , satisfies (1). Note that \dot{x} may fail to exist on a set $A \subset [0, \infty)$ of zero Lebesgue measure. Take S to be the set $[0, \infty) \setminus A$. We say that

$$d := \lim_{\substack{t \rightarrow \beta \\ S}} \frac{\dot{x}}{\|\dot{x}\|},$$

when the limit exists, is a *tangential direction* of x at $0 \in \mathbb{R}^n$. The notation $\lim_{\substack{t \rightarrow \beta \\ S}}$ means that the limit is taken for $t \in S$.

We are now in a position to propose our problem.

Conjecture: Suppose that $f(0) = 0$ and let x be a solution of (1) such that $x(t) \rightarrow 0$, as $t \rightarrow \beta$. Then there exists a unique tangential direction.

2 MOTIVATION AND HISTORY OF THE PROBLEM

This problem has been stated, for the first time, in the smooth case, that is, in the situation where f is a real analytic function on an open neighborhood $U_0 \subset \mathbb{R}^n$ of a point x_0 , and x is a maximal curve of (1) with ∇f , the gradient of f , replacing the generalized gradient of f and $x(t) \rightarrow x_0$, as $t \rightarrow \beta$. Under this conditions, R. Thom asked whether the tangent of $x(t)$ at x_0 was well-defined. This was later named the *conjecture of the gradient*, see, for example, [4, 5, 6].

We now show that this was a natural question to ask. Assuming that f is an analytic function as above and that $x_0 = 0$ and $f(0) = 0$, Lojasiewicz proved in [8, p. 92] that there exists $0 < \theta < 1$ such that

$$|\nabla f(x)| \geq |f(x)|^\theta, \quad \text{for } x \in U_0.$$

This result is known as *Lojasiewicz Inequality* and is the main tool in the proof of the next stated result. For an account on this see, for example, [7] and [9].

Theorem (Lojasiewicz): Let $A = f^{-1}(0) \cap U_0$. Then $\beta = +\infty$ and if $x(t)$ tends toward A , then $x(t)$ tends to a unique point of A .

(A simple proof of this theorem is provided in [3]).

Since, from the theorem above, we see that a maximal trajectory x lives in the whole interval $[0, \infty)$ and approximates a unique point in the inverse image of 0 by f , it is natural to ask if the tangent of $x(t)$ in the limit point is also unique. This was precisely what R. Thom conjectured and became the well-known *gradient conjecture*.

In this work, we propose an extension of this conjecture to the nonsmooth case.

3 KNOWN RESULTS AND REMARKS

The gradient conjecture, as it is known in the regular case, is equivalent to fact that the integral curves of ∇f have tangent in all points of $\omega(x)$. Partial results on the conjecture of the gradient was given in [3], [11], and [9]. The first proof of the general regular case was given in [4] and a simpler modified proof has appeared in [6]. Actually, it has been proved a stronger result that states that the radial projection of $x(t)$ from $x(0)$ into the sphere S^{n-1} has finite length. The arguments of the proof rely on the Łojasiewicz Inequality. The new conjecture of the gradient is stated in the nonsmooth setting and is called *the extended gradient conjecture*. As far as we know, no result has appeared in this direction. We reckon that a simple extension of the standard techniques used to prove the regular case is not enough. It will be necessary to come up with new ideas to prove this conjecture if it happens to be true.

BIBLIOGRAPHY

- [1] F. H. Clarke, *Optimization and Nonsmooth Analysis*, Wiley Interscience, New York 1983; reprinted as vol. 5 of Classics in Applied Mathematics, SIAM, Philadelphia, PA, 1990; Russian translation, Nauka, Moscow, 1988.
- [2] F. H. Clarke, Yu. S. Ledyaev, R. J. Stern, P. R. Wolenski, *Nonsmooth Analysis and Control Theory*, GTM 178, New York, Springer-Verlag, 1998.
- [3] F. Ichikawa, "Thom's conjecture on singularities of gradient vector fields," Kodai Math. J. 15, pp. 134-140, 1992.
- [4] K. Kurdyka, T. Mostowski, *The Gradient Conjecture of R. Thom*, preprint, 1996 (revised 1999). http://www.lama.univ-savoie.fr/sitelama/Membres/pages_web/KURDIKA/index.html.
- [5] K. Kurdyka, "On the gradient conjecture of R. Thom," *Seminari di Geometria 1998-1999*, Università di Bologna, Istituto di Geometria, Dipartimento di Matematica, pp. 143-151, 2000.
- [6] K. Kurdyka, T. Mostowski, A. Parusinski, "Proof of the gradient conjecture of R. Thom," *Annals of Mathematics*, 152 (2000), pp. 763-792.
- [7] S. Łojasiewicz, "Une propriété topologique des sous-ensembles analytiques réels," *Colloques Internationaux du C.N.R.S. #117, les équations aux dérivées partielles*, Paris 25-30 juin (1962), pp. 87-89.
- [8] S. Łojasiewicz, *Ensembles semi-analytiques*, IHES preprint, 1965.

- [9] R. Moussu, "Sur la dynamique des gradients. Existence de variétés invariants," *Math. Ann.*, 307 (1997), pp. 445-460.
- [10] F. Sanz, "Non oscillating solutions of analytic gradient vector fields," *Ann. Inst. Fourier*, Grenoble 48 (4) 1998, pp. 1045-1067.
- [11] H. Xing Lin, *Sur la structure des champs de gradients de fonctions analytiques réelles*, Ph. D. Thèse, Université Paris VII, 1992.

Problem 3.5

Optimal transaction costs from a Stackelberg perspective

Geert Jan Olsder

Faculty of Information Technology and Systems
Delft University of Technology
P.O. Box 5031, 2600 GA Delft
The Netherlands
`g.j.olsder@its.tudelft.nl`

1 DESCRIPTION OF THE PROBLEM

The problem to be considered is

$$\dot{x} = f(x, u), \quad x(0) = x_0, \quad (1)$$

$$\max_u J_F = \max_u \{q(x(T)) + \int_0^T g(x, u) dt - \int_0^T \gamma(u(t)) dt\}, \quad (2)$$

$$\max_{\gamma(\cdot)} J_L = \max_{\gamma(\cdot)} \int_0^T \gamma(u(t)) dt, \quad (3)$$

with f , g and q being given functions, the state $x \in R^n$, the control $u \in R$, and $\gamma(\cdot)$ is a scalar function which maps R into R . The problem concerns a dynamic game problem in which u is the decision variable of one player called the Follower, and the function γ is up to the choice of the other player called the Leader. An essential feature of the problem is that the Leader's profit (3) is a direct loss for the Follower in (2). The Leader lives as a parasite on the Follower. In the next section, a more concrete motivation will be given. The function γ must be chosen subject to the constraints

$$\gamma(0) = 0, \quad \gamma(\cdot) \geq 0$$

and if at all possible it must be nondecreasing with respect to $|u|$, and possibly also $\gamma(u) = \gamma(-u)$. By means of the notation introduced and the names of the players it should be clear that the problem formulated is a (special kind of) Stackelberg game [2]. The Leader announces the function γ that thus becomes known to the Follower who subsequently chooses u . Thus the optimal u is a function of $\gamma(\cdot)$.

2 MOTIVATION AND HISTORY OF THE PROBLEM

For $n = 1$, i.e., $x \in R$, which we assume henceforth, an interpretation of this model is that $x(t)$ represents the Follower's wealth at time t . This Follower is an investor and who would like to maximize

$$\int_0^T g(x, u) dt + q(x(T)). \quad (4)$$

The term $q(x(T))$ in this criterion is a function of the wealth of the investor at the final time T and the term $\int_0^T g(x, u) dt$ represents the consumption during the time interval $[0, T]$. The decision variable $u(t)$ denotes the transactions with the bank at time t (e.g., selling or buying stocks). To be more precise, $u(t)$ denotes a transaction density, i.e., during the time interval $[t, t + dt]$ the number of transactions equals $u(t)dt$. For $u = 0$, no transactions take place and the bank does not earn money (because $\gamma(0) = 0$). Transactions cost money and we assume that the bank (i.e., the Leader) wants to maximize these transaction costs as indicated by (3). These costs are subtracted from (4) and hence the ultimate criterion of the Follower is given by (2). The restrictions posed on γ (nondecreasing with respect to $|u|$ and $\gamma(0) = 0$) now have a clear meaning. The higher the number of transactions (either buying or selling, one being related to a positive u , the other one to a negative u), the higher the costs.

Equation (1) is supposed to tell how the wealth x evolves in time. Usually, such models are represented by stochastic differential equations, but due to the complexity of the problem, we restrict ourselves to a less realistic deterministic differential function.

3 AVAILABLE RESULTS AND BACKGROUND

Problems with transaction costs have been studied before, see e.g., [1, 3, 4], but never from the point of view of the bank to maximize these costs. The problem as stated is a difficult one, see [7] for some first solution attempts. The principal difficulty is that composed functions are involved, i.e., one function is the argument of another [6]. Hence, we will also consider the following static problem, which is simpler than the time-dependent one:

$$\max_u (q(u) - \gamma(u)), \quad \max_{\gamma(\cdot)} \gamma(u),$$

subject to $\gamma(\cdot) \geq 0$ and $\gamma(0) = 0$ and possibly also $\gamma(u)$ nondecreasing with respect to $|u|$. With the same interpretation as before, the investor is secured of a minimum value $q(0)$ by playing $u = 0$. Therefore, he will only take u -values into consideration for which $q(u) \geq q(0)$. This static problem is a special case of the so-called inverse Stackelberg problem as it was introduced in [5] and a solution method is known, see chapter 7 of [2].

To start with, in a *conventional* Stackelberg game, there are two players, called Leader and Follower respectively, each having their own cost function

$$J_L(u_L, u_F), \quad J_F(u_L, u_F),$$

where $u_F, u_L \in \mathbb{R}$. Each player wants to choose his own decision variable in such a way as to maximize his own criterion. Without giving an equilibrium concept, the problem as stated so far is not well defined. Such an equilibrium concept could, for instance, be one named after Nash or Pareto. In the Stackelberg equilibrium concept, one player, the Leader, announces his decision u_L , which is subsequently known to the other player, the Follower. With this knowledge, the Follower chooses his u_F . Hence, u_F becomes a function of u_L , written as

$$u_F = l_F(u_L),$$

which is determined through the relation

$$\min_{u_F} J_F(u_L, u_F) = J_F(u_L, l_F(u_L)),$$

provided that this minimum exists and is a singleton for each possible choice u_L of the Leader. The function $l_F(\cdot)$ is sometimes called a reaction function. Before the Leader announces his decision u_L , he will realize how the Follower will react and hence the Leader chooses u_L such as to minimize

$$J_L(u_L, l_F(u_L)).$$

In an *inverse* Stackelberg game, the Leader does not announce his choice u_L ahead of time, as above, but instead a function $\gamma_L(u_F)$. Think (as another motivating example) of the Leader being the government and of the Follower as a citizen. The government states how much income tax the citizen has to pay and this tax will depend on the income u_F of the citizen. It is up to the citizen as to how much money to earn (by working harder or not) and thus he can choose u_F . The income tax the government will receive equals $\gamma_L(u_F)$, where the "rule for taxation" $\gamma_L(\cdot)$ was made known ahead of time.

BIBLIOGRAPHY

- [1] M. Akian, J. L. Menaldi, and A. Sulem, "On an investment-consumption model with transaction costs," *SIAM J. Control and Optim.*, vol. 34 pp. 329-364, 1996.
- [2] T. Başar and G. J. Olsder, *Dynamic Noncooperative Game Theory*, SIAM, Philadelphia, 1999.
- [3] P. Bernhard, "A robust control approach to option pricing including transaction costs," In: *Annals of the ISDG 7*, A. Nowak, ed., Birkhäuser, 2002.

- [4] E. R. Grannan and G. H. Swindle, "Minimizing transaction costs of option hedging strategies," *Mathematical finance*, vol. 6, no. 4, 341–364, 1996.
- [5] Y.-C. Ho, P. B. Luh and G. J. Olsder, "A control-theoretic view on incentives," *Automatica*, vol. 18, pp. 167-179, 1982.
- [6] M. Kuczma, *Functional Equations in a Single Variable*, Polish Scientific Publishers, 1968.
- [7] G. J. Olsder, "Differential game-theoretic thoughts on option pricing and transaction costs," *International Game Theory Review*, vol. 2, pp. 209-228, 2000.

Problem 3.6

Does cheap control solve a singular nonlinear quadratic problem?

Yuri V. Orlov

Electronics Department
CICESE Research Center
Ensenada, BC 22860
Mexico
yorlov@cicese.mx

1 DESCRIPTION OF THE PROBLEM

A standard control synthesis for affine systems

$$\dot{x} = f(x) + g(x)u, \quad x \in R^n, u \in R^m \quad (1)$$

under degenerate performance criterion

$$J(u) = \int_0^\infty x^T(t)Px(t)dt, \quad P = P^T > 0 \quad (2)$$

depending on the state vector $x(t)$ only, replaces this singular optimization problem by its regularization through ε -approximation

$$J_\varepsilon(u) = \int_0^\infty [x^T(t)Px(t) + \varepsilon u^T(t)Ru(t)]dt, \quad \varepsilon > 0, \quad R = R^T > 0 \quad (3)$$

of this criterion with small (cheap) penalty on the control input u . Hereafter, functions f, g are assumed sufficiently smooth, and all quantities in (1) through (3) are assumed to have compatible dimensions.

The optimal control synthesis corresponding to (2) is then obtained as a limit as $\varepsilon \rightarrow 0$ of the optimal control law u_ε^0 corresponding to (3). Since only particular approximation is taken while other approximations are possible as well there is no guarantee that the original performance criterion is minimized by the control law obtained via this procedure.

An open problem that arises here is to prove that

$$\inf_u J(u) = \lim_{\varepsilon \rightarrow 0} \inf_u J_\varepsilon(u) \quad (4)$$

or present a counterexample of system (1) where the limiting relation (4) is not satisfied.

2 MOTIVATION AND HISTORY OF THE PROBLEM

The above problem is well-understood in the linear case when system (1) is specified as follows:

$$\dot{x} = Ax + Bu, \quad x \in R^n, u \in R^m. \quad (5)$$

Under the stabilizability and detectability conditions the linear system (5) driven by the cheap control u_ε^0 exhibits an initial fast transient followed by a slow motion on a singular arc (see, e.g., [3, section 6] and references therein). In the limit $\varepsilon \rightarrow 0$, a singular perturbation analysis reveals that the stable fast modes decay instantaneously as if they would be driven by the impulsive component of the controller minimizing the degenerate performance criterion (2).

This feature, however, does not admit a straightforward extension to the system in question because in contrast to the linear system (5), an instantaneous impulse response of the affine system (1), generally speaking, depends on an approximation of the impulse [2]. Thus, it might happen that the original performance (2) is not minimized through the ε -approximation (3) of this criterion.

3 AVAILABLE RESULTS

A distribution-oriented variational analysis [1] of the singular nonlinear quadratic problem (1), (2), admitting both integrable and impulsive inputs, reveals that the infimum of the degenerate criterion (2) is typically attained by a controller with impulsive behavior at the initial time moment. In that case, an instantaneous impulse response of the closed-loop system does not depend on an approximation of the impulse if and only if the affine system (1) satisfies the Frobenius condition, i.e., the distribution spanned by the columns of $g(x)$ is involutive (see [2] for details).

Motivated by these arguments, the author suspects that the limiting relation (4) holds whenever system (1) satisfies the Frobenius condition, and a counterexample of system (1), violating (4), is indeed possible if the Frobenius condition is not imposed on the system anymore.

BIBLIOGRAPHY

- [1] Y. Orlov, "Necessary optimality conditions of generalized control actions 1,2," *Automation and Remote Control*, vol. 44, no. 7, pp. 868-877; vol. 44, no. 8, pp. 998-1105, 1984.
- [2] Y. Orlov, "Instantaneous impulse response of nonlinear systems," *IEEE Trans. Aut. Control*, vol. 45, no. 5, pp. 999-1001, 2000.

- [3] V. R. Saksena, J. O'Reilly, and P. V. Kokotovic, "Singular perturbations and time-scale methods in control theory: Survey 1976-1983," *Automatica*, vol. 20, no. 3, pp. 273-293, 1984.

Problem 3.7

Delta-Sigma modulator synthesis

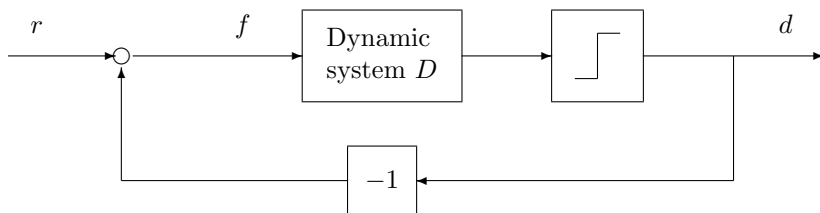
Anders Rantzer

Dept. of Automatic Control
Lund Institute of Technology
P.O. Box 118
SE-221 00 LUND, Sweden
`rantzer@control.lth.se`

1 DESCRIPTION OF THE PROBLEM

Delta-Sigma modulators are among the key components in modern electronics. Their main purpose is to provide cheap conversion from analog to digital signals. In the figure below, the analog signal r with values in the interval $[-1, 1]$ is supposed to be approximated by the digital signal d that takes only two values, -1 and 1 . One can not expect good approximation at all frequencies. Hence, the dynamic system D should be chosen to minimize the error f in a given frequency range $[\omega_1, \omega_2]$.

There is a rich literature on Delta-Sigma modulators. See [2, 1] and references therein. The purpose of this note is to reach a broad audience by focusing on the central mathematical problem.



To make a precise problem formulation, we need to introduce some notation:

Notation: The signal space $\ell[0, \infty]$ is the set of all sequences $\{f(k)\}_{k=0}^{\infty}$ such that $f(k) \in [-1, 1]$ for $k = 0, 1, 2, \dots$. A map $D : \ell[0, \infty] \rightarrow \ell[0, \infty]$ is called a *causal dynamic system* if for every $u, v \in \ell[0, \infty]$ such that $u(k) = v(k)$ for $k \leq T$ it holds that $[D(u)](k) = [D(v)](k)$ for $k \leq T$. Define also the

function

$$\operatorname{sgn}(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{else} \end{cases}$$

Problem: Given $r \in \ell[0, \infty]$ and a causal dynamic system D , define $d, f \in \ell[0, \infty]$ by

$$\begin{cases} d(k+1) = \operatorname{sgn}([D(f)](k)), & d(0) = 0 \\ f(k) = r(k) - d(k) \end{cases}$$

and find a causal dynamic system D such that regardless of the reference input r , the error signal f becomes small in a prespecified frequency interval $[\omega_1, \omega_2]$.

The problem formulation is intentionally left vague on the last line. The size of f can be measured in many different ways. One option is to require a uniform bound on

$$\limsup_{T \rightarrow \infty} \frac{1}{T} \left| \sum_{k=0}^T e^{-ik\omega} f(k) \right|$$

for all $\omega \in [\omega_1, \omega_2]$ and all reference signals $r \in \ell[0, \infty]$.

Another option is to allow D to be stochastic system and put a bound on the spectral density of f in the frequency interval. This would be consistent with the wide-spread practice to add a stochastic “dithering signal” before the nonlinearity in order to avoid undesired periodic orbits.

2 AVAILABLE RESULTS

The simplest and best understood case is where

$$\begin{cases} x(k+1) = x(k) + f(k) \\ f(k) = r(k) - \operatorname{sgn}(x(k)) \end{cases}$$

In this case, it is easy to see that the set $x \in [-2, 2]$ is invariant, so with

$$F_T(z) = \sum_{k=0}^T z^{-k} f(k) \qquad X_T(z) = \sum_{k=0}^T z^{-k} x(k)$$

it holds that

$$\begin{aligned}
 \frac{1}{T} \int_0^{\omega_0} |F_T(e^{i\omega})|^2 d\omega &= \frac{1}{T} \int_0^{\omega_0} |(e^{i\omega} - 1)X_T(e^{i\omega})|^2 d\omega \\
 &= \frac{1}{T} \int_0^{\omega_0} [2(1 - \cos \omega) |X_T(e^{i\omega})|^2] d\omega \\
 &\leq 2(1 - \cos \omega_0) \frac{1}{T} \int_0^{\pi} |X_T(e^{i\omega})|^2 d\omega \\
 &= 2(1 - \cos \omega_0) \frac{\pi}{T} \sum_{k=0}^T x(k)^2 \\
 &\leq 8\pi(1 - \cos \omega_0)
 \end{aligned}$$

which clearly bounds the error f at low frequencies.

Many modifications using higher order dynamics have been suggested in order to further reduce the error. However, there is still a strong demand for improvements and a better understanding of the nonlinear dynamics. The following two references are suggested as entries to the literature on Δ - Σ -modulators:

BIBLIOGRAPHY

- [1] James A. Cherry, *Continuous Time Delta-Sigma Modulators for High-Speed A/D Conversion: Theory, Practice & Fundamental Performance Limits*, Kluwer, 1999.
- [2] S. R. Norsworthy, R. Schreier, and G. C. Temes, *Delta-Sigma Data Converters*, IEEE Press, New York, 1997.

Problem 3.8

Determining of various asymptotics of solutions of
nonlinear time-optimal problems via right ideals in
the moment algebra

G. M. Sklyar

Szczecin University

Wielkopolska str. 15, 70-451 Szczecin, Poland;

Kharkov National University

Svoboda sqr. 4, 61077 Kharkov

Ukraine

sklar@sus.univ.szczecin.pl, sklyar@univer.kharkov.ua

S. Yu. Ignatovich

Kharkov National University

Svoboda sqr. 4, 61077 Kharkov, Ukraine

bob@online.kharkiv.com

1 MOTIVATION AND HISTORY OF THE PROBLEM

The time-optimal control problem is one of the most natural and at the same time difficult problems in the optimal control theory.

For linear systems, the maximum principle allows to indicate a class of optimal controls. However, the explicit form of the solution can be given only in a number of particular cases [1-3]. At the same time [4], an arbitrary linear time-optimal problem with analytic coefficients can be approximated (in a neighborhood of the origin) by a certain linear problem of the form

$$\dot{x}_i = -t^{q_i} u, \quad i = 1, \dots, n, \quad q_1 < \dots < q_n, \quad x(0) = x^0, x(\theta) = 0,$$

$$|u| \leq 1, \theta \rightarrow \min. \quad (1)$$

In the nonlinear case, the careful analysis is required for any particular system [5, 6]. However, in a number of cases the time-optimal problem for a nonlinear system can be approximated by a linear problem of the form (1) [7]. We recall this result briefly. Consider the time-optimal problem in the

following statement

$$\dot{x} = a(t, x) + ub(t, x), \quad a(t, 0) \equiv 0, \quad x(0) = x^0, \quad x(\theta) = 0, \quad |u| \leq 1, \quad \theta \rightarrow \min, \quad (2)$$

where a, b are real analytic in a neighborhood of the origin in \mathbb{R}^{n+1} . Let us denote by (θ_{x^0}, u_{x^0}) the solution of this problem.

Denote by R_a, R_b the operators acting as $R_a d(t, x) = d_t(t, x) + d_x(t, x) \cdot a(t, x)$, $R_b d(t, x) = d_x(t, x) \cdot b(t, x)$ for any vector function $d(t, x)$ analytic in a neighborhood of the origin in \mathbb{R}^{n+1} and let $\text{ad}_{R_a}^{m+1} R_b = [R_a, \text{ad}_{R_a}^m R_b]$, $m \geq 0$; $\text{ad}_{R_a}^0 R_b = R_b$, where $[\cdot, \cdot]$ is the operator commutator. Denote $E(x) \equiv x$.

Theorem 1: The conditions $\text{rank}\{\text{ad}_{R_a}^j R_b E(x)|_{x=0}\}_{j \geq 0} = n$ and

$$[\text{ad}_{R_a}^{m_1} R_b, \dots [\text{ad}_{R_a}^{m_{k-1}} R_b, \text{ad}_{R_a}^{m_k} R_b] \dots] E(x)|_{x=0} \in \text{Lin} \left\{ \text{ad}_{R_a}^j R_b E(x)|_{x=0} \right\}_{j=0}^{m-2} \quad (3)$$

for any $k \geq 2$ and $m_1, \dots, m_k \geq 0$, where $m = m_1 + \dots + m_k + k$, hold if and only if there exist a nonsingular transformation Φ of a neighborhood of the origin in \mathbb{R}^n , $\Phi(0) = 0$, and a linear time-optimal problem of the form (1), which approximates problem (2) in the following sense

$$\frac{\theta_{\Phi(x^0)}}{\theta_{x^0}^{Lin}} \rightarrow 1, \quad \frac{1}{\theta} \int_0^\theta |u_{x^0}^{Lin}(t) - u_{\Phi(x^0)}(t)| dt \rightarrow 0 \quad \text{as} \quad x^0 \rightarrow 0,$$

where $(\theta_{x^0}^{Lin}, u_{x^0}^{Lin})$ denotes the solution of (1) and $\theta = \min\{\theta_{\Phi(x^0)}, \theta_{x^0}^{Lin}\}$.

That means that if the conditions of theorem 1 are not satisfied, then the asymptotic behavior of the solution of the nonlinear time-optimal problem differs from asymptotics of solutions of all linear problems.

In order to formulate the next result, let us give the representation of the system in the form of a series of nonlinear power moments [7]. We assume the initial point x^0 is steered to the origin in the time θ by the control $u(t)$ w.r.t. system (2). Then under our assumptions for rather small θ one has

$$x^0 = \sum_{m=1}^{\infty} \sum_{m_1 + \dots + m_k = m} v_{m_1 \dots m_k} \xi_{m_1 \dots m_k}(\theta, u), \quad (4)$$

where $\xi_{m_1 \dots m_k}(\theta, u) = \int_0^\theta \int_0^{\tau_1} \dots \int_0^{\tau_{k-1}} \prod_{j=1}^k \tau_j^{m_j} u(\tau_j) d\tau_k \dots d\tau_2 d\tau_1$ are *nonlinear power moments* and

$$v_{m_1 \dots m_k} = \frac{(-1)^k}{m_1! \dots m_k!} \text{ad}_{R_a}^{m_1} R_b \text{ad}_{R_a}^{m_2} R_b \dots \text{ad}_{R_a}^{m_k} R_b E(x)|_{x=0}.$$

We say that $\text{ord}(\xi_{m_1 \dots m_k}) = m_1 + \dots + m_k + k$ is the order of $\xi_{m_1 \dots m_k}$.

Theorem 1 means that there exists a transformation Φ which reduces (4) to

$$(\Phi(x^0))_i = \xi_{q_i}(\theta, u) + \rho_i, \quad i = 1, \dots, n,$$

where ρ_i includes power moments of order greater than $q_i + 1$ only while the representation (4) for the linear system (1) obviously has the form

$$x_i^0 = \xi_{q_i}(\theta, u), \quad i = 1, \dots, n.$$

That is the linear moments that correspond to the linear time-optimal problem (1) form the *principal part* of the series in representation (4) as $\theta \rightarrow 0$.

When condition (3) is not satisfied, one can try to find a *nonlinear* system which has rather simply form and approximates system (2) in the sense of time optimality. In [8] we claim the following result.

Consider the linear span \mathcal{A} of all nonlinear moments $\xi_{m_1 \dots m_k}$ over \mathbb{R} as a *free algebra* with the basis $(\xi_{m_1 \dots m_k} : k \geq 1, m_1, \dots, m_k \geq 0)$ and the product $\xi_{m_1 \dots m_k} \xi_{n_1 \dots n_s} = \xi_{m_1 \dots m_k n_1 \dots n_s}$. Introduce the inner product in \mathcal{A} assuming the basis $(\xi_{m_1 \dots m_k})$ to be orthonormal. Consider also the *Lie algebra* L over \mathbb{R} generated by the elements $(\xi_m)_{m=0}^\infty$ with the commutator $[\ell_1, \ell_2] = \ell_1 \ell_2 - \ell_2 \ell_1$. Introduce further the graded structure $\mathcal{A} = \sum_{m=1}^\infty \mathcal{A}_m$ putting $\mathcal{A}_m = \text{Lin}\{\xi_{m_1 \dots m_k} : \text{ord}(\xi_{m_1 \dots m_k}) = m\}$.

Consider now a system of the form (2). The series in (4) naturally defines the linear mapping $v : \mathcal{A} \rightarrow \mathbb{R}^n$ by the rule $v(\xi_{m_1 \dots m_k}) = v_{m_1 \dots m_k}$. Further we assume the system (2) to be n -dimensional, i.e. $\dim v(L) = n$. Note that the form of coefficients $v_{m_1 \dots m_k}$ of the series in (4) implies the following property of the mapping v : the equality $v(\ell) = 0$ for $\ell \in L$ implies $v(\ell x) = 0$ for any $x \in \mathcal{A}$. That means that any system of the form (2) generates a *right ideal* in the algebra \mathcal{A} . We introduce the right ideal in the following way.

Consider the sequence of subspaces $D_r = v(L \cap (\mathcal{A}_1 + \dots + \mathcal{A}_r)) \subset \mathbb{R}^n$, and put $r_0 = \min\{r : \dim D_r = n\}$. For any $r \leq r_0$ consider a subspace P_r of all elements $y \in L \cap \mathcal{A}_r$ such that $v(y) \in D_{r-1}$ (we assume $D_0 = \{0\}$). Then put $J = \sum_{r=1}^{r_0} P_r(\mathcal{A} + \mathbb{R})$. Let J^\perp be the orthogonal complement of J . In the next theorem L_{J^\perp} denotes the projection of the Lie algebra L on J^\perp .

Theorem 2: (A) Let system (2) be n -dimensional, $\tilde{\ell}_1, \dots, \tilde{\ell}_n$ be a basis of $\sum_{r=1}^{r_0} (L_{J^\perp} \cap \mathcal{A}_r)$ such that $\text{ord}(\tilde{\ell}_i) \leq \text{ord}(\tilde{\ell}_j)$ as $i < j$. Then there exists a nonsingular analytic transformation Φ of a neighborhood of the origin that reduces (4) to the following form

$$(\Phi(x^0))_i = \tilde{\ell}_i + \rho_i, \quad i = 1, \dots, n,$$

where ρ_i contains moments of order greater than $\text{ord}(\tilde{\ell}_i)$ only. Moreover, there exists a control system of the form

$$\dot{x} = ub^*(t, x), \quad (5)$$

such that representation (4) for this system is of the form

$$x_i^0 = \tilde{\ell}_i, \quad i = 1, \dots, n. \quad (6)$$

(B) Suppose there exists an open domain $\Omega \subset \mathbb{R}^n \setminus \{0\}$, $0 \in \overline{\Omega}$, such that

i) the time-optimal problem for system (5) with representation (6) has a unique solution $(\theta_{x^0}^*, u_{x^0}^*(t))$ for any $x^0 \in \overline{\Omega}$;

ii) the function $\theta_{x^0}^*$ is continuous for $x^0 \in \overline{\Omega}$;

iii) denote $K = \{u_{x^0}^*(t\theta_{x^0}^*) : x^0 \in \overline{\Omega}\}$ and suppose that the following condition holds: when considering K as a set in the space $L_2(0, 1)$, the weak convergence of a sequence of elements from K implies the strong convergence.

Then the time-optimal problem for system (5) approximates problem (2) in the domain Ω in the following sense: there exists a set of pairs $(\tilde{\theta}_{x^0}, \tilde{u}_{x^0}(t))$, $x^0 \in \Omega$, such that the control $\tilde{u}_{x^0}(t)$ steers the point $\Phi(x^0)$ to the origin in the time $\tilde{\theta}_{x^0}$ w.r.t. system (2) and

$$\frac{\theta_{\Phi(x^0)}}{\theta_{x^0}^*} \rightarrow 1, \quad \frac{\tilde{\theta}_{x^0}}{\theta_{x^0}^*} \rightarrow 1, \quad \frac{1}{\theta} \int_0^\theta |u_{x^0}^*(t) - \tilde{u}_{x^0}(t)| dt \rightarrow 0 \quad \text{as } x^0 \rightarrow 0, \quad x^0 \in \Omega,$$

where $\theta_{\Phi(x^0)}$ is the optimal time for problem (2) and $\theta = \min\{\tilde{\theta}_{x^0}, \theta_{x^0}^*\}$.

Remark 1: If there exists the autonomous system $\dot{x} = a(x) + ub(x)$ such that its representation (4) is of the form (6) and the origin belongs to the interior of the controllability set then the function $\theta_{x^0}^*$ is continuous in a neighborhood of the origin [9]. Further, if time-optimal controls for system (5) are bang-bang then they satisfy condition iii) of theorem 2.

Remark 2: Consider any $r_0 \geq 0$ and an arbitrary sequence of subspaces $M = \{M_r\}_{r=1}^{r_0}$, $M_r \subset L \cap \mathcal{A}_r$, such that $\sum_{r=1}^{r_0} (\dim(L \cap \mathcal{A}_r) - \dim M_r) = n$. Put $J_M = \sum_{r=1}^{r_0} M_r(\mathcal{A} + \mathbb{R})$. We denote by \mathcal{J} the set of all such ideals. For any $J \in \mathcal{J}$ one can construct a control system of the form (5) such that its representation (4) is of the form (6).

2 FORMULATION OF THE PROBLEM.

Thus, the steering problem $\dot{x} = a(t, x) + ub(t, x)$, $x(\theta) = 0$, where $a(t, 0) \equiv 0$, generates the right ideal in the algebra \mathcal{A} , which defines system (5), and, under conditions i)–iii) of theorem 2, describes the asymptotics of the solution of time-optimal problem (2). The question is: *if any system of the form (5) having the representation of the form (6) satisfies conditions i)–iii) of theorem 2. The positive answer means that all possible asymptotics of solutions of the time-optimal problems (2) are represented as asymptotics of solutions of the time-optimal problems for systems (5) with representations of the form (6).*

In other words, if any system of the form (5) having the representation of the form (6) satisfies conditions i)–iii) of theorem 2, then time-optimal problems (2) induce the same structure in the algebra \mathcal{A} as steering problems to the origin under the constraint $|u| \leq 1$, namely, the set of right ideals \mathcal{J} . If this is not the case, then the next problem is to describe constructively the class of such systems.

BIBLIOGRAPHY

- [1] V. I. Korobov and G. M. Sklyar, "Time optimality and the power moment problem," *Math. USSR Sb.*, 62, pp. 185-205, 1989.

- [2] V. I. Korobov and G. M. Sklyar, "Time optimality and the trigonometric moment problem," *Math. USSR Izv.*, 35, pp. 203-220, 1990.
- [3] V. I. Korobov and G. M. Sklyar, "Markov power min-problem with periodic gaps," *J. Math. Sci.*, 80, pp. 1559-1581, 1996.
- [4] G. M. Sklyar and S. Yu. Ignatovich, "A classification of linear time-optimal control problems in a neighborhood of the origin," *J. Math. Anal. Applic.*, 203, pp. 791-811, 1996.
- [5] A. Bressan, "The generic local time-optimal stabilizing controls in dimension 3," *SIAM J. Control Optimiz.*, 24, pp. 177-190, 1986.
- [6] B. Bonnard, "On singular extremals in the time minimal control problem in \mathbb{R}^3 ," *SIAM J. Control Optimiz.*, 23, pp. 794-80, 1985.
- [7] G. M. Sklyar and S. Yu. Ignatovich, "Moment approach to nonlinear time optimality," *SIAM J. Control Optimiz.*, 38, pp. 1707-1728, 2000.
- [8] G. M. Sklyar and S. Yu. Ignatovich, "Approximation in the sense of time optimality via nonlinear moment problems," *SIAM J. Control Optimiz.*
- [9] V. I. Korobov, "On continuous dependence of the solution of the optimal control problem with free time on initial data," *Differen. Uravn.*, 7, pp. 1120-1123, 1971, (in Russian).

Problem 3.9

Dynamics of principal and minor component flows

U. Helmke and S. Yoshizawa

Department of Mathematics
University of Würzburg
Am Hubland, D-97074 Würzburg
Germany
`helmke@mathematik.uni-wuerzburg.de`

R. Evans, J.H. Manton and I.M.Y. Mareels

Department of Electrical and Electronic Engineering
The University of Melbourne
Victoria 3010
Australia
`jon@ee.mu.oz.au`

Stochastic subspace tracking algorithms in signal processing and neural networks are often analyzed by studying the associated matrix differential equations. Such gradient-like nonlinear differential equations have an intricate convergence behavior that is reminiscent of matrix Riccati equations. In fact, these types of systems are closely related. We describe a number of open research problems concerning the dynamics of such flows for principal and minor component analysis.

1 DESCRIPTION OF THE PROBLEM

Principal component analysis is a widely used method in neural networks, signal processing, and statistics for extracting the dominant eigenvalues of the covariance matrix of a sequence of random vectors. In the literature, various algorithms for principal component and principal subspace analysis have been proposed along with some, but in many aspects incomplete, theoretical analyses of them. The analysis is usually based on stochastic approximation techniques and commonly proceeds via the so-called Ordinary Differential Equation (ODE) method, i.e., by associating an ODE whose convergence properties reflect that of the stochastic algorithm; see e.g., [7]. In the sequel, we consider some of the relevant ODEs in more detail and pose some open

problems concerning the dynamics of the flows.

In order to state our problems in precise mathematical terms, we give a formal definition of a principal and minor component flow.

Definition[PSA/MSA Flow]: *A normalized subspace flow for a covariance matrix C is a matrix differential equation $\dot{X} = f(X)$ on $\mathbb{R}^{n \times p}$ with the following properties:*

1. *Solutions $X(t)$ exist for all $t \geq 0$ and have constant rank.*
2. *If X_0 is orthonormal, then $X(t)$ is orthonormal for all t .*
3. *$\lim_{t \rightarrow \infty} X(t) = X_\infty$ exists for all full rank initial conditions X_0 .*
4. *X_∞ is an orthonormal basis matrix of a p -dimensional eigenspace of C .*

The subspace flow is called a PSA (principal subspace) or MSA (minor subspace) flow, if, for generic initial conditions, the solutions $X(t)$ converge for $t \rightarrow \infty$ to an orthonormal basis of the eigenspace that is spanned by the first p dominant or minor eigenvectors of C , respectively.

In the neural network and signal processing literature, a number of such principal subspace flows have been considered. The best-known example of a PSA flow is Oja's flow [9, 10]

$$\dot{X} = (I - XX')CX. \quad (1)$$

Here $C = C' > 0$ is the $n \times n$ covariance matrix and X is an $n \times p$ matrix. Actually, it is nontrivial to prove that this cubic matrix differential equation is indeed a PSA in the above sense and thus, generically, converges to a dominant eigenspace basis. Another, more general example of a PSA flow is that introduced by [12, 13] and [17]:

$$\dot{X} = CXN - XNX'CX \quad (2)$$

Here $N = N' > 0$ denotes an arbitrary diagonal $k \times k$ matrix with distinct eigenvalues. This system is actually a joint generalization of Oja's flow (1) and of Brockett's [1] gradient flow on orthogonal matrices

$$\dot{X} = [C, XNX']X \quad (3)$$

For symmetric matrix diagonalisation, see also [6]. In [19], Oja's flow was re-derived by first proposing the gradient flow

$$\dot{X} = (C(I - XX') + (I - XX')C)X \quad (4)$$

and then omitting the first term $C(I - XX')X$ because $C(I - XX')X = CX(I - X'X) \rightarrow 0$, a consequence of both terms in (4) forcing X to the invariant manifold $\{X : X'X = I\}$. Interestingly, it has recently been realized [8] that (4) has certain computational advantages compared with (1), however, a rigorous convergence theory is missing. Of course, these three systems are just prominent examples from a bigger list of potential PSA

flows. One open problem in most of the current research is a lack of a full convergence theory, establishing pointwise convergence to the equilibria. In particular, a solution to the following three problems would be highly desirable. The first problem addresses the qualitative analysis of the flows.

Problem 1. Develop a complete phase portrait analysis of (1), (2) and (4). In particular, prove that the flows are PSA, determine the equilibria points, their local stability properties and the stable and unstable manifolds for the equilibrium points.

The previous systems are useful for principal component analysis, but they cannot be used immediately for minor component analysis. Of course, one possible approach might be to apply any of the above flows with C replaced by C^{-1} . Often this is not reasonable though, as in most applications the covariance matrix C is implemented by recursive estimates and one does not want to invert these recursive estimates on line. Another alternative could be to put a negative sign in front of the equations. But this does not work either, as the minor component equilibrium point remains unstable. In the literature, therefore, other approaches to minor component analysis have been proposed [2, 3, 5], but without a complete convergence theory.¹ Moreover, a guiding geometric principle that allows for the systematic construction of minor component flows is missing. The key idea here seems to be an appropriate concept of duality between principal and minor component analysis.

Conjecture 1. Principal component flows are dual to minor component flows, via an involution in matrix space $\mathbb{R}^{n \times p}$, that establishes a bijective correspondence between solutions of PSA flows and MSA flows, respectively. If a PSA flow is actually a gradient flow for a cost function f , as is the case for (1), (2) and (4), then the corresponding dual MSA flow is a gradient flow for the Legendre dual cost function f^* of f .

When implementing these differential equations on a computer, suitable discretizations are to be found. Since we are working in unconstrained Euclidean matrix space $\mathbb{R}^{n \times p}$, we consider Euler step discretizations. Thus, e.g., for system (1) consider

$$X_{t+1} = X_t - s_t(I - X_t X_t') C X_t, \quad (5)$$

with suitably small step sizes. Such Euler discretization schemes are widely used in the literature, but usually without explicit step-size selections that guarantee, for generic initial conditions, convergence to the p dominant orthonormal eigenvectors of A . A further challenge is to obtain step-size selections that achieve quadratic convergence rates (e.g., via a Newton-type approach).

¹It is remarked that the convergence proof in [5] appears flawed; they argue that because $\frac{d \operatorname{vec} Q}{dt} = G(t) \operatorname{vec} Q$ for some matrix $G(t) < 0$ then $Q \rightarrow 0$. However, counterexamples are known [15] where $G(t)$ is strictly negative definite (with constant eigenvalues) yet Q diverges.

Problem 2. Develop a systematic convergence theory for discretisations of the flows, by specifying step-size selections that imply global as well as local quadratic convergence to the equilibria.

2 MOTIVATION AND HISTORY

Eigenvalue computations are ubiquitous in Mathematics and Engineering Sciences. In applications, the matrices whose eigenvectors are to be found are often defined in a recursive way, thus demanding recursive computational methods for eigendecomposition. Subspace tracking algorithms are widely used in neural networks, regression analysis, and signal processing applications for this purpose. Subspace tracking algorithms can be studied by replacing the stochastic, recursive algorithm through an averaging procedure by a nonlinear ordinary differential equation. Similarly, new subspace tracking algorithms can be developed by starting with a suitable ordinary differential equation and then converting it to a stochastic approximation algorithm [7]. Therefore, understanding the dynamics of such flows is paramount to the continuing development of recursive eigendecomposition techniques.

The starting point for most of the current work in principal component analysis and subspace tracking has been Oja's system from neural network theory. Using a simple Hebbian law for a single perceptron with a linear activation function, Oja [9, 10] proposed to update the weights according to

$$X_{t+1} = X_t - s_t(I - X_t X_t') u_t u_t' X_t. \quad (6)$$

Here X_t denotes the $n \times p$ weight matrix and u_t the input vector of the perceptron, respectively. By applying the ODE method to this system, Oja arrives at the differential equation (1). Here $C = E(u_t u_t')$ is the covariance matrix. Similarly, the other flows, (2) and (4), have analogous interpretations.

In [9, 11] it is shown for $p = 1$ that (1) is a PSA flow, i.e., it converges for generic initial conditions to a normalised dominant eigenvector of C . In [11] the system (1) was studied for arbitrary values of p and it was conjectured that (1) is a PSA flow. This conjecture was first proven in [18], assuming positive definiteness of C . Moreover, in [18, 4], explicit initial conditions in terms of intersection dimensions for the dominant eigenspace with the initial subspace were given, such that the flow converges to a basis matrix of the p -dimensional dominant eigenspace. This is reminiscent of Schubert type conditions in Grassmann manifolds.

Although the Oja flow serves as a principal *subspace* method, it is not useful for principal *component* analysis because it does not converge in general to a basis of eigenvectors. Flows for principal component analysis such as (2) have been first studied in [14, 12, 13, 17]. However, pointwise convergence to the equilibria points was not established. In [16] a Lyapunov function for the Oja flow (1) was given, but without recognizing that (1) is actually a gradient

flow. There have been confusing remarks in the literature claiming that (1) cannot be a gradient system as the linearization is not a symmetric matrix. However, this is due to a misunderstanding of the concept of a gradient. In [20] it is shown that (2), and in particular (1), is actually a gradient flow for the cost function $f(X) = 1/4\text{tr}(AXNX')^2 - 1/2\text{tr}(A^2XD^2X')$ and a suitable Riemannian metric on $\mathbb{R}^{n \times p}$. Moreover, starting from any initial condition in $\mathbb{R}^{n \times p}$, pointwise convergence of the solutions to a basis of k independent eigenvectors of A is shown together with a complete phase portrait analysis of the flow. First steps toward a phase portrait analysis of (4) are made in [8].

3 AVAILABLE RESULTS

In [12, 13, 17] the equilibrium points of (2) were computed together with a local stability analysis. Pointwise convergence of the system to the equilibria is established in [20] using an early result by Lojasiewicz on real analytic gradient flows. Thus these results together imply that (2), and hence (1), is a PSA. An analogous result for (4) is forthcoming; see [8] for first steps in this direction. Sufficient conditions for initial matrices in the Oja flow (1) to converge to a dominant subspace basis are given in [18, 4], but not for the other, unstable equilibria, nor for system (2). A complete characterization of the stable/unstable manifolds is currently unknown.

BIBLIOGRAPHY

- [1] R. W. Brockett, "Dynamical systems that sort lists, diagonalise matrices, and solve linear programming problems," *Linear Algebra Appl.*, 146:79–91, 1991.
- [2] T. Chen, "Modified Oja's algorithms for principal subspace and minor subspace extraction," *Neural Processing Letters*, 5:105–110, April 1997.
- [3] T. Chen, S. Amari, and Q. Lin, "A unified algorithm for principal and minor component extraction," *Neural Networks*, 11:385–390, 1998.
- [4] T. Chen, Y. Hua, and W.-Y. Yan, "Global convergence of Oja's subspace algorithm for principal component extraction," *IEEE Transactions on Neural Networks*, 9(1):58–67, 1998.
- [5] S. C. Douglas, S.-Y. Kung, and S. Amari, "A self-stabilized minor subspace rule," *IEEE Signal Processing Letters*, 5(12):328–330, December 1998.
- [6] U. Helmke and J. B. Moore, *Optimization and Dynamical Systems*, Springer-Verlag, 1994.

- [7] H. J. Kushner and G. G. Yin, *Stochastic Approximation Algorithms and Applications*, Springer, 1997.
- [8] J. H. Manton, I. M. Y. Mareels, and S. Attallah, "An analysis of the fast subspace tracking algorithm NOja," In: *IEEE Conference on Acoustics, Speech and Signal Processing*, Orlando, Florida, May 2002.
- [9] E. Oja, "A simplified neuron model as a principal component analyzer," *Journal of Mathematical Biology*, 15:267–273, 1982.
- [10] E. Oja, "Neural networks, principal components, and subspaces," *International Journal of Neural Systems*, 1:61–68, 1989.
- [11] E. Oja and J. Karhunen, "On stochastic approximation of the eigenvectors and eigenvalues of the expectation of a random matrix," *Journal of Mathematical Analysis and Applications*, 106:69–84, 1985.
- [12] E. Oja, H. Ogawa, and J. Wangviwattana, "Principal component analysis by homogeneous neural networks, Part I: The weighted subspace criterion," *IEICE Transactions on Information and Systems*, 3:366–375, 1992.
- [13] E. Oja, H. Ogawa, and J. Wangviwattana, "Principal component analysis by homogeneous neural networks, Part II: Analysis and extensions of the learning algorithms," *IEICE Transactions on Information and Systems*, 3:376–382, 1992.
- [14] T. D. Sanger, "Optimal unsupervised learning in a single-layer linear feedforward network," *Neural Networks*, 2:459–473, 1989.
- [15] J. L. Willems, *Stability Theory of Dynamical Systems*, Studies in Dynamical Systems, London, Nelson, 1970.
- [16] J. L. Wyatt and I. M. Elfadel, "Time-domain solutions of Oja's equations," *Neural Computation*, 7:915–922, 1995.
- [17] L. Xu, "Least mean square error recognition principle for self organizing neural nets," *Neural Networks*, 6:627–648, 1993.
- [18] W.-Y. Yan, U. Helmke, and J. B. Moore, "Global analysis of Oja's flow for neural networks," *IEEE Transactions on Neural Networks*, 5(5):674–683, September 1994.
- [19] B. Yang, "Projection approximation subspace tracking", *IEEE Transactions on Signal Processing*, 43(1):95–107, January 1995.
- [20] S. Yoshizawa, U. Helmke, and K. Starkov, "Convergence analysis for principal component flows," *International Journal of Applied Mathematics and Computer Science*, 11:223–236, 2001.

PART 4

Discrete Event, Hybrid Systems

Problem 4.1

\mathcal{L}_2 -induced gains of switched linear systems

João P. Hespanha¹

Dept. of Electrical and Computer Engineering
University of California, Santa Barbara
USA
hespanha@ece.ucsb.edu

1 SWITCHED LINEAR SYSTEMS

In the 1999 collection of Open Problems in Mathematical Systems and Control Theory, we proposed the problem of computing input-output gains of switched linear systems. Recent developments provided new insights into this problem leading to new questions.

A *switched linear system* is defined by a parameterized family of realizations $\{(A_p, B_p, C_p, D_p) : p \in \mathcal{P}\}$, together with a family of piecewise constant *switching signals* $\mathcal{S} := \{\sigma : [0, \infty) \rightarrow \mathcal{P}\}$. Here we consider switched systems for which all the matrices A_p , $p \in \mathcal{P}$ are Hurwitz. The corresponding switched system is represented by

$$\dot{x} = A_\sigma x + B_\sigma u, \quad y = C_\sigma x + D_\sigma u, \quad \sigma \in \mathcal{S} \quad (1)$$

and by a *solution* to (1), we mean a pair (x, σ) for which $\sigma \in \mathcal{S}$ and x is a solution to the time-varying system

$$\dot{x} = A_{\sigma(t)}x + B_{\sigma(t)}u, \quad y = C_{\sigma(t)}x + D_{\sigma(t)}u, \quad t \geq 0. \quad (2)$$

Given a set of switching signals \mathcal{S} , we define the \mathcal{L}_2 -induced gain of (1) by

$$\inf\{\gamma \geq 0 : \|y\|_2 \leq \gamma \|u\|_2, \forall u \in \mathcal{L}_2, x(0) = 0, \sigma \in \mathcal{S}\},$$

where y is computed along solutions to (1). The \mathcal{L}_2 -induced gain of (1) can be viewed as a “worst case” energy amplification gain for the switched system, *over all possible inputs and switching signals* and is an important tool to study the performance of switched systems, as well as the stability of interconnections of switched systems.

¹This material is based upon work supported by the National Science Foundation under Grant No. ECS-0093762.

2 PROBLEM DESCRIPTION

We are interested here in families of switching signals for which consecutive discontinuities are separated by no less than a positive constant called the *dwell-time*. For a given $\tau_D > 0$, we denote by $\mathcal{S}[\tau_D]$ the set of piecewise constant switching signals with interval between consecutive discontinuities no smaller than τ_D . The general problem that we propose is the computation of the function $\mathbf{g} : [0, \infty) \rightarrow [0, \infty]$ that maps each *dwell-time* τ_D with the \mathcal{L}_2 -induced gain of (1), for the set of dwell-time switching signals $\mathcal{S} := \mathcal{S}[\tau_D]$. Until recently? little more was known about \mathbf{g} other than the following:

1. \mathbf{g} is monotone decreasing
2. \mathbf{g} is bounded below by

$$\mathbf{g}_{\text{static}} := \sup_{p \in \mathcal{P}} \|C_p(sI - A_p)^{-1}B_p + D_p\|_{\infty},$$

where $\|T\|_{\infty} := \sup_{\Re[s] \geq 0} \|T(s)\|$ denotes the \mathcal{H}_{∞} -norm of a transfer matrix T . We recall that $\|T\|_{\infty}$ is numerically equal to the \mathcal{L}_2 -induced gain of any linear time-invariant system with transfer matrix T .

Item 1 is a trivial consequence of the fact that given two dwell-times $\tau_{D_1} \leq \tau_{D_2}$, we have that $\mathcal{S}[\tau_{D_1}] \supset \mathcal{S}[\tau_{D_2}]$. Item 2 is a consequence of the fact that every set $\mathcal{S}[\tau_D]$, $\tau_D > 0$ contains all the constant switching signals $\sigma = p$, $p \in \mathcal{P}$. It was shown in [2] that the lower-bound $\mathbf{g}_{\text{static}}$ is strict and in general there is a gap between $\mathbf{g}_{\text{static}}$ and

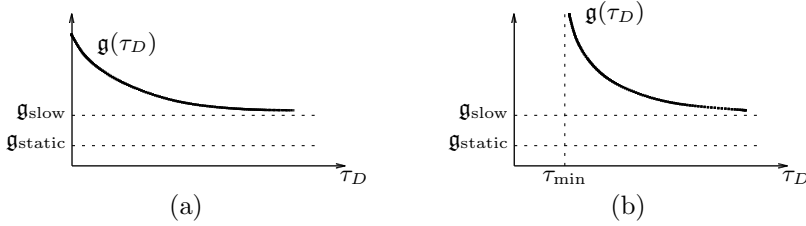
$$\mathbf{g}_{\text{slow}} := \lim_{\tau_D \rightarrow \infty} \mathbf{g}[\tau_D].$$

This means that even switching arbitrarily seldom, one may not be able to recover the \mathcal{L}_2 -induced gains of the “unswitched systems.” In [2] a procedure was given to compute \mathbf{g}_{slow} . Opposite to what had been conjectured, \mathbf{g}_{slow} is realization dependent and cannot be determined just from the transfer functions of the systems being switched.

The function \mathbf{g} thus looks roughly like the ones shown in figure 4.1.1, where (a) corresponds to a set of realizations that remains stable for arbitrarily fast switching and (b) to a set that can exhibit unstable behavior for sufficiently fast switching [3]. In (b), the scalar τ_{\min} denotes the smallest dwell-time for which instability can occur for some switching signal in $\mathcal{S}[\tau_{\min}]$.

Several important basic questions remain open:

1. Under what conditions is \mathbf{g} bounded? This is really a stability problem whose general solution has been eluding researchers for a while now (cf., the survey paper [3] and references therein).
2. In case \mathbf{g} is unbounded (case (b) in figure 4.1.1), how to compute the position of the vertical asymptote? Or, equivalently, what is the smallest dwell-time τ_{\min} for which one can have instability?

Figure 4.1.1 \mathcal{L}_2 -induced gain versus the dwell-time.

3. Is g a convex function? Is it smooth (or even continuous)?

Even if direct computation of g proves to be difficult, answers to the previous questions may provide indirect methods to compute tight bounds for it. They also provide a better understanding of the trade-off between switching speed and induced gain. As far as we know, currently only very coarse upper-bounds for g are available. These are obtained by computing a conservative upper-bound τ_{upper} for τ_{\min} and then an upper-bound for g that is valid for every dwell-time larger than τ_{upper} (cf., e.g., [4, 5]). These bounds do not really address the trade-off mentioned above.

BIBLIOGRAPHY

- [1] J. P. Hespanha and A. S. Morse, "Input-output gains of switched linear systems," In: *Open Problems in Mathematical Systems Theory and Control*, V. D. Blondel, E. D. Sontag, M. Vidyasagar, and J. C. Willems, eds., London: Springer-Verlag, 1999.
- [2] J. P. Hespanha, "Computation of root-mean-square gains of switched linear systems," presented at the *Fifth Hybrid Systems: Computation and Control Workshop*, Mar. 2002.
- [3] D. Liberzon and A. S. Morse, "Basic problems in stability and design of switched systems," *IEEE Contr. Syst. Mag.*, vol. 19, pp. 59–70, Oct. 1999.
- [4] J. P. Hespanha and A. S. Morse, "Stability of switched systems with average dwell-time," In: *Proc. of the 38th Conf. on Decision and Contr.*, pp. 2655–2660, Dec. 1999.
- [5] G. Zhai, B. Hu, K. Yasuda, and A. N. Michel, "Disturbance attenuation properties of time-controlled switched systems," submitted to publication, 2001.

Problem 4.2

The state partitioning problem of quantised systems

Jan Lunze

Institute of Automation and Computer Control

Ruhr-University Bochum

D-44780 Bochum

Germany

Lunze@esr.ruhr-uni-bochum.de

1 DESCRIPTION OF THE PROBLEM

Consider a continuous system whose state can only be accessed through a quantizer. The quantizer is defined by a partition of the state space. The system generates an event if the system trajectory crosses the boundary between adjacent partitions.

The problem concerns the prediction of the event sequence generated by the system for a given initial event. As the initial event does not define the initial system state unambiguously but only restricts the initial state to a partition boundary, when predicting the system behavior the bundle of all state trajectories have to be considered that start on this partition boundary. The question to be answered is: under what conditions on the vector field of the system and the state partition is the event sequence unique?

In more detail, consider the continuous-variable system

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}(t)), \quad \mathbf{x}(0) = \mathbf{x}_0 \quad (1)$$

with the state $\mathbf{x} \in \mathcal{X} \subseteq \mathcal{R}^n$. The vector field \mathbf{f} satisfies a Lipschitz condition so that eqn. (1) has, for all $\mathbf{x}_0 \in \mathcal{X}$, a unique solution.

The state space \mathcal{X} is partitioned into N disjoint sets $\mathcal{Q}_x(i)$ ($i = 1, 2, \dots, N$) that satisfy the conditions

$$\mathcal{X} = \bigcup_{i=1}^N \mathcal{Q}_x(i) \quad \text{and} \quad \mathcal{Q}_x(i) \cap \mathcal{Q}_x(j) = \emptyset \text{ for } i \neq j.$$

The set

$$\mathcal{Q} = \{\mathcal{Q}_x(i) : i = 1, 2, \dots, N\}$$

is called a state quantization. The quantized state is denoted by $[\mathbf{x}]$ and defined by

$$[\mathbf{x}] = i \Leftrightarrow \mathbf{x} \in Q_x(i). \quad (2)$$

The change of the quantized state is called an *event*, where the event e_{ij} occurs at time \bar{t} if the relations

$$[\mathbf{x}(\bar{t} + \delta t)] = i \text{ and } [\mathbf{x}(\bar{t} - \delta t)] = j$$

hold for small $\delta t > 0$. Hence, at time \bar{t} the state \mathbf{x} is on the boundary between the state partitions $Q_x(i)$ and $Q_x(j)$

$$\mathbf{x}(\bar{t}) \in \delta Q_x(i) \cap Q_x(j),$$

where δQ_x denotes the hull of Q_x . The system (1) together with the quantization Q is called the *quantized system*.

For given initial state \mathbf{x}_0 the system (1) generates, for the time interval $[0, T]$, a unique state trajectory $\mathbf{x}(\mathbf{x}_0, t)$ and, hence, a unique event sequence

$$E = (e_0, e_1, \dots, e_H) = \text{Quant}(\mathbf{x}(\mathbf{x}_0, t)),$$

which formally can be represented as the result of the operator Quant applied to the state trajectory. H is the number of events generated by the system within the time interval $[0, T]$. The following considerations concern only those initial events e_0 for which the quantized system generates an event sequence with $H > 1$.

If instead of the initial state \mathbf{x}_0 only the initial event $e_0 = e_{ij}$ is given, the initial state is only known to lie on the boundary $\delta Q_x(i) \cap Q_x(j)$ between the state partitions $Q_x(i)$ and $Q_x(j)$. Consequently, the bundle of trajectories starting in all these initial states have to be considered. These trajectories yield the set

$$\mathcal{E}(e_0) = \{E = \text{Quant}(\mathbf{x}(\mathbf{x}_0, t)) \text{ for } \mathbf{x}_0 \in \delta Q_x(i) \cap Q_x(j)\}$$

of event sequences. If the set \mathcal{E} has more than one element, the quantized system is nondeterministic in the sense that the knowledge of the initial event e_0 is not sufficient to predict the future event sequence unambiguously. On the other hand, the quantized system is called to be *deterministic* if the set $\mathcal{E}(e_0)$ is a singleton for all possible initial events e_0 .

In order to define the events precisely, the state partition should satisfy the following assumptions:

- A1.** The trajectories do not lie in the hypersurfaces that represent the partition boundaries.
- A2.** The system cannot generate an infinite number of events in a finite time interval.
- A3.** No fix-point of the vector field \mathbf{f} lie on a partition boundary.

These assumptions can be satisfied by appropriately defining the state partitions for the given vector field \mathbf{f} .

State partitioning problem. Find conditions under which the quantised system is deterministic.

This problem can be reformulated in two versions:

Problem: For given vector field \mathbf{f} , find a partition of the state space such that the quantized system is deterministic.

Problem B: For given vector field \mathbf{f} and a state quantization \mathcal{Q} , test whether the quantized system is deterministic.

Both formulations have their engineering relevance. Where problem A concerns the practical situation in which a state partition has to be selected, problem B refers to the test of the determinism of the system for given partition.

The problem stated so far is, possibly, too general in two respects. First, the problem for testing the determinism of the system should be as simple as possible. For a given partition consisting of N disjoint sets, Problem B can be solved by considering all trajectory bundles that start on all partition boundaries. Here, the characterization of classes of vector fields \mathbf{f} and partitioning methods is interesting for which the complexity of the test is constant or grows only linearly with N . Second, for problem A it is interesting to find partitions that can be distinguished with only a few measurements. For example, rectangular partitions are interesting from a practical viewpoint which result from separate quantizations of all n state variables x_i .

Nonautonomous systems.

The problem can be extended to nonautonomous quantized systems

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}(t), \mathbf{u}(t)), \quad \mathbf{x}(0) = \mathbf{x}_0 \quad (3)$$

$$\mathbf{y}(t) = \mathbf{g}(\mathbf{x}(t), \mathbf{u}(t)) \quad (4)$$

with input $\mathbf{u} \in \mathcal{U} \subseteq \mathcal{R}^m$ and output $\mathbf{y} \in \mathcal{Y} \subseteq \mathcal{R}^r$. The functions \mathbf{f} and \mathbf{g} satisfy a Lipschitz condition so that eqns. (3), (4) have, for all $\mathbf{x}_0 \in \mathcal{X}$ and $\mathbf{u}(t)$, a unique solution. The output quantizer is defined by a partition of the output space \mathcal{Y} into the sets $\mathcal{Q}_y(i)$ where the quantized output $[\mathbf{y}]$ is defined analogously to equation (2). The event sequence E is now defined in terms of the events that the output signal \mathbf{y} generates. The system is considered with the quantized input $[\mathbf{u}]$. An injector associates with each input a unique element of the finite discrete set

$$\mathcal{U} = \{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_M\}$$

such that $\mathbf{u}(t) = \mathbf{u}_i$ if $[\mathbf{u}(t)] = i$. Again, a change of the quantized input value is called an (input) event. It is assumed that the input and output events occur synchronously. This assumption fixes the time instances in which the input changes its value. It is motivated by the fact that in closed-loop

systems a supervisor defines the quantized input in the same time instant in which an output event occurs.

Here the state partitioning problem includes also to define an output partition and an input set \mathcal{U} such that the quantized system is deterministic for all input sequences.

2 MOTIVATION AND HISTORY OF THE PROBLEM

The problem results from hybrid systems, whose simplest form is a continuous-variable system with discrete inputs. Many technological systems that are controlled by programmable logic controllers (PLC) have a continuous state space and are controlled by discrete inputs. The contrast of the continuous state and the discrete input does not matter because many systems are designed in such a way that any accessible input results in an unambiguous state or output event.

For example, the (simplified) state space of a lift has the state variables “vehicle position,” and “door position” both of which are quantized where the vehicle position refers to the floor in which it stops and the two discrete door position are called “open” or “closed.” For the performance of this system, only the events are important, which refer to the beginning and the end of the presence of the vehicle or the door in one of these positions. As the PLC can only switch on or off, the motors of the vehicle or the door and it is programmed so that the next command is given only after the next output event has occurred, every new input event is followed by exactly one output event (unless the system is faulty). So, the lift is a continuous-variable system (3), (4) with quantized input and output, that is deterministic.

In this case, the solution to the state partitioning problem is simple. The determinism of the quantized system results from the fact that the system trajectories are parallel to the coordinate axes of the state space for all accessible inputs and the quantization refers to separate intervals of both state variables. So, the end point of any movement initiated by a PLC command is a *point* in the state space and every trajectory of the closed-loop system results in precisely one output event.

In a more general setting, continuous-variable systems are dealt with as quantized systems for process supervision tasks. Then the system is not designed to behave like a discrete-event system but has a continuous state space. The quantizers are introduced deliberately to reduce the information to be processed. For example, alarm messages show that a certain signal has exceeded a threshold. The state partitioning problem asks for the a choice of discrete sensors such that the system behavior is deterministic.

As the third motivation for the state partitioning problem, hybrid systems theory concerns dynamical systems with continuous-variable and discrete-event subsystems. The interfaces between both parts are the quantizer and

the injector introduced above that transform the discrete output signal of the discrete subsystem into a real-valued input signal of the continuous subsystem and vice versa. The problem occurs under what condition the overall hybrid system has a deterministic input-output behavior if only the discrete inputs and outputs of the discrete subsystem are considered. The main source of nondeterminism results from the quantization of the signal space of the continuous subsystem, which again leads to the state partitioning problem.

In all these situations, the discrete behavior of a continuous system is considered. In the literature on fault diagnosis and verification of discrete control algorithms the hybrid nature of the closed-loop system is removed by using a discrete-event representation of the quantized system. As in many practical situations the quantizers can be chosen, the state partitioning problem asks for guidelines of this selection. For a deterministic discrete behavior, a deterministic model can be used to describe the quantised system. If, however, the discrete behavior is non-deterministic, a nondeterministic model like a nondeterministic or stochastic automaton or a Petri net has to be used. Several ways for determining such models for a given quantized system have been elaborated recently ([3], [4], [6], [7], [8], [9]).

3 AVAILABLE RESULTS

The first result on the state partitioning problem concerns discrete-time systems (rather than continuous-time systems) with quantized state space. Reference [4] gives a necessary and sufficient condition for the determinism of the discrete behavior for linear autonomous systems with a state space partition that regularly decomposes each state variable into intervals of the same size. In [5] it has been shown how state partitions can be generated by mapping a given initial set $\mathcal{Q}_x(1)$ by the model (1) that is used with reversed time axis.

For the problem stated here, only preliminary results are available. If the system trajectories are, like in the lift example, parallel to the coordinate axes of the state space and the quantization boundaries define rectangular cells whose axes are parallel to the coordinate axes, the discrete behavior is deterministic. This situation is encountered if, for example, the state variables are decoupled and controlled by separate inputs. Hence, the model can be decomposed into

$$\begin{aligned}\dot{x}_i &= f_i(x_i, u_i) \\ y_i &= g_i(x_i),\end{aligned}$$

which corresponds again to the simple lift example. Another example is an undamped oscillator with a state partition that decomposes the state space into the two half-planes. Then the fix-point lies on the partition boundary (and, thus, violates assumption A3). However, the oscillator generates, for

each initial state, a unique (alternating) event sequence.

Results on symbolic dynamics are closely related to the problem stated here (cf. [1], [2]). A bundle of trajectories (or flows) is considered, which generate a symbolic output if some partition boundary is crossed. The partition is called Markovian if all trajectories of the bundle cross the same partition boundary and, hence, generate the same symbol. In the terminology used there, the problem posed here asks the question how to find *Markovian* partitions.

BIBLIOGRAPHY

- [1] A. Lasota, M. C. Mackey, *Chaos, Fractals and Noise*, Springer-Verlag, New York, 1994.
- [2] D. Lind, B. Marcus, *An Introduction to Symbolic Dynamics and Coding*, Cambridge University Press, 1995.
- [3] J. Lunze, "A Petri net approach to qualitative modelling of continuous dynamical systems," *Systems Analysis, Modelling, Simulation*, 9, pp. 885-903, 1992.
- [4] J. Lunze, "Qualitative modelling of linear dynamical systems with quantized state measurements," *Automatica*, 30, pp. 417-431, 1994.
- [5] J. Lunze, B. Nixdorf, J. Schröder, "On the nondeterminism of discrete-event representations of continuous-variables systems," *Automatica*, 35, pp. 395-408, 1999.
- [6] H.A. Preisig, M.J.H. Pijpers, M. Weiss, "A discrete modelling procedure for continuous processes based on state-discretisation," *2n MATHMOD*, Vienna, pp. 189-194, 1997.
- [7] J. Raisch, S. O'Young, "A totally ordered set of discrete abstractions for a given hybrid or continuous system," pp. 342-360 In: P. Antsaklis, W. Kohn, A. Nerode, S. Sastry, eds., *Hybrid Systems II*, Springer-Verlag, Berlin, 1995.
- [8] J. Schröder, *Modelling, State Observation and Diagnosis of Quantised Systems*, Springer-Verlag, Heidelberg, 2002.
- [9] O. Stursberg, S. Kowalewski, S. Engell, "Generating timed discrete models," *2n MATHMOD*, Vienna, pp. 203-207, 1997.

Problem 4.3

Feedback control in flowshops

S. P. Sethi

School of Management

The University of Texas at Dallas

Box 830688, Mail Station JO4.7

Richardson, TX 75083

USA

sethi@utdallas.edu

Q. Zhang

Department of Mathematics

University of Georgia

Athens, GA 30602

USA

qingz@math.uga.edu

1 DESCRIPTION OF THE PROBLEM

Consider a manufacturing system producing a single finished product using m machines in tandem that are subject to breakdown and repair. We are given a finite-state Markov chain $\alpha(\cdot) = (\alpha_1(\cdot), \dots, \alpha_m(\cdot))$ on a probability space (Ω, \mathcal{F}, P) , where $\alpha_i(t)$, $i = 1, \dots, m$, is the capacity of the i -th machine at time t . We use $u_i(t)$ to denote the input rate to the i -th machine, $i = 1, \dots, m$, and $x_i(t)$ to denote the number of parts in the buffer between the i -th and $(i+1)$ -th machines, $i = 1, \dots, m-1$. Finally, the surplus is denoted by $x_m(t)$. The dynamics of the system can then be written as follows:

$$\dot{x}(t) = Au(t) + Bz, x(0) = x, \quad (1)$$

where z is the rate of demand and

$$A = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \text{ and } B = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ -1 \end{pmatrix}.$$

Since the number of parts in the internal buffers cannot be negative, we impose the state constraints $x_i(t) \geq 0, i = 1, \dots, m-1$. To formulate the problem precisely, let $S = [0, \infty)^{m-1} \times (-\infty, \infty) \subset \mathbb{R}^m$ denote the state constraint domain. For $\alpha = (\alpha_1, \dots, \alpha_m) \geq 0$, let

$$U(\alpha) = \{u = (u_1, \dots, u_m) : 0 \leq u_i \leq \alpha_i, i = 1, \dots, m\},$$

and for $x \in S$, let

$$U(x, \alpha) = \{u : u \in U(\alpha); x_i = 0 \Rightarrow u_i - u_{i+1} \geq 0, i = 1, \dots, m-1\}.$$

Let $\mathcal{M} = \{\alpha^1, \dots, \alpha^p\}$ for a given integer $p \geq 1$, where $\alpha^j = (\alpha_1^j, \dots, \alpha_m^j)$ with α_i^j denoting the possible capacity states of the i -th machine, $i = 1, \dots, m$. Let the σ -algebra $\mathcal{F}_t = \sigma\{\alpha(s) : 0 \leq s \leq t\}$.

Definition 1: A control $u(\cdot)$ is *admissible* with respect to the initial state $x \in S$ and $\alpha \in \mathcal{M}$ if: (i) $u(\cdot)$ is $\{\mathcal{F}_t\}$ -adapted, (ii) $u(t) \in U(\alpha(t))$ for all $t \geq 0$, and (iii) the corresponding state process $x(t) = (x_1(t), \dots, x_m(t)) \in S$ for all $t \geq 0$.

Let $\mathcal{A}(x, \alpha)$ denote the set of admissible controls.

Definition 2: A function $u(x, \alpha)$ is called a feedback control, if (i) for any given initial x , the equation (1) has a unique solution; and (ii) $u(\cdot) = \{u(t) = u(x(t), \alpha(t)), t \geq 0\} \in \mathcal{A}(x, \alpha)$.

The problem is to find an admissible control $u(\cdot)$ that minimizes

$$J(x, \alpha, u(\cdot)) = E \int_0^\infty e^{-\rho t} G(x(t), u(t)) dt, \quad (2)$$

where $G(x, u)$ defines the cost of surplus x and production u , α is the initial value of $\alpha(t)$, and $\rho > 0$ is the discount rate. We assume that $G(x, u) \geq 0$ is jointly convex and locally Lipschitz.

The value function is then defined as

$$v(x, \alpha) = \inf_{u(\cdot) \in \mathcal{A}(x, \alpha)} J(x, \alpha, u(\cdot)). \quad (3)$$

The optimal control of this problem was considered in [1] using HJB equations with directional derivatives. It is shown that there exists a unique optimal control. In addition, a verification theorem associated with the HJB equations is obtained. However, these HJB equations are difficult to solve numerically, especially when the state space of \mathcal{M} is large. In this case, it is desirable to derive an approximate solution instead. We consider the case when $\alpha(\cdot)$ jumps rapidly. In particular, we assume $\alpha(t) = \alpha^\varepsilon(t) \in \mathcal{M}$, $t \geq 0$, to be a Markov chain with the generator

$$Q^\varepsilon = \frac{1}{\varepsilon} \tilde{Q} + \hat{Q},$$

where $\tilde{Q} = (\tilde{q}_{ij})$ and $\hat{Q} = (\hat{q}_{ij})$ are generator matrices and \tilde{Q} is weakly irreducible. Here ε is a small parameter. We use \mathcal{P}^ε to denote our control problem. As ε gets smaller and smaller, one expects that \mathcal{P}^ε approaches to a

limiting problem. To obtain such limiting problem, let $\nu = (\nu^1, \dots, \nu^p)$ denote the equilibrium distribution of \tilde{Q} . We consider the class of deterministic controls defined below.

Definition 3: For $x \in S$, let $\mathcal{A}^0(x)$ denote the set of the following measurable controls

$$U(\cdot) = (u^1(\cdot), \dots, u^p(\cdot)) = ((u_1^1(\cdot), \dots, u_m^1(\cdot)), \dots, (u_1^p(\cdot), \dots, u_m^p(\cdot)))$$

such that $0 \leq u_i^j(t) \leq \alpha_i^j$ for all $t \geq 0$, $i = 1, \dots, n$ and $j = 1, \dots, p$, and the corresponding solutions $x(\cdot)$ of the system

$$\dot{x}(t) = A \sum_{j=1}^p \nu^j u^j(t) + Bz, \quad x(0) = x$$

satisfy $x(t) \in S$ for all $t \geq 0$.

The objective of the limiting problem is to choose a control $U(\cdot) \in \mathcal{A}^0(x)$ that minimizes

$$J^0(x, U(\cdot)) = \int_0^\infty e^{-\rho t} \sum_{j=1}^p \nu^j G(x(t), u^j(t)) dt.$$

We use \mathcal{P}^0 to denote the limiting problem and $v^0(x)$ the corresponding value function.

2 MOTIVATION AND HISTORY OF THE PROBLEM

It is shown in [1] that the value function $v^\varepsilon(x, \alpha)$ converges to $v^0(x)$ as $\varepsilon \rightarrow 0$. The limiting problem is much easier to solve. The goal is to use the solution of the limiting problem to construct a control for the original problem that is nearly optimal.

3 AVAILABLE RESULTS

The idea is to use an optimal (or a near optimal) control to construct a control for the original problem \mathcal{P}^ε . The main difficulty is how to construct an *admissible* control for \mathcal{P}^ε in a way that still guarantees the asymptotic optimality as ε goes to zero. Partial results were obtained using a “lifting” and “modification” approach. This was applied to open-loop controls; see [1]. Construction of an asymptotic optimal feedback control remains open. A resolution of this problem would perhaps also apply to a more complicated jobshop considered in [1].

BIBLIOGRAPHY

- [1] S. P. Sethi and Q. Zhang, *Hierarchical Decision Making in Stochastic Manufacturing Systems*, Birkhäuser, Boston, 1994.

Problem 4.4

Decentralized control with communication between controllers

Jan H. van Schuppen

CWI

P.O. Box 94079, 1090 GB Amsterdam

The Netherlands

`J.H.van.Schuppen@cwi.nl`

1 DESCRIPTION OF THE PROBLEM

Problem 1: Decentralized control with communication between controllers

Consider a control system with inputs from r different controllers. Each controller has partial observations of the system and the partial observations of each pair of controllers is different. The controllers are allowed to exchange online information on their partial observations, state estimates, or input values, but there are constraints on the communication channels between each tuple of controllers. In addition, there is specified a control objective.

The problem is to synthesize r controllers and a communication protocol for each directed tuple of controllers, such that when the controllers all use their received communications the control objective is met as well as possible.

The problem can be considered for a discrete-event system in the form of a generator, for a timed discrete-event system, for a hybrid system, for a finite-dimensional linear system, for a finite-dimensional Gaussian system, etc. In each case, the communication constraint has to be chosen and a formulation has to be proposed on how to integrate the received communications into the controller.

Remarks on problem

(1) The constraints on the communication channels between controllers are essential to the problem. Without it, every controller communicates all his/her partial observations to all other controllers and one obtains a control problem with a centralized controller, albeit one where each controller carries out the same control computations.

(2) The complexity of the problem is large, for control of discrete-event systems it is likely to be undecidable. Therefore, the problem formulation has to be restricted. Note that the problem is analogous to human communication in groups, firms, and organizations and that the communication problems in such organizations are effectively solved on a daily basis. Yet there is scope for a fundamental study of this problem also for engineering control systems. The approach to the problem is best focused on the formulation and analysis of simple control laws and on the formulation of necessary conditions.

(3) The basic underlying problem seems to be: what information of a controller is so essential in regard to the control purpose that it has to be communicated to other controllers? A system theoretic approach is suitable for this.

(4) The problem will also be useful for the development of hierarchical models. The information to be communicated has to be dealt with at a global level, the information that does not need to be communicated can be treated at the local level.

To assist the reader with the understanding of the problem, the special cases for discrete-event systems and for finite-dimensional linear systems are stated below.

Problem 2: Decentralized control of a discrete-event system with communication between supervisors

Consider a discrete-event system in the form of a generator and $r \in \mathbb{Z}_+$ supervisors:

$$\begin{aligned}
 G &= (Q, E, f, q_0), \quad Q, \text{ the state set, } q_0 \in Q, \text{ the initial state,} \\
 &\quad E, \text{ the event set, } f : Q \times E \rightarrow Q, \text{ the transition function,} \\
 L(G) &= \{s \in E^* \mid f(q_0, s) \text{ is defined}\}, \\
 &\quad \forall k \in \mathbb{Z}_r = \{1, 2, \dots, r\}, \text{ a partition, } E = E_{c,k} \cup E_{uc,k}, \\
 E_{cp,k} &= \{E_e \subseteq E \mid E_{uc,k} \subseteq E_e\}, \\
 &\quad \forall k \in \mathbb{Z}_r, \text{ a partition, } E = E_{o,k} \cup E_{uo,k}, \quad p_k : E^* \rightarrow E_{o,k}^*, \quad \forall k \in \mathbb{Z}_r, \\
 &\quad \text{an event is enabled if it is enabled by all supervisors,} \\
 &\quad \{v_k : p_k(L(G)) \rightarrow E_{cp,k}, \forall k \in \mathbb{Z}_r\}, \\
 &\quad \text{the set of supervisors based on partial observations,} \\
 L_r, L_a &\subseteq L(G), \text{ required and admissible language, respectively.}
 \end{aligned}$$

The problem or better, a variant of it, is to determine a set of subsets of the event set that represent the events to be communicated by each supervisor

to the other supervisors and a set of supervisors,

$$\forall (i, j) \in \mathbb{Z}_r \times \mathbb{Z}_r, E_{o,i,j} \subseteq E_{o,i}, p_{i,j} : E \rightarrow E_{o,i,j},$$

the set of supervisors based on partial observations and on communications,

$$\{v_k(p_k(s), \{p_{j,k}(s), \forall j \in \mathbb{Z}_+ \setminus \{k\}\})\} \mapsto E_{cp,k}, \forall k \in \mathbb{Z}_r\};$$

is such that the closed-loop language, $L(v_1 \wedge \dots \wedge v_r/G)$, satisfies

$$L_r \subseteq L(v_1 \wedge \dots \wedge v_r/G) \subseteq L_a, \text{ and the controlled system is nonblocking.}$$

Problem 3: Decentralized control of a finite-dimensional linear system with communication between controllers

Consider a finite-dimensional linear system with $r \in \mathbb{Z}_+$ input signals and r output signals,

$$\begin{aligned} \dot{x}(t) &= Ax(t) + \sum_{k=1}^r B_k u_k(t), \quad x(t_0) = x_0, \\ y_j(t) &= C_j x(t) + \sum_{k=1}^r D_{j,k} u_k(t), \quad \forall j \in \mathbb{Z}_r = \{1, 2, \dots, r\}, \\ y_{s,j}(t) &= C_j(v_{s,j}(t))x(t), \end{aligned}$$

where $y_{j,s}$ represents the communication signal from Controller s to Controller j , where $v_{s,j}$ is the control input of Controller s for the communication to Controller j , and where the dimensions of the state, the input signals, the output signals, and of the matrices have been omitted. The i th controller observes output y_i and provides to the system input u_i . Suppose that Controller 2 communicates some components of his observed output signal to Controller 1. Can the system then be stabilized? How much can a quadratic cost be lowered by doing so? The problem becomes different if the communications from Controller 2 to Controller 1 are not continuous but are spaced periodically in time. How should the period be chosen for stability or for a cost minimization? The period will have to take account of the feedback achievable time constants of the system. A further restriction on the communication channel is to impose that messages can carry at most a finite number of bits. Then quantization is required. For a recent work on quantization in the context of control see, [17].

2 MOTIVATION

The problem is motivated by control of networks: for example, of communication networks, of telephone networks, of traffic networks, firms consisting of many divisions, etc. Control of traffic on the internet is a concrete example. In such networks, there are local controllers at the nodes of the network,

each having local information about the state of the network but no global information.

Decentralized control is used because it is technologically demanding and economically expensive to convey all observed informations to other controllers. Yet it is often possible to communicate information at a cost. This viewpoint has not been considered much in control theory. In the trade-off, the economic costs of communication have to be compared with the gains for the control objectives. This was already remarked on in the context of team theory a long time ago. But this has not been used in control theory till recently. The current technological developments make the communication relatively cheap and therefore the trade-off has shifted toward the use of more communication.

3 HISTORY OF THE PROBLEM

The decentralized control problem with communication between supervisors was formulated by the author of this paper around 1995. The plan for this problem is older, though, but there are no written records. With Kai C. Wong a necessary and sufficient condition was derived (see [20]) for the case of two controllers with asymmetric communication. The aspect of the problem that asks for the minimal information to be communicated was not solved in that paper. Subsequent research has been carried out by many researchers in control of discrete-event systems, including George Barrett, Rene Boel, Rami Debouk, Stephane Lafortune, Laurie Ricker, Karen Rudie, Demos Teneketzis; see [1, 2, 3, 4, 5, 11, 12, 13, 14, 15, 16, 19]. Besides the control problem, the corresponding problem for failure diagnosis has also been analyzed; see [6, 7, 8, 9]. The problem for failure diagnosis is simpler than that for control due to the fact that there is no relation of the diagnosing via the input to the future observations. The problem for timed discrete-event systems has been formulated also because in communication networks time delays due to communication need to be taken into account.

There are relations of the problem with team theory; see [10]. There are also relations with the asymptotic agreement problem in distributed estimation; see [18]. There are also relations of the problem to graph models and Bayesian belief networks where computations for large scale systems are carried out in a decentralized way.

4 APPROACH

Suggestions follow for the solution of the problem. Approaches are: (1) Exploration of simple algorithms. (2) Development of fundamental properties of control laws.

An example of a simple algorithm is the IEEE 802.11 protocol for wireless communication. The protocol prescribes stations when they can transmit and when not. All stations are in competition with each other for the available broadcasting time on a particular frequency. The protocol does not have a theoretical analysis and was not designed via a control synthesis procedure. Yet it is a beautiful example of a decentralized control law with communication between supervisors. The alternating bit protocol is another example. In a recent paper, S. Morse has analyzed another algorithm for decentralized control with communication based on a model for a school of fishes.

A more fundamental study will have to be directed at structural properties. Decentralized control theory is based on the concept of Nash equilibrium from game theory and on the concept of person-by-person optimality from team theory. The computation of an equilibrium is difficult because it is the solution of a fixpoint equation in function space. However, properties of the control law may be derived from the equilibrium equation, as is routinely done for optimal control problems.

Consider then the problem for a particular controller: it regards as the combined system the plant with the other controllers being fixed. The controller then faces the problem of designing a control law for the combined system. However, due to communication with other supervisors, it can in addition select components of the state vector of the combined system for its own observation process. A question then is which components to select. This approach leads to a set of equations, which, combined with those for other controllers, have to be solved.

Special cases of which the solution may point to generalizations are the case of two controllers with asymmetric communication and the case of three controllers. For larger number of controllers graph theory may be exploited but it is likely that simple algorithms will carry the day.

Constraints can be formulated in terms of information-like quantities as information rate, but this seems most appropriate for decentralized control of stochastic systems. Constraints can also be based on complexity theory as developed in computer science, where computations are counted. This case can be extended to counting bits of information.

BIBLIOGRAPHY

- [1] G. Barrett, *Modeling, Analysis and Control of Centralized and Decentralized Logical Discrete-Event Systems*, Ph.D. thesis, University of Michigan, Ann Arbor, 2000.
- [2] G. Barrett and S. Lafortune, "A novel framework for decentralized supervisory control with communication," In: *Proc. 1998 IEEE Systems, Man, and Cybernetics Conference*, New York, IEEE Press, 1998.

- [3] G. Barrett and S. Lafortune, "On the synthesis of communicating controllers with decentralized information structures for discrete-event systems," In: *Proceedings IEEE Conference on Decision and Control*, pp. 3281–3286, New York, IEEE Press, 1998.
- [4] G. Barrett and S. Lafortune, "Some issues concerning decentralized supervisory control with communication," In: *Proceedings 38th IEEE Conference on Decision and Control*, pp. 2230–2236, New York, IEEE Press, 1999.
- [5] G. Barrett and S. Lafortune, "Decentralized supervisory control with communicating controllers," *IEEE Trans. Automatic Control*, 45:1620–1638, 2000.
- [6] R. K. Boel and J. H. van Schuppen, "Decentralized failure diagnosis for discrete-event systems with costly communication between diagnosers," In: *Proceedings of International Workshop on Discrete Event Systems (WODES2002)*, pp. 175–181, Los Alamitos, IEEE Computer Society, 2002.
- [7] R. Debouk, *Failure diagnosis of decentralized discrete-event systems*, Ph.D. thesis, University of Michigan, Ann Arbor, 2000.
- [8] R. Debouk, S. Lafortune, and D. Teneketzis, "Coordinated decentralized protocols for failure diagnosis of discrete-event systems," *Discrete Event Dynamics Systems*, 10:33–86, 2000.
- [9] R. Debouk, S. Lafortune, and D. Teneketzis, "Coordinated decentralized protocols for failure diagnosis of discrete event systems," Report CGR-97-17, College of Engineering, University of Michigan, Ann Arbor, 1998.
- [10] R. Radner, "Allocation of a scarce resource under uncertainty: An example of a team," In: C. B. McGuire and R. Radner, eds, *Decision and organization*, pp. 217 – 236. North-Holland, Amsterdam, 1972.
- [11] S. Ricker and K. Rudie, "Know means no: Incorporating knowledge into decentralized discrete-event control," In: *Proc. 1997 American Control Conference*, 1997.
- [12] S. L. Ricker, *Knowledge and Communication in Decentralized Discrete-Event Control*, Ph.D. thesis, Queen's University, Department of Computing and Information Science, August 1999.
- [13] S. L. Ricker and G. Barrett, "Decentralized supervisory control with single-bit communications," In: *Proceedings of American Control Conference (ACC01)*, pp. 965–966, 2001.

- [14] S. L. Ricker and K. Rudie, "Incorporating communication and knowledge into decentralized discrete-event systems," In: *Proceedings 38th IEEE Conference on Decision and Control*, pp. 1326–1332, New York, IEEE Press, 1999.
- [15] S. L. Ricker and J. H. van Schuppen, "Asynchronous communication in timed discrete event systems," In: *Proceedings of the American Control Conference (ACC2001)*, pp. 305–306, 2001.
- [16] S. L. Ricker and J. H. van Schuppen, "Decentralized failure diagnosis with asynchronous communication between supervisors," In: *Proceedings of the European Control Conference (ECC2001)*, pp. 1002–1006, 2001.
- [17] S. C. Tatikonda, *Control under communication constraints*, Ph.D. thesis, Department of Electrical Engineering and Computer Science, MIT, Cambridge, MA, 2000.
- [18] D. Teneketzis and P. Varaiya, "Consensus in distributed estimation with inconsistent beliefs," *Systems & Control Lett.*, 4:217–221, 1984.
- [19] J. H. van Schuppen, "Decentralized supervisory control with information structures," In: *Proceedings International Workshop on Discrete Event Systems (WODES98)*, pp. 36–41, London, IEE, 1998.
- [20] K. C. Wong and J. H. van Schuppen, "Decentralized supervisory control of discrete-event systems with communication," In: *Proceedings International Workshop on Discrete Event Systems 1996 (WODES96)*, pp. 284–289, London, IEE, 1996.

PART 5

Distributed Parameter Systems

Problem 5.1

Infinite dimensional backstepping for nonlinear parabolic PDEs

Andras Balogh and Miroslav Krstic

Department of MAE

University of California at San Diego

La Jolla, CA 92093-0411

USA

abalogh@ucsd.edu and krstic@ucsd.edu

1 INTRODUCTION

This note explores an approach to global stabilization of boundary controlled nonlinear PDEs by a technique inspired by finite dimensional backstepping/feedback linearization. Solution of the problem presented herein would be of enormous significance because these are the only truly constructive and systematic techniques in finite dimension.

We consider nonlinear parabolic PDEs of the form

$$u_t(x, t) = \varepsilon u_{xx}(x, t) + f(u(x, t)) \quad (1)$$

for $x \in (0, 1)$, $t > 0$, with boundary conditions

$$u(0, t) = 0, \quad (2)$$

$$u(1, t) = \alpha_1(u), \quad (3)$$

initial condition

$$u(x, 0) = u_0(x), \quad x \in [0, 1],$$

and under the assumption

$$\varepsilon > 0, \quad f \in C^\infty(\mathbb{R}).^1 \quad (4)$$

The task is to derive a nonlinear (feedback) functional $\alpha_1 : C([0, 1]) \rightarrow \mathbb{R}$ that stabilizes the trivial solution $u(x, t) \equiv 0$ in an appropriate way. An infinite dimensional version of backstepping was introduced in [2] that solves

¹The smoothness requirement is explained after formula (18).

this problem for $f(u) = \lambda u$ with $\lambda > 0$ arbitrarily large. Superlinear nonlinearities can imply finite time blow-up for the uncontrolled case [6, 7, 9, 10]. However, numerical results in a series of papers by Boskovic and Krstic [3, 4, 5] show promise for the applicability of the infinite dimensional backstepping to nonlinear problems, at least for finite-grid discretizations. In this note, we present the open problem of convergence of nonlinear backstepping schemes as the discretization grid becomes infinitely refined. Note that this problem is different from the question of controllability [1, 8].

2 BACKSTEPPING TRANSFORMATION

We look for a coordinate transformation of the form

$$w = u - \alpha(u), \quad (5)$$

where $\alpha : C([0, 1]) \rightarrow C([0, 1])$ is a nonlinear operator to be found, that transforms system (1)–(3) into the exponentially stable system

$$w_t(x, t) = \varepsilon w_{xx}(x, t), \quad x \in (0, 1), \quad t > 0, \quad (6)$$

with boundary conditions

$$w(0, t) = 0, \quad (7)$$

$$w(1, t) = 0. \quad (8)$$

Once transformation (5) is found, it is realized through the stabilizing boundary feedback control (3) with $\alpha_1(u) = \alpha(u)|_{x=1}$.

In order to find (5) in a constructive way, we first discretize in space (1)–(3), then we develop a stabilizing coordinate transformation for the semi-discretized system. The main question of showing that the discretization converges to an infinite dimensional transformation is open in the case of functions $f(u)$ that are nonlinear.

We define $u_i^n = u(ih, t)$ for $i, j = 0, 1, \dots, n+1$, $n = 1, 2, \dots$ where $h = 1/(n+1)$, and the finite difference discretization of the rest of the functions is defined the same way. The discretized version of coordinate transformation (5) now has the form

$$\mathbf{w}^n = (\mathbf{I} - \alpha^n)(\mathbf{u}^n) \quad n = 1, 2, \dots \quad (9)$$

where α^n is an n -vector valued function of \mathbf{u}^n and

$$\mathbf{w}^n = [w_0^n, w_1^n, \dots, w_{n+1}^n]^T, \quad (10)$$

$$\mathbf{u}^n = [u_0^n, u_1^n, \dots, u_{n+1}^n]^T. \quad (11)$$

The discretized form of system (1)–(3) is

$$u_0^n = 0, \quad (12)$$

$$\dot{u}_i^n = \varepsilon \frac{u_{i+1}^n - 2u_i^n + u_{i-1}^n}{h^2} + f(u_i^n), \quad i = 1, \dots, n, \quad (13)$$

$$u_{n+1}^n = \alpha_n^n(u_1^n, u_2^n, \dots, u_n^n). \quad (14)$$

with the convention of $\alpha_0^n = 0$. The discretized form of system (6)–(8) is

$$w_0^n = 0, \quad (15)$$

$$\dot{w}_i^n = \varepsilon \frac{w_{i+1}^n - 2w_i^n + w_{i-1}^n}{h^2}, \quad i = 1, 2, \dots, n, \quad (16)$$

$$w_{n+1}^n = 0. \quad (17)$$

Combining (16), (9) and (13), and solving for α_i^n , we obtain the final form of the recursive formula for the transformation:

$$\begin{aligned} \alpha_i^n = & -\frac{h^2}{\varepsilon} f(u_i^n) + 2\alpha_{i-1}^n - \alpha_{i-2}^n \\ & + \sum_{j=1}^{i-1} \frac{\partial \alpha_{i-1}^n}{\partial u_j} \left(u_{j+1}^n - 2u_j^n + u_{j-1}^n + \frac{h^2}{\varepsilon} f(u_j^n) \right) \end{aligned} \quad (18)$$

for $i = 1, 2, \dots, n$. This recursive formula contains the functions $f(u)$ (which is nonlinear in general,) and it involves differentiation. As a result, as $n \rightarrow \infty$, eventually infinite smoothness of the function f is required. A few values of α_i^n :

$$\alpha_0^n = 0, \quad (19)$$

$$\alpha_1^n = -\frac{h^2}{\varepsilon} f(u_1^n), \quad (20)$$

$$\alpha_2^n = -\frac{h^2}{\varepsilon} f(u_2^n) - 2\frac{h^2}{\varepsilon} f(u_1^n) - \frac{h^2}{\varepsilon} f'(u_1^n) \left(u_2^n - 2u_1^n + \frac{h^2}{\varepsilon} f(u_1^n) \right), \quad (21)$$

$$\begin{aligned} \alpha_3^n = & -\frac{h^2}{\varepsilon} f(u_3^n) - 2\frac{h^2}{\varepsilon} f(u_2^n) - 3\frac{h^2}{\varepsilon} f(u_1^n) \\ & - 2\frac{h^2}{\varepsilon} f'(u_1^n) \left(u_2^n - 2u_1^n + \frac{h^2}{\varepsilon} f(u_1^n) \right) \\ & + \left(-\frac{h^2}{\varepsilon} f''(u_1^n) \left(u_2^n - 2u_1^n + \frac{h^2}{\varepsilon} f(u_1^n) \right) - \left(\frac{h^2}{\varepsilon} f'(u_1^n) \right)^2 \right) \\ & \cdot \left(u_2^n - 2u_1^n + \frac{h^2}{\varepsilon} f(u_1^n) \right) \\ & - \left(\frac{h^2}{\varepsilon} f'(u_2^n) + \frac{h^2}{\varepsilon} f'(u_1^n) \right) \left(u_3^n - 2u_2^n + \frac{h^2}{\varepsilon} f(u_2^n) \right) \end{aligned} \quad (22)$$

3 OPEN PROBLEM

Using the above backstepping approach, the problem of finding the coordinate transformation (5) and the corresponding stabilizing boundary control (3) requires two steps.

1. Find assumptions on the nonlinear function f that ensures the convergence of the discretized coordinate transformation (18) to a (nonlinear) operator α in order to obtain the feedback boundary control law (5).
2. Establish the bounded invertibility of operator $I - \alpha$ (see equation (5)) in appropriate function spaces.

4 KNOWN LINEAR RESULT

For the linear case $f(u) = \lambda u$ we have the following result [2].

Theorem 1: For any $\lambda \in \mathbb{R}$ and $\varepsilon, c > 0$ there exists a function $k_1 \in L_\infty(0, 1)$ such that for any $u_0 \in L_\infty(0, 1)$ the unique classical solution $u(x, t) \in C^1((0, \infty); C^2(0, 1))$ of system (1)–(3) with boundary feedback control

$$\alpha_1(u) = \int_0^1 k_1(\xi) u(\xi, t) d\xi \quad (23)$$

is exponentially stable in the $L_2(0, 1)$ and maximum norms with decay rate c . The precise statements of stability properties are the following: there exists a positive constant M^2 such that for all $t > 0$

$$\|u(t)\| \leq M \|u_0\| e^{-ct} \quad (24)$$

and

$$\max_{x \in [0, 1]} |u(t, x)| \leq M \sup_{x \in [0, 1]} |u_0(x)| e^{-ct}. \quad (25)$$

In this linear case, the transformation is a bounded linear operator $\alpha : L_1 \rightarrow L_1$ in the form of $\alpha(u) = \int_0^x k(x, \xi) u(\xi) d\xi$ with integral kernel $k \in L_\infty([0, \infty] \times [0, \infty])$. The boundary control is $\alpha_1(u) = \int_0^1 k(1, \xi) u(\xi) d\xi$. The explicit form of α_i is

$$\alpha_i^n = \sum_{j=1}^i k_{i,j}^n u_j^n, \quad i = 1, \dots, n, \quad (26)$$

where

$$\begin{aligned} k_{i,i-j}^n = & - \binom{i}{j+1} \left(\frac{(c+\lambda)}{\varepsilon(n+1)^2} \right)^{j+1} \\ & - (i-j) \sum_{l=1}^{[j/2]} \frac{1}{l} \binom{j-l}{l-1} \binom{i-l}{j-2l} \left(\frac{(c+\lambda)}{\varepsilon(n+1)^2} \right)^{j-2l+1} \end{aligned} \quad (27)$$

for $i = 1, \dots, n, j = 1, \dots, i$.

² M grows with c, λ and $1/\varepsilon$.

5 NUMERICAL RESULTS

In the nonlinear case, we need at least the uniform boundedness of sequences $\{\alpha_i^n(u)\}_{i=1}^n \subset \mathbb{R}$ as $n \rightarrow \infty$ for all u from some reasonable function space. We used Mathematica and MuPAD to calculate $\alpha_n^n(u)$ symbolically using the recursive relationship (18) and then to evaluate it for several different functions $u(x)$ and for different nonlinear functions $f(u)$. Since we found no qualitative difference between results corresponding to functions $u(x)$ of the same size, we present here only the results for functions of the form $u(x) = p \sin(\pi x)$ with different values of p . The symbolic calculation becomes extremely demanding computationally for increasing values of n . We were able to evaluate α_n^n for values up to $n = 9$ or $n = 10$ depending on the complexity of the nonlinear function $f(u)$. The results are collected below in two tables.

1. In the case of $f(u) = u \ln(1 + u^2)$, we have superlinearity $\frac{f(u)}{u} \xrightarrow{u \rightarrow \infty} \infty$, but the condition $\int_b^\infty \frac{du}{f(u)} < \infty$, which is necessary for finite time blow up (see, e.g., [9]) is not satisfied for any $b > 0$. Also, the zero solution of equation (1) is locally stable. The value $p = 1.5$ corresponds to an initial value for which the open-loop solution converges to zero. As the corresponding column in the table below shows, the control operator α_n^n converges to a finite value. For $p = 2$ the uncontrolled solution of (1) does not converge to zero, but still α_n^n converges to a finite value. For larger values of p , the convergence is not obvious from the calculations, but the concavity of the function graphs (decreasing rates of change in the values of α_n^n) suggest that we have convergence for increasing values of n with a decreasing rate of convergence as the size of the initial function is increased.

α_n^n for $f(u) = u \ln(1 + u^2)$				
n	$p = 1.5$	$p = 2$	$p = 5$	$p = 10$
1	-4.4	-8.0	-40.7	-115.3
2	-4.5	-11.0	-97.4	-356.2
3	-4.4	-11.6	-141.1	-615.1
4	-4.3	-12.3	-178.4	-867.1
5	-4.3	-12.6	-209.0	-1099.1
6	-4.2	-12.8	-233.4	-1301.5
7	-4.2	-13.0	-252.5	-1472.6
8	-4.2	-13.1	-267.6	-1615.4
9	-4.2	-13.2	-279.5	-1733.6

2. For $f(u) = u^2$ solutions corresponding to large initial data exhibit finite time blow-up. In fact, all of the present p values correspond to initial functions that result in finite time blow-up. However, for $p = 1.5$

and $p = 2$, the control values seem to converge as the table below shows. For larger values ($p = 5$ and $p = 10$), numerical calculations suggest fast divergence.

α_n^n for $f(u) = u^2$				
n	$p = 1.5$	$p = 2$	$p = 5$	$p = 10$
1	-5.6	-10.0	-62.5	-250.0
2	-7.2	-16.1	-221.3	-1687.0
3	-7.6	-18.6	-402.0	-4974.2
4	-8.0	-21.1	-637.3	-11202.1
5	-8.2	-22.6	-926.7	-22798.3
6	-8.3	-23.8	-1244.8	-41999.6
7	-8.3	-24.6	-1578.1	-70862.2
8	-8.4	-25.3	-1915.4	-111498.4
9	-8.4	-25.8	-2247.4	-165709.2
10	-8.5	-26.1	-2567.5	-234811.7

BIBLIOGRAPHY

- [1] S. Anita and V. Barbu, "Null controllability of nonlinear convective heat equations," *ESAIM COCV*, vol. 5, pp. 157–173, 2000.
- [2] A. Balogh and M. Krstic, "Infinite dimensional backstepping-style feedback transformations for a heat equation with an arbitrary level of instability," *European Journal of Control*, 2002.
- [3] D. Boskovic and M. Krstic, "Nonlinear stabilization of a thermal convection loop by state feedback," *Automatica*, vol. 37, pp. 2033–2040, 2001.
- [4] D. Boskovic and M. Krstic, "Backstepping control of chemical tubular reactors," *Computers and Chemical Engineering*, vol. 26, pp. 1077–1085, 2002.
- [5] D. Boskovic and M. Krstic, "Stabilization of a solid propellant rocket instability by state feedback," *Int. J. Robust and Nonlinear Control*, in press.
- [6] L.A. Caffarrelli and A. Friedman, "Blow-up of solutions of nonlinear heat equations," *J. Math. Anal. Appl.*, 129, pp. 409–419, 1988.
- [7] M. Chipot, M. Fila, and P. Quittner, "Stationary solutions, blow up and convergence to stationary solutions for semilinear parabolic equations with nonlinear boundary conditions," *Acta Math. Univ. Comenianae*, vol. LX, no. 1, pp. 35–103, 1991.

- [8] E. Fernandez-Cara, “Null controllability of the semilinear heat equation,” *ESAIM COCV*, vol. 2, pp. 87–103, 1997.
- [9] A. A. Lacey, “Mathematical analysis of thermal runaway for spatially inhomogeneous reactions,” *SIAM J. Appl. Math.*, vol. 43, no. 6, pp. 1350–1366, 1983.
- [10] S. Seo, “Blowup of solutions to heat equations with nonlocal boundary conditions,” *Kobe J. Math.*, vol. 13, pp. 123–132, 1996.

Problem 5.2

The dynamical Lamé system with boundary control: on the structure of reachable sets

M. I. Belishev¹

Dept. of the Steklov Mathematical Institute (POMI)

Fontanka 27

St. Petersburg 191011

Russia

belishev@pdmi.ras.ru

1 MOTIVATION

The questions posed below come from dynamical inverse problems for the hyperbolic systems with boundary control. These questions arise in the framework of the BC-method, which is an approach to inverse problems based on their relations to the boundary control theory [1], [2].

2 GEOMETRY

Let $\Omega \subset \mathbf{R}^3$ be a bounded domain with the smooth (enough) boundary Γ ; λ, μ, ρ smooth functions (Lamé parameters) satisfying $\rho > 0$, $\mu > 0$, $3\lambda + 2\mu > 0$ in $\bar{\Omega}$.

The parameters determine two metrics in $\bar{\Omega}$

$$dl_\alpha^2 = \frac{|dx|^2}{c_\alpha^2}, \quad \alpha = p, s$$

where

$$c_p := \left(\frac{\lambda + 2\mu}{\rho} \right)^{\frac{1}{2}}, \quad c_s := \left(\frac{\mu}{\rho} \right)^{\frac{1}{2}}$$

are the velocities of p - (pressure) and s - (shear) waves; let dist_α be the corresponding distances.

¹Supported by the RFBR grant No. 02-01-00260.

The distant functions (eikonals)

$$\tau_\alpha(x) := \text{dist}_\alpha(x, \Gamma), \quad x \in \bar{\Omega}$$

determine the subdomains

$$\Omega_\alpha^T := \{x \in \Omega \mid \tau_\alpha(x) < T\}, \quad T > 0$$

and the values $T_\alpha := \inf\{T > 0 \mid \Omega_\alpha^T = \Omega\}$, which are the times it takes for α -waves moving from Γ to fill the whole of Ω . The relation $c_s < c_p$ implies $\tau_p < \tau_s$, $\Omega_s^T \subset \Omega_p^T$, and $T_s > T_p$. If $T < T_s$ then

$$\Delta\Omega^T := \Omega_p^T \setminus \bar{\Omega}_s^T$$

is a nonempty open set.

If $T > 0$ is 'not too large', the vector fields

$$\nu_\alpha := \frac{\nabla \tau_\alpha}{|\nabla \tau_\alpha|}$$

are regular and satisfy $\nu_p(x) \cdot \nu_s(x) > 0$, $x \in \Omega_p^T$. Due to the latter, each vector field (\mathbf{R}^3 -valued function) $u = u(x)$ may be represented in the form

$$u(x) = u(x)_p + u(x)_s, \quad x \in \Omega_p^T \quad (*)$$

with $u(x)_p \parallel \nu_p(x)$ and $u(x)_s \perp \nu_s(x)$.

3 LAME SYSTEM. CONTROLLABILITY

Consider the dynamical system

$$u_{i \, tt} = \rho^{-1} \sum_{j,k,l=1}^3 \partial_j c_{ijkl} \partial_l u_k \quad (i = 1, 2, 3) \quad \text{in } \Omega \times (0, T);$$

$$u|_{t=0} = u_t|_{t=0} = 0 \quad \text{in } \Omega;$$

$$u = f \quad \text{on } \Gamma \times [0, T],$$

($\partial_j := \frac{\partial}{\partial x^j}$) where c_{ijkl} is the elasticity tensor of the Lamé model:

$$c_{ijkl} = \lambda \delta_{ij} \delta_{kl} + \mu (\delta_{ik} \delta_{jl} + \delta_{il} \delta_{jk});$$

let $u = u^f(x, t) = \{u_i^f(x, t)\}_{i=1}^3$ be the solution (wave).

Denote $\mathcal{H} := L_{2,\rho}(\Omega; \mathbf{R}^3)$ (with measure ρdx); $\mathcal{H}_\alpha^T := \{y \in \mathcal{H} \mid \text{supp } y \subset \bar{\Omega}_\alpha^T\}$. As was shown in [3], the map $f \mapsto u^f$ is continuous from $L_2(\Gamma \times [0, T]; \mathbf{R}^3)$ into $C([0, T]; \mathcal{H})$. By virtue of this and due to the finiteness of the wave velocities, the reachable set

$$\mathcal{U}^T := \{u^f(\cdot, T) \mid f \in L_2(\Gamma \times [0, T]; \mathbf{R}^3)\}$$

is embedded into \mathcal{H}_p^T . As was proved in the same paper, the relation

$$\text{clos } \mathcal{U}^T \supset \mathcal{H}_s^T$$

is valid for *any* $T > 0$, i.e., an approximate controllability always holds in the subdomain Ω_s^T filled with the shear waves, whereas the elements of the defect subspace

$$\mathcal{N}^T := \mathcal{H}_p^T \ominus \text{clos } \mathcal{H}_s^T$$

('unreachable states') can be supported only in $\Delta\Omega^T$. On the other hand, it is not difficult to show the examples with $\mathcal{N}^T \neq \{0\}$, $T < T_s$.

4 PROBLEMS AND HYPOTHESES

The open problem is to characterize the defect subspace N^T . The following is the reasonable hypotheses.

- The defect space is always nontrivial: $N^T \neq \{0\}$ for $T < T_s$ in the general case (not only in examples). Let us note that, due to the standard ‘controllability-observability’ duality, this property would mean that in *any* inhomogeneous isotropic elastic media there exist *the slow waves* whose forward front propagates with the velocity c_s .
- In the subdomain $\Delta\Omega^T$, where the elements of the defect subspace are supported, the pressure component of the wave (see (*)) determines its shear component through a linear operator: $u^f(\cdot, T)_s = K^T[u^f(\cdot, T)_p]$ in $\Delta\Omega^T$. If this holds, the question is to describe the operator K^T .
- The decomposition (*) diagonalizes the principal part of the Lamé system.

The progress in these questions would be of great importance for the inverse problems of the elasticity theory that is now the most difficult and challenging class of dynamical inverse problems.

BIBLIOGRAPHY

- [1] M. I. Belishev, “Boundary control in reconstruction of manifolds and metrics (the BC-method),” *Inv.Prob.*, 13(5):R1–R45, 1997.
- [2] M. I. Belishev and A. K. Glasman, “Dynamical inverse problem for the Maxwell system: Recovering the velocity in the regular zone (the BC-method),” *St.Petersburg Math. J.*, 12 (2):279–316, 2001.
- [3] M. I. Belishev and I. Lasiecka, “The dynamical Lamé system: Regularity of solutions, boundary controllability and boundary data continuation,” *ESAIM COCV*, 8:143–167, 2002.

Problem 5.3

Null-controllability of the heat equation in unbounded domains

Sorin Micu

Facultatea de Matematică-Informatică

Universitatea din Craiova

Al. I. Cuza 13, Craiova, 1100

Romania

sd.micu@yahoo.com

Enrique Zuazua

Departamento de Matemáticas, Facultad de Ciencias

Universidad Autónoma

Cantoblanco, 28049, Madrid

Spain

enrique.zuazua@uam.es

1 DESCRIPTION OF THE PROBLEM

Let Ω be a smooth domain of \mathbb{R}^n with $n \geq 1$. Given $T > 0$ and $\Gamma_0 \subset \partial\Omega$, an open non-empty subset of the boundary of Ω , we consider the linear heat equation:

$$\begin{cases} u_t - \Delta u = 0 & \text{in } Q \\ u = v 1_{\Sigma_0} & \text{on } \Sigma \\ u(x, 0) = u_0(x) & \text{in } \Omega, \end{cases} \quad (1)$$

where $Q = \Omega \times (0, T)$, $\Sigma = \partial\Omega \times (0, T)$ and $\Sigma_0 = \Gamma_0 \times (0, T)$ and where 1_{Σ_0} denotes the characteristic function of the subset Σ_0 of Σ .

In (1) $v \in L^2(\Sigma)$ is a *boundary control* that acts on the system through the subset Σ_0 of the boundary and $u = u(x, t)$ is the *state*.

System (1) is said to be *null-controllable* at time T if for any $u_0 \in L^2(\Omega)$ there exists a control $v \in L^2(\Sigma_0)$ such that the solution of (1) satisfies

$$u(x, T) = 0 \text{ in } \Omega. \quad (2)$$

This article is concerned with the null-controllability problem of (1) when the domain Ω is unbounded.

2 MOTIVATION AND HISTORY OF THE PROBLEM

We begin with the following well-known result

Theorem 1. When Ω is a bounded domain of class C^2 system (1) is null-controllable for any $T > 0$.

We refer to D. L. Russell [12] for some particular examples treated by means of moment problems and Fourier series and to A. Fursikov and O. Yu. Imanuvilov [3] and G. Lebeau and L. Robbiano [7] for the general result covering any bounded smooth domain Ω and open, nonempty subset Γ_0 of $\partial\Omega$. Both the approaches of [3] and [7] are based on the use of Carleman inequalities.

However, in many relevant problems the domain Ω is unbounded. We address the following question: if Ω is an *unbounded* domain, is system (1) null-controllable for some $T > 0$?

None of the approaches mentioned above apply in this situation. In fact, very particular cases being excepted (see the following section), there exist no results on the null-controllability of the heat equation (1) when Ω is unbounded.

The approach described in [6] and [9] is also worth mentioning. In this article it is proved that, for any $T > 0$, the heat equation has a fundamental solution that is C^∞ away from the origin and with support in the strip $0 \leq t \leq T$. This fundamental solution, of course, grows very fast as $|x|$ goes to infinity. As a consequence of this, a boundary controllability result may be immediately obtained in any domain Ω with controls distributed all along its boundary. Note, however, that when the domain is unbounded the solutions and controls obtained in this way grow too fast as $|x| \rightarrow \infty$ and, therefore, these are not solutions in the classical sense. In fact, in the frame of unbounded domains, one has to be very careful in defining the class of admissible controlled solutions. When imposing, for instance, the classical integrability conditions at infinity, one is imposing additional restrictions that may determine the answer to the controllability problem. This is indeed the case, as we shall explain.

There is a weaker notion of controllability property. It is the so-called approximate controllability property. System (1) is said to be *approximately controllable* in time T if for any $u_0 \in L^2(\Omega)$ the set of reachable states, $R(T; u_0) = \{u(T) : u \text{ solution of (1) with } v \in L^2(\Sigma_0)\}$, is dense in $L^2(\Omega)$.

With the aid of classical backward uniqueness results for the heat equation (see, for instance, J.L. Lions and E. Malgrange [8] and J.M. Ghidaglia [4]), it can be seen that null-controllability implies approximate controllability. The approximate control problem for the *semilinear* heat equation in general unbounded domains was addressed in [13] where an approximation method was developed. The domain Ω was approximated by bounded domains (essentially by $\Omega \cap B_R$, B_R being the ball of radius R) and the approximate control in the unbounded domain Ω was obtained as limit of the approximate

control on the approximating bounded domain $\Omega \cap B_R$. But this approach does not apply in the context of the null-control problem.

However, taking into account that approximate controllability holds, it is natural to analyze whether null-controllability holds as well.

In [1] it was proved that the null-controllability property holds even in unbounded domains if the control is supported in a subdomain that only leaves a bounded set uncontrolled. Obviously, this result is very close to the case in which the domain Ω is bounded and does not answer to the main issue under consideration of whether heat processes are null-controllable in unbounded domains.

3 AVAILABLE RESULTS

To our knowledge, in the context of unbounded domains Ω and the boundary control problem, only the particular case of the half-space has been considered:

$$\begin{aligned}\Omega &= \mathbb{R}_+^n = \{x = (x', x_n) : x' \in \mathbb{R}^{n-1}, x_n > 0\} \\ \Gamma_0 &= \partial\Omega = \mathbb{R}^{n-1} = \{(x', 0) : x' \in \mathbb{R}^{n-1}\}\end{aligned}\tag{3}$$

(see [10] for $n = 1$ and [11] for $n > 1$).

According to the results in [10] and [11], the situation is completely different to the case of bounded domains. In fact a simple argument shows that the null controllability result which holds for the case Ω bounded is no longer true. Indeed, the null-controllability of (1) with initial data in $L^2(\mathbb{R}_+^n)$ and boundary control in $L^2(\Sigma)$ is equivalent to an *observability inequality* for the adjoint system

$$\begin{cases} \varphi_t + \Delta\varphi = 0 & \text{on } Q \\ \varphi = 0 & \text{on } \Sigma. \end{cases}\tag{4}$$

More precisely, it is equivalent to the existence of a positive constant $C > 0$ such that

$$\|\varphi(0)\|_{L^2(\mathbb{R}_+^n)}^2 \leq C \int_{\Sigma} \left| \frac{\partial\varphi}{\partial x_n} \right|^2 dx' dt\tag{5}$$

holds for every smooth solution of (4).

When Ω is bounded, Carleman inequalities provide the estimate (5) and, consequently, null-controllability holds (see, for instance, [3]). In the case of a half-space, by using a translation argument, it is easy to see that (5) does not hold (see [11]).

In the case of bounded domains, by using Fourier series expansion, *the control problem may be reduced to a moment problem*. However, Fourier series cannot be used directly in \mathbb{R}_+^n . Nevertheless, it was observed by M. Escobedo and O. Kavian in [2] that, on suitable similarity variables and at the

appropriate scale, solutions of the heat equation on conical domains may be indeed developed in Fourier series on a weighted L^2 -space. This idea was used in [10] and [11] to study the null-controllability property when Ω is given by (3).

Firstly, we use similarity variables and weighted Sobolev spaces to develop the solutions in Fourier series. A sequence of one-dimensional controlled systems like those studied in [10] is obtained. Each of these systems is equivalent to a moment problem of the following type: given $S > 0$ and $(a_n)_{n \geq 1}$ (depending on the Fourier coefficients of the initial data u_0) find $f \in L^2(0, S)$ such that

$$\int_0^S f(s)e^{ns}ds = a_n, \quad \forall n \geq 1. \quad (6)$$

This moment problem turns out to be critical since it concerns the family of real exponential functions $\{e^{\lambda_n s}\}_{n \geq 1}$ with $\lambda_n = n$, in which the usual summability condition on the inverses of the exponents, $\sum_{n \geq 1} \frac{1}{\lambda_n} < \infty$, does not hold. It was proved that, if the sequence $(a_n)_{n \geq 1}$ has the property that, for any $\delta > 0$, there exists $C_\delta > 0$ such that

$$|a_n| \leq C_\delta e^{\delta n}, \quad \forall n \geq 1, \quad (7)$$

problem (6) has a solution if and only if $a_n = 0$ for all $n \geq 1$.

Since $(a_n)_{n \geq 1}$ depend on the Fourier coefficients of the initial data, the following negative controllability result for the one-dimensional systems is obtained:

Theorem 2. When Ω is the half line, there is no nontrivial initial datum u_0 belonging to a negative Sobolev space that is null-controllable in finite time with L^2 boundary controls.

This negative result was complemented by showing that there exist initial data with exponentially growing Fourier coefficients for which null-controllability holds in finite time with L^2 -controls.

We mention that in [10] and [11] we are dealing with solutions defined in the sense of transposition, and therefore the solutions in [6] and [9] that grow and oscillate very fast at infinity are excluded.

4 OPEN PROBLEMS

As we have already mentioned, the null-controllability property of (1) when Ω is unbounded and different from a half-line or half-space is still open.

The approach based on the use of the similarity variables may still be used in general conical domains. But, due to the lack of orthogonality of the traces of the normal derivatives of the eigenfunctions, the corresponding moment problem is more complex and remains to be solved.

When Ω is a general unbounded domain, the similarity transformation does not seem to be of any help since the domain one gets after transformation depends on time. Therefore, a completely different approach seems to be needed when Ω is not conical. However, one may still expect a bad behavior of the null-control problem. Indeed, assume for instance that Ω contains \mathbb{R}_+^n . If one is able to control to zero in Ω an initial data u_0 by means of a boundary control acting on $\partial\Omega \times (0, T)$, then, by restriction, one is able to control the initial data $u_0|_{\mathbb{R}_+^n}$ with the control being the restriction of the solution in the larger domain $\Omega \times (0, T)$ to $\mathbb{R}^{n-1} \times (0, T)$. A careful development of this argument and of the result it may lead to remains to be done.

ACKNOWLEDGEMENTS

The first author was partially supported by Grant PB96-0663 of DGES (Spain) and Grant A3/2002 of CNCSIS (Romania). The second author was partially supported by Grant PB96-0663 of DGES (Spain) and the TMR network of the EU "Homogenization and Multiple Scales (HMS2000)."

BIBLIOGRAPHY

- [1] V. Cabanillas, S. de Menezes and E. Zuazua, "Null controllability in unbounded domains for the semilinear heat equation with nonlinearities involving gradient terms," *J. Optim. Theory Applications*, 110 (2) (2001), 245-264.
- [2] M. Escobedo and O. Kavian, "Variational problems related to self-similar solutions of the heat equation," *Nonlinear Anal. TMA*, 11 (1987), 1103-1133.
- [3] A. Fursikov and O. Yu. Imanuvilov, "Controllability of evolution equations," *Lecture Notes Series #34*, Research Institute of Mathematics, Global Analysis Research Center, Seoul National University, 1996.
- [4] J. M. Ghidaglia, "Some backward uniqueness results," *Nonlinear Anal. TMA*, 10 (1986), 777-790.
- [5] O. Yu. Imanuvilov and M. Yamamoto, "Carleman estimate for a parabolic equation in Sobolev spaces of negative order and its applications," In: *Control of Nonlinear Distributed Parameter Systems*, G. Chen et al. eds., Marcel-Dekker, 2000, pp. 113-137.
- [6] B. F. Jones, Jr., "A fundamental solution of the heat equation which is supported in a strip," *J. Math. Anal. Appl.*, 60 (1977), 314-324.
- [7] G. Lebeau and L. Robbiano, "Contrôle exact de l'équation de la chaleur," *Comm. P.D.E.*, 20 (1995), 335-356.
- [8] J. L. Lions and E. Malgrange, "Sur l'unicité rétrograde dans les problèmes mixtes paraboliques," *Math. Scan.*, 8 (1960), 277-286.

- [9] W. Littman, "Boundary control theory for hyperbolic and parabolic partial differential equations with constant coefficients," *Annali Scuola Norm. Sup. Pisa*, Serie IV, 3 (1978), 567-580.
- [10] S. Micu and E. Zuazua, "On the lack of null-controllability of the heat equation on the half-line," *Trans. AMS*, 353 (2001), 1635-1659.
- [11] S. Micu and E. Zuazua, "On the lack of null-controllability of the heat equation on the half-space," *Portugalia Matematica*, 58 (2001), 1-24.
- [12] D. L. Russell, "Controllability and stabilizability theory for linear partial differential equations. Recent progress and open questions," em *SIAM Rev.*, 20 (1978), 639-739.
- [13] L. de Teresa and E. Zuazua, "Approximate controllability of the heat equation in unbounded domains," *Nonlinear Anal. TMA*, 37 (1999), 1059-1090.

Problem 5.4

Is the conservative wave equation regular?

George Weiss

Dept. of Electrical and Electronic Engineering

Imperial College London

Exhibition Road

London SW7 2BT

UK

G.Weiss@imperial.ac.uk

1 DESCRIPTION OF THE PROBLEM

We consider an infinite-dimensional system described by the wave equation on an n -dimensional domain, with mixed boundary control and mixed boundary observation, which has been analyzed (as an example for a certain class of conservative linear systems) in [13]. A somewhat simpler version of this system has appeared (also as an example) in the paper [11, section 7] and a related system has been discussed in [5].

We assume that $\Omega \subset \mathbb{R}^n$ is a bounded domain with Lipschitz boundary Γ , as defined in Grisvard [3]. This means that, locally, after a suitable rotation of the orthogonal coordinate system, the boundary is the graph of a Lipschitz function defined on an open set in \mathbb{R}^{n-1} . Such a boundary admits corners and edges. Γ_0 and Γ_1 are nonempty open subsets of Γ such that $\Gamma_0 \cap \Gamma_1 = \emptyset$ and $\overline{\Gamma_0} \cup \overline{\Gamma_1} = \Gamma$. We denote by x the space variable ($x \in \overline{\Omega}$). A function $b \in L^\infty(\Gamma_1)$ is given, which intuitively expresses how strongly the input signal acts on different parts of the active boundary Γ_1 . We assume that $b(x) \neq 0$ for almost every $x \in \Gamma_1$. The equations of the system are

$$\left\{ \begin{array}{ll} \ddot{z}(x, t) = \Delta z(x, t) & \text{on } \Omega \times [0, \infty), \\ z(x, t) = 0 & \text{on } \Gamma_0 \times [0, \infty), \\ \frac{\partial}{\partial \nu} z(x, t) + |b(x)|^2 \dot{z}(x, t) = \sqrt{2} \cdot b(x) u(x, t) & \text{on } \Gamma_1 \times [0, \infty), \\ \frac{\partial}{\partial \nu} z(x, t) - |b(x)|^2 \dot{z}(x, t) = \sqrt{2} \cdot b(x) y(x, t) & \text{on } \Gamma_1 \times [0, \infty), \\ z(x, 0) = z_0(x), \quad \dot{z}(x, 0) = w_0(x) & \text{on } \Omega, \end{array} \right. \quad (1)$$

where u is the input function and y is the output function. The functions

z_0 and w_0 are the initial state of the system. The part Γ_0 of the boundary is just reflecting waves, while inputs and outputs act through the part Γ_1 .

For every $g \in \mathcal{H}^1(\Omega)$ we denote by γg the Dirichlet trace of g on Γ (for $g \in C^1(\bar{\Omega}) \subset \mathcal{H}^1(\Omega)$ this would simply be the restriction of g to Γ). We regard γg as an element of $L^2(\Gamma)$. We define the Hilbert space

$$\mathcal{H}_{\Gamma_0}^1(\Omega) = \{g \in \mathcal{H}^1(\Omega) \mid \gamma g = 0 \text{ on } \Gamma_0\}, \quad \|g\|_{\mathcal{H}^1} = \|\nabla g\|_{L^2}.$$

Proposition 1. *The equations (1) determine a well-posed linear system Σ with input space $U = L^2(\Gamma_1)$, output space $Y = L^2(\Gamma_1)$ and state space*

$$X = \mathcal{H}_{\Gamma_0}^1(\Omega) \times L^2(\Omega).$$

For the precise meaning of a well-posed linear system we refer to [8, 9, 6]. These papers use the same notation and terminology that we use here, but their references will indicate other works in which equivalent definitions can be found. We give a short explanation of what well-posedness means in our case. If we take $x(0) = [z_0 \ w_0]^T \in X$, $u \in L^2([0, \infty); U)$ and we solve the equations (1) on the time interval $[0, \infty)$, then we get $x(\tau) = [z(\tau) \ \dot{z}(\tau)]^T \in X$ for every $\tau \geq 0$. $x(\tau)$ is called the state of the system at time τ . Moreover, if we denote the restriction of y to $[0, \tau]$ by $\mathbf{P}_\tau y$, then $\mathbf{P}_\tau y \in L^2([0, \tau]; Y)$. (Note that in our particular case, $U = Y$.) We can introduce four families of bounded operators \mathbb{T}, Φ, Ψ , and \mathbb{F} indexed by $\tau \geq 0$ such that for every such τ ,

$$x(\tau) = \mathbb{T}_\tau x(0) + \Phi_\tau \mathbf{P}_\tau u, \quad \mathbf{P}_\tau y = \Psi_\tau x(0) + \mathbb{F}_\tau \mathbf{P}_\tau u.$$

Thus, for every $\tau \geq 0$, the operator matrix

$$\Sigma_\tau = \begin{bmatrix} \mathbb{T}_\tau & \Phi_\tau \\ \Psi_\tau & \mathbb{F}_\tau \end{bmatrix}$$

defines a bounded operator from $X \times L^2([0, \tau]; U)$ to $X \times L^2([0, \tau]; Y)$. This is the essential feature of a well-posed linear system. In fact, in [8, 9, 6], Σ is defined as the family of operators Σ_τ . For a well-posed linear system, the family \mathbb{T} is a strongly continuous semigroup of operators acting on X . Proposition 1 was proved in [13, section 7], together with the following:

Proposition 2. *The system Σ from Proposition 1 is conservative.*

The fact that Σ is conservative means that the operators Σ_τ are unitary. In particular, the fact that Σ_τ is isometric means that we have

$$\|x(\tau)\|^2 - \|x(0)\|^2 = \int_0^\tau \|u(t)\|^2 dt - \int_0^\tau \|y(t)\|^2 dt,$$

which can be interpreted as an energy balance equation. For background on conservative systems, we refer to [1, 2, 4, 7, 12, 13].

The system Σ has, like every conservative system, a transfer function \mathbf{G} that is in the Schur class. This means that \mathbf{G} is analytic on the open right

half-plane \mathbb{C}_0 and $\|\mathbf{G}(s)\| \leq 1$ for all $s \in \mathbb{C}_0$. For the simple proof of this fact, see [13, theorem 1.3 and proposition 4.5]. The boundary values $\mathbf{G}(i\omega)$ can be defined for almost every $\omega \in \mathbb{R}$ as nontangential limits, and we have

$$\mathbf{G}(i\omega)^* \mathbf{G}(i\omega) = \mathbf{G}(i\omega) \mathbf{G}(i\omega)^* = I$$

for almost every $\omega \in \mathbb{R}$ (i.e., \mathbf{G} is inner and co-inner). This follows from [10, proposition 2.1] or alternatively from [7, corollary 7.3].

Recall that a well-posed linear system with input space U , output space Y , and transfer function \mathbf{G} is called *regular* if for every $v \in U$, the limit

$$\lim_{s \rightarrow +\infty, s \in \mathbb{R}} \mathbf{G}(s)v = Dv$$

exists. In this case, the operator $D \in \mathcal{L}(U, Y)$ is called the feedthrough operator of the system (see [8, 9, 6] for further details). For regular linear systems, the theories of local representation, feedback and dynamic stabilization are much simpler than for well-posed linear systems.

Conjecture. *The system Σ from Proposition 1 is regular and its feedthrough operator is zero.*

Consider the particular situation when Ω is one-dimensional: $\Omega = (0, 1)$, $\Gamma_0 = \{0\}$, $\Gamma_1 = \{1\}$ and $U = Y = \mathbb{C}$. Now the function b becomes a nonzero number, and without loss of generality we may take $b = 1$. It is easy to see that the input signal enters the domain at $x = 1$, propagates along the domain (with unit speed) until it gets reflected at $x = 0$ and then it propagates back to exit (as the output signal) at $x = 1$. If the initial state is zero, then for $t \geq 2$ we have $y(t) = u(t-2)$, so that the transfer function is $\mathbf{G}(s) = e^{-2s}$. Note that \mathbf{G} is indeed inner and it is regular with feedthrough operator zero.

The author thinks that he can prove the conjecture in the following particular case: the active boundary Γ_1 can be partitioned into a finite union of open subsets that are either planar (i.e., an open subset of an $n-1$ dimensional hyperplane) or spherical (i.e., an open subset of an $n-1$ dimensional sphere). The idea is to construct solutions of (1), which locally (near a boundary point) look like a planar or spherical wave moving into the domain Ω (the initial state is zero) and locally (in time and space), u is a step function. Then locally (in time and space) y is zero, which proves the claim, due to the equivalent characterization of regularity via the step response, see [8, theorem 5.8].

BIBLIOGRAPHY

- [1] D. Z. Arov and M. A. Nudelman, "Passive linear stationary dynamical scattering systems with continuous time," *Integral Equations and Operator Theory*, 24 (1996), pp. 1–43.

- [2] J. A. Ball, "Conservative dynamical systems and nonlinear Livsic-Brodskii nodes," *Operator Theory: Advances and Applications*, 73 (1994), pp. 67–95.
- [3] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Pitman, Boston, 1985.
- [4] B. M. J. Maschke and A. J. van der Schaft, "Portcontrolled Hamiltonian representation of distributed parameter systems," *Proc. of the IFAC Workshop on Lagrangian and Hamiltonian Methods for Nonlinear Control*, N. E. Leonard and R. Ortega, eds., Princeton University Press, 2000, pp. 28–38.
- [5] A. Rodriguez-Bernal and E. Zuazua, "Parabolic singular limit of a wave equation with localized boundary damping," *Discrete and Continuous Dynamical Systems*, 1 (1995), pp. 303–346.
- [6] O. J. Staffans and G. Weiss, "Transfer functions of regular linear systems. Part II: The system operator and the Lax-Phillips semigroup," *Trans. American Math. Society*, 354 (2002), pp. 3229–3262.
- [7] O. J. Staffans and G. Weiss, "Transfer functions of regular linear systems. Part III: Inversions and duality," submitted.
- [8] G. Weiss, "Transfer functions of regular linear systems. Part I: Characterizations of regularity," *Trans. American Math. Society*, 342 (1994), pp. 827–854.
- [9] G. Weiss, "Regular linear systems with feedback," *Mathematics of Control, Signals and Systems*, 7 (1994), pp. 23–57.
- [10] G. Weiss, "Optimal control of systems with a unitary semigroup and with colocated control and observation," *Systems and Control Letters*, 48 (2003), pp. 329–340.
- [11] G. Weiss and R. Rebarber, "Optimizability and estimatability for infinite-dimensional linear systems," *SIAM J. Control and Optimization*, 39 (2001), pp. 1204–1232.
- [12] G. Weiss, O. J. Staffans and M. Tucsnak, "Well-posed linear systems: A survey with emphasis on conservative systems," *Applied Mathematics and Computer Science*, 11 (2001), pp. 101–127.
- [13] G. Weiss and M. Tucsnak, "How to get a conservative well-posed linear system out of thin air. Part I: well-posedness and energy balance," *ESAIM COCV*, vol. 9, pp. 247–74, 2003.

Problem 5.5

Exact controllability of the semi-linear wave equation

Xu Zhang

Departamento de Matemática Aplicada

Universidad Complutense

28040 Madrid

Spain

and

School of Mathematics

Sichuan University

Chengdu 610064

China

xuzhang@fudan.edu

Enrique Zuazua

Departamento de Matemáticas, Facultad de Ciencias

Universidad Autónoma

28049 Madrid

Spain

enrique.zuazua@uam.es

1 DESCRIPTION OF THE PROBLEM

Let $T > 0$ and $\Omega \subset \mathbf{R}^n$ ($n \in \mathbf{N}$) be a bounded domain with a $C^{1,1}$ boundary $\partial\Omega$. Let ω be a proper subdomain of Ω and denote the characteristic function of the set ω by χ_ω . Fix a nonlinear function $f \in C^1(\mathbf{R})$.

We are concerned with the exact controllability of the following semilinear wave equation:

$$\begin{cases} y_{tt} - \Delta y + f(y) = \chi_\omega(x)u(t, x) & \text{in } (0, T) \times \Omega, \\ y = 0 & \text{on } (0, T) \times \partial\Omega, \\ y(0) = y_0, \quad y_t(0) = y_1 & \text{in } \Omega. \end{cases} \quad (1)$$

In (1), $(y(t, \cdot), y_t(t, \cdot))$ is the *state* and $u(t, \cdot)$ is the *control* that acts on the system through the subset ω of Ω .

In what follows, we choose the *state space* and the *control space* as $H_0^1(\Omega) \times L^2(\Omega)$ and $L^2((0, T) \times \Omega)$, respectively. Of course, the choice of these spaces is

not unique. But this one is very natural in the context of the wave equation. The space $H_0^1(\Omega) \times L^2(\Omega)$ is often referred to as the *energy space*.

The exact (internal) controllability problem for (1) (in $H_0^1(\Omega) \times L^2(\Omega)$) may be formulated as follows: for any given $(y_0, y_1), (z_0, z_1) \in H_0^1(\Omega) \times L^2(\Omega)$, to find (if possible) a control $u \in L^2((0, T) \times \Omega)$ such that the weak solution y of (1) satisfies

$$y(T) = z_0 \quad \text{and} \quad y_t(T) = z_1 \quad \text{in } \Omega. \quad (2)$$

The exact (boundary) controllability problem of (1) can be formulated similarly. In that case, the control u enters on the system through the boundary conditions. This produces extra technical difficulties. The main open problem on the controllability of this semilinear wave equation we shall describe here arises in both cases. We prefer to present it in the case where the control acts on the internal subdomain ω to avoid unnecessary technical difficulties.

First of all, it is well-known that when f grows too fast, the solution of (1) may blow up. In the presence of blow-up phenomena, as a consequence of the finite speed of propagation of solutions of (1), the exact controllability of (1) does not hold unless $\omega = \Omega$ ([13]). This exception means that the control acts on the system everywhere in Ω in which case the effect of nonlinearity may be suppressed easily. Therefore, we suppose that

(H1) *The nonlinearity $f \in C^1(\mathbf{R})$ is such that (1) admits a global weak solution $y \in C([0, T]; H_0^1(\Omega)) \cap C^1([0, T]; L^2(\Omega))$ for any given $(y_0, y_1) \in H_0^1(\Omega) \times L^2(\Omega)$ and $u \in L^2((0, T) \times \Omega)$.*

There are two classes of conditions on f guaranteeing that (H1) holds. The first one, which will be called *mild growth condition*, amounts to requesting that $f \in C^1(\mathbf{R})$ grows “mildly” at infinity (see [2] and [3]), i.e.,

$$\overline{\lim}_{|x| \rightarrow \infty} \left| \int_0^x f(s) ds \right| \left[|x| \prod_{k=1}^{\infty} \log_k(e_k + x^2) \right]^{-2} < \infty, \quad (3)$$

where the iterated logarithm function \log_j is defined by the formulas

$$\log_0 s = s \quad \text{and} \quad \log_{j+1} s = \log(\log_j s), \quad j = 0, 1, 2, \dots,$$

the number e_j is defined by the equations $\log_j e_j = 1$.

It is obvious that any globally Lipschitz continuous function f satisfies (3). But, of course, (3) allows f to grow in a slight superlinear way at infinity.

The second one, which will be called *good sign growth condition*, is the class of functions $f \in C^1(\mathbf{R})$ that grow fast at infinity but satisfy a “good-sign” condition, i.e., there exist constants $L > 0$, $p \in (1, n/(n-2)]$ if $n \geq 3$ and $p \in (1, \infty)$ if $n = 1, 2$, such that

$$|f(r) - f(s)| \leq L(1 + |r|^{p-1} + |s|^{p-1})|r - s|, \quad \forall r, s \in \mathbf{R} \quad (4)$$

and

$$\int_0^x f(s) ds \geq -Lx^2 \quad \text{as } |x| \rightarrow \infty. \quad (5)$$

A typical example is

$$f(u) = u^3 \quad \text{for } n = 1, 2, 3. \quad (6)$$

On the other hand, it is well-known that, even in the linear case where $f \equiv 0$, some conditions on the controllability time T and the geometry of the set ω where the control applies are needed in order to guarantee the exact controllability property. Thus, we assume that

(H2) T and ω are such that (1) with $f \equiv 0$ is exactly controllable.

There are also two classes of conditions on T and ω guaranteeing that (H2) holds. The first one, which we will call the *classical multiplier condition*, is when ω is a neighborhood of a subset of the boundary of the form $\Gamma(x_0) = \{x \in \partial\Omega : (x - x_0) \cdot \nu(x) > 0\}$ for some $x_0 \in \mathbf{R}^n$, where $\nu(x)$ is the unit outward normal vector to $\partial\Omega$ at x , and $T > 2 \max\{|x - x_0| : x \in \Omega \setminus \omega\}$. This is the typical situation one encounters when applying the multiplier technique ([8]). The second one is when T and Ω satisfy the so-called *Geometric Control Condition* introduced in [1].

We have the following

Open Problem: Do (H1) and (H2) imply the exact controllability of (1)?

The above problem can also be formulated in the more general case in which the nonlinearity is of the form $f(t, x, y, y_t, \nabla y)$. Of course, the problem is even more difficult in that case and new phenomena may occur due to the strong dissipative effects that terms of the form $|u_t|^{p-1}u_t$ may produce. Thus, we shall focus in the case $f = f(y)$. This open problem will be made more precise below.

2 AVAILABLE RESULTS AND OPEN PROBLEMS

Nonlinearities with mild growth condition

For the one space dimensional case, by combining the sidewise energy estimates for the $1 - d$ wave equations and the fixed point technique, Zuazua ([13]) obtained the following result:

Theorem 1: Assume $n = 1$ and $\Omega = (0, 1)$. Let (a, b) be a (proper) subinterval of $(0, 1)$, $T > 2 \max(a, 1 - b)$ and

$$\lim_{|x| \rightarrow \infty} |f(x)| |x|^{-1} \log^{-2} |x| = 0. \quad (7)$$

Then (1) is exact controllable.

Later on, based on a method due to Émanuilov ([5]), Cannarsa, Komornik, and Loreti ([2]) improved theorem 1 by relaxing the growth condition on f . The main result in [2] says that the same conclusion in theorem 1 holds if the condition (7) on f is replaced by (3). The growth condition (3) on f is

sharp (since solutions of (1) may blow up whenever f grows faster than (3) at infinity and f has the bad sign).

For the higher dimensional case, Li and Zhang ([7]) proved the following result:

Theorem 2: *Let ω be a neighborhood of $\partial\Omega$, $T > \text{diam}(\Omega \setminus \omega)$ and*

$$\lim_{|x| \rightarrow \infty} f(x)|x|^{-1} \log^{-1/2} |x| = 0. \quad (8)$$

Then (1) is exactly controllable.

A special case of theorem 2 is when f is globally Lipschitz continuous, which gives the main result of Zuazua in [12]. The main result in [12] was generalized to an abstract setting by Lasiecka and Triggiani ([6]) using a global version of Inverse Function theorem and was extended in [9] to the case when T and ω satisfy the classical multiplier condition.

It is natural to conjecture that the same conclusion in Theorem 2 holds under the growth condition (3) on f as in one dimension. But this is by now an open problem. On the other hand, whether the same conclusion in theorem 2 holds for more general conditions on T and ω , say the classical multiplier condition or Geometric Control Condition, is also an open problem. Especially, when T and ω satisfy the Geometric Control Condition, the exact controllability problem for (1) is open even for globally Lipschitz continuous nonlinearities.

Nonlinearities with good sign and superlinear growth at infinity

In this case, there are no global exact controllability results in the literature. However, using a fixed point argument, Zuazua proved the following local exact controllability results for (1) ([10]):

Theorem 3: *Let (H2) hold, $f \in C^1(\mathbf{R})$ satisfy (4) and $f(0) = 0$. Then there is a $\delta > 0$ such that for any (y_0, y_1) and (z_0, z_1) in $H_0^1(\Omega) \times L^2(\Omega)$ with $|(y_0, y_1)|_{H_0^1(\Omega) \times L^2(\Omega)} + |(z_0, z_1)|_{H_0^1(\Omega) \times L^2(\Omega)} \leq \delta$, there is a control $u \in L^2((0, T) \times \Omega)$, such that (2) holds.*

Combining theorem 3 and the stabilization results for the semilinear wave equations with “good-sign” condition on the nonlinearity ([11] and [4]), it is easy to show that

Theorem 4: *Let T_0 and ω satisfy the classical multiplier condition and f satisfy (4)–(5). Then for any (y_0, y_1) and (z_0, z_1) in $H_0^1(\Omega) \times L^2(\Omega)$, there exist a time $T \geq T_0$ and a control $u(\cdot) \in L^2((0, T) \times \Omega)$, such that (2) holds.*

Note that the controllability time T in theorem 4 depends on (y_0, y_1) and (z_0, z_1) . According to [11], one can obtain explicit bounds on T . However, whether T may be chosen to be uniform, i.e., independent of the data (y_0, y_1) and (z_0, z_1) , is an open problem even for the nonlinearity in (6) for $n = 1$.

This is certainly one of the main open problems in the context of controllability of nonlinear PDE.

ACKNOWLEDGEMENTS

This work was supported in part by the grants PB96-0663 of the DGES (Spain), the EU TMR Project "Homogenization and Multiple Scales," a Foundation for the Authors of Excellent Ph.D. Theses in China, and the NSF of China (19901024).

BIBLIOGRAPHY

- [1] C. Bardos, G. Lebeau, and J. Rauch, "Sharp sufficient conditions for the observation, control, and stabilization of waves from the boundary," *SIAM J. Control Optim.*, 30, pp. 1024-1065, 1992.
- [2] P. Cannarsa, V. Komornik, and P. Loreti, "One-sided and internal controllability of semilinear wave equations with infinitely iterated logarithms," Preprint.
- [3] T. Cazenave and A. Haraux, "Équations d'évolution avec non linéarité logarithmique," *Ann. Fac. Sci. Toulouse*, 2, pp. 21-51, 1980.
- [4] B. Dehman, G. Lebeau, and E. Zuazua, "Stabilization and control for the subcritical semilinear wave equation," Preprint, 2002.
- [5] O. Yu. Émanuilov, "Boundary controllability of semilinear evolution equations," *Russian Math. Surveys*, 44, pp. 183-184, 1989.
- [6] I. Lasiecka and R. Triggiani, "Exact controllability of semilinear abstract systems with application to waves and plates boundary control problems," *Appl. Math. Optim.*, 23, pp. 109-154, 1991.
- [7] L. Li and X. Zhang, "Exact controllability for semilinear wave equations," *J. Math. Anal. Appl.*, 250, pp. 589-597, 2000.
- [8] J. L. Lions, *Contrôlabilité exacte, perturbations et systèmes distribués, Tome 1*, Rech. Math. Appl. 8, Masson, Paris, 1988.
- [9] X. Zhang, "Explicit observability estimate for the wave equation with potential and its application," *R. Soc. Lond. Proc. Ser. A Math. Phys. Eng. Sci.*, 456, pp. 1101-1115, 2000.
- [10] E. Zuazua, "Exact controllability for the semilinear wave equation," *J. Math. Pures Appl.*, 69, pp. 1-31, 1990.
- [11] E. Zuazua, "Exponential decay for semilinear wave equations with localized damping," *Comm. Partial Differential Equations*, 15, pp. 205-235, 1990.
- [12] E. Zuazua, "Exact boundary controllability for the semilinear wave equation," In: *Nonlinear partial differential equations and their applica-*

- tions, Collège de France Seminar, vol. X (Paris, 1987-1988), pp. 357-391, Pitman Res. Notes Math. Ser., 220, Longman Sci. Tech., Harlow, 1991.
- [13] E. Zuazua, "Exact controllability for semilinear wave equations in one space dimension," *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 10, pp. 109-129, 1993.

Problem 5.6

Some control problems in electromagnetics and fluid dynamics

Lorella Fatone

Dipartimento di Matematica Pura ed Applicata
Università di Modena e Reggio Emilia
Via Campi 213/b, 41100 Modena (MO)
Italy
`fatone.lorella@unimo.it`

Maria Cristina Recchioni

Istituto di Teoria delle Decisioni e Finanza Innovativa (DE.F.IN.)
Università di Ancona
Piazza Martelli 8, 60121 Ancona (AN)
Italy
`recchioni@posta.econ.unian.it`

Francesco Zirilli

Dipartimento di Matematica “G. Castelnuovo”
Università di Roma “La Sapienza”
Piazzale Aldo Moro 2, 00185 Roma
Italy
`f.zirilli@caspur.it`

1 INTRODUCTION

In recent years, as a consequence of the dramatic increases in computing power and of the continuing refinement of the numerical algorithms available, the numerical treatment of control problems for systems governed by partial differential equation; see, for example, [1], [3], [4], [5], [8]. The importance of these mathematical problems in many applications in science and technology cannot be overemphasized.

The most common approach to a control problem for a system governed by partial differential equations is to see the problem as a constrained non-linear optimization problem in infinite dimension. After discretization the

problem becomes a finite dimensional constrained nonlinear optimization problem that can be attacked with the usual iterative methods of nonlinear optimization, such as Newton or quasi-Newton methods. Note that the problem of the convergence, when the “discretization step goes to zero,” of the solutions computed in finite dimension to the solution of the infinite dimensional problem is a separate question and must be solved separately. When this approach is used an objective function evaluation in the nonlinear optimization procedure involves the solution of the partial differential equations that govern the system. Moreover, the evaluation of the gradient or Hessian of the objective function involves the solution of some kind of sensitivity equations for the partial differential equations considered. The nonlinear optimization procedure that usually involves function, gradient and Hessian evaluation is computationally very expensive.

This fact is a serious limitation to the use of control problems for systems governed by partial differential equations in real situations. However the approach previously described is very straightforward and does not use any of the special features present in every system governed by partial differential equations. So that, at least in some special cases, it should be possible to improve on this straightforward approach.

The purpose of this paper is to point out a problem, see [6], [2], where a new approach, that greatly improves on the previously described one, has been introduced and to suggest some other problems where, hopefully, similar improvements can be obtained. In particular, we propose two control problems of great relevance in several applications in science and technology and we suggest the (open) question of characterizing the optimal solution of these control problems as the solution of suitable systems of partial differential equations. If this question has an affirmative answer, high performance algorithms can be developed to solve the control problems proposed. Note that in [6], [2] this characterization has been made for some control problems in acoustics, thanks to the use of the Pontryagin maximum principle, and has permitted to develop high performance algorithms to solve these control problems. Moreover, we suggest the (open) question of using effectively the dynamic programming method to derive closed loop control laws for the control problems considered. For effective use of the dynamic programming method, we mean the possibility of computing a closed loop control law at approximately the same computational cost of solving the original problem when no control strategy is involved.

In section 2 we summarize the results obtained in [6], [2], and in section 3 we present two problems that we believe can be approached in a way similar to the one described in [6], [2].

2 PREVIOUS RESULTS

In [6], [2] a furtivity problem in time dependent acoustic obstacle scattering is considered. An obstacle of known acoustic impedance is hit by a known incident acoustic field. When hit by the incident acoustic field, the obstacle generates a scattered acoustic field. To make the obstacle furtive means to “minimize” the scattered field. The furtivity effect is obtained circulating on the boundary of the obstacle a “pressure current” that is a quantity whose physical dimension is: pressure divided by time. The problem consists in finding the optimal “pressure current” that “minimizes” the scattered field and the “size” of the pressure current employed. The mathematical model used to study this problem is a control problem for the wave equation, where the control function (i.e., the pressure current) influences the state variable (i.e., the scattered field) through a boundary condition imposed on the boundary of the obstacle, and the cost functional depends explicitly from both the state variable and the control function. Introducing an auxiliary variable and using the Pontryagin maximum principle (see [7]) in [6], [2] it is shown that the optimal control of this problem can be obtained from the solution of a system of two coupled wave equations. This system of wave equations is equipped with suitable initial, final, and boundary conditions. Thanks to this ingenious construction the solution of the optimal control problem can be obtained solving the system of wave equations without the necessity of going through the iterations implied in general by the nonlinear optimization procedure. This fact avoids many of the difficulties, that have been mentioned above, present in the general case. Finally, the system of wave equations is solved numerically using a highly parallelizable algorithm based on the operator expansion method (for more details, see [6], [2] and the references therein). Some numerical results obtained with this algorithm on simple test problems can be seen in the form of computer animations in the websites: <http://www.econ.unian.it/recchioni/w6>, <http://www.econ.unian.it/recchioni/w8>. In the following section, we suggest two problems where will be interesting to carry out a similar analysis.

3 TWO CONTROL PROBLEMS

Let \mathbf{R} be the set of real numbers, $\mathbf{x} = (x_1, x_2, x_3)^T \in \mathbf{R}^3$ (where the superscript T means transposed) be a generic vector of the three-dimensional real Euclidean space \mathbf{R}^3 , and let (\cdot, \cdot) , $\|\cdot\|$ and $[\cdot, \cdot]$ denote the Euclidean scalar product, the Euclidean vector norm and the vector product in \mathbf{R}^3 , respectively.

The first problem suggested is a “masking” problem in time-dependent electromagnetic scattering. Let $\Omega \subset \mathbf{R}^3$ be a bounded simply connected open set (i.e., the obstacle) with locally Lipschitz boundary $\partial\Omega$. Let $\overline{\Omega}$ denote the closure of Ω and $\mathbf{n}(\mathbf{x}) = (n_1(\mathbf{x}), n_2(\mathbf{x}), n_3(\mathbf{x}))^T \in \mathbf{R}^3$, $\mathbf{x} \in \partial\Omega$ be the

outward unit normal vector in \mathbf{x} for $\mathbf{x} \in \partial\Omega$. Note that $\mathbf{n}(\mathbf{x})$ exists almost everywhere in \mathbf{x} for $\mathbf{x} \in \partial\Omega$. We assume that the obstacle Ω is characterized by an electromagnetic boundary impedance $\chi > 0$. Note that $\chi = 0$ ($\chi = +\infty$) corresponds to consider a perfectly conducting (insulating) obstacle. Let $\mathbf{R}^3 \setminus \Omega$ be filled with a homogeneous isotropic medium characterized by a constant electric permittivity $\epsilon > 0$, a constant magnetic permeability $\nu > 0$, zero electric conductivity, zero free charge density, and zero free current density.

Let $(\mathcal{E}^i(\mathbf{x}, t), \mathcal{B}^i(\mathbf{x}, t)), (\mathbf{x}, t) \in \mathbf{R}^3 \times \mathbf{R}$ (where \mathcal{E}^i is the electric field and \mathcal{B}^i is the magnetic induction field) be the incoming electromagnetic field propagating in the medium filling $\mathbf{R}^3 \setminus \Omega$ and satisfying the Maxwell equations (1)-(3) in $\mathbf{R}^3 \times \mathbf{R}$. Let $(\mathcal{E}^s(\mathbf{x}, t), \mathcal{B}^s(\mathbf{x}, t)), (\mathbf{x}, t) \in (\mathbf{R}^3 \setminus \bar{\Omega}) \times \mathbf{R}$ be the electromagnetic field scattered by the obstacle Ω when hit by the incoming field $(\mathcal{E}^i(\mathbf{x}, t), \mathcal{B}^i(\mathbf{x}, t)), (\mathbf{x}, t) \in \mathbf{R}^3 \times \mathbf{R}$. The scattered electric field \mathcal{E}^s and the scattered magnetic induction field \mathcal{B}^s satisfy the following equations:

$$\left(\text{curl} \mathcal{E}^s + \frac{\partial \mathcal{B}^s}{\partial t} \right) (\mathbf{x}, t) = \mathbf{0}, \quad (\mathbf{x}, t) \in (\mathbf{R}^3 \setminus \bar{\Omega}) \times \mathbf{R}, \quad (1)$$

$$\left(\text{curl} \mathcal{B}^s - \frac{1}{c^2} \frac{\partial \mathcal{E}^s}{\partial t} \right) (\mathbf{x}, t) = \mathbf{0}, \quad (\mathbf{x}, t) \in (\mathbf{R}^3 \setminus \bar{\Omega}) \times \mathbf{R}, \quad (2)$$

$$\text{div} \mathcal{B}^s(\mathbf{x}, t) = 0, \quad \text{div} \mathcal{E}^s(\mathbf{x}, t) = 0, \quad (\mathbf{x}, t) \in (\mathbf{R}^3 \setminus \bar{\Omega}) \times \mathbf{R}, \quad (3)$$

$$\begin{aligned} & [\mathbf{n}(\mathbf{x}), \mathcal{E}^s(\mathbf{x}, t)] - c\chi[\mathbf{n}(\mathbf{x}), [\mathbf{n}(\mathbf{x}), \mathcal{B}^s(\mathbf{x}, t)]] = \\ & -[\mathbf{n}(\mathbf{x}), \mathcal{E}^i(\mathbf{x}, t)] + c\chi[\mathbf{n}(\mathbf{x}), [\mathbf{n}(\mathbf{x}), \mathcal{B}^i(\mathbf{x}, t)]], \quad (\mathbf{x}, t) \in \partial\Omega \times \mathbf{R}, \end{aligned} \quad (4)$$

$$\mathcal{E}^s(\mathbf{x}, t) = O\left(\frac{1}{r}\right), \quad [\mathcal{B}^s(\mathbf{x}, t), \hat{\mathbf{x}}] - \frac{1}{c} \mathcal{E}^s(\mathbf{x}, t) = o\left(\frac{1}{r}\right), \quad r \rightarrow +\infty, \quad t \in \mathbf{R}, \quad (5)$$

where $\mathbf{0} = (0, 0, 0)^T$, $c = 1/\sqrt{\epsilon\nu}$, $r = \|\mathbf{x}\|$, $\mathbf{x} \in \mathbf{R}^3$, $\hat{\mathbf{x}} = \frac{\mathbf{x}}{\|\mathbf{x}\|}$, $\mathbf{x} \neq \mathbf{0}$, $\mathbf{x} \in \mathbf{R}^3$, $O(\cdot)$ and $o(\cdot)$ are the Landau symbols, and $\text{curl} \cdot$ and $\text{div} \cdot$ denote the curl and the divergence operator of \cdot with respect to the \mathbf{x} variables respectively.

A classical problem in electromagnetics consists in the recognition of the obstacle Ω through the knowledge of the incoming electromagnetic field and of the scattered field $(\mathcal{E}^s(\mathbf{x}, t), \mathcal{B}^s(\mathbf{x}, t)), (\mathbf{x}, t) \in (\mathbf{R}^3 \setminus \bar{\Omega}) \times \mathbf{R}$ solution of (1)-(5). In the above situation, Ω plays a “passive” (“static”) role. We want to make the obstacle Ω “active” (“dynamic”) in the sense that, thanks to a suitable control function chosen in a proper way, the obstacle itself tries to react to the incoming electromagnetic field producing a scattered field that looks like the field scattered by a preassigned obstacle D (the “mask”) with impedance χ' . We suggest to consider the following control problem:

Problem 1: Electromagnetic “Masking” Problem: *Given an incoming electromagnetic field $(\mathcal{E}^i, \mathcal{B}^i)$, an obstacle Ω and its electromagnetic boundary impedance χ , and given an obstacle D such that $D \subseteq \Omega$ with electromagnetic boundary impedance χ' , choose a vector control function ψ defined on*

the boundary of the obstacle $\partial\Omega$ for $t \in \mathbf{R}$ and appearing in the boundary condition satisfied by the scattered electromagnetic field on $\partial\Omega$, in order to minimize a cost functional that measures the “difference” between the electromagnetic field scattered by Ω , i.e., $(\mathcal{E}^s, \mathcal{B}^s)$, and the electromagnetic field scattered by D , i.e., $(\mathcal{E}_D^s, \mathcal{B}_D^s)$, when Ω and D respectively are hit by the incoming field $(\mathcal{E}^i, \mathcal{B}^i)$, and the “size” of the vector control function employed.

The control function ψ has the physical dimension of an electric field and the action of the optimal control electric field on the boundary of the obstacle makes the obstacle “active” (“dynamic”) and able to react to the incident electromagnetic field to become “unrecognizable,” that is “ Ω will do its best to appear as his mask D .”

The second control problem we suggest to consider is a control problem in fluid dynamics. Let us consider an obstacle $\Omega_t, t \in \mathbf{R}$, that is a rigid body, assumed homogeneous, moving in \mathbf{R}^3 with velocity $\tilde{\mathbf{v}} = \tilde{\mathbf{v}}(\mathbf{x}, t), (\mathbf{x}, t) \in \Omega_t \times \mathbf{R}$. Moreover for $t \in \mathbf{R}$ the obstacle $\Omega_t \subset \mathbf{R}^3$ is a bounded simply connected open set. For $t \in \mathbf{R}$ let $\boldsymbol{\xi} = \boldsymbol{\xi}(t)$ be the position of the center of mass of the obstacle Ω_t . The motion of the obstacle is completely described by the velocity $\mathbf{w} = \mathbf{w}(\boldsymbol{\xi}, t), t \in \mathbf{R}$ of the center of mass of the obstacle (i.e., $\mathbf{w} = \frac{d\boldsymbol{\xi}}{dt}, t \in \mathbf{R}$), the angular velocity $\boldsymbol{\omega} = \boldsymbol{\omega}(\boldsymbol{\xi}, t), t \in \mathbf{R}$ of the obstacle around the instantaneous rotation axis going through the center of mass $\boldsymbol{\xi} = \boldsymbol{\xi}(t), t \in \mathbf{R}$ and the necessary initial conditions. Note that the velocities of the points belonging to the obstacle $\tilde{\mathbf{v}}(\mathbf{x}, t), (\mathbf{x}, t) \in \Omega_t \times \mathbf{R}$ can be expressed in terms of $\mathbf{w}(\boldsymbol{\xi}, t), \boldsymbol{\omega}(\boldsymbol{\xi}, t), t \in \mathbf{R}$. Let $\mathbf{R}^3 \setminus \Omega_t, t \in \mathbf{R}$ be filled with a Newtonian incompressible viscous fluid of viscosity η . We assume that both the density of the fluid and the temperature are constant. For example, $\Omega_t, t \in \mathbf{R}$ can be a submarine or an airfoil immersed in an incompressible viscous fluid. Let $\mathbf{v} = (v_1, v_2, v_3)^T$ and p be the velocity field and the pressure field of the fluid, respectively, \mathbf{f} be the density of the external forces per mass unit acting on the fluid, and $\mathbf{v}_{-\infty}$ be an assigned solenoidal vector field. We assume that in the limit $t \rightarrow -\infty$ the body Ω_t is at rest in the position $\Omega_{-\infty}$. Under these assumptions, we have that in the reference frame given by $\mathbf{x} = (x_1, x_2, x_3)^T$ the following system of Navier-Stokes equations holds:

$$\frac{\partial \mathbf{v}}{\partial t}(\mathbf{x}, t) + (\mathbf{v}(\mathbf{x}, t), \nabla) \mathbf{v}(\mathbf{x}, t) - \eta \Delta \mathbf{v}(\mathbf{x}, t) + \nabla p(\mathbf{x}, t) = \mathbf{f}(\mathbf{x}, t), \quad (6)$$

$$(\mathbf{x}, t) \in (\mathbf{R}^3 \setminus \bar{\Omega}_t) \times \mathbf{R},$$

$$\operatorname{div} \mathbf{v}(\mathbf{x}, t) = 0, (\mathbf{x}, t) \in (\mathbf{R}^3 \setminus \bar{\Omega}_t) \times \mathbf{R}, \quad (7)$$

$$\lim_{t \rightarrow -\infty} \mathbf{v}(\mathbf{x}, t) = \mathbf{v}_{-\infty}(\mathbf{x}), \mathbf{x} \in \mathbf{R}^3 \setminus \bar{\Omega}_{-\infty}, \mathbf{v}(\mathbf{x}, t) = \tilde{\mathbf{v}}(\mathbf{x}, t), (\mathbf{x}, t) \in \partial\Omega_t \times \mathbf{R}. \quad (8)$$

In (6) we have $\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3} \right)^T$ and

$$(\mathbf{v}, \nabla) \mathbf{v} = \left(\sum_{j=1}^3 v_j \frac{\partial v_1}{\partial x_j}, \sum_{j=1}^3 v_j \frac{\partial v_2}{\partial x_j}, \sum_{j=1}^3 v_j \frac{\partial v_3}{\partial x_j} \right)^T.$$

The boundary condition in (8) requires that the fluid velocity \mathbf{v} and the velocity of the obstacle $\tilde{\mathbf{v}}$ are equal on the boundary of the obstacle for $t \in \mathbf{R}$. We want to consider the problem associated to the choice of a maneuver $\mathbf{w}(\boldsymbol{\xi}, t)$, $\boldsymbol{\omega}(\boldsymbol{\xi}, t)$, $t \in \mathbf{R}$ connecting two given states that minimizes the work done by the obstacle Ω_t , $t \in \mathbf{R}$ against the fluid going from the initial state to the final state, and the “size” of the maneuver employed. Note that in this context a maneuver connecting two given states is made of two functions $\mathbf{w}(\boldsymbol{\xi}, t)$, $\boldsymbol{\omega}(\boldsymbol{\xi}, t)$, $t \in \mathbf{R}$ such that $\lim_{t \rightarrow \pm\infty} \mathbf{w}(\boldsymbol{\xi}, t) = \mathbf{w}^\pm$ and $\lim_{t \rightarrow \pm\infty} \boldsymbol{\omega}(\boldsymbol{\xi}, t) = \boldsymbol{\omega}^\pm$, where \mathbf{w}^\pm and $\boldsymbol{\omega}^\pm$ are preassigned. The couple $(\mathbf{w}^-, \boldsymbol{\omega}^-)$ is the initial state and the couple $(\mathbf{w}^+, \boldsymbol{\omega}^+)$ is the final state. For simplicity, we have assumed $(\mathbf{w}^-, \boldsymbol{\omega}^-) = (\mathbf{0}, \mathbf{0})$. We formulate the following problem:

Problem 2: “Drag” Optimization Problem: *Given a rigid obstacle Ω_t , $t \in \mathbf{R}$ moving in a Newtonian fluid characterized by a viscosity η and the initial condition and forces acting on the fluid, and given the initial state $(\mathbf{0}, \mathbf{0})$ and the final state $(\mathbf{w}^+, \boldsymbol{\omega}^+)$, choose a maneuver connecting these two states in order to minimize a cost functional that measures the work that the obstacle Ω_t , $t \in \mathbf{R}$ must exert on the fluid to make the maneuver, and the “size” of the maneuver employed.*

From the previous considerations, several problems arise. The first one is connected with the question of formulating problem 1 and problem 2 as control problems. In [2] we suggest a possible formulation of a furtivity problem similar to problem 1 as a control problem. In particular, the open question that we suggest is how problem 1 and problem 2 should be formulated as control problems whose optimal solutions can be determined solving suitable systems of partial differential equations via an ingenious way of using the Pontryagin maximum principle as done in [2], [6]. The relevance of this formulation lies in the fact that avoids computationally expensive iterative procedures to solve the control problems considered. Moreover, a second open question is the derivation of closed loop control laws at an affordable computational cost for the control problems associated to Problem 1 and Problem 2.

Furthermore many variations of problem 1 and 2 can be considered. For example in problem 1 we have assumed, for simplicity, that the “mask” is a passive obstacle, that is $(\mathcal{E}_D^s(\mathbf{x}, t), \mathcal{B}_D^s(\mathbf{x}, t))$, $(\mathbf{x}, t) \in (\mathbf{R}^3 \setminus \overline{D}) \times \mathbf{R}$ is the solution of problem (1)-(5) when Ω, χ are replaced with D, χ' , respectively. In a more general situation also the “mask” can be an active obstacle. Finally, problem 1 and 2 are examples of control problems for systems governed by the Maxwell equations and the Navier-Stokes equations, respectively. Many other examples relevant in several application fields involving different systems of partial differential equations can be considered.

BIBLIOGRAPHY

- [1] T. S. Angell, A. Kirsch, and R. E. Kleinman, "Antenna control and optimization," *Proceedings of the IEEE*, 79, pp. 1559-1568, 1991.
- [2] L. Fatone, M. C. Recchioni, and F. Zirilli, "Furtivity and masking problems in time dependent acoustic obstacle scattering," forthcoming in *Third ISAAC Congress Proceedings*.
- [3] J. W. He, R. Glowinski, R. Metcalfe, A. Nordlander, and J. Periaux, "Active control and drag optimization for flow past a circular cylinder," *Journal of Computational Physics*, 163, pp. 83-117, 2000.
- [4] J. E. Lagnese, "A singular perturbation problem in exact controllability of the Maxwell system," *ESAIM Control, Optimization and Calculus of Variations*, 6, pp. 275-289, 2001.
- [5] J. Luniley and P. Blossery, "Control of turbulence," *Annual Review of Fluid Mechanics*, 30, pp. 311-327, 1998.
- [6] F. Mariani, M. C. Recchioni, and F. Zirilli, "The use of the Pontryagin maximum principle in a furtivity problem in time-dependent acoustic obstacle scattering," *Waves in Random Media*, 11, pp. 549-575, 2001.
- [7] L. Pontriaguine, V. Boltianski, R. Gamkréldzé, and F. Micktchenko, *Théorie Mathématique des Processus Optimaux*, Editions Mir, Moscow, 1974.
- [8] S. S. Sritharan, "An introduction to deterministic and stochastic control of viscous flow," In: *Optimal Control of Viscous Flow*, S. S. Sritharan eds., SIAM, Philadelphia, pp. 1-42, 1998.

PART 6

Stability, Stabilization

Problem 6.1

Copositive Lyapunov functions

M. K. Camlibel

Department of Mathematics

University of Groningen

P.O. Box 800, 9700 AV Groningen

The Netherlands

k.camlibel@math.rug.nl

J. M. Schumacher

Department of Econometrics and Operations Research

Tilburg University

P.O. Box 90153, 5000 LE Tilburg

The Netherlands

j.m.schumacher@kub.nl

1 PRELIMINARIES

The following notational conventions and terminology will be in force. Inequalities for vectors are understood component-wise. Given two matrices M and N with the same number of columns, the notation $\text{col}(M, N)$ denotes the matrix obtained by stacking M over N . Let M be a matrix. The *submatrix* M_{JK} of M is the matrix whose entries lie in the rows of M indexed by the set J and the columns indexed by the set K . For square matrices M , M_{JJ} is called a *principal submatrix* of M . A symmetric matrix M is said to be *non-negative (nonpositive) definite* if $x^T M x \geq 0$ ($x^T M x \leq 0$) for all x . It is said to be *positive (negative) definite* if the equalities hold only for $x = 0$. Sometimes, we write $M > 0$ ($M \geq 0$) to indicate that M is positive definite (non-negative definite), respectively. We say that a square matrix M is *Hurwitz* if its eigenvalues have negative real parts. A pair of matrices (A, C) is *observable* if the corresponding system $\dot{x} = Ax$, $y = Cx$ is observable, equivalently if $\text{col}(C, CA, \dots, CA^{n-1})$ is of rank n where A is of order n .

2 MOTIVATION

Lyapunov stability theory is one of the ever green topics in systems and control. For (finite dimensional) linear systems, the following theorem is very well-known.

Theorem 1:[3, Theorem 1.2]: The following conditions are equivalent.

1. The system $\dot{x} = Ax$ is asymptotically stable.
2. The Lyapunov equation $A^T P + PA = Q$ has a positive definite symmetric solution P for any negative definite symmetric matrix Q .

As a refinement, we can replace the last statement by

2'. *The Lyapunov equation $A^T P + PA = Q$ has a positive definite symmetric solution P for any nonpositive definite symmetric matrix Q such that the pair (A, Q) is observable.*

An interesting application is to the stability of the so-called *switched systems*. Consider the system

$$\dot{x} = A_\sigma x \tag{1}$$

where the *switching signal* $\sigma : [0, \infty) \rightarrow \{1, 2\}$ is a piecewise constant function. We assume that it has a finite number of discontinuities over finite time intervals in order to rule out infinitely fast switching. A strong notion of stability for the system (1) is the requirement of stability for arbitrary switching signals.

The dynamics of (1) coincides with one of the linear subsystems if the switching signal is constant, i.e., there are no switchings at all. This leads us to an obvious necessary condition: stability of each subsystem. Another extreme case would emerge if there exists a common Lyapunov function for the subsystems. Indeed, such a Lyapunov function would immediately prove the stability of (1). An earlier paper [8] pointed out the importance of commutation relations between A_1 and A_2 in finding a common Lyapunov function. More precisely, it has been shown that if A_1 and A_2 are Hurwitz and commutative then they admit a common Lyapunov function. In [1, 6], the commutation relations of subsystems are studied further in a Lie algebraic framework and sufficient conditions for the existence of a common Lyapunov function are presented. Notice that the results of [1] are stronger than those in [6]. However, we prefer to restate [6, Theorem 2] for simplicity.

Theorem 2: If A_i is a Hurwitz matrix for $i = 1, 2$ and the Lie algebra $\{A_1, A_2\}_{LA}$ is solvable then there exists a positive definite matrix P such that $A_i^T P + PA_i < 0$ for $i = 1, 2$.

So far, we quoted some known results. Our main goal is to pose two open problems that can be viewed as extensions of Theorems 2 and 2 for a class of piecewise linear systems. More precisely, we will consider systems of the form

$$\dot{x} = A_i x \quad \text{for } C_i x \geq 0 \quad i = 1, 2. \quad (2)$$

Here the cones $\mathcal{C}_i = \{x \mid C_i x \geq 0\}$ do not necessarily cover the whole x -space. We assume that

- a. there exists a (possibly discontinuous) function f such that (2) can be described by $\dot{x} = f(x)$ for all $x \in \mathcal{C}_1 \cup \mathcal{C}_2$, and
- b. for each initial state $x_0 \in \mathcal{C}_1 \cup \mathcal{C}_2$, there exists a unique solution x in the sense of Carathéodory, i.e., $x(t) = x_0 + \int_0^t f(x(\tau)) d\tau$.

A natural example[b.] of such piecewise linear dynamics is a linear complementarity system (see [9]) of the form

$$\dot{x} = Ax + Bu, \quad y = Cx + Du$$

$$\{(u(t) \geq 0) \text{ and } (y(t) \geq 0) \text{ and } (u(t) = 0 \text{ or } y(t) = 0)\} \text{ for all } t \geq 0$$

where $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times 1}$, $C \in \mathbb{R}^{1 \times n}$, and $D \in \mathbb{R}$. If $D > 0$ this system can be put into the form of (2) with $A_1 = A$, $C_1 = C$, $A_2 = A - BD^{-1}C$, and $C_2 = -C$. Equivalently, it can be described by

$$\dot{x} = f(x) \quad (3)$$

where $f(x) = Ax$ if $Cx \geq 0$ and $f(x) = (A - BD^{-1}C)x$ if $Cx \leq 0$. Note that f is Lipschitz continuous and hence (3) admits a unique (continuously differentiable) solution x for all initial states x_0 .

One way of studying the stability of the system (2) is simply to utilize Theorem 2. However, there are some obvious drawbacks:

- i. It requires positive definiteness of the common Lyapunov function whereas the positivity on a cone is enough for the system (2).
- ii. It considers *any* switching signal whereas the initial state determines *the* switching signal in (2).

In the next section, we focus on ways of eliminating the conservatism mentioned in 2.

3 DESCRIPTION OF THE PROBLEMS

First, we need to introduce some nomenclature. A matrix M is said to be *copositive* (*strictly copositive*) with respect to a cone \mathcal{C} if $x^T M x \geq 0$ ($x^T M x > 0$) for all nonzero $x \in \mathcal{C}$. We use the notation $M \stackrel{\mathcal{C}}{\succcurlyeq} 0$ and $M \stackrel{\mathcal{C}}{\succ} 0$

respectively for copositivity and strict copositivity. When the cone \mathcal{C} is clear from the context we just write \succ or \prec .

The first problem that we propose calls for an extension of Theorem 2 for linear dynamics restricted to a cone.

Problem 1: Let a square matrix A and a cone $\mathcal{C} = \{x \mid Cx \geq 0\}$ be given. Determine necessary and sufficient conditions for the existence of a symmetric matrix P such that $P \succ 0$ and $A^T P + P A \prec 0$.

An immediate necessary condition for the existence of such a matrix P is that the matrix A should not have any eigenvectors in the cone \mathcal{C} corresponding to its positive eigenvalues.

Once Problem 3 solved, it would be natural to take a step further by attempting to extend Theorem 2 to the systems (2). In other words, it would be natural to attack the following problem.

Problem 2: Let two square matrices A_1, A_2 , and two cones $\mathcal{C}_1 = \{x \mid C_1 x \geq 0\}$, $\mathcal{C}_2 = \{x \mid C_2 x \geq 0\}$ be given. Determine sufficient conditions for the existence of a symmetric matrix P such that $P \stackrel{\mathcal{C}_i}{\succ} 0$ and $A_i^T P + P A_i \stackrel{\mathcal{C}_i}{\prec} 0$ for $i = 1, 2$.

4 ON COPOSITIVE MATRICES

This last section discusses copositive matrices in order to provide a starting point for further investigation of the proposed problems.

The class of copositive matrices occurs in optimization theory and particularly in the study of the linear complementarity problem [2]. We quote from [4] the following theorem which provides a characterization of copositive matrices.

Theorem 2. A symmetric matrix M is (strictly) copositive with respect to the cone $\{x \mid x \geq 0\}$ if and only if every principal submatrix of M has no eigenvector $v > 0$ with associated eigenvalue $(\lambda \leq 0) \lambda < 0$.

Since the number of principal submatrices of a matrix of order n is roughly 2^n , this result has a practical disadvantage. In fact, Murty and Kabadi [7] showed that testing for copositivity is NP-complete. An interesting subclass of copositive matrices are the ones that are equal to the sum of a nonnegative definite matrix and a non-negative matrix. This class of matrices is studied in [5] where a relatively more tractable algorithm has been presented for checking if a given matrix belongs to the class or not.

BIBLIOGRAPHY

- [1] A. A. Agrachev and D. Liberzon, "Lie-algebraic stability criteria for switched systems," *SIAM J. Control Optim.*, 40(1):253–269, 2001.
- [2] R. W. Cottle, J.-S. Pang, and R. E. Stone, *The Linear Complementarity Problem*, Academic Press, Inc., Boston, 1992.
- [3] Z. Gajić and M. T. J. Qureshi, *Lyapunov Matrix Equation in System Stability and Control*, volume 195 of *Mathematics in Science and Engineering*, Academic Press, San Diego, 1995.
- [4] W. Kaplan, "A test for copositive matrices," *Linear Algebra Appl.*, 313:203–206, 2000.
- [5] W. Kaplan, "A copositivity probe," *Linear Algebra Appl.*, 337:237–251, 2001.
- [6] D. Liberzon, J. P. Hespanha, and A. S. Morse, "Stability of switched systems: A Lie-algebraic condition," *Systems Control Lett.*, 37:117–122, 1999.
- [7] K. Murty and S. N. Kabadi, "Some NP-complete problems in quadratic and nonlinear programming," *Math. Programming*, 39:117–129, 1987.
- [8] K. S. Narendra and J. Balakrishnan, "A common Lyapunov function for stable LTI systems with commuting A -matrices," *IEEE Trans. Automatic Control*, 39:2469–2471, 1994.
- [9] A. J. van der Schaft and J. M. Schumacher, "Complementarity modelling of hybrid systems," *IEEE Transactions on Automatic Control*, 43(4):483–490, 1998.

Problem 6.2

The strong stabilization problem for linear time-varying systems

Avraham Feintuch

Department of Mathematics
Ben-Gurion University of the Negev
Beer-Sheva
Israel
abie@math.bgu.ac.il

1 DESCRIPTION OF THE PROBLEM

I will formulate the strong stabilization problem in the formalism of the operator theory of systems. In this framework, a linear system is a linear transformation L acting on a Hilbert space H that is equipped with a natural time structure, which satisfies the standard physical realizability condition known as causality. To simplify the formulation, we choose H to be the sequence space $l^2[0, \infty) = \{ \langle x_0, x_1, \dots \rangle : x_i \in \mathbb{C}^n, \sum \|x_i\|^2 < \infty \}$ and denote by P_n the truncation projection onto the subspace generated by the first n vectors $\{e_0, \dots, e_n\}$ of the standard orthonormal basis on H . Causality of L is expressed as $P_n L = P_n L P_n$ for all non-negative integers n . A linear system L is *stable* if it is a bounded operator on H . A fundamental issue that was studied in both classical and modern control theory was that of internal stabilization of unstable systems by feedback. It is generally acknowledged that the paper of Youla et al. [2] was a landmark event in this study and in fact the issue of strong stabilization was first raised there. It was quickly seen [5] that while this paper restricted itself to the classical case of rational transfer functions its ideas were given to abstraction to much more general frameworks. We briefly describe the one relevant to our discussion.

For a linear system L , its graph $G(L)$ is the range of the operator $\begin{bmatrix} I \\ L \end{bmatrix}$ defined on the domain $\mathcal{D}(L) = \{x \in H : Lx \in H\}$. $G(L)$ is a subspace of $H \oplus H$. The operator $\begin{bmatrix} I & C \\ L & -I \end{bmatrix}$ defined on $\mathcal{D}(L) \oplus \mathcal{D}(C)$ is called the

feedback system $\{L, C\}$ with plant L and compensator C , and $\{L, C\}$ is stable if it has a bounded causal inverse. L is *stabilizable* if there exists a causal linear system C (not necessarily stable) such that $\{L, C\}$ is stable.

The analogue of the result of Youla et al. which characterizes all stabilizable linear systems and parametrizes all stabilizers was given by Dale and Smith [4]:

Theorem 1.[[6], p. 103] : Suppose L is a linear system and there exist causal stable systems $M, N, X, Y, \hat{M}, \hat{N}, \hat{X}, \hat{Y}$ such that (1) $G(L) =$

$$\text{Ran} \begin{bmatrix} M \\ N \end{bmatrix} = \text{Ker} \begin{bmatrix} -\hat{N} & \hat{M} \end{bmatrix}, (2) \begin{bmatrix} M & -\hat{X} \\ N & \hat{Y} \end{bmatrix} = \begin{bmatrix} Y & X \\ -\hat{N} & \hat{M} \end{bmatrix}^{-1}.$$

Then

(1) L is stabilizable

(2) C stabilizes L if and only if

$$G(C) = \text{Ran} \begin{bmatrix} \hat{Y} - NQ \\ \hat{X} + MQ \end{bmatrix} = \text{Ker} \begin{bmatrix} -(X + Q\hat{M}) & Y - Q\hat{N} \end{bmatrix}, \text{ where } Q$$

varies over all stable linear systems.

The Strong Stabilization Problem is:

Suppose L is stabilizable. Can internal stability be achieved with C itself a stable system? In such a case, L is said to be strongly stabilizable.

Theorem 2.[[6], p.108]: A linear system L with property (1), (2) of Theorem 1 is stabilized by a stable C if and only if $\hat{M} + \hat{N}C$ is an invertible operator. Equivalently, a stable C stabilizes L if and only if $M + CN$ is an invertible operator (by an invertible operator we mean that its inverse is also bounded).

It is not hard to show that in fact the same C works in both cases; i.e., $M + CN$ is invertible if and only if $\hat{M} + \hat{N}C$ is invertible. So here is the precise mathematical formulation of the problem:

Given causal stable systems M, N, X, Y such that $XM + YN = I$. Does there exist a causal stable system C such that $M + CN$ is invertible?

2 MOTIVATION AND HISTORY OF THE PROBLEM

The notion of strong internal stabilization was introduced in the classical paper of Youla et al. [2] and was solved for rational transfer functions. Another formulation was given in [1]. An approach to the classical problem from the point of view described here was first given in [9]. Recently sufficient conditions for the existence of strongly stabilizing controllers were formulated from the point of view of H^∞ control problems. The latest such effort is [7].

It is of interest to write that our formulation of the strong stabilization problem connects it to an equivalent problem in Banach algebras, the question of 1-stability of a Banach algebra: given a pair of elements $\{a, b\}$ in a Ba-

nach algebra B which satisfies the Bezout identity $xa + yb = 1$ for some $x, y \in B$, does there exist $c \in B$: $a + cb$ is a unit? This was shown to be the case for $B = H^\infty$ by Treil [8] and this proves that every stabilizable scalar time-invariant system is strongly stabilizable over the complex number field. The matrix analogue to Treil's result is not known. It is interesting that the Banach algebra $B(H)$ of all bounded linear operators on a given Hilbert space H is not 1-stable [3]. Our strong stabilization problem is the question whether nest algebras are 1-stable.

BIBLIOGRAPHY

- [1] B. D. O. Anderson, "A note on the Youla-Bongiorno-Lu condition," *Automatica* 12 (1976), 387-388.
- [2] J. J. Bongiorno, C. N. Lu, D. C. Youla, "Single-loop feedback stabilization of linear multivariable plants," *Automatica* 10 (1974), 159-173.
- [3] G. Corach, A. Larotonda, "Stable range in Banach algebras," *J. Pure and Applied Alg.* 32 (1984), 289-300.
- [4] W. Dale, M. Smith, "Stabilizability and existence of system representations for discrete-time, time-varying systems," *SIAM J. Cont. and Optim.* 31 (1993), 1538-1557.
- [5] C. A. Desoer, R. W. Liu, J. Murray, R. Saecks, "Feedback system design: the factorial representation approach to analysis and synthesis," *IEEE Trans. Auto. Control* AC-25 (1980), 399-412.
- [6] A. Feintuch, "Robust Control Theory on Hilbert Space," *Applied Math. Sciences* 130, Springer, 1998.
- [7] H. Ozbay, M. Zeren, "On the strong stabilization and stable H^∞ -controller design problems for MIMO systems," *Automatica* 36 (2000), 1675-1684.
- [8] S. Treil, "The stable rank of the algebra H^∞ equals 1," *J. Funct. Anal.* 109 (1992), 130-154.
- [9] M. Vidyasagar, *Control System Synthesis: A Factorization Approach*, M.I.T. Press, 1985.

Problem 6.3

Robustness of transient behavior

Diederich Hinrichsen, Elmar Plischke,
and Fabian Wirth

Zentrum für Technomathematik

Universität Bremen

28334 Bremen

Germany

{dh, elmar, fabian}@math.uni-bremen.de

1 DESCRIPTION OF THE PROBLEM

By definition, a system of the form

$$\dot{x}(t) = Ax(t), \quad t \geq 0 \tag{1}$$

($A \in \mathbb{K}^{n \times n}$, $\mathbb{K} = \mathbb{R}, \mathbb{C}$) is exponentially stable if and only if there are constants $M \geq 1$, $\beta < 0$ such that

$$\|e^{At}\| \leq Me^{\beta t}, \quad t \geq 0. \tag{2}$$

The respective roles of the two constants in this estimate are quite different. The exponent $\beta < 0$ determines the long-term behavior of the system, whereas the factor $M \geq 1$ bounds its short-term or transient behavior. In applications large transients may be unacceptable. This leads us to the following stricter stability concept.

Definition 1: Let $M \geq 1$, $\beta < 0$. A matrix $A \in \mathbb{K}^{n \times n}$ is called (M, β) -stable if (2) holds.

Here $\beta < 0$ and $M \geq 1$ can be chosen in such a way that (M, β) -stability guarantees both an acceptable decay rate and an acceptable transient behavior.

For any $A \in \mathbb{K}^{n \times n}$ let $\gamma(A)$ denote the spectral abscissa of A , i.e., the maximum of the real parts of the eigenvalues of A . It is well-known that $\gamma(A) < 0$ implies exponential stability. More precisely, for every $\beta > \gamma(A)$ there exists a constant $M \geq 1$ such that (2) is satisfied. However, it is unknown how to determine the minimal value of M such that (2) holds for a given $\beta \in (\gamma(A), 0)$.

Problem 1:

- a) Given $A \in \mathbb{K}^{n \times n}$ and $\beta \in (\gamma(A), 0)$, determine analytically the minimal value $M_\beta(A)$ of $M \geq 1$ for which A is (M, β) -stable.
- b) Provide easily computable formulas for upper and lower bounds for $M_\beta(A)$ and analyze their conservatism.

Associated to this problem is the design problem for linear control systems of the form

$$\dot{x} = Ax + Bu, \quad (3)$$

where $(A, B) \in \mathbb{K}^{n \times n} \times \mathbb{K}^{n \times m}$. Assume that a desired transient and stability behavior for the closed loop is prescribed by given constants $M \geq 1, \beta < 0$, then the pair (A, B) is called (M, β) -*stabilizable* (by state feedback), if there exists an $F \in \mathbb{K}^{m \times n}$ such that $A - BF$ is (M, β) -stable.

Problem 2:

- a) Given constants $M \geq 1, \beta < 0$, characterize the set of (M, β) -stabilizable pairs $(A, B) \in \mathbb{K}^{n \times n} \times \mathbb{K}^{n \times m}$.
- b) Provide a method for the computation of (M, β) -stabilizing feedbacks F for (M, β) -stabilizable pairs (A, B) .

In order to account for uncertainties in the model, we consider systems described by

$$\dot{x} = A_\Delta x = (A + D\Delta E)x,$$

where $A \in \mathbb{K}^{n \times n}$ is the nominal system matrix, $D \in \mathbb{K}^{n \times \ell}$ and $E \in \mathbb{K}^{q \times n}$ are given structure matrices, and $\Delta \in \mathbb{K}^{\ell \times q}$ is an unknown perturbation matrix for which only a bound of the form $\|\Delta\| \leq \delta$ is assumed to be known.

Problem 3:

- a) Given $A \in \mathbb{K}^{n \times n}$, $D \in \mathbb{K}^{n \times \ell}$ and $E \in \mathbb{K}^{q \times n}$, determine analytically the (M, β) -*stability radius* defined by

$$r_{(M, \beta)}(A; D, E) = \inf \left\{ \|\Delta\| \in \mathbb{K}^{\ell \times q}, \exists \tau > 0 : \|e^{(A + D\Delta E)\tau}\| \geq Me^{\beta\tau} \right\}. \quad (4)$$

- b) Provide an algorithm for the calculation of this quantity.
- c) Determine easily computable upper and lower bounds for $r_{(M, \beta)}(A; D, E)$.

The two previous problems can be thought of as steps towards the following final problem.

Problem 4: Given a system $(A, B) \in \mathbb{K}^{n \times n} \times \mathbb{K}^{n \times m}$, a desired transient behavior described by $M \geq 1, \beta < 0$, and matrices $D \in \mathbb{K}^{n \times \ell}$, $E \in \mathbb{K}^{q \times n}$ describing the perturbation structure,

- a) characterize the constants $\gamma > 0$ for which there exists a state feedback matrix such that

$$r_{(M,\beta)}(A - BF; D, E) \geq \gamma. \quad (5)$$

- b) Provide a method for the computation of feedback matrices F such that (5) is satisfied.

2 MOTIVATION AND HISTORY OF THE PROBLEM

Stability and stabilization are fundamental concepts in linear systems theory and in most design problems exponential stability is the minimal requirement that has to be met. From a practical point of view, however, the transient behavior of a system may be of equal importance and is often one of the criteria that decides on the quality of a controller in applications. As such, the notion of (M, β) -stability is related to such classical design criteria as “overshoot” of system responses. The question of how far transients move away from the origin is of interest in many situations; for instance, if certain regions of the state space are to be avoided in order to prevent saturation effects.

A similar problem occurs if linear design is performed as a local design for a nonlinear system. In this case, large transients may result in a small domain of attraction. For an introduction to the relation of the constant M with estimates of the domain of attraction, we refer to [4, Chapter 5]. The solution of Problem 4 and also of the other problems would provide a way to design local linear feedbacks with good local estimates for the domain of attraction without having to resort to the knowledge of Lyapunov functions. While the latter method is excellent if a Lyapunov function is known, it is also known that it may be quite hard to find them or if quadratic Lyapunov functions are used then the obtainable estimates may be far from optimal, see section 3.

Apart from these motivations from control the relation between domains of attraction and transient behavior of linearizations at fixed points is an active field in recent years motivated by problems in mathematical physics, in particular, fluid dynamics; see [1, 10] and references therein. Related problems occur in the study of iterative methods in numerical analysis; see e.g., [3].

We would like to point out that the problems discussed in this note give *pointwise* conditions in time for the bounds and are therefore different from criteria that can be formulated via integral constraints on the positive time axis. In the literature, such integral criteria are sometimes also called bounds on the transient behavior; see e.g., [9] where interesting results are obtained for this case.

Stability radii with respect to asymptotic stability of linear systems were introduced in [5] and there is a considerable body of literature investigating

this problem. The questions posed in this note are an extension of the available theory insofar as the transient behavior is neglected in most of the available results on stability radii.

3 AVAILABLE RESULTS

A number of results are available for problem 1. Estimates of the transient behavior involving either quadratic Lyapunov functions or resolvent inequalities are known but can be quite conservative or intractable. Moreover, for many of the available estimates, little is known on their conservatism.

The Hille-Yosida Theorem [8] provides an equivalent description of (M, β) -stability in terms of the norm of powers of the resolvent of A . Namely, A is (M, β) -stable if and only if for all $n \in \mathbb{N}$ and all $\alpha \in \mathbb{R}$ with $\alpha > \beta$ it holds that

$$\|(\alpha I - A)^{-n}\| \leq \frac{M}{(\alpha - \beta)^n}.$$

A characterization of M as the minimal eccentricity of norms that are Lyapunov functions of (1) is presented in [7]. While these conditions are hard to check, there is a classical, easily verifiable, sufficient condition using quadratic Lyapunov functions. Let $\beta \in (\gamma(A), 0)$, if $P > 0$ satisfies the Lyapunov inequality

$$A^*P + PA \leq 2\beta P < 0,$$

and has condition number $\kappa(P) := \|P\|\|P^{-1}\| \leq M^2$ then A is (M, β) -stable. The existence of $P > 0$ satisfying these conditions may be posed as an LMI-problem [2]. However, it can be shown that if $\beta < 0$ is given and the spectral bound of A is below β then this method is necessarily conservative, in the sense that the best bound on M obtainable in this way is strictly larger than the minimal bound. Furthermore, experiments show that the gap between these two bounds can be quite large. In this context, note that the problem cannot be solved by LMI techniques since the characterization of the optimal M for given β is not an algebraic problem.

There is a large number of further upper bounds available for $\|e^{At}\|$. These are discussed and compared in detail in [4, 11], see also the references therein. A number of these bounds is also valid in the infinite-dimensional case.

For problem 2, sufficient conditions are derived in [7] using quadratic Lyapunov functions and LMI techniques. The existence of a feedback F such that

$$P(A - BF) + (A - BF)^*P \leq 2\beta P \quad \text{and} \quad \kappa(P) = \|P\|\|P^{-1}\| \leq M^2, \quad (6)$$

or, equivalently, the solvability of the associated LMI problem, is characterized in geometric terms. This, however, only provides a sufficient condition under which Problem 2 can be solved. But the LMI problem (6) is far from being equivalent to Problem 2.

Concerning problem 3 differential Riccati equations were used to derive bounds for the (M, β) –stability radius in [6]. Suppose there exist positive definite Hermitian matrices P^0, Q, R of suitable dimensions such that the differential Riccati equation

$$\dot{P} - (A - \beta I)P - P(A - \beta I)^* - E^*QE - PDRD^*P = 0 \quad (7)$$

$$P(0) = P^0 \quad (8)$$

has a solution on \mathbb{R}_+ which satisfies

$$\bar{\sigma}(P(t))/\underline{\sigma}(P^0) \leq M^2, \quad t \geq 0.$$

Then the structured (M, β) –stability radius is at least

$$r_{(M, \beta)}(A; D, E) \geq \sqrt{\underline{\sigma}(Q)\underline{\sigma}(R)}, \quad (9)$$

where $\bar{\sigma}(X)$ and $\underline{\sigma}(X)$ denote the largest and smallest singular value of X . However, it is unknown how to choose the parameters P^0, Q, R in an optimal way and it is unknown whether equality can be obtained in (9) by an optimal choice of P^0, Q, R .

To the best of our knowledge, no results are available dealing with problem 4.

BIBLIOGRAPHY

- [1] J. S. Baggett and L. N. Trefethen, “Low-dimensional Models of Subcritical Transition to Turbulence” *Physics of Fluids* 9:1043–1053, 1997.
- [2] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in Systems and Control Theory*, volume 15 of *Studies in Applied Mathematics*, SIAM, Philadelphia, 1994.
- [3] T. Braconnier and F. Chaitin-Chatelin, “Roundoff induces a chaotic behavior for eigensolvers applied on highly nonnormal matrices,” M.-O. Bristeau et al. (eds.), *Computational Science for the 21st Century*. Symposium, Tours, France, May 5–7, 1997. Chichester: John Wiley & Sons. 43-52 (1997).
- [4] M. I. Gil’, *Stability of Finite and Infinite Dimensional Systems*, Kluwer Academic Publishers, Boston, 1998.
- [5] D. Hinrichsen and A. J. Pritchard, “Stability radius for structured perturbations and the algebraic Riccati equation,” *Systems & Control Letters*, 8:105–113, 1986.
- [6] D. Hinrichsen, E. Plischke, and A. J. Pritchard, “Liapunov and Riccati equations for practical stability,” In: *Proc. European Control Conf. ECC-2001*, Porto, Portugal, (CD-ROM), pp. 2883–2888, 2001.

- [7] D. Hinrichsen, E. Plischke, and F. Wirth; “State Feedback Stabilization with Guaranteed Transient Bounds,” In: *Proceedings of MTNS-2002*, Notre Dame, IN, USA, August, 2002.
- [8] A. Pazy, *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Springer-Verlag, New York, 1983.
- [9] A. Saberi, A. A. Stoorvogel, and P. Sannuti, *Control of Linear Systems with Regulation and Input Constraints*, Springer-Verlag, London, 2000.
- [10] L. N. Trefethen, “Pseudospectra of linear operators,” *SIAM Review*, 39(3):383–406, 1997.
- [11] K. Veselić, “Bounds for exponentially stable semigroups,” *Lin. Alg. Appl.*, 358:195–217, 2003.

Problem 6.4

Lie algebras and stability of switched nonlinear systems

Daniel Liberzon

Coordinated Science Laboratory

University of Illinois at Urbana-Champaign

Urbana, IL 61801

USA

liberzon@uiuc.edu

1 PRELIMINARY DESCRIPTION OF THE PROBLEM

Suppose that we are given a family f_p , $p \in P$ of continuously differentiable functions from R^n to R^n , parameterized by some index set P . This gives rise to the *switched system*

$$\dot{x} = f_\sigma(x), \quad x \in R^n \tag{1}$$

where $\sigma : [0, \infty) \rightarrow P$ is a piecewise constant function of time, called a *switching signal*. Impulse effects (state jumps), infinitely fast switching (chattering), and Zeno behavior are not considered here. We are interested in the following problem: find conditions on the functions f_p , $p \in P$ which guarantee that the switched system (1) is asymptotically stable, uniformly over the set of all possible switching signals. If this property holds, we will refer to the switched system simply as being *stable*. It is clearly necessary for each of the subsystems $\dot{x} = f_p(x)$, $p \in P$ to be asymptotically stable—which we henceforth assume—but simple examples show that this condition alone is not sufficient.

The problem posed above naturally arises in the stability analysis of switched systems in which the switching mechanism is either unknown or too complicated to be explicitly taken into account. This problem has attracted considerable attention and has been studied from various angles (see [7] for references). Here we explore a particular research direction, namely, the role of commutation relations among the subsystems being switched. In the following sections, we provide an overview of available results on this topic and delineate the open problem more precisely.

2 AVAILABLE RESULTS: LINEAR SYSTEMS

In this section, we concentrate on the case when the subsystems are linear. This results in the *switched linear system*

$$\dot{x} = A_\sigma x, \quad x \in R^n. \quad (2)$$

We assume throughout that $\{A_p : p \in P\}$ is a compact set of stable matrices.

To understand how commutation relations among the linear subsystems being switched play a role in the stability question for the switched linear system (2), consider first the case when P is a finite set and the matrices commute pairwise: $A_p A_q = A_q A_p$ for all $p, q \in P$. Then it not hard to show by a direct analysis of the transition matrix that the system (2) is stable. Alternatively, in this case one can construct a quadratic common Lyapunov function for the family of linear subsystems $\dot{x} = A_p x$, $p \in P$ as shown in [10], which is well-known to lead to the same conclusion.

A useful object that reveals the nature of commutation relations is the *Lie algebra* g generated by the matrices A_p , $p \in P$. This is the smallest linear subspace of $R^{n \times n}$ that contains these matrices and is closed under the *Lie bracket* operation $[A, B] := AB - BA$ (see, e.g., [11]). Beyond the commuting case, the natural classes of Lie algebras to study in the present context are *nilpotent* and *solvable* ones. A Lie algebra is nilpotent if all Lie brackets of sufficiently high order vanish. Solvable Lie algebras form a larger class of Lie algebras, in which all Lie brackets of sufficiently high order having a certain structure vanish.

If P is a finite set and g is a nilpotent Lie algebra, then the switched linear system (2) is stable; this was proved in [4] for the discrete-time case. The system (2) is still stable if g is solvable and P is not necessarily finite (as long as the compactness assumption made at the beginning of this section holds). The proof of this more general result, given in [6], relies on the facts that matrices in a solvable Lie algebra can be simultaneously put in the triangular form (Lie's Theorem) and that a family of linear systems with stable triangular matrices has a quadratic common Lyapunov function.

It was subsequently shown in [1] that the switched linear system (2) is stable if the Lie algebra g can be decomposed into a sum of a solvable ideal and a subalgebra with a compact Lie group. Moreover, if g fails to satisfy this condition, then it can be generated by families of stable matrices giving rise to stable as well as to unstable switched linear systems, i.e., the Lie algebra alone does not provide enough information to determine whether or not the switched linear system is stable (this is true under the additional technical requirement that $I \in g$).

By virtue of the above results, one has a complete characterization of all matrix Lie algebras g with the property that every set of stable generators for g gives rise to a stable switched linear system. The interesting and rather surprising discovery is that this property depends only on the structure of g as a Lie algebra, and not on the choice of a particular matrix represen-

tation of g . Namely, Lie algebras with this property are precisely the Lie algebras that admit a decomposition of the kind described earlier. Thus, in the linear case, the extent to which commutation relations can be used to distinguish between stable and unstable switched systems is well understood. Lie-algebraic sufficient conditions for stability are mathematically appealing and easily checkable in terms of the original data (it has to be noted, however, that they are not robust with respect to small perturbations in the data and therefore highly conservative).

3 OPEN PROBLEM: NONLINEAR SYSTEMS

We shall now turn to the general nonlinear situation described by equation (1). Linearizing the subsystems and applying the results described in the previous section together with Lyapunov's indirect method, it is not hard to obtain Lie-algebraic conditions for local stability of the system (1). This was done in [6, 1]. However, the problem we are posing here is to investigate how the structure of the Lie algebra generated by the original nonlinear vector fields f_p , $p \in P$ is related to stability properties of the switched system (1). Taking higher-order terms into account, one may hope to obtain more widely applicable Lie-algebraic stability criteria for switched nonlinear systems.

The first step in this direction is the result proved in [8] that if the set P is finite and the vector fields f_p , $p \in P$ commute pairwise, in the sense that

$$[f_p, f_q](x) := \frac{\partial f_q(x)}{\partial x} f_p(x) - \frac{\partial f_p(x)}{\partial x} f_q(x) = 0 \quad \forall x \in R^n, \quad \forall p, q \in P$$

then the switched system (1) is (globally) stable. In fact, commutativity of the flows is all that is needed, and the continuous differentiability assumption on the vector fields can be relaxed. If the subsystems are exponentially stable, a construction analogous to that of [10] can be applied in this case to obtain a local common Lyapunov function; see [12].

A logical next step is to study switched nonlinear systems with nilpotent or solvable Lie algebras. One approach would be via simultaneous triangularization, as done in the linear case. Nonlinear versions of Lie's Theorem, which provide Lie-algebraic conditions under which a family of nonlinear systems can be simultaneously triangularized, are developed in [3, 5, 9]. However, as demonstrated in [2], the triangular structure alone is not sufficient for stability in the nonlinear context. Additional conditions that can be imposed to guarantee stability are identified in [2], but they are coordinate-dependent and so cannot be formulated in terms of the Lie algebra. Moreover, the results on simultaneous triangularization described in the papers mentioned above require that the Lie algebra have full rank, which is not true in the case of a common equilibrium. Thus an altogether new approach seems to be required.

In summary, the main open question is this:

Q: *Which structural properties (if any) of the Lie algebra generated by a noncommuting family of asymptotically stable nonlinear vector fields guarantee stability of every corresponding switched system? For example, when does nilpotency or solvability of the Lie algebra imply stability?*

To begin answering this question, one may want to first address some special classes of nonlinear systems, such as homogeneous systems or systems with feedback structure. One may also want to restrict attention to finite-dimensional Lie algebras.

A more general goal of this paper is to point out the fact that Lie algebras seem to be directly connected to stability of switched systems and, in view of the well-established theory of the former and high theoretical interest as well as practical importance of the latter, there is a need to develop a better understanding of this connection. It may also be useful to pursue possible relationships with Lie-algebraic results in the controllability literature (see [1] for a brief preliminary discussion on this matter).

BIBLIOGRAPHY

- [1] A. A. Agrachev and D. Liberzon, "Lie-algebraic stability criteria for switched systems," *SIAM J. Control Optim.*, 40:253–269, 2001.
- [2] D. Angeli and D. Liberzon, "A note on uniform global asymptotic stability of nonlinear switched systems in triangular form," In: *Proc. 14th Int. Symp. on Mathematical Theory of Networks and Systems (MTNS)*, 2000.
- [3] P. E. Crouch, "Dynamical realizations of finite Volterra series," *SIAM J. Control Optim.*, 19:177–202, 1981.
- [4] L. Gurvits, "Stability of discrete linear inclusion," *Linear Algebra Appl.*, 231:47–85, 1995.
- [5] M. Kowski, "Nilpotent Lie algebras of vectorfields," *J. Reine Angew. Math.*, 388:1–17, 1988.
- [6] D. Liberzon, J. P. Hespanha, and A. S. Morse, "Stability of switched systems: a Lie-algebraic condition," *Systems Control Lett.*, 37:117–122, 1999.
- [7] D. Liberzon, *Switching in Systems and Control*. Birkhäuser, Boston, 2003.

- [8] J. L. Mancilla-Aguilar, "A condition for the stability of switched nonlinear systems," *IEEE Trans. Automat. Control*, 45:2077–2079, 2000.
- [9] A. Marigo, "Constructive necessary and sufficient conditions for strict triangularizability of driftless nonholonomic systems'," In: *Proc. 38th IEEE Conf. on Decision and Control*, pages 2138–2143, 1999.
- [10] K. S. Narendra and J. Balakrishnan, "A common Lyapunov function for stable LTI systems with commuting A -matrices," *IEEE Trans. Automat. Control*, 39:2469–2471, 1994.
- [11] H. Samelson, *Notes on Lie Algebras*. Van Nostrand Reinhold, New York, 1969.
- [12] H. Shim, D. J. Noh, and J. H. Seo, "Common Lyapunov function for exponentially stable nonlinear systems," In: *Proc. 4th SIAM Conference on Control and its Applications*, 1998.

Problem 6.5

Robust stability test for interval fractional order linear systems

Ivo Petráš

Department of Informatics and Process Control
BERG Faculty, Technical University of Košice
B. Němcovej 3, 042 00 Košice
Slovak Republic
ivo.petras@tuke.sk

YangQuan Chen

Center for Self-Organizing and Intelligent System
Dept. of Electrical and Computer Engineering
Utah State University
Logan, UT-84322-4160
USA
yqchen@ieee.org

Blas M. Vinagre

Dept. of Electronic and Electromechanical Engineering
Industrial Engineering School, University of Extremadura
Avda. De Elvas s/n, 06071-Badajoz
Spain
bvinagre@unex.es

1 DESCRIPTION OF THE PROBLEM

Recently, a robust stability test procedure is proposed for linear time-invariant fractional order systems (LTI FOS) of commensurate orders with parametric interval uncertainties [6]. The proposed robust stability test method is based on the combination of the argument principle method [2] for LTI FOS and the celebrated Kharitonov's edge theorem. In general, an LTI FOS can be described by the differential equation or the corresponding transfer

function of noncommensurate real orders [7] of the following form:

$$G(s) = \frac{b_m s^{\beta_m} + \dots + b_1 s^{\beta_1} + b_0 s^{\beta_0}}{a_n s^{\alpha_n} + \dots + a_1 s^{\alpha_1} + a_0 s^{\alpha_0}} = \frac{Q(s^{\beta_k})}{P(s^{\alpha_k})}, \quad (1)$$

where α_k, β_k ($k = 0, 1, 2, \dots$) are real numbers and without loss of generality they can be arranged as $\alpha_n > \dots > \alpha_1 > \alpha_0$, $\beta_m > \dots > \beta_1 > \beta_0$. The coefficients a_k, b_k ($k = 0, 1, 2, \dots$) are uncertain constants within a known interval.

It is well-known that an integer order LTI system is stable if all the roots of the characteristic polynomial $P(s)$ are negative or have negative real parts if they are complex conjugate (e.g., [1]). This means that they are located on the left of the imaginary axis of the complex s -plane. When dealing with noncommensurate order systems (or, in general, with fractional order systems) it is important to bear in mind that $P(s^\alpha)$, $\alpha \in \mathbb{R}$ is a multivalued function of s , the domain of which can be viewed as a Riemann surface (see e.g., [4]).

A question of robust stability test procedure and proof of its validity for general type of the LTI FOS described by (1) is still open.

2 MOTIVATION AND HISTORY OF THE PROBLEM

For the LTI FOS with no uncertainty, the existing stability test (or check) methods for dynamic systems with integer-orders such as Routh table technique, cannot be directly applied. This is due to the fact that the characteristic equation of the LTI FOS is, in general, not a polynomial but a pseudo-polynomial function of the fractional powers of s .

Of course, being the characteristic equation a function of a complex variable, stability test based on the argument principle can be applied. On the other hand, it has been shown, by several authors and by using several methods, that for the case of LTI FOS of commensurate order, a geometrical method based on the argument of the roots of the characteristic equation (a polynomial in this particular case) can be used for the stability check in the BIBO (bounded-input bounded-output) sense (see, e.g., [3]).

In the particular case of *commensurate order* systems, it holds that $\alpha_k = \alpha k, \beta_k = \alpha k, (0 < \alpha < 1), \forall k \in \mathbb{Z}$, and the transfer function has the following form

$$G(s) = K_0 \frac{\sum_{k=0}^M b_k (s^\alpha)^k}{\sum_{k=0}^N a_k (s^\alpha)^k} = K_0 \frac{Q(s^\alpha)}{P(s^\alpha)} \quad (2)$$

With $N > M$ the function $G(s)$ becomes a proper rational function in the complex variable s^α and can be expanded in partial fractions of the form

$$G(s) = K_0 \left[\sum_{i=1}^N \frac{A_i}{s^\alpha + \lambda_i} \right], \quad (3)$$

where λ_i ($i = 1, 2, \dots, N$) are the roots of the polynomial $P(s^\alpha)$ or the system poles that are assumed to be simple. Stability condition can then be stated that [2, 3]:

A commensurate order system described by a rational transfer function (2) is stable if $|\arg(\lambda_i)| > \alpha \frac{\pi}{2}$, with λ_i the i -th root of $P(s^\alpha)$.

For the LTI FOS with commensurate order where system poles are in general complex conjugate, the stability condition can be expressed as follows [2, 3]:

A commensurate order system described by a rational transfer function $G(\sigma) = \frac{Q(\sigma)}{P(\sigma)}$, where $\sigma = s^\alpha$, $\alpha \in \mathbb{R}^+$, ($0 < \alpha < 1$), is stable if $|\arg(\sigma_i)| > \alpha \frac{\pi}{2}$, with σ_i the i -th root of $P(\sigma)$.

The *robust stability* test procedure for the LTI FOS of commensurate orders with parametric interval uncertainties can be divided into the following steps:

- **step 1:** Rewrite the LTI FOS $G(s)$ of the commensurate order α , to the equivalence system $H(\sigma)$, where transformation is: $s^\alpha \rightarrow \sigma$, $\alpha \in \mathbb{R}^+$;
- **step 2:** Write the interval polynomial $P(\sigma, q)$ of the equivalence system $H(\sigma)$, where interval polynomial is defined as

$$P(\sigma, q) = \sum_{i=0}^n [q^-, q^+] \sigma^i;$$

- **step 3:** For interval polynomial $P(\sigma, q)$, construct four Kharitonov's polynomials:

$$p^{--}(\sigma), p^{-+}(\sigma), p^{+-}(\sigma), p^{++}(\sigma);$$

- **step 4:** Test the four Kharitonov's polynomials whether they satisfy the stability condition: $|\arg(\sigma_i)| > \alpha \frac{\pi}{2}$, $\forall \sigma \in \mathbb{C}$, with σ_i the i -th root of $P(\sigma)$;

Note that for low-degree polynomials, less Kharitonov's polynomials are to be tested:

- Degree 5: $p^{--}(\sigma), p^{-+}(\sigma), p^{+-}(\sigma)$;
- Degree 4: $p^{+-}(\sigma), p^{++}(\sigma)$;
- Degree 3: $p^{+-}(\sigma)$.

We demonstrated this technique for the robust stability check for the LTI FOS with parametric interval uncertainties through some worked-out illustrative examples in [6]. In [6] the time-domain analytical expressions are available and therefore the time-domain and the frequency-domain stability test results (see also [5]) can be cross-validated.

3 AVAILABLE RESULTS

For general LTI FOS, if the coefficients are uncertain but are known to lie within known intervals, how to generalize the robust stability test result by Kharitonov's well-known edge theorem? This is definitely a new research topic.

The main future research objectives could be:

- A proof of validity of the *robust stability* test procedure for the LTI FOS of commensurate orders with parametric interval uncertainties.
- An algebraic method and an exact proof for the stability investigation for the LTI FOS of noncommensurate orders with known parameters.
- A *robust stability* test procedure of LTI FOS of noncommensurate orders with parametric interval uncertainties.

BIBLIOGRAPHY

- [1] R. C. Dorf and R. H. Bishop, *Modern Control Systems*, Addison-Wesley Publishing Company, 1990.
- [2] D. Matignon, "Stability result on fractional differential equations with applications to control processing," In: *IMACS - SMC Proceeding*, July, Lille, France, pp. 963-968, 1996.
- [3] D. Matignon, "Stability properties for generalized fractional differential systems," In: *Proceeding of Fractional Differential Systems: Models, Methods and Applications*, vol. 5, pp. 145-158, 1998.
- [4] D. A. Pierre, "Minimum Mean-Square-Error Design of Distributed Parameter Control Systems," *ISA Transactions*, vol. 5, pp. 263-271, 1966.
- [5] I. Petráš and Ľ. Dorčák, "The Frequency Method for Stability Investigation of Fractional Control Systems," *J. of SACTA*, vol. 2, no. 1-2, pp. 75-85, 1999.
- [6] I. Petráš, Y. Q. Chen, B. M. Vinagre, "A robust stability test procedure for a class of uncertain LTI fractional order systems," In: *Proc. of ICC2002*, May 27-30, Beskydy, pp. 247-252, 2002.
- [7] I. Podlubny, *Fractional Differential Equations*, Academic Press, San Diego, 1999.

Problem 6.6

Delay independent and delay dependent Aizerman problem

Vladimir Răşvan

Department of Automatic Control

University of Craiova

13 A.I.Cuza Street

1100 Craiova

Romania

`vrasvan@automation.ucv.ro`

1 INTRODUCTION

The half-century old problem of Aizerman consists in a comparison of the absolute stability sector with the Hurwitz sector of stability for the linearized system. While the first has been shown to be, generally speaking, smaller than the second one, this comparison still serves as a test for the sharpness of sufficient stability criteria as Liapunov function or Popov inequality. On the other hand, there are now very popular for linear time delay systems two types of sufficient stability criteria: delay-independent and delay-dependent. The present paper suggests a comparison of these criteria with the corresponding ones for nonlinear systems with sector restricted nonlinearities. In this way, a problem of Aizerman type is suggested for systems with delay. Some examples are analyzed.

2 A SIMPLE EXAMPLE. STATEMENT OF THE PROBLEM.

Consider the simple time delay equation

$$\dot{x} + a_0x(t) + a_1x(t - \tau) = 0, \tau > 0 \quad (1)$$

with a_0 , a_1 , x scalars. It is a well-known fact [7, 9, 10] that exponential stability of (1) is ensured provided the following inequalities hold:

$$1 + a_0\tau > 0, \quad -a_0\tau < a_1\tau < \psi(a_0\tau) \quad (2)$$

where $\psi(\xi)$ is obtained by eliminating the parameter λ between the two equalities below

$$\xi = -\frac{\lambda}{\tan \lambda}, \quad \psi = \frac{\lambda}{\sin \lambda} \quad (3)$$

Since these conditions contain the time delay τ such property is called *delay-dependent stability*. If one is interested in exponential stability conditions that hold for any delay $\tau > 0$, this property, called *delay-independent stability* is ensured provided the simple inequalities

$$a_0 > 0, \quad |a_1| < a_0 \quad (4)$$

are fulfilled. It can be shown [10] that $\psi(\xi) > \xi$ for $\xi > 0$, hence the fulfilment of (4) implies the fulfilment of (2).

Let us follow the way of Barbashin [6] to introduce a *stability problem in the nonlinear case*: given system (1) for $a_0 > 0$, if we replace $a_0 x$ by $\varphi(x)$ where $\varphi(x)x > 0$, the equilibrium at the origin of the nonlinear time delay system should be globally asymptotically stable provided

$$\frac{\varphi(\sigma)}{\sigma} > |a_1| \quad (5)$$

for the delay-independent stability, or provided

$$\frac{\varphi(\sigma)}{\sigma} > \max \left\{ -a_1, \frac{1}{\tau} \psi^{-1}(a_1 \tau) \right\} \quad (6)$$

in the delay-dependent case.

We may view the above problem in a more general setting and state it as follows:

Problem: *Given the delay-(in)dependent exponential stability conditions for some time delay linearized system, are they valid in the case when the nonlinear system with a sector restricted nonlinearity, i.e., satisfying*

$$\underline{\varphi}\sigma^2 < \varphi(\sigma)\sigma < \overline{\varphi}\sigma^2 \quad (7)$$

is considered instead of the linear one, or have they to be strengthened?

It is clear that we have gathered here both the delay-independent and delay-dependent cases, thus defining a stability problem in two different cases. This problem is called *Aizerman problem*, stated here as *delay-dependent* (Aizerman problem) and *delay-independent* (Aizerman problem).

Since this problem in the ODE (ordinary differential equations) setting is not only well-known but also quite well-studied, a short state of the art could be useful.

3 THE PROBLEM OF THE ABSOLUTE STABILITY. THE PROBLEMS OF AIZERMAN AND KALMAN

Exactly 60 years ago a paper of B. V. Bulgakov [8] considered, apparently for the first time, a problem of global asymptotic stability for the zero equi-

librium of a feedback control system composed of a linear dynamic part and a nonlinear static element

$$\dot{x} = Ax - b\varphi(c^*x) \quad (8)$$

where x, b, c are n -dimensional vectors, A is a $n \times n$ matrix and $\varphi: \mathbb{R} \rightarrow \mathbb{R}$ is a continuous function. The only additional assumption about φ was its location in some sector as defined by (7), where the inequalities may be non-strict. In this very first paper, only conditions for the absence of self-sustained oscillations were obtained but in another, more famous paper of Lurie and Postnikov [17] global asymptotic stability conditions were obtained for a system (8) of 3d order with $\varphi(\sigma)$ satisfying $\varphi(\sigma)\sigma > 0$, i.e., satisfying (7) with $\underline{\varphi} = 0, \bar{\varphi} = +\infty$. The conditions obtained using a suitably chosen Liapunov function of the form “quadratic form of the state variables plus an integral of the nonlinearity” were in fact valid for the whole class of nonlinear functions defined by $\varphi(\sigma)\sigma > 0$. Later this was called *absolute stability* but it is obviously a *robust stability problem* since it deals with the uncertainty on the nonlinear function defined by (7). We shall not insist more on this problem and we shall concentrate on another one, connected with it, stated by M. A. Aizerman [1, 2]. This last problem is on (8) and its linearized version

$$\dot{x} = Ax - bhc^*x \quad (9)$$

i.e., system (8) with $\varphi(\sigma) = h\sigma$. It is known that the necessary and sufficient conditions of asymptotic stability for (9) will require h to be restricted to some interval (\underline{h}, \bar{h}) called the *Hurwitz sector*. On the other hand, for system (8) the absolute stability problem is stated: find conditions of global asymptotic stability of the zero equilibrium for all functions satisfying (7). All functions include the linear ones hence the class of systems defined by (8) is larger than the class of systems defined by (9). Consequently the sector $(\underline{\varphi}, \bar{\varphi})$ from (7) may be at most as large as the Hurwitz sector (\underline{h}, \bar{h}) . *The Aizerman problem asks simply: do these sectors always coincide? The Aizerman conjecture assumed: yes.*

The first counter-example to this conjecture has been produced by Krasovskii [16] in the form of a 2nd order system of special form. The most celebrated counterexample is a 3rd order system and belongs to Pliss [21]. Today we know that the conjecture of Aizerman does not hold in general. Nevertheless the problem itself stimulated interesting research that could be summarized as seeking necessary and sufficient conditions for absolute stability.

A straightforward application of these studies is checking of the sharpness for “traditional” absolute stability criteria: the Liapunov function and the frequency domain inequality of Popov. In fact this is nothing more but comparison of the absolute stability sector with the Hurwitz sector. One can mention here the results of Voronov [26] and his co-workers on what they called “stability in the Hurwitz sector.”

Other noteworthy results belong to Pyatnitskii who found necessary and sufficient conditions of absolute stability connected to a special variational

problem and to N. E. Barabanov (e.g., [4]). Among the results of Barabanov we would like to mention those concerned with the so-called Kalman problem and conjecture topics that deserve some particular attention. In his paper [15] R. E. Kalman replaced the class of nonlinear functions defined by (7) by the class of differentiable functions with slope restrictions

$$\underline{\gamma} < \varphi'(\sigma) < \bar{\gamma} \quad (10)$$

The *Kalman problem asks*: do coincide the intervals $(\underline{\gamma}, \bar{\gamma})$ and (\underline{h}, \bar{h}) the last one being previously defined by the inequalities of Hurwitz? The answer to this question is also negative but its story is not quite straightforward. A good reference is the paper of Barabanov [3] and we would like to follow some of the presentation there: the only counterexample known up to that paper had been published by Fitts [11] and the authors of a well-known and cited monograph in the field (Narendra and Taylor, [18]) were citing it as a basic argument for the negative answer to Kalman conjecture. In fact there was no proof in the paper of Fitts but just a simulation: a specific linear subsystem had been adopted, a specific nonlinearity also and self-sustained periodic oscillations were computed for various values of a system's parameter. In his important paper Barabanov [3] was able to prove rigorously the following:

- the answer to the problem of Kalman is positive for all 3d order systems; it follows that the system of Pliss counter-example is absolutely stable within the Hurwitz sector provided the class of the nonlinear functions is defined by (10) instead of (7);
- the counterexample given by Fitts is not correct at least for some subset of its parameters as is follows by simple application of the Brockett Willems frequency domain inequality for absolute stability of systems with slope restricted nonlinearity.

Moreover, the paper of Barabanov provides an algorithm of finding systems with a non-trivial periodic solution; in this way, a procedure is given for constructing counterexamples to the two conjectures discussed above. Obviously, the technique of Barabanov seems an echo of the pioneering paper of Bulgakov [8], but we shall insist no more on this subject.

4 STABILITY AND ABSOLUTE STABILITY OF THE SYSTEMS WITH TIME DELAY

A. We shall consider for simplicity only the case of the systems described by functional differential equations of delayed type (according to the well-known classification of these equations; see, for instance, Bellman and Cooke [7]) and we shall restrict ourselves to the single delay case. In the linear case, the system is described by

$$\dot{x} = A_0x(t) + A_1x(t - \tau), \tau > 0 \quad (11)$$

Exponential stability of this system is ensured by the location in the LHP (left-hand plane) of the roots of the characteristic equation

$$\det(\lambda I - A_0 - A_1 e^{-\lambda\tau}) = 0 \quad (12)$$

where the LHS (left-hand side) is a quasipolynomial. We have here the *Routh-Hurwitz problem for quasipolynomials*. This problem has been studied since the first applications of (11); the basic results are to be found in the paper of Pontryagin [22] and in the memoir of Chebotarev and Meiman [9]. A valuable reference is the book of Stepan [25]. From this topic, we shall recall the following. Starting from their algebraic intuition Chebotarev and Meiman pointed out that, according to Sturm theory, the Routh-Hurwitz conditions for quasi-polynomials have to be expressed as a finite number of inequalities that might be transcendental. The detailed analysis performed in their memoir for the 1st and 2nd degree quasipolynomials showed two types of inequalities: one of them contained only algebraic inequalities, while the other contained also transcendental inequalities; the first ones correspond to stability for arbitrary values of the delay τ , while the second ones put some limitations on the values of $\tau > 0$ for which exponential stability of (11) holds. The system described by (1) and conditions (2), (3) and (4) are good illustrations of this. The aspect is quite transparent in the examples analysis performed throughout author's book [23] as well as throughout the book of Stepan [25]. We may see here the difference operated between what will be called later *delay-independent* and *delay-dependent stability*.

This difference will become important after the publication of the paper of Hale et al. [13], which will be assimilated by the control community after its incorporation in the 3d edition of Hale's monograph, authorized by Hale and Verduyn Lunel [14]. There are by now dozens of references concerning *delay-dependent and delay-independent Routh-Hurwitz problem* for (11); we send the reader to the books of S. I. Niculescu [19, 20] with their rich reference lists.

A special case of (2) that is in fact the underlying topic of most references cited in [19, 20] is stability for small delays.

As shown in [10] the stability inequalities are given by

$$a_1 + a_0 > 0, \quad 0 \leq \tau < \frac{\arccos\left(-\frac{a_0}{a_1}\right)}{\sqrt{a_1^2 - a_0^2}} \quad (13)$$

provided $a_1 > |a_0|$ (otherwise (4) holds and stability is delay-independent). In fact most recent research defines delay-dependent stability as above, i.e., as preservation of stability for small delays (a better name would be "delay robust stability" since, according to a paper of Jaroslav Kurzweil, "small delays don't matter").

B. Since linear blocks with delay are usual in control, introduction of systems with sector restricted nonlinearities (7) is only natural. The most suitable references on this problem are the monographs of A. Halanay [12] and of the

author [23]. If we restrict ourselves again to the case of delayed type with a single delay, then a model problem could be the system

$$\dot{x} = A_0 x(t) + A_1 x(t - \tau) - b\varphi(c_0^* x(t) + c_1^* x(t - \tau)) \quad (14)$$

where x, b, c_0, c_1 are n -vectors and A_0, A_1 are $n \times n$ matrices; the nonlinear function $\varphi(\sigma)$ satisfies the sector condition (7).

Following author's book [23] we shall consider a scalar version of (14):

$$\dot{x} + a_0 x(t) + \varphi(x(t) + c_1 x(t - \tau)) = 0 \quad (15)$$

where $\varphi(\sigma)\sigma > 0$. Assume that $a_0 > 0$ and apply the frequency domain inequality of Popov for $\bar{\varphi} = +\infty$:

$$\operatorname{Re}(1 + j\omega\beta)H(j\omega) > 0, \quad \forall \omega \geq 0 \quad (16)$$

Since

$$H(s) = \frac{1 + c_1 e^{-\tau s}}{s + a_0}$$

the frequency domain inequality reads

$$\frac{(a_0^2 + \omega^2\beta)(1 + c_1 \cos \omega\tau) + \omega(a_0\beta - 1) \sin \omega\tau}{a_0^2 + \omega^2} > 0$$

By choosing the Popov parameter $\beta = a_0^{-1}$ the above inequality becomes

$$1 + c_1 \cos \omega\tau > 0, \quad \forall \omega \geq 0, \quad (17)$$

which cannot hold for $\forall \omega$ but only with $|c_1| < 1$. The frequency domain inequality of Popov prescribes in this case a *delay-independent absolute stability*.

5 BACK TO THE EXAMPLE

We have stated a delay-independent and a delay-dependent Aizerman problem for systems with time delay in a rather general setting that could include rather general systems of differential equations with deviated argument while we chose the starting system as a very simple one, of the delayed type. In the following, we shall illustrate the solving of a specific problem for the initial example.

Consider, for instance, the delay-independent Aizerman problem defined above, for system (1) replaced by

$$\dot{x} + a_1 x(t - \tau) + \varphi(x(t)) = 0 \quad (18)$$

where $\varphi(\sigma)\sigma > 0$. Taking into account that (4) suggests $\varphi(\sigma) > |a_1|\sigma$ we introduce a new nonlinear function

$$f(\sigma) = \varphi(\sigma) - |a_1|\sigma$$

and obtain the transformed system (*via* a sector rotation):

$$\dot{x} + |a_1|x(t) + a_1x(t - \tau) + f(x(t)) = 0 \quad (19)$$

For this system, we apply the frequency domain inequality of Popov for $\bar{\varphi} = +\infty$, i.e., inequality (16); here

$$H(s) = \frac{1}{s + |a_1| + a_1e^{-s\tau}} \quad (20)$$

and the frequency domain inequality reduces to

$$\beta\omega^2 - (\beta a_1 \sin \omega\tau)\omega + |a_1| + a_1 \cos \omega\tau \geq 0 \quad (21)$$

which is fulfilled provided the free Popov parameter β is chosen from

$$0 < \beta |a_1| < 2 \quad (22)$$

(more details concerning manipulation of the frequency domain inequality for time delay systems may be found in author's book [23]).

It follows that (19) is absolutely stable for the nonlinearities satisfying $f(\sigma)\sigma > 0$ i.e. $\varphi(\sigma)\sigma > |a_1|\sigma^2$: the just stated delay-independent Aizerman problem for (1) and (18) has been answered positively.

6 CONCLUDING REMARKS

Since the class of systems with time delays is considerably larger than the class of systems described by ordinary differential equations, we expect various settings of Aizerman (or Kalman) problems. The case of the equations of neutral type that express propagation phenomena was not yet analyzed from this point of view even if the absolute stability has been considered for such systems (see author's book [23]). Such a variety of systems and problems should be stimulating for the development of the tools of analysis.

It is a known fact that the frequency domain inequalities are better suited for delay-independent results, as well as the mostly used Liapunov-Krasovskii functionals leading to finite dimensional LMIs (see e.g., the cited books of Niculescu [19, 20]); the Liapunov-Krasovskii approach has nevertheless some "opening" to delay-dependent results and it is worth trying to apply it in solving the delay-dependent Aizerman problem. The algebraic approach suggested by the memoir of Chebotarev and Meiman [9] could be also applied as well as the (non)-existence of self-sustained oscillations that sends back to Bulgakov and Pliss.

As in the case without delay the statement and solving of the Aizerman problems could be rewarding from at least two points of view: extension of the class of the systems having an "almost linear behavior" [5, 24] and refinement of analysis tools by testing the "sharpness" of the sufficient conditions.

BIBLIOGRAPHY

- [1] M. A. Aizerman, "On convergence of the control process under large deviations of the initial conditions," (in Russian) *Avtom. i telemekh. vol. VII*, no. 2-3, pp. 148-167, 1946.
- [2] M. A. Aizerman, "On a problem concerning stability "in the large" of dynamical systems," (in Russian) *Usp.Mat.Nauk t.4*, no. 4, pp. 187-188, 1949.
- [3] N. E. Barabanov, "About the problem of Kalman," (in Russian) *Sib.Mat.J. vol.XXIX*, no. 3, pp. 3-11, 1988.
- [4] N. E. Barabanov, "On the problem of Aizerman for non-stationary systems of 3d order," (in Russian) *Differ.uravn. vol.29*, no. 10, pp. 1659-1668, 1992.
- [5] I. Barbălat and A. Halanay, "Conditions de comportement presque linéaire dans la thorie des oscillations," *Rev.Roum.Sci. Techn.-Electrotechn. et Energ. vol. 29*, no. 2, pp. 321-341, (1974).
- [6] E. A. Barbashin, *Introduction to stability theory*, (in Russian) Nauka Publ.House, Moscow, 1967.
- [7] R. E. Bellman and K. L. Cooke, *Differential Difference Equations*, Acad.Press, N.Y., 1963.
- [8] B. V. Bulgakov, "Self-sustained oscillations of control systems," (in Russian) *DAN SSSR vol. 37*, no. 9, pp. 283-287, 1942.
- [9] N. G. Chebotarev and N. N. Meiman, "The Routh-Hurwitz problem for polynomials and entire functions," (in Russian) *Trudy Mat.Inst. V. A. Steklov, vol. XXVI*, 1949.
- [10] L. E. El'sgol'ts and S. B. Norkin, *Introduction to the theory and applications of differential equations with deviating arguments*(in Russian), Nauka Publ.House, Moscow, 1971; English version by Acad. Press, 1973.
- [11] R. E. Fitts, "Two counter-examples to Aizerman's conjecture," *IEEE Trans. vol. AC-11*, no. 3, pp. 553-556, July 1966.
- [12] A. Halanay, *Differential Equations. Stability. Oscillations. Time Lags*, Acad.Press, N.Y., 1966.
- [13] J. K. Hale, E. F. Infante and F. S. P. Tsen, "Stability in linear delay equations," *J. Math. Anal. Appl. vol. 105*, pp. 533-555, 1985.
- [14] J. K. Hale and S. Verduyn Lunel, *Introduction to Functional Differential Equations*, Springer Verlag, 1993.

- [15] R. E. Kalman, "Physical and mathematical mechanisms of instability in nonlinear automatic control systems," *Trans.ASME*, vol. 79, no. 3, pp. 553-566, April 1957.
- [16] N. N. Krasovskii, "Theorems concerning stability of motions determined by a system of two equations," (in Russian) *Prikl. Mat. Mekh. (PMM)*, vol. XVI, no. 5, pp. 547-554, 1952.
- [17] A. I. Lurie and V. N. Postnikov, "On the theory of stability for control systems," (in Russian) *Prikl. Mat. Mekh. (PMM)*, vol. VIII, no. 3, pp. 246-248, 1944.
- [18] K. S. Narendra and J. H. Taylor, *Frequency domain stability criteria*, Acad.Press, N.Y., 1973.
- [19] S. I. Niculescu, *Systèmes a retard*, Diderot, Paris, 1997.
- [20] S. I. Niculescu, "Delay effects on stability: A robust control approach," *Lecture Notes in Control and Information Sciences*, no. 269, Springer Verlag, 2001.
- [21] V. A. Pliss, *Some Problems of the Theory of Stability of Motion in the Large* (in Russian), Leningrad State Univ. Publ. House, Leningrad, (1958).
- [22] L. S. Pontryagin, "On zeros of some elementary transcendental functions," (in Russian) *Izv. AN SSSR Ser. Matem.* vol. 6, no. 3, pp. 115-134, 1942, with an Appendix published in *DAN SSSR*, vol. 91, no. 6, pp. 1279-1280, 1953.
- [23] Vl. Răsvan, *Absolute stability of automatic control systems with time delay* (in Romanian), Editura Academiei, Bucharest, 1975; Russian version by Nauka Publ. House, Moscow, 1983.
- [24] Vl. Răsvan, "Almost linear behavior in systems with sector restricted nonlinearities," *Proc. Romanian Academy Series A: Math., Phys., Techn. Sci, Inform. Sci.*, vol. 2, no. 3, pp. 127-135, 2002.
- [25] G. Stepan, "Retarded dynamical systems: stability and characteristic function," *Pitman Res. Notes in Math.* vol. 210, Longman Scientific and Technical, 1989.
- [26] A. A. Voronov, *Stability, controllability, observability* (in Russian), Nauka Publ.House, Moscow, 1979.

Problem 6.7

Open problems in control of linear discrete multidimensional systems

Li Xu

Dept. of Electronics and Information Systems
Akita Prefectural University
Honjo, Akita 015-0055
Japan
xuli@akita-pu.ac.jp

Zhiping Lin

School of EEE
Nanyang Technological University
Singapore 639798
Republic of Singapore
EZPlin@ntu.edu.sg

Jiang-Qian Ying

Faculty of Regional Studies
Gifu University
1-1 Yanagido, Gifu 501-1193
Japan
ying@cc.gifu-u.ac.jp

Osami Saito

Faculty of Engineering
Chiba University
Inageku, Chiba 263-8522
Japan

Yoshihisa Anazawa

Dept. of Electronics and Information Systems
Akita Prefectural University
Honjo, Akita 015-0055
Japan

1 INTRODUCTION

This chapter summarizes several open problems closely related to the following control problems in linear discrete multidimensional (nD , $n \geq 2$) systems:

- output feedback stabilizability and stabilization,
- strong stabilizability and stabilization, or, equivalently, simultaneous stabilizability and stabilization of two given nD systems,
- regulation and tracking control.

Though some of the open problems presented here have been scattered in the literature (see e.g., [7, 13, 24, 26] and the references therein), it seems that they have not received sufficient attention, and were even occasionally mistaken as known results. The purpose of this chapter is twofold: first, to clear up such confusions and to call for more efforts for the solution to these existing open problems; and second, to propose some related new open problems.

2 DESCRIPTION OF THE PROBLEMS

Let $\mathbf{R}[\mathbf{z}]$, where $\mathbf{z} \triangleq (z_1, \dots, z_n)$, be the set of nD polynomials in the variables z_1, \dots, z_n with coefficients in the field of real numbers \mathbf{R} ; $\mathbf{R}(\mathbf{z})$ the set of nD rational functions over \mathbf{R} ; $\mathbf{R}_s[\mathbf{z}]$, $\mathbf{R}_s(\mathbf{z})$ the set of (structurally) stable nD polynomials and rational functions, respectively, i.e., nD polynomials having no zeros in $\bar{U}^n \triangleq \{\mathbf{z} \in \mathbf{C}^n : |z_1| \leq 1, \dots, |z_n| \leq 1\}$ and nD rational functions whose denominators belong to $\mathbf{R}_s[\mathbf{z}]$. Similarly, let $\mathbf{C}[\mathbf{z}]$ be the set of nD polynomials over the field of complex numbers \mathbf{C} , etc.

Problem 1: Let $a_1(\mathbf{z}), \dots, a_M(\mathbf{z}) \in \mathbf{R}[\mathbf{z}]$ be given. Let \mathcal{J} denote the ideal generated by $a_1(\mathbf{z}), \dots, a_M(\mathbf{z})$, and $\mathcal{V}(\mathcal{J})$ the algebraic variety of \mathcal{J} , i.e., $\mathcal{V}(\mathcal{J}) = \{\mathbf{z} \in \mathbf{C}^n : a_i(\mathbf{z}) = 0, i = 1, \dots, M\}$. Suppose that $\mathcal{V}(\mathcal{J}) \cap \bar{U}^n = \emptyset$. Find a constructive method to obtain $h_1(\mathbf{z}), \dots, h_M(\mathbf{z}) \in \mathbf{R}[\mathbf{z}]$ such that

$$a_1(\mathbf{z})h_1(\mathbf{z}) + \dots + a_M(\mathbf{z})h_M(\mathbf{z}) \neq 0 \quad \text{in } \bar{U}^n \quad (1)$$

or, equivalently, to obtain $\tilde{h}_1(\mathbf{z}), \dots, \tilde{h}_\beta(\mathbf{z}) \in \mathbf{R}_s(\mathbf{z})$ such that

$$a_1(\mathbf{z})\tilde{h}_1(\mathbf{z}) + \dots + a_\beta(\mathbf{z})\tilde{h}_\beta(\mathbf{z}) = 1 \quad (2)$$

This problem can be reduced to problem 1', in the sense that once the following problem is solved, problem 1 can be solved easily using the Gröbner basis approach [6, 11, 23].

Problem 1': Under the assumption that $\mathcal{V}(\mathcal{J}) \cap \bar{U}^n = \emptyset$, find a constructive method to obtain a polynomial $s(\mathbf{z})$ such that $s(\mathbf{z}) \in \mathbf{R}_s[\mathbf{z}]$ and $s(\mathbf{z})$ vanishes on $\mathcal{V}(\mathcal{J})$.

Problem 2: Let $g(\mathbf{z}), a_2(\mathbf{z}), \dots, a_M(\mathbf{z}) \in \mathbf{R}[\mathbf{z}]$ be given. Suppose that it is known that there exist $h_2(\mathbf{z}), \dots, h_M(\mathbf{z}) \in \mathbf{R}_s(\mathbf{z})$ such that

$$g(\mathbf{z}) + \sum_{i=2}^M a_i(\mathbf{z})h_i(\mathbf{z}) \neq 0 \quad \text{in } \bar{U}^n. \quad (3)$$

Find a constructive method to obtain such $h_2(\mathbf{z}), \dots, h_M(\mathbf{z})$.

Problem 3: Let $D(\mathbf{z}) \in \mathbf{R}^{m \times m}[\mathbf{z}]$, $N(\mathbf{z}) \in \mathbf{R}^{m \times l}[\mathbf{z}]$ be given. Denote by $\alpha_1(\mathbf{z}), \dots, \alpha_M(\mathbf{z})$ the $m \times m$ minors of $[D(\mathbf{z}) \ N(\mathbf{z})]$ with $M \triangleq \binom{m+l}{m}$ and $\alpha_1(\mathbf{z}) = \det D(\mathbf{z})$. Suppose that $D(\mathbf{z})$ and $N(\mathbf{z})$ are minor left coprime (MLC), i.e., $\alpha_1(\mathbf{z}), \dots, \alpha_M(\mathbf{z})$ have no nonunit common factors over $\mathbf{R}[\mathbf{z}]$ [30]. Suppose that some $h_2(\mathbf{z}), \dots, h_M(\mathbf{z}) \in \mathbf{R}_s(\mathbf{z})$ have been found such that

$$\det D(\mathbf{z}) + \sum_{i=2}^M \alpha_i(\mathbf{z})h_i(\mathbf{z}) \neq 0 \quad \text{in } \bar{U}^n, \quad (4)$$

show whether or not there exists a matrix $C(\mathbf{z}) \in \mathbf{R}_s^{l \times m}(\mathbf{z})$ such that

$$\det(D(\mathbf{z}) + N(\mathbf{z})C(\mathbf{z})) \neq 0 \quad \text{in } \bar{U}^n, \quad (5)$$

and further find a constructive method to obtain such a $C(\mathbf{z})$ when its existence is known.

Problem 4: Let $D(\mathbf{z}) \in \mathbf{R}^{m \times m}[\mathbf{z}]$, $N(\mathbf{z}) \in \mathbf{R}^{l \times m}[\mathbf{z}]$ be given. Show the condition for the existence of $X(\mathbf{z}) \in \mathbf{R}_s^{m \times m}(\mathbf{z})$, $Y(\mathbf{z}) \in \mathbf{R}_s^{m \times l}(\mathbf{z})$ such that

$$D(\mathbf{z})X(\mathbf{z}) + Y(\mathbf{z})N(\mathbf{z}) = I, \quad (6)$$

and further find a constructive method to obtain $X(\mathbf{z})$, $Y(\mathbf{z})$ when the existence is known.

3 MOTIVATIONS

Since the beginning of 1970s, growing interests have led to a considerable number of contributions to the theory of n D systems. This is, of course, mainly due to the diversity of the actual and potential applications of n D systems theory embracing n D signal processing, variable-parameter and lumped-distributed network synthesis, delay-differential systems, linear systems of partial difference and differential equations, iterative learning control systems, linear multipass processes, etc. (see, e.g., the books of [5, 10], the special issues of [2, 3, 15, 18] and the references therein). As it is well-known, the generalization of the conventional one-dimensional (1D) systems theory

to its n D counterpart is nontrivial because of many deep and substantial differences between the two. Despite of the tremendous progress made in the past three decades, there are still many open problems in the area of n D systems, either theoretically challenging or practically important or both, remaining to be tackled. In this chapter, we are mainly concerned with open problems in the area of n D control systems, although some of these problems are also closely related to n D signal processing, as to be discussed shortly.

An n D MIMO (multi-input-multi-output) system $P(\mathbf{z}) \in \mathbf{R}^{m \times l}(\mathbf{z})$ is said to be output feedback (structurally) stabilizable if there is a compensator $C(\mathbf{z}) \in \mathbf{R}^{l \times m}(\mathbf{z})$ such that the closed-loop transfer matrix $H(\mathbf{z})$ defined below is (structurally) stable, i.e., each entry of $H(\mathbf{z})$ is in $\mathbf{R}_s(\mathbf{z})$:

$$H(\mathbf{z}) = \begin{bmatrix} (I + PC)^{-1} & -P(I + CP)^{-1} \\ C(I + PC)^{-1} & (I + CP)^{-1} \end{bmatrix}. \quad (7)$$

If $C(\mathbf{z})$ itself can be further chosen to be stable, $P(\mathbf{z})$ is said to be strongly stabilizable. It can be shown that two unstable systems can be simultaneously stabilized by a single compensator if a certain system constructed from the two given ones is strongly stabilizable [21, 20].

Consider an n D system given by a left matrix fractional description (MFD) $P(\mathbf{z}) = D(\mathbf{z})^{-1}N(\mathbf{z})$ with $D(\mathbf{z}) \in \mathbf{R}^{m \times m}[\mathbf{z}]$ and $N(\mathbf{z}) \in \mathbf{R}^{m \times l}[\mathbf{z}]$. For simplicity, suppose that $D(\mathbf{z})$ and $N(\mathbf{z})$ are MLC. Then, $P(\mathbf{z})$ is stabilizable if and only if

$$\mathcal{V}(\mathcal{J}) \cap \bar{U}^n = \emptyset \quad (8)$$

where $\mathcal{V}(\mathcal{J})$ is the algebraic variety of the ideal \mathcal{J} generated by the $m \times m$ minors $\alpha_1(\mathbf{z}), \dots, \alpha_M(\mathbf{z})$ of $[D(\mathbf{z}) \ N(\mathbf{z})]$ as defined in problem 3 [11, 12, 19, 23]. This condition is equivalent [4] to that there exist $h_1(\mathbf{z}), \dots, h_M(\mathbf{z}) \in \mathbf{R}[\mathbf{z}]$ such that

$$\sum_{i=1}^M \alpha_i(\mathbf{z})h_i(\mathbf{z}) \neq 0 \quad \text{in } \bar{U}^n. \quad (9)$$

Further, it has been shown that, once $h_1(\mathbf{z}), \dots, h_M(\mathbf{z})$ satisfying (9) have been found, a stabilizing compensator $C(\mathbf{z}) = Y(\mathbf{z})X(\mathbf{z})^{-1} \in \mathbf{R}^{l \times m}(\mathbf{z})$ with $X(\mathbf{z}) \in \mathbf{R}^{m \times m}[\mathbf{z}]$, $Y(\mathbf{z}) \in \mathbf{R}^{l \times m}[\mathbf{z}]$ can be constructed [12, 23].

Therefore, the stabilizability for a given $P(\mathbf{z})$ is equivalent to the condition of (8) or the solvability of (9), while the stabilization problem, i.e., the problem of designing a stabilizing compensator, for a stabilizable $P(\mathbf{z})$ is reduced to the problem of constructing $h_1(\mathbf{z}), \dots, h_M(\mathbf{z})$ in (9), which is just what has been described in Problem 1. As mentioned previously, problem 1 can be further reduced to problem 1'. In addition to the above-mentioned stabilization problem, problem 1 also plays an essential role in n D signal processing, such as the design of n D filter banks (see, e.g., [1, 7, 17]).

For the strong stabilizability and stabilization problems, it is further required that $C(\mathbf{z}) = Y(\mathbf{z})X(\mathbf{z})^{-1} \in \mathbf{R}_s^{l \times m}(\mathbf{z})$, which is equivalent to requiring that

$\det X(\mathbf{z}) \in \mathbf{R}_s[\mathbf{z}]$ [26]. It is then easy to see [28] that a necessary condition for $P(\mathbf{z})$ to be strongly stabilizable is that there exist $e_2(\mathbf{z}), \dots, e_M(\mathbf{z})$ such that

$$\det D(\mathbf{z}) + \sum_{i=2}^M \alpha_i(\mathbf{z}) e_i(\mathbf{z}) \neq 0 \quad \text{in } \bar{U}^n \quad (10)$$

where the assumption $\alpha_1(\mathbf{z}) = \det D(\mathbf{z})$ is used. This condition has been shown to be also sufficient for SIMO (single-input-multi-output) and MISO (multi-input-single-output) n D systems [28], and for these special cases, if $e_2(\mathbf{z}), \dots, e_M(\mathbf{z})$ satisfying (10) can be found, a stable stabilizing compensator $C(\mathbf{z})$ can then be constructively obtained. However, the sufficiency of this condition for a general MIMO n D system is still unknown and the problem for constructing a stable stabilizing compensator for a general MIMO n D system is still open, even if $e_2(\mathbf{z}), \dots, e_M(\mathbf{z})$ have been obtained. Problem 2 corresponds to the solution problem of (10), while problem 3 relates to the strong stabilizability and stabilization of a general MIMO n D system. It is clear that the solution of problem 2 is assumed to be a precondition for problem 3.

Another important issue in feedback system design is the tracking and disturbance rejection problems. It can be shown that equation (6) plays a central role for various types of regulation and tracking problems (see, e.g., [21, 22, 25]). So, problem 4 is related to the solvability and solution of regulation and tracking problems of n D MIMO systems.

4 AVAILABLE RESULTS

Problem 1 and problem 1': The test of the solvability condition $\mathcal{V}(\mathcal{J}) \cap \bar{U}^n = \emptyset$ and the solutions of problem 1 and problem 1' for the case $n = 2$ can be found in [11, 23] and the references therein. For the case $n \geq 3$, if \mathcal{J} is of zero dimensional, i.e., $\mathcal{V}(\mathcal{J})$ consists of only a finite number of points, the solvability test and solution construction can then be carried out by utilizing the Gröbner basis approach [6, 24]. For some other special cases when $n \geq 3$, see [14]. Another solution method has been suggested in [4] by using analytic function theory. However, we believe that this method is not constructive. Further, as the determination of whether or not $\mathcal{V}(\mathcal{J}) \cap \bar{U}^n = \emptyset$ can be formulated as a typical quantifier elimination problem, it may be possible to solve it by Cylindrical Algebraic Decomposition (CAD) techniques developed in the field of computer algebra [8, 9].

Problem 2: Problem 2 is much more complicated and difficult than its 1D counterpart (see e.g., [21]). To solve this problem, we have to follow two steps: first, to see if there exist $h_2(\mathbf{z}), \dots, h_M(\mathbf{z}) \in \mathbf{C}_s(\mathbf{z})$, and then to see if there exist $h_2(\mathbf{z}), \dots, h_M(\mathbf{z}) \in \mathbf{R}_s(\mathbf{z})$, such that (3) holds. It is also

interesting to note that in contrast to the 1D case, Problem 2 may possess no solution on $\mathbf{R}_s(\mathbf{z})$, even if it has a solution on $\mathbf{C}_s(\mathbf{z})$ [26, 28, 29].

Necessary and sufficient conditions for the existence of $h_2(\mathbf{z}), \dots, h_M(\mathbf{z}) \in \mathbf{C}_s(\mathbf{z})$ and $\mathbf{R}_s(\mathbf{z})$, respectively, have been shown in [26, 28], which can be verified by the CAD based method [27].

Problem 3: Problem 3 has been considered and solved for the cases when $m = 1$ or $l = 1$ (i.e., for SIMO and MISO n D systems) [28], and for the case when $D(\mathbf{z})$ and $N(\mathbf{z})$ satisfy certain conditions given in [16].

Problem 4: Necessary and sufficient solvability conditions and constructive solution procedures for the case $n = 2$ can be found in [22].

BIBLIOGRAPHY

- [1] S. Basu, "Multi-dimensional filter banks and wavelets: A system theoretic perspective," *J. Franklin Inst.*, 335B(8), pp. 1367-1409, 1998.
- [2] S. Basu and B. C. Lévy, Eds., "Special Issue on Multidimensional Filter Banks and Wavelets," *Multidimensional Systems and Signal Processing*, 7/8, 1996/7.
- [3] S. Basu and M. N. S. Swamy, Eds., "Special Issue on Multidimensional Signals and Systems," *IEEE Trans. Circuits Syst. I*, 49, 2002.
- [4] C. A. Berenstein and D. C. Struppa, "1-Inverses for polynomial matrices of non-constant rank," *Systems & Control Letters*, 6, pp. 309-314, 1986.
- [5] N. K. Bose, *Applied Multidimensional Systems Theory*, New York: Van Nostrand Reinhold, 1982.
- [6] B. Buchberger, "Gröbner bases: An algorithmic method in polynomial ideal theory," in *Multidimensional Systems Theory: Progress, Directions and Open Problems* N. K. Bose, ed., Dordrecht: Reidel, pp. 184-232, 1985.
- [7] C. Charoenlarnnopparut and N. K. Bose, "Gröbner Bases for Problem Solving in Multidimensional Systems," *Multidimensional Systems and Signal Processing*, 12(3/4), pp. 365-376, 2001.
- [8] G. Collins, "Quantifier elimination for real closed fields by cylindrical algebraic decomposition," *LNCS*, 33, pp. 134-183, 1975.
- [9] G. Collins, H. Hong, "Partial cylindrical algebraic decomposition and quantifier elimination," *J. Symbolic Computation*, 12, pp. 299-328, 1991.

- [10] K. Galkowski and J. Wood, Eds., *Multidimensional Signals, Circuits and Systems*, Taylor & Francis, 2001.
- [11] J. P. Guiver and N. K. Bose, "Causal and weakly causal 2-D filters with applications in stabilizations," In: *Multidimensional Systems Theory: Progress, Directions and Open Problems*, N. K. Bose, ed., Dordrecht: Reidel, pp. 52–100, 1985.
- [12] Z. Lin, Feedback Stabilizability of MIMO n -D Linear Systems, *Multidimensional Systems and Signal Processing*, 9, pp. 149–172, 1998.
- [13] Z. Lin, "Output feedback stabilizability and stabilization of linear n D systems," In: *Multidimensional Signals, Circuits and Systems*, K. Galkowski and J. Wood, eds., Taylor & Francis, pp. 59 - 76, 2001.
- [14] Z. Lin, J. Lam, K. Galkowski and S. Xu, "A Constructive Approach to Stabilizability and Stabilization of a Class of n D Systems," *Multidimensional Systems and Signal Processing*, 12(3/4), pp. 329–344, 2001.
- [15] Z. Lin and L. Xu, Eds., "Special Issue on Applications of Gröbner Bases to Multidimensional Systems and Signal Processing," *Multidimensional Systems and Signal Processing*, 12, 2001.
- [16] Z. Lin, J. Q. Ying, L. Xu, "An Algebraic Approach to Strong Stabilizability of Linear n D MIMO Systems," *IEEE Trans. Automat. Contr.*, 47(9), pp. 1510 - 1514, 2002.
- [17] H. Park, T. Kalker, M. Vetterli, "Gröbner Bases and Multidimensional FIR Multirate Systems," *Multidimensional Systems and Signal Processing*, 8(1/2), pp. 11–30, 1997.
- [18] E. Rogers and P. Rocha, Eds., "Recent Progress in Multidimensional Control Theory and Applications," *Multidimensional Systems and Signal Processing*, 11, 2000.
- [19] S. Shankar, V. R. Sule, "Algebraic Geometric Aspects of Feedback Stabilization," *SIAM J. Contr. Optim.*, 30, pp. 11–30, 1992.
- [20] S. Shankar, "An obstruction to the simultaneous stabilization of two n -D plants," *Acta Applicandae Mathematicae*, 36, pp. 289–301, 1994.
- [21] M. Vidyasagar, *Control System Synthesis: A Factorization Approach*, Cambridge, MA: MIT Press, 1985.
- [22] L. Xu, O. Saito, and K. Abe. "Bilateral Polynomial Matrix Equations in Two Indeterminates," *Multidimensional Systems and Signal Processing*, 1(4), pp. 363–37, 1990.
- [23] L. Xu, O. Saito, and K. Abe. "Output Feedback Stabilizability and Stabilization Algorithms for 2D systems," *Multidimensional Systems and Signal Processing*, 5(1), pp. 41–60, 1994.

- [24] L. Xu, J. Q. Ying, O. Saito, "Feedback Stabilization for a Class of MIMO n -D Systems by Gröbner Basis Approach," In: *Abstract of First International Workshop on Multidimensional Systems*, pp. 88-90, Poland, 1998.
- [25] L. Xu, O. Saito, J. Q. Ying, "2D Feedback System Design: The Tracking and Disturbance Rejection Problems," *Proceedings of ISCAS99*, V, pp. 13-16, Orlando, USA, 1999.
- [26] J. Q. Ying, "Conditions for strong stabilizabilities of n -dimensional systems," *Multidimensional Systems and Signal Processing*, 9, pp. 125-148, 1998.
- [27] J. Q. Ying, L. Xu, Z. Lin, "A computational method for determining strong stabilizability of n -D systems", *J. Symbolic Computation*, 27, pp. 479-499, 1999.
- [28] J. Q. Ying, "On the strong stabilizability of MIMO n -dimensional linear systems," *SIAM J. Contr. Optim.*, 38, pp. 313-335, 2000.
- [29] J. Q. Ying, Z. Lin, L. Xu, "Some algebraic aspects of the strong stabilizability of time-delay linear systems," *IEEE Trans. Automat. Contr.*, 46, pp. 454-457, 2001.
- [30] D. C. Youla and G. Gnani, "Notes on n -dimensional system theory," *IEEE Trans. Circuits Syst.*, 26, pp. 105-111, 1979.

Problem 6.8

An open problem in adaptative nonlinear control theory

Leonid S. Zhiteckij

Int. Centre of Inform. Technologies and Systems

Institute of Cybernetics

40 Prospect Akademika Glushkova

MSP03680 Kiev 187

Ukraine

zls@d310.icyb.kiev.ua

1 STATEMENT OF THE PROBLEM

We deal with the problem of globally stable adaptive control for discrete-time, time-invariant, nonlinear, but linearly parameterized (LP) systems described by the difference equation

$$y_t = \theta^T \varphi(x_{i-1}) + bu_{t-1} + v_t, \quad (1)$$

where $y_t : \mathbb{Z}^+ \rightarrow \mathbb{R}$ and $u_t : \mathbb{Z}^+ \rightarrow \mathbb{R}$ are the measurable output and control input, respectively, and $v_t : \mathbb{Z}^+ \rightarrow \mathbb{R}$ is the unmeasured disturbance (the integer t denotes the discrete time). $\theta \in \mathbb{R}^d$ and $b \in \mathbb{R}$ are the unknown parameter vector and scalar ($d \geq 1$). $f(\cdot) : \mathbb{R}^N \rightarrow \mathbb{R}^d$ represents a known nonlinear vector function depending on the vector $x_{t-1}^T = [y_{t-1}, \dots, y_{t-N}]$ of N past outputs. Its growth is given by

$$\|\varphi(x)\| = O(\|x\|^\beta) \quad \text{as} \quad \|x\| \rightarrow \infty. \quad (2)$$

Assume that v_t is upper bounded by some finite η , i.e.,

$$\|v_t\|_\infty \leq \eta < \infty, \quad (3)$$

where $\|v_t\|_\infty := \sup_{0 \leq t < +\infty} |v_t|$ denotes the l_∞ -norm of v_t . To regulate y_t around zero, we choose the well-known certainty equivalence (CE) feedback control law

$$u_t = -b_t^{-1} \theta_t^T \varphi(x_t), \quad (4)$$

where b_t and θ_t are the estimates of unknown b and θ that are to be updated online by using either gradient or least squares (LS) based algorithms. These

classical recursive adaptation algorithms may be written in a general form as

$$\bar{\theta}_t = \bar{\theta}_{t-1} + \alpha_t \Omega_\theta(P_t, y_t, \bar{\varphi}(x_{t-1})) \quad (5)$$

$$P_t = P_{t-1} - \Omega_P(P_{t-1}, \bar{\varphi}(x_{t-1}), \alpha_t) \quad (6)$$

with

$$\alpha_t = \Omega_\alpha(y_t, \varphi(x_{t-1})), \quad \alpha_t \geq 0, \quad (7)$$

where $\bar{\theta}_t^T = [\theta_t^T, b_t]$, $\bar{\varphi}_t^T = [\varphi_t^T, u_t]$ are the extended vectors, P_t is a positive definite $(d+1) \times (d+1)$ matrix and $\Omega_\theta : \mathbb{R}^{(d+1) \times (d+1)} \times \mathbb{R} \times \mathbb{R}^{d+1} \rightarrow \mathbb{R}^{d+1}$, $\Omega_P : \mathbb{R}^{(d+1) \times (d+1)} \times \mathbb{R}^{d+1} \times \mathbb{R} \rightarrow \mathbb{R}^{(d+1) \times (d+1)}$ and $\Omega_\alpha : \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}$.

Definition. System (1) is said to be globally stabilizable if there exists an adaptive feedback control of the form (4)-(7) such that

$$\limsup_{t \rightarrow \infty} |y_t| < \infty$$

for any initial $x_0 \in \mathbb{R}^N$, $\bar{\theta}_0 \in \mathbb{R}^{d+1}$, $\alpha_0 \in \mathbb{R}^+$, some $P_0 > 0$ and a given sequence of the disturbances $\{v_t\}$ satisfying (3).

Now, we formulate the problem as follows: determine the triple $(\Omega_\theta, \Omega_P, \Omega_\alpha)$ such that the adaptive feedback control (4)-(7) will ensure the global stabilizability of system (1) for the given class of $\{v_t\} \in l_\infty$ provided that $\varphi(x)$ belongs to a given class of nonlinearities having a growth rate (2) with some β satisfying $1 < \beta < \beta^*$, where β^* needs to be evaluated. The problem stated thus generalizes the problem solved in [1] and [4] to the bounded disturbance case. This is an open and difficult problem in the adaptive control theory. To the best of the author's knowledge, there are no available results solving it for $\|v_t\|_\infty \neq 0$, whereas the solution to its continuous-time counterpart is known.

2 MOTIVATION

In contrast to the adaptive control of nonlinear continuous-time systems, where substantial breakthroughs in the theoretical area have been achieved by the middle of the 1990s (see, e.g., [3], [5], etc.), very few similar works are available in the literature that address the global stable adaptive control design for discrete-time systems with nonlinearities [1], [2], [4], [6]-[8]. One of the inherent difficulties of discrete-time adaptive control is that the Lyapunov stability techniques typically exploited in the continuous time case may not be straightforwardly extended to its discrete-time counterpart, as detailed in [4], [6], [8]. It has been shown in Section II of [4], and in [8] and [9] that the so-called Key Technical Lemma, which has played a key role in analyzing the adaptation algorithms of type (5)-(7) applied to linear discrete-time systems, can be used to derive the stabilizability properties of adaptive nonlinear LP systems with a nonlinearity whose growth rate (2)

is linear ($\beta = 1$). Unfortunately, this stability analysis tool is no longer valid if $\varphi(x)$ has a growth rate faster than linear, i.e., $\beta > 1$ (see, e.g., [4], [8]). In such a situation, the following questions naturally arise: Can the linear growth restriction $\beta = 1$ be relaxed without going to the instability of closed loop? What are the limitations of gradient and LS based algorithms? An answer to these questions can be partially found in recent works [2],[7] dealing with a similar problem in the stochastic framework. Although the results of [2], [7] shed some light on restrictions that must be imposed on β to achieve global stability, however, the question of how they might be extended to the nonstochastic case, where $\{v_t\} \in l_\infty$, has not been resolved as yet.

3 RELATED RESULTS

The first step allowing to relax the linear growth condition with respect to $\|\varphi(x)\|$ has been made by Kanellakopoulos [4] who dealt with the scalar one-parametric disturbance-free system of form (1) ($N = 1, d = 1, v_t \equiv 0$) provided that the gain b is known and equal to 1. In Section III of [8] it has been established that the LS algorithm (5), (6) with the nonlinear gain determined in (7) as $\alpha_t = 1 + \varphi^2(x_t)$ can be used to adaptively stabilize system (1) for any smooth nonlinearity $\varphi(x) : \mathbb{R} \rightarrow \mathbb{R}$ independently of its growth rate. To derive this global stability result, Kanellakopoulos employed the Lyapunov function

$$V_t = \ln(1 + x_t^2) + cP_t^{-1}\tilde{\theta}_t^2 + P_t^2$$

with some $c > 0$, where $\tilde{\theta}_t = \theta - \theta_t$ is the parameter error vector. The adaptive control of system (1) with no disturbance and $b = 1$ has also been studied by Guo and Wei [1]. In contrast to [4], these authors used the standard ($\alpha_t \equiv 1$) LS based algorithm of form (5), (6). By exploiting a new theoretical tool based on some boundedness properties of $\{\det P_t^{-1}\}$, they have proved that if $d = 1$, then the closed-loop adaptive system (1),(5)-(7) is globally stable whenever $\beta < 8$. It has been also established for the multiparameter case ($d > 1$) that the global stability condition is $\beta d < 4$ (see theorem 3 of [4]). The fundamental limitations of the standard LS-based adaptive control applied to system (1) with $d = 1$ and $b = 1$ in the presence of stochastic $\{v_t\}$ have been established by Guo [2] who proved that a globally stabilizing adaptive LS-based controller can be designed if and only if $\beta < 4$. Recently, Xie and Guo [7] showed that if $d \gg 1$, then the linear growth restriction ($\beta = 1$) cannot be essentially relaxed in general to globally stabilize system (1) subjected to a Gaussian white noise $\{v_t\}$, unless additional conditions on number d and the structure of $\varphi(\cdot)$ are imposed (see Remark 3 of [7]). It seems that a new theoretical tool should be devised to solve the problem formulated above.

BIBLIOGRAPHY

- [1] L. Guo and C. Wei, "Global stability/instability of LS-based discrete-time adaptive nonlinear control," In: *Proc. 13th IFAC World Congress*, San Francisco, CA, USA, Vol. K, pp. 277-282, 1996.
- [2] L. Guo, "On critical stability of discrete-time adaptive Nonlinear control," *IEEE Trans. Automat. Contr.*, 42, pp. 1488-1499, 1997.
- [3] I. Kanellakopoulos, P. V. Kokotovic and A. S. Morse, "Systematic design of adaptive controllers for feedback linearizable systems," *IEEE Trans. Automat. Contr.*, 36, pp. 1242-1253, 1991.
- [4] I. Kanellakopoulos, "Discrete-time adaptive nonlinear system," *IEEE Trans. Automat. Contr.*, 39, pp. 2362-2365, 1994.
- [5] M. Krstic, I. Kanellakopoulos, and P. V. Kokotovic, *Nonlinear and Adaptive Control Design*, N.Y.: Wiley, 1995.
- [6] Y. Song and J. W. Grizzle, "Adaptive output-feedback control of a class of discrete time nonlinear systems," In: *Proc. 1993 Amer. Contr. Conf.*, San Francisco, USA, pp. 1359-1364, 1993.
- [7] L.-L. Xie and L. Guo, "Fundamental limitations of discrete-time adaptive nonlinear control," *IEEE Trans. Automat. Contr.*, 44, pp. 1777-1782, 1999.
- [8] P.-C. Yeh and P. V. Kokotovic, "Adaptive control of a class of nonlinear discrete-time systems," *Int. J. Control*, 62, pp. 303-324, 1995.
- [9] L. S. Zhiteckij, "Singularity-free stable adaptive control of a class of nonlinear discrete-time systems," In: *Proc. 15th IFAC World Congress*, Barcelona, Spain, 2002.

Problem 6.9

Generalized Lyapunov theory and its omega-transformable regions

Sheng-Guo Wang

College of Engineering

University of North Carolina at Charlotte

Charlotte, NC 28223-0001

USA

swang@uncc.edu

1 DESCRIPTION OF THE PROBLEM

The open problem discussed here is a Generalized Lyapunov Theory and its Ω -transformable regions. First, we provide the definition of the Ω -transformable regions and its degrees. Then the open problem is presented and discussed.

Definition 1: (Gutman & Jury 1981) A region

$$\Omega_v = \{(x, y) \mid f(\lambda, \lambda^*) = f(x + jy, x - jy) = f_{xy}(x, y) < 0\} \quad (1)$$

is Ω -transformable if any two points $\alpha, \beta \in \Omega_v$ imply $Re[f(\alpha, \beta^*)] < 0$, where function $f(\lambda, \lambda^*) = f_{xy}(x, y) = 0$ is the boundary function of the region Ω_v and v is the degree of the function f . Otherwise, the region Ω_v is non- Ω -transformable.

It is noticed that a region on one side of a line and a region within a circle in the plane both are Ω -transformable regions. However, some regions are non- Ω -transformable regions.

Open Problem: (Generalized Lyapunov Theory) Consider a matrix $A \in C^{n \times n}$ and any Ω -transformable region Ω_v described by $f_{xy}(x, y) = f(\lambda, \lambda^*) < 0$ with its boundary equation $f_{xy}(x, y) = f(\lambda, \lambda^*) = 0$, where v (any positive integer number) is the degree of the boundary function f and

$$f(\lambda, \lambda^*) = \sum_{p+q \leq v, p, q=1}^v c_{pq} \lambda^p \lambda^{*q}, \quad \lambda = x + jy \quad (2)$$

λ is a point on the complex plane. For the eigenvalues of the matrix A to lie in Ω_v , it is necessary and sufficient that given any positive definite

(*p.d.*) Hermitian matrix Q , there exists a unique *p.d.* Hermitian matrix P satisfying the Generalized Laypunov Equation (GLE)

$$\sum_{p+q \leq v, p, q=1}^v c_{pq} A^p P A^{*q} = -Q \quad (3)$$

Strictly speaking, the above open problem is for Ω -transformable regions with degree v greater than two. However, in order to let the problem be more general, we present the genrerl Lyapunov equation for any positive integer v .

2 MOTIVATION AND HISTORY OF THE PROBLEM

The Lyapunov theory is well known for Hurwitz stability and Schur stability, i.e., the continuous-time system Lyapunov theory and the discrete-time system Lyapunov theory, respectively. The above generalized Lyapunov theory (GLT) takes both continuous-time system and discrete-time system Lyapunov theories as its special cases. Furthermore, it is well known that the closed-loop system poles determine the system stability and nature, and dominate the system response and performance. Thus, when we consider the performance, we need the closed-loop system poles, i.e., the closed-loop system matrix eigenvalues, within a specific region. Various engineering applications and performance requirements need a consideration to locate the system poles within various specific regions. The GLT provides a necessary and sufficient condition to these problems as the Lyapunov theory to the stability problems.

Here, let us briefly review the history of the classical Lyapunov theory as follows. Its significance is to provide a necessary and sufficient condition for matrix eigenvalues to lie in the left-half plane by the Lyapunov equation for the continuous-time systems.

Lyapunov Theory (continuous-time): For the eigenvalues of matrix A to lie in the left half plane, i.e., matrix A is Hurwitz stable, it is necessary and sufficient that given any positive definite (*p.d.*) Hermitian matrix Q , there exists a unique *p.d.* Hermitian matrix P satisfying the following Laypunov Equation (LE)

$$AP + PA^* = -Q \quad (4)$$

For discrete-time systems, the interest on the system stability is to check if the system matrix eigenvalues lie within the unit disk. The corresponding Lyapunov theory for the discrete-time systems is as follows:

Lyapunov Theory (discrete-time): For the eigenvalues of matrix A to lie in the unit-disk, i.e., matrix A is Schur stable, it is necessary and sufficient that given any *p.d.* Hermitian matrix Q , there exists a unique *p.d.* Hermitian matrix P satisfying the following LE

$$APA^* - P = -Q \quad (5)$$

It is clear that the Lyapunov theory for the Hurwitz stability and the Schur stability is a special case of the Generalized Lyapunov Theory described in the open problem with their specific Ω -transformable regions of the left half-plane and the unit disk, respectively. The degree v of the left-half plane is one, and the degree v of the unit disk is two.

For the system performance, we may check the system pole clustering in specific interested general Ω -transformable regions by the GLT.

With respect to the robust control, we need the robust performance in addition to the robust stability. Thus, a robust pole clustering, or robust root clustering, or robust Gamma stability, as called in the literature (Ackermann, Kaesbauer & Muench 1991, Barmish 1994, Wang & Shieh 1994a,b, Yedavalli 1993, among others), is needed. The approach for the robust pole clustering is first to define the region boundary function for the system performance. Then the General Lyapunov Theory will be very useful for us to determine the robust pole clustering in general Ω -transformable regions for the system robust performance, which is similar to our dealing with the system stability and robust stability via the Lyapunov theory.

Also, the more general regions may be very interesting in the study of discrete-time systems, where the transient behavior is hard to specify in terms of common simple regions. In other areas, such as multidimensional digital filters and multidimensional systems, the Ω -transformable regions and its related GLT, as well as non- Ω -transformable regions, will be further useful. The non- Ω -transformable regions also identify the GLT as invalid in the regions.

All these considerations constitute the motivation to investigate the open problem GLT and its Ω -transformable regions.

3 AVAILABLE RESULTS

This section describes some related available results.

Theorem 1: (GLT: Gutman & Jury 1981) Let $A \in C^{n \times n}$ and consider any Ω -transformable Ω_v in Equation (1) with its boundary function f , where $v = 1, 2$ and

$$f(\lambda, \lambda^*) = \sum_{p+q \leq v, p, q=1}^v c_{pq} \lambda^p \lambda^{*q} \quad (6)$$

For the eigenvalues of A to lie in Ω_v , it is necessary and sufficient that given any $p.d.$ Hermitian matrix Q , there exists a unique $p.d.$ Hermitian matrix P satisfying the GLE

$$\sum_{p+q \leq v, p, q=1}^v c_{pq} A^p P A^{*q} = -Q \quad (7)$$

Notice that the GLT is proved and valid for Ω -transformable regions with $v = 1, 2$ (Gutman & Jury 1981). For Ω -transformable regions with $v \geq 3$, the GLT is only a conjecture so far.

On the other hand, it is also noticed that the GLT is not valid for non- Ω -transformable regions as pointed in Gutman & Jury 1981 and Wang 1996. In Wang (1996), a counterexample shows that the GLT is not valid for non- Ω -transformable regions.

Furthermore, notice from Gutman & Jury 1981 that Γ -transformable regions proposed by Kalman (1969) and Ω -transformable regions do not cover each other. Γ -transformable regions are originally a rational mapping from the upper half-plane (UHP) or the left half-plane (LHP) into the unit circle, identical to the region proposed by Hermite (1856) (see Gutman and Jury 1981). Strictly speaking, a region Γ_v is

$$\Gamma_v = \{(x, y) \mid |\psi(s)|^2 - |\phi(s)|^2 < 0, s = x + jy\} \quad (8)$$

that is mapped from the unit disk $\{w \mid |w| < 1\}$ by the rational function $w = \frac{\psi(s)}{\phi(s)}$, $s = x + jy$, with v being the degree of the (x, y) polynomial in Equation (8).

By applying the GLT, Horng, Horng and Chou (1993) and Yedavalli (1993) discussed robust pole clustering in Ω -transformable regions with degrees one and two. However, Wang and Shieh (1994a,b) used a Rayleigh principle approach to analyze the robust pole clustering in general Ω -regions, described as

$$\Omega_v = \{(x, y) \mid f(\lambda, \lambda^*) = \sum_{p+q \leq v, p, q=1}^v c_{pq} \lambda^p \lambda^{*q} < 0, \lambda = x + jy\} \quad (9)$$

which they called *Hermitian regions* or *general Ω regions*, including both Ω -transformable and non- Ω -transformable regions, as well as Γ regions.

It is well-known that if a region is Ω -transformable, its complement is not Ω -transformable. On the other hand, the complement of a Γ -transformable region is also a Γ -transformable region. However, the general Ω -regions or Hermitian regions include all of them.

Notice that the general Ω regions (Wang and Shieh 1994a,b) do not need to satisfy the condition in Definition 1 of Ω -transformable regions. Wang and Yedavalli (1997) discussed eigenvectors and robust pole clustering in general subregions Ω of complex plane for uncertain matrices. Wang (1999, 2000, and 2003) discussed robust pole clustering in a good ride quality region of aircraft, a specific non- Ω -transformable region.

However, so far the related researches have not provided any solution to the above open problem. Therefore, the above open problem remains an open problem thus far.

BIBLIOGRAPHY

- [1] J. Ackermann, D. Kaesbauer, and R. Muench, "Robust Gamma stability analysis in a plant parameter space," *Automatica*, 27, pp.75-85, 1991.
- [2] B. R. Barmish, *New Tools for Robustness of Linear Systems*, Macmillan Publishing Comp., New York, 1994.
- [3] S. Gutman and E. I. Jury, "A general theory for matrix root-clustering in subregions of the complex plane," *IEEE Trans. Automatic Control*, 26(4), pp.853-863, 1981.
- [4] C. Hermite, "On the number of roots of an algebraic equation contained between given limits," *J. Reine Angew. Math.*, 52, pp. 39-51, 1856; also in *Int. J. Control*, 26, pp. 183-195, 1977.
- [5] I.-R. Horng, H.-Y. Horng and J.-H. Chou, "Eigenvalue clustering in subregions of the complex plane for interval dynamic systems," *Int. J. Systems Science*, 24, pp. 901-914, 1993.
- [6] R. E. Kalman, "Algebraic characterization of polynomials whose zeros lie in a certain algebraic domain," *Proc. Nat. Acad. Sci.*, 64, pp.818-823, 1969.
- [7] S.-G. Wang, "Comments on perturbation bounds for root-clustering of linear systems in a specified second order subregion," *IEEE Trans. Automatic Control*, 41, pp. 766-767, 1996.
- [8] S.-G. Wang, "Robust pole clustering in a good ride quality region of aircraft for structured uncertain matrices," *Proc. the 14th World Congress of IFAC*, Vol. P, pp. 277-282, Beijing, China, 1999.
- [9] S.-G. Wang, "Analysis of robust pole clustering in a good ride quality region for uncertain matrices," *Proc. the 39th IEEE CDC*, pp. 4203-4208, Sydney, Australia, 2000.
- [10] S.-G. Wang, "Robust pole clustering in a good ride quality region of aircraft for matrices with structured uncertainties," *Automatica*, 39(3), pp.525-532, 2003.
- [11] S.-G. Wang and L. Shieh, "Robustness of linear quadratic regulators with regional-pole constraints for uncertain linear systems," *Int. J. Control Theory & Advanced Technology*, 10(4), pp.737-769, 1994a.
- [12] S.-G. Wang and L. Shieh, "A general theory for analysis and design of robust pole clustering in subregions of the complex plane," *Proc. 1994 American Control Conf.*, pp. 627-631, Baltimore, USA, 1994b.

- [13] S.-G. Wang and R. Yedavalli, "Eigenvectors and robust pole clustering in general subregions of complex plane for uncertain matrices," *Proc. 36th IEEE Conf. Decision and Control*, pp. 2121-2126, San Diego, CA, 1997.
- [14] R. K. Yedavalli, "Robust root clustering for linear uncertain system using generalized Lyapunov theory," *Automatica*, 29(1), pp. 237-240, 1993.

Problem 6.10

Smooth Lyapunov characterization of measurement to error stability

Brian P. Ingalls

Control and Dynamical Systems
California Institute of Technology
Pasadena, CA 91125
USA
ingalls@cds.caltech.edu

Eduardo D. Sontag

Department of Mathematics
Rutgers University
Piscataway, NJ 08854
USA
sontag@math.rutgers.edu

1 DESCRIPTION OF THE PROBLEM

Consider the system

$$\dot{x}(t) = f(x(t), u(t)) \tag{1}$$

with two output maps

$$y(t) = h(x(t)), \quad w(t) = g(x(t)),$$

with states $x(t) \in \mathbb{R}$ and controls u measurable essentially bounded functions into \mathbb{R}^m . Assume that the function $f : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ is locally Lipschitz, and that the system is forward complete. Assume that the output maps $h : \mathbb{R}^n \rightarrow \mathbb{R}^{p_y}$ and $g : \mathbb{R}^n \rightarrow \mathbb{R}^{p_w}$ are locally Lipschitz.

The Euclidean norm in a space \mathbb{R}^k is denoted simply by $|\cdot|$. If z is a function defined on a real interval containing $[0, t]$, $\|z\|_{[0, t]}$ is the sup norm of the restriction of z to $[0, t]$, that is $\|z\|_{[0, t]} = \text{ess sup } \{|z(t)| : t \in [0, t]\}$.

A function $\gamma : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ is of class \mathcal{K} (denoted $\gamma \in \mathcal{K}$) if it is continuous, positive definite, and strictly increasing; and is of class \mathcal{K}_∞ if in addition it

is unbounded. A function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ is of class \mathcal{KL} if for each fixed $t \geq 0$, $\beta(\cdot, t)$ is of class \mathcal{K} and for each fixed $s \geq 0$, $\beta(s, t)$ decreases to zero as $t \rightarrow \infty$.

The following definitions are given for a forward complete system with two output channels as in (1). The outputs y and w are considered as error and measurement signals, respectively.

Definition: We say that the system (1) is *input-measurement to error stable* (IMES) if there exist $\beta \in \mathcal{KL}$ and $\gamma_1, \gamma_2 \in \mathcal{K}$ so that

$$|y(t)| \leq \max\{\beta(|x(0)|, t), \gamma_1(\|w\|_{[0,t]}), \gamma_2(\|u\|_{[0,t]})\}$$

for each solution of (1), and all $t \geq 0$.

Open Problem: Find a (if possible, smooth) Lyapunov characterization of the IMES property.

2 MOTIVATION AND HISTORY OF THE PROBLEM

The input-measurement to error stability property is a generalization of input to state stability (ISS). Since its introduction in [11], the ISS property has been extended in a number of ways. One of these is to a notion of output stability, input to output stability (IOS), in which the magnitude of an output signal is asymptotically bounded by the input. Another is to a detectability notion: input-output to state stability (IOSS). In this case, the size of the state is asymptotically bounded by the input and output.

In these two concepts, the outputs play distinct roles. In IOS the output is to be kept small, e.g., an error. In IOSS the output provides information about the size of the state, e.g., a measurement. This leads one to consider a system with two output channels: an error and a measurement. The notions of IOS and IOSS can be combined to yield (IMES), a property in which the error is asymptotically bounded by the input and a measurement. This partial detectability notion is a direct generalization of IOS and IOSS (and ISS). It constitutes the key concept needed in order to fully extend regulator theory to a global nonlinear context, and was introduced in [12], where it was called “input measurement to output stability” (IMOS).

One of the most useful results on ISS is its characterization in terms of the existence of an appropriate smooth Lyapunov function [13]. As the IOS and IOSS properties were introduced, they too were characterized in terms of Lyapunov functions (in [16, 17] and [7, 14, 15], respectively). A Lyapunov characterization of IMES would include both of these results, as well as the original characterization of ISS. For applications of Lyapunov functions to ISS and related properties, see, for instance, [1, 4, 5, 6, 8, 9, 10].

3 AVAILABLE RESULTS

In an attempt to determine a Lyapunov characterization for IMES, one might hope to fashion a proof along the same lines as that for the IOSS characterization given in [7]. Such an attempt has been made, with preliminary results reported in [3]. In that paper, the MES property (i.e., IMES for a system with no input) is addressed. The relation between MES and a secondary property, stability in three measures (SIT), is described, and the following (discontinuous) Lyapunov characterization for SIT is given.

Definition: We say that the system (1) is *measurement to error stable* (MES) if there exist $\beta \in \mathcal{KL}$ and $\gamma_1 \in \mathcal{K}$ so that

$$|y(t)| \leq \max\{\beta(|x(0)|, t), \gamma_1(\|w\|_{[0,t]})\}$$

for each solution of (1), and all $t \geq 0$.

Definition: Let $\rho \in \mathcal{K}$. We say that the system (1) satisfies the *stability in three measures* (SIT) property (with gain ρ) if there exists $\beta \in \mathcal{KL}$ so that for any solution of (1), if there exists $t_1 > 0$ so that $|y(t)| > \rho(|w(t)|)$ for all $t \in [0, t_1]$, then

$$|y(t)| \leq \beta(|x(0)|, t) \quad \forall t \in [0, t_1].$$

The MES property implies the SIT property. The converse does not hold in general, but is true under additional assumptions on the system.

Definition: Let $\rho \in \mathcal{K}$. We say that a lower semicontinuous function $V : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$ is a *lower semicontinuous SIT-Lyapunov function* for system (1) with gain ρ if

- there exist $\alpha_1, \alpha_2 \in \mathcal{K}_\infty$ so that

$$\alpha_1(|h(\xi)|) \leq V(\xi) \leq \alpha_2(|\xi|), \quad \forall \xi \text{ so that } |h(\xi)| > \rho(|g(\xi)|),$$

- there exists $\alpha_3 : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 0}$ continuous positive definite so that for each ξ so that $|h(\xi)| > \rho(|g(\xi)|)$,

$$\zeta \cdot v \leq -\alpha_3(V(\xi)) \quad \forall \zeta \in \partial_D V(\xi), \quad \forall v \in F(\xi). \quad (2)$$

(Here ∂_D denotes a viscosity subgradient.)

Theorem: Let a system of the form (1) and a function $\rho \in \mathcal{K}$ be given. The following are equivalent.

- The system satisfies the SIT property with gain ρ .
- The system admits a lower semicontinuous SIT-Lyapunov function with gain ρ .

- iii. The system admits a lower semicontinuous exponential decay SIT-Lyapunov function with gain ρ .

Further details are available in [3] and [2].

BIBLIOGRAPHY

- [1] R. A. Freeman and P. V. Kokotović, *Robust Nonlinear Control Design, State-Space and Lyapunov Techniques*, Birkhauser, Boston, 1996.
- [2] B. Ingalls, *Comparisons of Notions of Stability for Nonlinear Control Systems with Outputs*, Ph.D. thesis, Rutgers University, New Brunswick, New Jersey, USA, 2001. Available at www.cds.caltech.edu/~ingalls.
- [3] B. Ingalls, E. D. Sontag, and Y. Wang, "Measurement to error stability: a notion of partial detectability for nonlinear systems," submitted.
- [4] A. Isidori, *Nonlinear Control Systems II*, Springer-Verlag, London, 1999.
- [5] H. K. Khalil, *Nonlinear Systems, Second Edition*, Prentice-Hall, Upper Saddle River, NJ, 1996.
- [6] P. Kokotović and M. Arcak, "Constructive nonlinear control: Progress in the 90s," Invited Plenary Talk, IFAC Congress, In: *Proc. 14th IFAC World Congress*, the Plenary and Index Volume, pp. 49–77, Beijing, 1999.
- [7] M. Krichman, E. D. Sontag, and Y. Wang, "Input-output-to-state stability," *SIAM Journal on Control and Optimization* 39, 2001, pp. 1874–1928, 2001.
- [8] M. Krstić and H. Deng, *Stabilization of Uncertain Nonlinear Systems*, Springer-Verlag, London, 1998.
- [9] M. Krstić, I. Kanellakopoulos, and P. V. Kokotović, *Nonlinear and Adaptive Control Design*, John Wiley & Sons, New York, 1995.
- [10] R. Sepulchre, M. Jankovic, P. V. Kokotović, *Constructive Nonlinear Control*, Springer, 1997.
- [11] E. D. Sontag, "Smooth stabilization implies coprime factorization," *IEEE Transactions on Automatic Control* 34, 1989, pp. 435–443.
- [12] E. D. Sontag, "The ISS philosophy as a unifying framework for stability-like behavior," In: *Nonlinear Control in the Year 2000* (Volume 2) (Lecture Notes in Control and Information Sciences, A. Isidori, F. Lamnabhi-Lagarigue, and W. Respondek, eds.), Springer-Verlag, Berlin, 2000, pp. 443–468.

- [13] E. D. Sontag and Y. Wang, "On characterizations of the input-to-state stability property," *Systems & Control Letters* 24, pp. 351–359, 1995.
- [14] E. D. Sontag and Y. Wang, "Detectability of nonlinear systems," In: *Proceedings of the Conference on Information Sciences and Systems (CISS 96)*, Princeton, NJ, 1996, pp. 1031–1036.
- [15] E. D. Sontag and Y. Wang, "Output-to-state stability and detectability of nonlinear systems," *Systems & Control Letters* 29, pp. 279–290, 1997.
- [16] E. D. Sontag and Y. Wang, "Notions of input to output stability," *Systems & Control Letters* 38, pp. 351–359, 1999.
- [17] E. D. Sontag and Y. Wang, "Lyapunov characterizations of input to output stability," *SIAM Journal on Control and Optimization* 39, pp. 226–249, 2001.

PART 7

Controllability, Observability

Problem 7.1

Time for local controllability of a 1-D tank containing a fluid modeled by the shallow water equations

Jean-Michel Coron

Université Paris-Sud

Département de Mathématique

Bâtiment 425

91405 Orsay

France

Jean-Michel.Coron@math.u-psud.fr

1 DESCRIPTION OF THE PROBLEM

We consider a 1-D tank containing an inviscid incompressible irrotational fluid. The tank is subject to one-dimensional horizontal moves. We assume that the horizontal acceleration of the tank is small compared to the gravity constant and that the height of the fluid is small compared to the length of the tank. This motivates the use of the Saint-Venant equations [5] (also called shallow water equations) to describe the motion of the fluid; see, e.g., [2, Sec. 4.2]. After suitable scaling arguments, the length of the tank and the gravity constant can be taken to be equal to 1; see [1]. Then the dynamics equations considered are, see [3] and [1],

$$H_t(t, x) + (Hv)_x(t, x) = 0, \quad (1)$$

$$v_t(t, x) + \left(H + \frac{v^2}{2}\right)_x(t, x) = -u(t), \quad (2)$$

$$v(t, 0) = v(t, 1) = 0, \quad (3)$$

$$\frac{ds}{dt}(t) = u(t), \quad (4)$$

$$\frac{dD}{dt}(t) = s(t), \quad (5)$$

where

- $H(t, x)$ is the height of the fluid at time t and for $x \in [0, 1]$,
- $v(t, x)$ is the horizontal water velocity of the fluid *in a referential attached to the tank* at time t and for $x \in [0, 1]$ (in the shallow water

model, all the points on the same vertical have the same horizontal velocity),

- u is the horizontal acceleration of the tank in the absolute referential,
- s is the horizontal velocity of the tank,
- D is the horizontal displacement of the tank.

This is a control system, denoted Σ , where

- the state is $Y = (H, v, s, D)$,
- the control is $u \in \mathbb{R}$.

Still, by scaling arguments, we may assume that, for every steady state, H , which is then a constant function, is equal to 1; see [1]. One is interested in the local controllability of the control system Σ around the equilibrium point

$$(Y_e, u_e) := ((1, 0, 0, 0), 0).$$

Of course, the total mass of the fluid is conserved so that, for every solution of (1) to (3),

$$\frac{d}{dt} \int_0^1 H(t, x) dx = 0. \quad (6)$$

One gets (6) by integrating (1) on $[0, 1]$ and by using (3 together with an integration by parts.) Moreover, if H and v are of class \mathcal{C}^1 , it follows from (2) and (3) that

$$H_x(t, 0) = H_x(t, 1) \quad (= -u(t)). \quad (7)$$

Therefore we introduce the vector space E of functions $Y = (H, v, s, D) \in \mathcal{C}^1([0, 1]) \times \mathcal{C}^1([0, 1]) \times \mathbb{R} \times \mathbb{R}$ such that

$$H_x(0) = H_x(1), \quad (8)$$

$$v(0) = v(1) = 0, \quad (9)$$

and consider the affine subspace $\mathcal{Y} \subset E$ of $Y = (H, v, s, D) \in E$ satisfying

$$\int_0^1 H(x) dx = 1. \quad (10)$$

With these notations, we can define a trajectory of the control system Σ .

Definition of a trajectory: Let T_1 and T_2 be two real numbers satisfying $T_1 \leq T_2$. A function $(Y, u) = ((H, v, s, D), u) : [T_1, T_2] \rightarrow \mathcal{Y} \times \mathbb{R}$ is a trajectory of the control system Σ if

- (i) the functions H and v are of class \mathcal{C}^1 on $[T_1, T_2] \times [0, 1]$,

(ii) the functions s and D are of class \mathcal{C}^1 on $[T_1, T_2]$ and the function u is continuous on $[0, T]$,

(iii) the equations (1) to (5) hold for every $(t, x) \in [T_1, T_2] \times [0, 1]$.

For $w \in \mathcal{C}^1([0, 1])$, let

$$|w|_1 := \text{Max}\{|w(x)| + |w_x(x)|; x \in [0, 1]\}.$$

We now consider the following property of local controllability of Σ around (Y_e, u_e) .

Definition of $\mathcal{P}(T)$: Let $T > 0$. The control system Σ satisfies the property $\mathcal{P}(T)$ if, for every ϵ , there exists $\eta > 0$ such that, for every $Y_0 = (H_0, v_0, s_0, D_0) \in \mathcal{Y}$, and for every $Y_1 = (H_1, v_1, s_1, D_1) \in \mathcal{Y}$ such that

$$|H_0 - 1|_1 + |v_0|_1 + |H_1 - 1|_1 + |v_1|_1 + |s_0| + |s_1| + |D_0| + |D_1| < \eta,$$

there exists a trajectory

$$(Y, u) : [0, T] \rightarrow \mathcal{Y} \times \mathbb{R}, \quad t \mapsto ((H(t), v(t), s(t), D(t)), u(t))$$

of the control system Σ such that

$$Y(0) = Y_0 \text{ and } Y(T) = Y_1, \quad (11)$$

and, for every $t \in [0, T]$,

$$|H(t) - 1|_1 + |v(t)|_1 + |s(t)| + |D(t)| + |u(t)| < \epsilon. \quad (12)$$

Our open problem is to find for which $T > 0$ $\mathcal{P}(T)$ holds. We conjecture that $\mathcal{P}(T)$ holds if and only if $T > 2$.

2 MOTIVATION AND HISTORY OF THE PROBLEM

The problem of controllability of the system Σ has been raised by F. Dubois, N. Petit, and P. Rouchon in [3]. Let us recall that they have studied in this paper the controllability of the linearized control system around (Y_e, u_e) . This linearized control system is

$$(\Sigma_0) \left\{ \begin{array}{l} h_t + v_x = 0, \\ v_t + h_x = -u(t), \\ v(t, 0) = v(t, 1) = 0, \\ \frac{ds}{dt}(t) = u(t), \\ \frac{dD}{dt}(t) = s(t), \end{array} \right. \quad (13)$$

where the state is $(h, v, s, D) \in \mathcal{Y}_0$, with

$$\mathcal{Y}_0 := \left\{ (h, v, s, D) \in E; \int_0^1 h dx = 0 \right\},$$

and the control is $u \in \mathbb{R}$. It is proved in [3] that Σ_0 is not controllable. It is also proved in [3] that, even if Σ_0 is not controllable, for any

$T > 1$, one can move during the interval of time $[0, T]$ from any steady state $(h_0, v_0, s_0, D_0) := (0, 0, 0, D_0)$ to any steady state $(h_1, v_1, s_1, D_1) := (0, 0, 0, D_1)$ for the linear control system Σ_0 ; see also [4] when the tank has a nonstraight bottom. Unfortunately, this does not imply that the related property (move from $(H_0, v_0, s_0, D_0) := (0, 0, 0, D_0)$ to $(H_1, v_1, s_1, D_1) := (0, 0, 0, D_1)$) also holds for the nonlinear control system Σ , even if $|D_1 - D_0|$ is arbitrary small but not 0. In fact we conjecture that, for $\epsilon > 0$ small enough, even if $|D_1 - D_0|$ is arbitrarily small but not 0, one needs $T > 2$ to move from $(H_0, v_0, s_0, D_0) := (1, 0, 0, D_0)$ to $(H_1, v_1, s_1, D_1) := (1, 0, 0, D_1)$ for the nonlinear control system Σ if one requires (12).

3 AVAILABLE RESULTS

Clearly, $\mathcal{P}(T)$ implies $\mathcal{P}(T')$ for $T \leq T'$. Using the characteristics of the hyperbolic system (1)-(2), one easily sees that $\mathcal{P}(T)$ does not hold $T < 1$. It is proved in [1] that $\mathcal{P}(T)$ holds for T large enough. The method used in [1] requires, at least, $T > 2$.

BIBLIOGRAPHY

- [1] J.-M. Coron, "Local controllability of a 1-D tank containing a fluid modeled by the shallow water equations," reprint University Paris-Sud, 2002, accepted for publication in *ESAIM: COCV*.
- [2] L. Debnath, *Nonlinear Water Waves*, Academic Press, San Diego, 1994.
- [3] F. Dubois, N. Petit, and P. Rouchon, "Motion planning and nonlinear simulations for a tank containing a fluid," *ECC 99*.
- [4] N. Petit and P. Rouchon, "Dynamics and solutions to some control problems for water-tank systems," preprint, CIT-CDS 00-004, 2000, accepted for publication in *IEEE Transactions on Automatic Control*.
- [5] A.J.C.B. de Saint-Venant, "Théorie du mouvement non permanent des eaux, avec applications aux crues des rivières et à l'introduction des marées dans leur lit," *C.R. Acad. Sci. Paris*, 53 pp. 147-154, 1971.

Problem 7.2

A Hautus test for infinite-dimensional systems

Birgit Jacob

Fachbereich Mathematik

Universität Dortmund

D-44221 Dortmund

Germany

`birgit.jacob@math.uni-dortmund.de`

Hans Zwart

Department of Applied Mathematics

University of Twente

P.O. Box 217, 7500 AE Enschede

The Netherlands

`h.j.zwart@math.utwente.nl`

1 DESCRIPTION OF THE PROBLEM

We consider the abstract system

$$\dot{x}(t) = Ax(t), \quad x(0) = x_0, \quad t \geq 0 \quad (1)$$

$$y(t) = Cx(t), \quad t \geq 0, \quad (2)$$

on a Hilbert space H . Here A is the infinitesimal generator of an exponentially stable C_0 -semigroup $(T(t))_{t \geq 0}$ and by the solution of (1) we mean $x(t) = T(t)x_0$, which is the weak solution. If C is a bounded linear operator from H to a second Hilbert space Y , then it is straightforward to see that $y(\cdot)$ in (2) is well-defined and continuous. However, in many PDE's, rewritten in the form (1)-(2), C is only a bounded operator from $D(A)$ to Y ($D(A)$ denotes the domain of A), although the output is a well-defined (locally) square integrable function. In the following, C will always be a bounded operator from $D(A)$ to Y . Note that $D(A)$ is a dense subset of H . If the output is locally square integrable, then C is called an *admissible observation operator*, see Weiss [11]. It is not hard to see that since the C_0 -semigroup is exponentially stable, the output is locally square integrable if and only if it is square integrable. Using the uniform boundedness theorem, we see that the observation operator C is admissible if and only if there

exists a constant $L > 0$ such that

$$\int_0^\infty \|CT(t)x\|_Y^2 dt \leq L\|x\|_H^2, \quad x \in D(A). \quad (3)$$

Assuming that the observation operator C is admissible, system (1)-(2) is said to be *exactly observable* if there is a bounded mapping from the output trajectory to the initial condition, i.e., there exists a constant $l > 0$ such that

$$\int_0^\infty \|CT(t)x\|_Y^2 dt \geq l\|x\|_H^2, \quad x \in D(A). \quad (4)$$

Often the emphasis is on exact observability on a finite interval, which means that the integral in (4) is over $[0, t_0]$ for some $t_0 > 0$. However, for exponentially stable semigroups, both notions are equivalent, i.e., if (4) holds and the system is exponentially stable, then there exists a $t_0 > 0$ such that the system is exactly observable on $[0, t_0]$.

There is a strong need for easy verifiable equivalent conditions for exact observability. Based on the observability conjecture by Russell and Weiss [9] we now conjecture the following:

Conjecture Let A be the infinitesimal generator of an exponentially stable C_0 -semigroup on a Hilbert space H and let C be an admissible observation operator. Then system (1)-(2) is exactly observable if and only if

(C1) $(T(t))_{t \geq 0}$ is similar to a contraction, i.e., there exists a bounded operator S from H to H , which is boundedly invertible such that $(ST(t)S^{-1})_{t \geq 0}$ is a contraction semigroup; and

(C2) there exists a $m > 0$ such that

$$\|(sI - A)x\|_H^2 + |Re(s)|\|Cx\|_Y^2 \geq m|Re(s)|^2\|x\|_H^2 \quad (5)$$

for all complex s with negative real part, and for all $x \in D(A)$.

Our conjecture is a revised version of the (false) conjecture by Russell and Weiss; they did not require that the semigroup is similar to a contraction.

2 MOTIVATION AND HISTORY OF THE CONJECTURE

System (1)-(2) with $A \in \mathbb{C}^{n \times n}$ and $C \in \mathbb{C}^{p \times n}$ is observable if and only if

$$\text{rank} \begin{matrix} sI - A \\ C \end{matrix} = n \quad \text{for all } s \in \mathbb{C}. \quad (6)$$

This is known as the Hautus test, due to Hautus [2] and Popov [8]. If A is a stable matrix, then (6) is equivalent to condition (C2). Although there are some generalizations of the Hautus test to delay differential equations

(see, e.g., Klamka [6] and the references therein) the full generalization of the Hautus test to infinite-dimensional linear systems is still an open problem.

It is not hard to see that if (1)-(2) is exactly observable, then the semigroup is similar to a contraction, see Grabowski and Callier [1] and Levan [7].

Condition (C2) was found by Russell and Weiss [9]: they showed that this condition is necessary for exact observability, and, under the extra assumption that A is bounded, they showed that this condition also is sufficient.

Without the explicit usage of condition (C1) it was shown that condition (C2) implies exact observability if

- A has a Riesz basis of eigenfunctions, $\operatorname{Re}(\lambda_n) = -\rho_1$, $|\lambda_{n+1} - \lambda_n| > \rho_2$, where λ_n are the eigenvalues of A , and $\rho_1, \rho_2 > 0$, [9]; or if
- m in equation (5) is one, [1]; or if
- A is skew-adjoint and C is bounded, Zhou and Yamamoto [12]; or if
- A has a Riesz basis of eigenfunctions, and $Y = \mathbb{C}^p$, Jacob and Zwart [5].

Recently, we showed that (C2) is not sufficient in general, [4]. The C_0 -semigroup in our counterexample is an analytic semigroup, which is not similar to a contraction semigroup. The output space in the example is just the complex plane \mathbb{C} .

3 AVAILABLE RESULTS AND CLOSING REMARKS

In order to prove the conjecture it is sufficient to show that for exponentially stable contraction semigroups condition (C2) implies exact observability.

It is well-known that system (1)-(2) is exactly observable if and only if there exists a bounded operator L that is positive and boundedly invertible and satisfies the Lyapunov equation

$$\langle Ax_1, Lx_2 \rangle_H + \langle Lx_1, Ax_2 \rangle_H = \langle Cx_1, Cx_2 \rangle_Y, \quad \text{for all } x_1, x_2 \in D(A). \quad (7)$$

From the admissibility of C and the exponential stability of the semigroup, one easily obtains that equation (7) has a unique (non-negative) solution. Russell and Weiss [9] showed that Condition (C2) implies that this solution has zero kernel. Thus the Lyapunov equation (2) could be a starting point for a proof of the conjecture.

We have stated our conjecture for infinite-dimensional output spaces. However, it could be that it only holds for finite-dimensional output spaces.

If the output space Y is one-dimensional one could try to prove the conjecture using powerful tools like the Sz.-Nagy-Foias model theorem (see [10]). This tool was quite useful in the context of admissibility conditions for contraction

semigroups [3]. Based on this observation, it would be of great interest to check our conjecture for the right shift semigroup on $L_2(0, \infty)$.

We believe that exponential stability is not essential in our conjecture, and can be replaced by strong stability and infinite-time admissibility, see [5].

Note that our conjecture is also related to the left-invertibility of semigroups, see [1] and [4] for more details.

BIBLIOGRAPHY

- [1] P. Grabowski and F. M. Callier, "Admissible observation operators, semigroup criteria of admissibility," *Integral Equation and Operator Theory*, 25:182–198, 1996.
- [2] M. L. J. Hautus, "Controllability and observability conditions for linear autonomous systems," *Ned. Akad. Wetenschappen, Proc. Ser. A*, 72:443–448, 1969.
- [3] B. Jacob and J. R. Partington, "The Weiss conjecture on admissibility of observation operators for contraction semigroups," *Integral Equations and Operator Theory*, 40(2):231–243, 2001.
- [4] B. Jacob and H. Zwart, "Disproof of two conjectures of George Weiss," *Memorandum 1546*, Faculty of Mathematical Sciences, University of Twente, 2000. Available at <http://www.math.utwente.nl/publications/>
- [5] B. Jacob and H. Zwart, "Exact observability of diagonal systems with a finite-dimensional output operator," *Systems & Control Letters*, 43(2):101–109, 2001.
- [6] J. Klamka, *Controllability of Dynamical Systems*, Kluwer Academic Publisher, 1991.
- [7] N. Levan, "The left-shift semigroup approach to stability of distributed systems," *J. Math. Ana. Appl.*, 152: 354–367, 1990.
- [8] V. M. Popov, *Hyperstability of Control Systems*. Editura Academiei, Bucharest, 1966 (in Romanian); English trans. by Springer Verlag, Berlin, 1973.
- [9] D. L. Russell and G. Weiss, "A general necessary condition for exact observability," *SIAM J. Control Optim.*, 32(1):1–23, 1994.
- [10] B. Sz.-Nagy and C. Foias, *Harmonic analysis of operators on Hilbert spaces*, Amsterdam-London: North-Holland Publishing Company. XIII, 1970.

- [11] G. Weiss, "Admissible observation operators for linear semigroups," *Israel Journal of Mathematics*, 65:17–43, 1989.
- [12] Q. Zhou and M. Yamamoto, "Hautus condition on the exact controllability of conservative systems," *Int. J. Control*, 67(3):371–379, 1997.

Problem 7.3

Three problems in the field of observability

Philippe Jouan

Laboratoire R. Salem

CNRS UMR 6085

Université de Rouen

Mathématiques, site Colbert

76821 Mont-Saint-Aignan Cedex

France

Philippe.Jouan@univ-rouen.fr

1 INTRODUCTION.

Let X be a \mathcal{C}^∞ (resp. \mathcal{C}^ω), connected manifold. We consider on X the system

$$\Sigma = \begin{cases} \dot{x} = f(x, u) \\ y = h(x) \end{cases} \quad (1)$$

where $x \in X$, $u \in U = [0, 1]^m$, and $y \in \mathbb{R}^p$. The parametrized vector field f and the output function h are assumed to be \mathcal{C}^∞ (resp. \mathcal{C}^ω). In order to avoid certain complications, the state space X is assumed to be compact, but this assumption is not crucial (we can for instance assume that the vector field f vanishes out of a relatively compact open subset of X).

The three problems addressed herein concern observability and the existence of observers for such systems.

2 PROBLEM 1.

We first consider an uncontrolled system:

$$\Sigma_u = \begin{cases} \dot{x} = f(x) \\ y = h(x). \end{cases} \quad (2)$$

This system is assumed to be observable (in the following sense: the trajectories starting from two different initial states are distinguished by the output).

Whenever the n^{th} derivative of the output with respect to the vector field f is a \mathcal{C}^r -function of the output and the $n - 1$ previous ones it is possible to construct observers (see [5], [9]). More accurately, the injective mapping

$$\Phi = (h, L_f h, L_f^2 h, \dots, L_f^{n-1} h)$$

is used to “immerse” Σ_u into \mathbb{R}^{pn} where a \mathcal{C}^r -observer is designed. The observed data are the outputs of Σ_u together with their $(n - 1)^{th}$ first derivatives. The state of the system, being a continuous mapping of $(h, L_f h, L_f^2 h, \dots, L_f^{n-1} h)$, is thus estimated by the observer.

More generally a \mathcal{C}^r -observer for Σ_u is a system $\hat{\Sigma}$ defined in an open subset V of \mathbb{R}^n by

$$\hat{\Sigma} = \begin{cases} \dot{z} = F(z, y) \\ \hat{x} = \theta(z) \end{cases} \quad (3)$$

where F is a \mathcal{C}^r -vector field on V and θ is a continuous mapping from V into X such that

$$\forall x \in X, \quad \forall z \in V \quad \lim_{t \rightarrow +\infty} d(x(t), \hat{x}(t)) = 0$$

for any distance d on X compatible with the topology of X .

The first problem is:

Does the existence of a \mathcal{C}^r -observer for Σ_u imply the existence of an integer n such that the n^{th} derivative of the output is a \mathcal{C}^r -function of the output and the $n - 1$ previous ones? In an equivalent way does it exist a \mathcal{C}^r -function φ such that

$$L_f^n h = \varphi(h, L_f h, L_f^2 h, \dots, L_f^{n-1} h) ?$$

A positive answer to this question would imply that all the observability properties of an uncontrolled system are contained in the functional relation

$$L_f^n h = \varphi(h, L_f h, L_f^2 h, \dots, L_f^{n-1} h).$$

Notice that we already know that the kind of dependence between the n^{th} derivative of the output and the preceding ones, that is the kind of function φ , determines whether the system is linearizable, or linearizable modulo an output injection (see [2], [7]).

3 PROBLEM 2.

Once we know that a *controlled* system is observable in the weak sense of [4] (two different initial states have to be distinguished by the output for at least one input) a question arises naturally: which inputs are universal? (An input is universal if any two different initial states are distinguished by the output for *this* input, see [8].) An equivalent formulation is: for which inputs is the system observable?

For generic reasons, we consider controlled systems with more outputs than inputs: $p > m$.

Problem 2 is:

Is it true that the set of \mathcal{C}^∞ -systems (resp. \mathcal{C}^ω -systems) that are observable for every \mathcal{C}^∞ input contains an open (or better an open and dense) subset of the set of \mathcal{C}^∞ -systems (resp. the set of \mathcal{C}^ω -systems)? Both in the \mathcal{C}^∞ and \mathcal{C}^ω cases does observability for every \mathcal{C}^∞ -input imply observability for every L^∞ -input?

The following facts are known:

- i. The set of systems observable for every \mathcal{C}^∞ -input is dense in the set of \mathcal{C}^∞ or \mathcal{C}^ω -systems (see [3], [1]).
- ii. For a given bound, the set of systems observable for every \mathcal{C}^∞ -input whose $2\dim(X)$ first derivatives are bounded contains an open and dense subset of the set of \mathcal{C}^∞ or \mathcal{C}^ω -systems (see [3], [1]).
- iii. A \mathcal{C}^ω -system observable for every \mathcal{C}^∞ -input is observable for every L^∞ -input (see [3]).
- iv. In the single-input, control-affine, \mathcal{C}^∞ -case, the implication

$$\Sigma \text{ } \mathcal{C}^\infty\text{-observable} \implies \Sigma \text{ } L^\infty\text{-observable}$$
 is true for an open and dense subset of systems (see [6]).

Of course, a positive answer to this problem would mean that the property of being observable for every L^∞ -input is preserved under slight perturbations.

4 PROBLEM 3.

Since the set of systems observable for every L^∞ -input is residual (with more outputs than inputs), it is very interesting to design observers for them, particularly if this set contains an open subset.

At the present time, the more general construction of observers for nonlinear systems is the high gain one (see [3]). But the observers designed in this way have the default to make use of the derivatives of the input and cannot work if this last is only L^∞ . In some particular cases (linearizable systems, linearizable modulo an output injection systems, bilinear systems, uniformly observable systems ...) observers that works for every input are known but they cannot be generalized.

Problem 3 is therefore:

For systems observable for every L^∞ -input, find a general construction of an observer which works for every L^∞ -input.

Notice that if the system is “immersed” in \mathbb{R}^N (in a sense to make precise) the “immersion” must not depend on the input: in that case the image of

the vector field f in \mathbb{R}^N would depend upon the derivative of the input. In particular the mapping

$$\Phi = (h, L_f h, L_f^2 h, \dots, L_f^{n-1} h)$$

from X into \mathbb{R}^{pn} cannot be used because the vector field $f(x, u)$ depends upon u and so are $L_f h(x, u)$, $L_f^2 h(x, u)$, \dots , $L_f^{n-1} h(x, u)$.

BIBLIOGRAPHY

- [1] M. Balde and Ph. Jouan, "Genericity of observability of control-affine systems," *Control Optimization and Calculus of Variations*, vol. 3, 1998, 345-359.
- [2] M. Fliess and I. Kupka, "A finiteness criterion for nonlinear input-output differential systems," *SIAM J. on Control and Optimization*, 21 (1983), 5, 721-728.
- [3] J. P. Gauthier and I. Kupka, *Deterministic Observation Theory and Applications*, Cambridge University Press, 2001.
- [4] R. Hermann and A. J. Krener, "Nonlinear controllability and observability," *IEEE Trans. Aut. Contro*, AC-22 (1977), 728-740.
- [5] Ph. Jouan and J. P. Gauthier, "Finite singularities of nonlinear systems. Output stabilization, observability and observers," *Journal of Dynamical and Control Systems*, vol.2, no 2, 1996, 255-288.
- [6] Ph. Jouan, " C^∞ and L^∞ observability of single-input C^∞ -Systems," *Journal of Dynamical and Control Systems*, vol. 7, no 2, 2001, 151-169.
- [7] Ph. Jouan, "Immersion of nonlinear systems into linear systems modulo output injection," *submitted to SIAM J. on Control and Optimization*.
- [8] H. J. Sussmann, "Single-input observability of continuous-time systems," *Math. Syst. Theory* 12 (1979), 371-393.
- [9] X. Xia and M. Zeitz, "On nonlinear continuous observers," *Int. J. Control*, vol. 66, no 6, 1997, 943-954.

Problem 7.4

Control of the KdV equation

Lionel Rosier

Institut Elie Cartan

Université Henri Poincaré Nancy 1

B.P. 239

54506 Vandœuvre-lès-Nancy Cedex

France

rosier@iecn.u-nancy.fr

1 DESCRIPTION OF THE PROBLEM

The Korteweg-de Vries (KdV) equation is the simplest model for unidirectional propagation of small amplitude long waves in nonlinear dispersive systems. It occurs in various physical contexts (e.g., water waves, plasma physics, nonlinear optics). It reads

$$y_t + y_{xxx} + y_x + y y_x = 0, \quad t > 0, \quad x \in \Omega, \quad (1)$$

the subscripts denoting partial derivatives (e.g., $y_t = \frac{\partial y}{\partial t}$). The KdV equation has been intensively studied since the 1960s because of its fascinating properties (infinite set of conserved integral quantities, integrability, Kato smoothing effect, etc.). (See [5] and the references therein.)

Here, we are concerned with the boundary controllability of the KdV equation in the domain $\Omega = (0, +\infty)$. For any pair (a, b) with $0 \leq a < b \leq +\infty$ let $C_0^\infty(a, b)$ denote the space of functions of class C^∞ and with compact support in (a, b) . Given $T > 0$, $y_0 \in C_0^\infty(0, +\infty)$ and $h \in C_0^\infty(0, T)$, it is by now well known (see [1]) that the initial-boundary-value problem

$$\begin{cases} y_t + y_{xxx} + y_x + y y_x = 0, & 0 < t < T, \quad 0 < x < +\infty \\ y(t, 0) = h(t), & 0 < t < T \\ y(0, x) = y_0(x), & 0 < x < +\infty \end{cases} \quad (2)$$

admits a unique classical solution that is smooth. The boundary value h is the *input* of the system. Let $\mathcal{R}(y_0, T)$ denote the space of all *reachable* states from y_0 in time T ; that is,

$$\mathcal{R}(y_0, T) = \{y(T, \cdot); y \text{ fulfills (2) for some } h \in C_0^\infty(0, T)\}. \quad (3)$$

We are now in a position to state the problem of interest.

Open Problem: *Is it true that $0 \in \mathcal{R}(y_0, T)$ for T is large enough ?*

The main difficulty of the problem is that the domain is *unbounded*. Notice that it would also be of great interest to identify the closure of $\mathcal{R}(y_0, T)$ in $L^2(0, +\infty)$.

2 MOTIVATION AND HISTORY OF THE PROBLEM

The KdV equation has been first introduced in [4] to explain the emergence of long solitary waves, the so-called “solitons.” In this context, t stands for the elapsed time, x is the independent space variable, and $y = y(t, x)$ stands for the deviation of the fluid surface from the rest position. The above control problem may serve as a model for the control of the fluid surface in a shallow canal by means of a wavemaker. Indeed, taking Lagrangian coordinates, it is proved in [10] that the movement of the fluid surface is governed by (2), the speed of the moving boundary being roughly represented by the input h . Thus, the space $\mathcal{R}(0, T)$ stands for the set of waves that may be generated (from the rest position) by the wavemaker in time T .

A similar control problem is investigated in [7], but with a fluid model in which both the dispersive and nonlinear effects are neglected. In [2] the author uses the (nonlinear) shallow water equations as a fluid model to investigate the control of the fluid surface in a moving tank. These equations are appropriate in situations where the dispersive effects may be neglected, e.g., when the height of the fluid and the length of the tank are of the same order of magnitude. The shallow water equations have to be replaced by the KdV equation (or the Boussinesq system) when studying the propagation of *traveling waves*.

The above problem is important for the following reason. A lack of compactness, due to the fact that the domain is unbounded, prevents us from using the standard linearization procedure in the study of the controllability properties of (2). Therefore, a new approach (based on the inverse scattering?) has to be developed to investigate the (exact or approximate) controllability of the nonlinear KdV equation on the half line.

3 AVAILABLE RESULTS

The boundary controllability of the KdV equation has been investigated in numerous papers; see, e.g., [3], [8], [11] and [12]. In these papers, the domain $\Omega = (0, L)$ is bounded and the control is applied at the *right* endpoint, although the waves are expected to move from the left to the right. If the

control is applied at the left endpoint, and if the system is supplemented by the boundary conditions $y = y_x = 0$ at $x = L$, then it is proved in [10] that for any $T, L > 0$, $0 \in \mathcal{R}(y_0, T)$ for any initial state y_0 with a small enough $H^3(0, L)$ -norm. It means that a soliton moving to the right may be caught up and annihilated by a set of waves generated by the wave-maker. This result rests on a Carleman estimate for the linearized equation (i.e., (1) without the nonlinear term yy_x). When we look at the linearized equation on the unbounded domain $\Omega = (0, +\infty)$, then the controllability results are not so good, due to a lack of compactness. Indeed, it is proved in [9] that there exists a state $y_0 \in L^2(0, +\infty)$ for which any trajectory connecting y_0 to the null state does not belong to $L^\infty(0, T, L^2(0, +\infty))$ (that is, $\text{ess sup}_{0 < t < T} \int_0^{+\infty} y(t, x)^2 dx = +\infty$). It means that the *bad* behavior of the trajectory $y(t, x)$ as $x \rightarrow +\infty$ is the price to be paid to get the null-controllability. (A similar phenomenon has been observed in [6] for the heat equation.) Thus, the linearization procedure fails since we do not have any bound on $\|y(t, \cdot)\|_{L^2(0, +\infty)}$.

BIBLIOGRAPHY

- [1] J. Bona, R. Winther, "The Korteweg-de Vries equation, posed in a quarter-plane," *SIAM J. Math. Anal.*, 14, pp. 1056-1106, 1983.
- [2] J.-M. Coron, "Local controllability of a 1-D tank containing a fluid modeled by the shallow water equations," *A tribute to J. L. Lions, ESAIM Control Optim. Calc. Var.*, 8, pp. 513-554, 2002 (electronic).
- [3] E. Crépeau, "Exact boundary controllability of the Korteweg-de Vries equation around a non-trivial stationary solution," *Internat. J. Control* 74, pp. 1096-1106, 2001.
- [4] D. J. Korteweg and G. de Vries, "On the change of form of long waves advancing in a rectangular canal, and on a new type of long stationary waves," *Philos. Mag.*, 39, pp. 422-443, 1995.
- [5] C. E. Kenig, G. Ponce, and L. Vega, "The Cauchy problem for the Korteweg-de Vries equation in Sobolev spaces of negatives indices," *Duke math J.*, 71, pp. 1-21, 1993.
- [6] S. Micu, E. Zuazua, "On the lack of null-controllability of the heat equation on the half space," *Portugal. Math.*, 58, pp. 1-24, 2001.
- [7] S. Mottelet, "Controllability and stabilization of a canal with wave generators," *SIAM J. Control Optim.*, 38, pp. 711-735, 2000.
- [8] L. Rosier, "Exact boundary controllability for the Korteweg-de Vries equation on a bounded domain," *ESAIM Control Optim. Calc. Var.*, 2, pp. 33-55, 1997.

- [9] L. Rosier, "Exact boundary controllability for the linear Korteweg-de Vries equation on the half-line," *SIAM J. Control Optim.*; 39, pp. 331-351, 2000.
- [10] L. Rosier, "Control of the surface of a fluid by a wavemaker," in preparation.
- [11] D. L. Russell, B. Y. Zhang, "Exact controllability and stabilizability of the Korteweg-de Vries equation," *Trans. Amer. Math. Soc.*, 348, pp. 3643-3672, 1996.
- [12] B. Y. Zhang, "Exact boundary controllability of the Korteweg-de Vries equation," *SIAM J. Control Optim.*, 37, pp. 543-565, 1999.

PART 8

Robustness, Robust Control

Problem 8.1

H_∞ -norm approximation

A. C. Antoulas

Department of Electrical and Computer Engineering
Rice University
6100 South Main Street
Houston, TX 77005-1892
USA
aca@rice.edu

A. Astolfi

Department of Electrical and Electronic Engineering
Imperial College
Exhibition Road
SW7 2BT, London
United Kingdom
a.astolfi@imperial.ac.uk

1 DESCRIPTION OF THE PROBLEM

Let \mathcal{RH}_∞^m be the (Hardy) space of real-rational scalar¹ transfer functions of order m , bounded on the imaginary axis and analytic into the right-half complex plane. The optimal approximation problem in the H_∞ norm can be stated as follows.

(A^{*}) (Optimal Approximation in the H_∞ norm)

Given $G(s) \in \mathcal{RH}_\infty^N$ and an integer $n < N$ find,² if possible, $A^*(s) \in \mathcal{RH}_\infty^n$ such that

$$A^*(s) = \arg \min_{A(s) \in \mathcal{RH}_\infty^n} \|G(s) - A(s)\|_\infty. \quad (1)$$

For such a problem, let

$$\gamma_n^* = \min_{A(s) \in \mathcal{RH}_\infty^n} \|G(s) - A(s)\|_\infty,$$

then two further problems can be posed.

¹Similar considerations can be done for the nonscalar case.

²By *find* we mean *find an exact solution* or an algorithm converging to the exact solution.

- (D) (Optimal Distance problem in the H_∞ norm)
 Given $G(s) \in \mathcal{RH}_\infty^N$ and an integer $n < N$ find γ_n^* .
- (A) (Sub-optimal Approximation in the H_∞ norm)
 Given $G(s) \in \mathcal{RH}_\infty^N$, an integer $n < N$ and $\gamma > \gamma_n^*$ find $\tilde{A}(s) \in \mathcal{RH}_\infty^n$ such that

$$\gamma_n^* \leq \|G(s) - \tilde{A}(s)\|_\infty \leq \gamma.$$

The optimal H_∞ approximation problem can be formally posed as a constrained min-max problem. For, note that any function in \mathcal{RH}_∞^n can be put in a one-to-one correspondence with a point θ of some (open) set $\Omega \subset R^{2n}$, therefore the problem of computing γ_n^* can be posed as

$$\gamma_n^* = \min_{\theta \in \Omega} \max_{\omega \in R} \|G(j\omega) - A(j\omega, \theta)\|, \quad (2)$$

where $A(s) = A(s, \theta)$. The above formulation provides a brute force approach to the solution of the problem. Unfortunately, this method is not of any use in general, because of the complexity of the set Ω and because of the curse of dimensionality. However, the formulation (2) suggests that possible candidate solutions of the optimal approximation problem are the saddle points of the function

$$\|G(j\omega) - A(j\omega, \theta)\|,$$

which can be, in principle, computed using numerical tools. It would be interesting to prove (or disprove) that

$$\min_{\theta \in \Omega} \max_{\omega \in R} \|G(j\omega) - A(j\omega, \theta)\| = \max_{\omega \in R} \min_{\theta \in \Omega} \|G(j\omega) - A(j\omega, \theta)\|.$$

The solution method based on the computation of saddle points does not give any insight into the problem, neither exposes any systems theoretic interpretation of the optimal approximant. An interesting property of the optimal approximant is stated in the following simple fact, which can be used to rule out that a candidate approximant is optimal.

Fact: Let $A^*(s) \in \mathcal{RH}_\infty^n$ be such that equation (1) holds. Suppose

$$|W(j\omega^*) - A^*(j\omega^*)| = \gamma_n^*, \quad (3)$$

and

$$A(j\omega^*) \neq 0 \quad (4)$$

for $\omega^* = 0$. Then there exists a constant $\tilde{\omega} \neq \omega^*$ such that

$$|W(j\tilde{\omega}) - A^*(j\tilde{\omega})| = \gamma_n^*,$$

i.e., if the value γ_n^* is attained by the function $|W(j\omega) - A^*(j\omega)|$ at $\omega = 0$ it is also attained at some $\omega \neq 0$.

Proof: We prove the statement by contradiction. Suppose

$$|W(j\omega) - A^*(j\omega)| < \gamma_n^*, \quad (5)$$

for all $\omega \neq \omega^*$ and consider the approximant $\tilde{A}(s) = (1 + \lambda)A^*(s)$, with $\lambda \in \mathbb{R}$. By equation (5), condition (4) and by continuity with respect to λ and ω of

$$|W(j\omega) - \tilde{A}(j\omega)|,$$

there is a λ^* (sufficiently small) such that

$$\max_{\omega} |W(j\omega) - (1 + \lambda^*)A^*(j\omega)| < \gamma_n^*,$$

or, what is the same, it is possible to obtain an approximant that is better than $A^*(s)$, hence a contradiction. \triangleleft

It would be interesting to show that the above fact holds (or it does not hold) when $\omega^* \neq 0$.

2 AVAILABLE RESULTS AND POSSIBLE SOLUTION PATHS

Approximation and model reduction have always been central issues in system theory. For a recent survey on model reduction in the large-scale setting, we refer the reader to [1]. There are several results in this area. If the approximation is performed in the Hankel norm, then an explicit solution of the optimal approximation and model reduction problems has been given in [3]. Note that this procedure provides, as a byproduct, an upper bound for γ_n^* and a solution of the suboptimal approximation problem. If the approximation is performed in the H_2 norm, several results and numerical algorithms are available [4]. For approximation in the H_∞ norm a conceptual solution is given in [5]. Therein it is shown that the H_∞ approximation problem can be reduced to a Hankel norm approximation problem for an extended system (i.e., a system obtained from a state space realization of the original transfer function $G(s)$ by adding inputs and outputs). The extended system has to be constructed with the constraint that the corresponding Grammians P and Q satisfy

$$\lambda_{\min}(PQ) = (\gamma_n^*)^2 \quad \text{with multiplicity } N - n. \quad (6)$$

However, the above procedure, as also noted by the authors of [5], is not computationally viable, and presupposes the knowledge of γ_n^* . Hence the need for further study of the problem. In the recent paper [2], the decay rates of the Hankel singular values of stable, single-input single-output systems, are studied. Let $G(s) = \frac{p(s)}{q(s)}$ be the transfer function under consideration. The decay rate of the Hankel singular values is studied by introducing a new set of input/output system invariants, namely the quantities $\frac{p(s)}{q^*(s)}$, where $q(s)^* = q(-s)$, evaluated at the poles of $G(s)$. These results are expected to yield light into the structure of the above problem (6). Another paper of interest especially for the suboptimal approximation case, is [6]. In this paper the set, of all systems whose H_∞ norm is less than some positive number γ is parameterized. Thus the following problem can be posed: given

such a system with H_∞ norm less than γ , find conditions under which it can be decomposed in the sum of two systems, one of which is prespecified. Finally, there are two special classes of systems that may be studied to improve our insight into the general problem. The first class is composed of single-input single-output discrete-time stable systems. For such systems, an interesting related problem is the *Carathéodory-Fejér (CF)* approximation problem that is used for elliptic filters design. In [7] it is shown that in the scalar, discrete-time case, optimal approximants in the Hankel norm approach asymptotically optimal approximants in the H_∞ norm (the asymptotic behavior being with respect to $\epsilon \rightarrow 0$, where $|z| \leq \epsilon < 1$). The CF problem through the contribution of Adamjan-Arov-Krein and later Glover, evolved into what is nowadays called the *Hankel-norm approximation* problem. However, no asymptotic results have been shown to hold in the general case. The second special class is that of symmetric systems, that is, systems whose state space representation (C, A, B) satisfies $A = A'$ and $B = C'$. For instance, these systems have a positive definite Hankel operator and have further properties that can be exploited in the construction of approximants in the H_∞ sense.

BIBLIOGRAPHY

- [1] A. C. Antoulas, "Lectures on the approximation of large scale dynamical systems," *SIAM, Philadelphia*, 2002.
- [2] A. C. Antoulas, D. C. Sorensen, and Y. Zhou, "On the decay rate of Hankel singular values and related issues," *Systems and Control Letters*, 2002.
- [3] K. Glover, "All optimal Hankel-norm approximations of linear multivariable systems and their L^∞ error bounds," *International Journal of Control*, 39: 1115-1193, 1984.
- [4] X.-X. Huang, W.-Y. Yan and K. L. Teo, " H_2 near optimal model reduction," *IEEE Trans. Automatic Control*, 46: 1279-1285, 2001.
- [5] D. Kavranoğlu and M. Bettayeb, "Characterization of the solution to the optimal H_∞ model reduction problem," *Systems and Control Letters*, 20: 99-108, 1993.
- [6] H. G. Sage and M. F. de Mathelin, "Canonical H_∞ state space parametrization," *Automatica*, July 2000.
- [7] L. N. Trefethen, "Rational Chebyshev approximation on the unit disc," *Numerische Mathematik*, 37: 297-320, 1981.

Problem 8.2

Non-iterative computation of optimal value in H_∞ control

Ben M. Chen

Department of Electrical and Computer Engineering
National University of Singapore
Singapore 117576
Republic of Singapore
bmchen@nus.edu.sg

1 DESCRIPTION OF THE PROBLEM

We consider an n -th order generalized linear system Σ characterized by the following state-space equations:

$$\Sigma : \begin{cases} \dot{x} = A x + B u + E w \\ y = C_1 x + D_{11} u + D_1 w \\ h = C_2 x + D_2 u + D_{22} w \end{cases} \quad (1)$$

where x is the state, u is the control input, w is the disturbance input, y is the measurement output, and h is the controlled output of Σ . For simplicity, we assume that $D_{11} = 0$ and $D_{22} = 0$. We also let Σ_p be the subsystem characterized by the matrix quadruple (A, B, C_2, D_2) and Σ_q be the subsystem characterized by (A, E, C_1, D_1) .

The standard H_∞ optimal control problem is to find an internally stabilizing proper measurement feedback control law,

$$\Sigma_{\text{cmp}} : \begin{cases} \dot{v} = A_{\text{cmp}} v + B_{\text{cmp}} y \\ u = C_{\text{cmp}} v + D_{\text{cmp}} y \end{cases} \quad (2)$$

such that when it is applied to the given plant (1), the H_∞ -norm of the resulting closed-loop transfer matrix function from w to h , say $T_{hw}(s)$, is minimized. We note that the H_∞ -norm of an asymptotically stable and proper continuous-time transfer matrix $T_{hw}(s)$ is defined as

$$\|T_{hw}\|_\infty := \sup_{\omega \in [0, \infty)} \sigma_{\max}[T_{hw}(j\omega)] = \sup_{\|w\|_2=1} \frac{\|h\|_2}{\|w\|_2},$$

where w and h are, respectively, the input and output of $T_{hw}(s)$.

The infimum or the optimal value associated with the H_∞ control problem is defined as

$$\gamma^* := \inf \left\{ \|T_{hw}(\Sigma \times \Sigma_{\text{cmp}})\|_\infty \mid \Sigma_{\text{cmp}} \text{ internally stabilizes } \Sigma \right\}. \quad (3)$$

Obviously, $\gamma^* \geq 0$. In fact, when $\gamma^* = 0$, the problem is reduced to the well-known problem of H_∞ almost disturbance decoupling with measurement feedback and internal stability.

We note that in order to design a meaningful H_∞ control law for the given system (1), the designer should know before hand the infimum γ^* , which represents the best achievable level of disturbance attenuation. Unfortunately, the problem of a noniterative computation of this γ^* for general systems still remains unsolved in the open literature.

2 MOTIVATION AND HISTORY OF THE PROBLEM

Over the last two decades, we have witnessed a proliferation of literature on H_∞ optimal control since it was first introduced by Zames [20]. The main focus of the work has been on the formulation of the problem for robust multivariable control and its solution. Since the original formulation of the H_∞ problem in Zames [20], a great deal of work has been done on finding the solution to this problem. Practically all the research results of the early years involved a mixture of time-domain and frequency-domain techniques including the following: 1) *interpolation approach* (see, e.g., [13]); 2) *frequency domain approach* (see, e.g., [5, 8, 9]); 3) *polynomial approach* (see, e.g., [12]); and 4) *J-spectral factorization approach* (see, e.g., [11]). Recently, considerable attention has been focused on purely *time-domain methods* based on algebraic Riccati equations (ARE) (see, e.g., [6, 7, 10, 15, 16, 17, 18, 19, 21]). Along this line of research, connections are also made between H_∞ optimal control and differential games (see, e.g., [1, 14]).

It is noted that most of the results mentioned above are focusing on finding solutions to H_∞ control problems. Many of them assume that γ^* is known or simply assume that $\gamma^* = 1$. The computation of γ^* in the literature are usually done by certain iteration schemes. For example, in the regular case and utilizing the results of Doyle et al. [7], an iterative procedure for approximating γ^* would proceed as follows: one starts with a value of γ and determines whether $\gamma > \gamma^*$ by solving two “indefinite” algebraic Riccati equations and checking the positive semi-definiteness and stabilizing properties of these solutions. In the case when such positive semi-definite solutions exist and satisfy a *coupling condition*, then we have $\gamma > \gamma^*$ and one simply repeats the above steps using a smaller value of γ . In principle, one can approximate the infimum γ^* to within any degree of accuracy in this manner. However, this search procedure is exhaustive and can be very costly. More significantly, due to the possible high-gain occurrence as γ gets close to γ^* , numerical solutions for these H_∞ AREs can become highly sensitive and

ill-conditioned. This difficulty also arises in the *coupling condition*. Namely, as γ decreases, evaluation of the *coupling condition* would generally involve finding eigenvalues of stiff matrices. These numerical difficulties are likely to be more severe for problems associated with the singular case. Thus, in general, the iterative procedure for the computation of γ^* based on AREs is not reliable.

3 AVAILABLE RESULTS

There are quite a few researchers who have attempted to develop procedures for the determination of the value of γ^* without iterations. For example, Petersen [15] has solved the problem for a class of one-block regular case. Scherer [17, 18] has obtained a partial answer for state feedback problem for a larger class of systems by providing a computable candidate value together with algebraically verifiable conditions, and Chen and his co-workers [3, 4] (see also [2]) have developed a noniterative procedures for computing the value of γ^* for a class of systems (singular case) that satisfy certain geometric conditions.

To be more specific, we introduce the following two geometric subspaces of linear systems: Given an n -th order linear system Σ_* characterized by a matrix quadruple (A_*, B_*, C_*, D_*) , we define

- i. $\mathcal{V}^-(\Sigma_*)$, a weakly unobservable subspace, is the maximal subspace of \mathbb{R}^n which is $(A_* + B_*F_*)$ -invariant and contained in $\text{Ker}(C_* + D_*F_*)$ such that the eigenvalues of $(A_* + B_*F_*)|_{\mathcal{V}^-}$ are contained in \mathbb{C}^- , the open-left complex plane, for some constant matrix F_* ; and
- ii. $\mathcal{S}^-(\Sigma_*)$, a strongly controllable subspace, is the minimal $(A_* + K_*C_*)$ -invariant subspace of \mathbb{R}^n containing $\text{Im}(B_* + K_*D_*)$ such that the eigenvalues of the map which is induced by $(A_* + K_*C_*)$ on the factor space $\mathbb{R}^n/\mathcal{S}^-$ are contained in \mathbb{C}^- for some constant matrix K_* .

The problem of noniterative computation of γ^* has been solved by Chen and his co-workers [3, 4] (see also [2]) for a class of systems that satisfy the following conditions:

- i. $\text{Im}(E) \subset \mathcal{V}^-(\Sigma_P) + \mathcal{S}^-(\Sigma_P)$; and
- ii. $\text{Ker}(C_2) \supset \mathcal{V}^-(\Sigma_Q) \cap \mathcal{S}^-(\Sigma_Q)$,

together with some other minor assumptions. The work of Chen et al. involves solving a couple of algebraic Riccati and Lyapunov equations. The computation of γ^* is then done by finding the maximum eigenvalue of a resulting constant matrix.

It has been demonstrated by an example in Chen [2] that the noniterative computation of γ^* can be done for a larger class of systems, which do not

necessarily satisfy the above geometric conditions. It is believed that there are rooms to improve the existing results.

BIBLIOGRAPHY

- [1] T. Başar and P. Bernhard, *H_∞ Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach*, 2nd Ed., Birkhäuser, Boston, 1995.
- [2] B. M. Chen, *H_∞ Control and Its Applications*, Springer, London, 1998.
- [3] B. M. Chen, Y. Guo and Z. L. Lin, "Noniterative computation of infimum in discrete-time H_∞ -optimization and solvability conditions for the discrete-time disturbance decoupling problem," *International Journal of Control*, vol. 65, pp. 433-454, 1996.
- [4] B. M. Chen, A. Saberi, and U. Ly, "Exact computation of the infimum in H_∞ -optimization via output feedback," *IEEE Transactions on Automatic Control*, vol. 37, pp. 70-78, 1992.
- [5] J. C. Doyle, *Lecture Notes in Advances in Multivariable Control*, ONR-Honeywell Workshop, 1984.
- [6] J. C. Doyle and K. Glover, "State-space formulae for all stabilizing controllers that satisfy an H_∞ -norm bound and relations to risk sensitivity," *Systems & Control Letters*, vol. 11, pp. 167-172, 1988.
- [7] J. Doyle, K. Glover, 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, pp. 831-847, 1989.
- [8] B. A. Francis, *A Course in H_∞ Control Theory*, Lecture Notes in Control and Information Sciences, vol. 88, Springer, Berlin, 1987.
- [9] K. Glover, "All optimal Hankel-norm approximations of linear multivariable systems and their \mathcal{L}_∞ error bounds," *International Journal of Control*, vol. 39, pp. 1115-1193, 1984.
- [10] P. Khargonekar, I. R. Petersen, and M. A. Rotea, " H_∞ -optimal control with state feedback," *IEEE Transactions on Automatic Control*, vol. AC-33, pp. 786-788, 1988.
- [11] H. Kimura, *Chain Scattering Approach to H_∞ Control*, Birkhäuser, Boston, 1997.
- [12] H. Kwakernaak, "A polynomial approach to minimax frequency domain optimization of multivariable feedback systems," *International Journal of Control*, vol. 41, pp. 117-156, 1986.

- [13] D. J. N. Limebeer and B. D. O. Anderson, "An interpolation theory approach to H_∞ controller degree bounds," *Linear Algebra and its Applications*, vol. 98, pp. 347-386, 1988.
- [14] G. P. Papavassilopoulos and M. G. Safonov, "Robust control design via game theoretic methods," *Proceedings of the 28th Conference on Decision and Control*, Tampa, Florida, pp. 382-387, 1989.
- [15] I. R. Petersen, "Disturbance attenuation and H_∞ optimization: A design method based on the algebraic Riccati equation," *IEEE Transactions on Automatic Control*, vol. AC-32, pp. 427-429, 1987.
- [16] A. Saberi, B. M. Chen, and Z. L. Lin, "Closed-form solutions to a class of H_∞ optimization problem," *International Journal of Control*, vol. 60, pp. 41-70, 1994.
- [17] C. Scherer, " H_∞ control by state feedback and fast algorithm for the computation of optimal H_∞ norms," *IEEE Transactions on Automatic Control*, vol. 35, pp. 1090-1099, 1990.
- [18] C. Scherer, "The state-feedback H_∞ problem at optimality," *Automatica*, vol. 30, pp. 293-305, 1994.
- [19] G. Tadmor, "Worst-case design in the time domain: The maximum principle and the standard H_∞ problem," *Mathematics of Control, Signals and Systems*, vol. 3, pp. 301-324, 1990.
- [20] G. Zames, "Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms, and approximate inverses," *IEEE Transactions on Automatic Control*, vol. 26, pp. 301-320, 1981.
- [21] K. Zhou, J. Doyle, and K. Glover, *Robust and Optimal Control*, Prentice Hall, New York, 1996.

Problem 8.3

Determining the least upper bound on the achievable delay margin

Daniel E. Davison and Daniel E. Miller

Department of Electrical and Computer Engineering

University of Waterloo

Waterloo, Ontario N2L 3G1

Canada

`ddavison@kingcong.uwaterloo.c` and `millier@hobbes.uwaterloo.ca`

1 MOTIVATION AND PROBLEM STATEMENT

Control engineers have had to deal with time delays in control processes for decades and, consequently, there is a huge literature on the topic, e.g., see [1] or [2] for collections of recent results. Delays arise from a variety of sources, including physical transport delay (e.g., in a rolling mill or in a chemical plant), signal transmission delay (e.g., in an earth-based satellite control system or in a system controlled over a network), and computational delay (e.g., in a system which uses image processing). The problems posed here are concerned in particular with systems where the time delay is not known exactly: such uncertainty exists, for example, in a rolling mill system where the physical speed of the process may change day-to-day, or in a satellite control system where the signal transmission time between earth and the satellite changes as the satellite moves, or in a control system implemented on the internet where the time delay is uncertain because of unknown traffic load on the network.

Motivated by the above examples, we focus here on the simplest problem that captures the difficulty of control in the face of uncertain delay. Specifically, consider the classical linear time-invariant (LTI) unity-feedback control system with a known controller and with a plant that is known except for an uncertain output delay. Denote the plant delay by τ , the plant transfer function by $P(s) = P_0(s)\exp(-s\tau)$, and the controller by $C(s)$. Assume the feedback system is internally stable when $\tau = 0$. Let us define the *delay margin (DM)* to be the largest time delay such that, for any delay less than

or equal to this value, the closed-loop system remains internally stable:

$$DM(P_0, C) := \sup\{\bar{\tau} : \text{for all } \tau \in [0, \bar{\tau}], \text{ the feedback control system with} \\ \text{controller } C(s) \text{ and plant } P(s) = P_0(s)\exp(-s\tau) \text{ is} \\ \text{internally stable}\}.$$

Computation of $DM(P_0, C)$ is straightforward. Indeed, the Nyquist stability criterion can be used to conclude that the delay margin is simply the phase margin of the undelayed system divided by the gain crossover frequency of the undelayed system. Other techniques for computing the delay margin for LTI systems have also been developed, e.g., see [3], [4], [5], and [6], just to name a few.

In contrast to the problem of computing the delay margin when the controller is known, the *design* of a controller to achieve a prespecified delay margin is not straightforward, except in the trivial case where the plant is open-loop stable, in which case the zero controller achieves $DM(P_0, C) = \infty$. To the best of the authors' knowledge, there is no known technique for designing a controller to achieve a prespecified delay margin. Moreover, the fundamental question of whether or not there exists a finite upper bound on the delay margin that is achievable by a LTI controller has not even been addressed. Hence, there are three unsolved problems:

Problem 1: Does there exist an (unstable) LTI plant, P_0 , for which there is a finite upper bound on the delay margin that is achievable by a LTI controller? In other words, does there exist a P_0 for which

$$DM_{sup}(P_0) := \sup\{DM(P_0, C) : \text{the feedback control system with} \\ \text{controller } C(s) \text{ and plant } P_0(s) \text{ is} \\ \text{internally stable}\}$$

is finite?

Problem 2: If the answer to Problem 1 is affirmative, devise a computationally feasible algorithm that, given $P_0(s)$, computes $DM_{sup}(P_0)$ to a given prescribed degree of accuracy.

Problem 3: If the answer to Problem 1 is affirmative, devise a computationally feasible algorithm that, given $P_0(s)$ and a value T in the range $0 < T < DM_{sup}(P_0)$, constructs a $C(s)$ that satisfies $DM(P_0, C) \geq T$.

2 RELATED RESULTS

It is natural to attempt to use robust control methods to solve these problems (e.g., see [7] or [8]). That is, construct a plant uncertainty "ball" that includes all possible delayed plants, then design a controller to stabilize every

plant within that ball. To the best of the authors' knowledge, such techniques always introduce conservativeness, and therefore cannot be used to solve the problems stated above.

Alternatively, it has been established in the literature that there are upper bounds on the *gain margin* and *phase margin* if the plant has poles and zeros in the open right-half plane [9], [7]. These bounds are not conservative, but it is not obvious how to apply the same techniques to the delay margin problem.

As a final possibility, performance limitation integrals, such as those described in [10], may be useful, especially for solving Problem 1.

BIBLIOGRAPHY

- [1] M. S. Mahmoud, *Robust Control and Filtering for Time-Delay Systems*, New York: Marcel Dekker, 2000.
- [2] L. Dugard and E. Verriest, eds., *Stability and Control of Time-Delay Systems*, Springer-Verlag, 1997.
- [3] J. Zhang, C. R. Knospe, and P. Tsiotras, "Stability of linear time-delay systems: A delay-dependent criterion with a tight conservatism bound," In: *Proceedings of the American Control Conference*, Chicago, Illinois, pp. 1458–1462, June 2000.
- [4] J. Chen, G. Gu, and C. N. Nett, "A new method for computing delay margins for stability of linear delay systems," In: *Proceedings of the 33rd Conference on Decision and Control*, Lake Buena Vista, Florida, pp. 433–437, Dec. 1994.
- [5] J. Chiasson, "A method for computing the interval of delay values for which a differential-delay system is stable," *IEEE Transactions on Automatic Control*, vol. 33, pp. 1176–1178, Dec. 1988.
- [6] K. Walton and J. Marshall, "Direct method for TDS stability analysis," *IEE Proceedings, Part D*, vol. 134, pp. 101–107, Mar. 1987.
- [7] J. C. Doyle, B. A. Francis, and A. R. Tannenbaum, *Feedback Control Theory*, New York, NY: Macmillan, 1992.
- [8] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*, Upper Saddle River, NJ: Prentice Hall, 1996.
- [9] P. P. Khargonekar and A. Tannenbaum, "Non-euclidean metrics and the robust stabilization of systems with parameter uncertainty," *IEEE Transactions on Automatic Control*, vol. AC-30, pp. 1005–1013, Oct. 1985.

- [10] J. Freudenberg, R. Middleton, and A. Stefanopoulou, "A survey of inherent design limitations," In: *Proceedings of the American Control Conference*, Chicago, Illinois, pp. 2987–3001, June 2000.

Problem 8.4

Stable controller coefficient perturbation in floating point implementation

Jun Wu

National Key Laboratory of Industrial Control Technology
Institute of Advanced Process Control
Zhejiang University
Hangzhou 310027
P. R. China
jwu@iipc.zju.edu.cn

Sheng Chen

Department of Electronics and Computer Science
University of Southampton, Highfield
Southampton SO17 1BJ
United Kingdom
sqc@ecs.soton.ac.uk

1 DESCRIPTION OF THE PROBLEM

For real matrix $\mathbf{X} = [x_{ij}]$, denote

$$\|\mathbf{X}\|_{\max} = \max_{i,j} |x_{ij}|. \quad (1)$$

For real matrices $\mathbf{X} = [x_{ij}]$ and $\mathbf{Y} = [y_{ij}]$ of the same dimension, denote the Hadamard product of \mathbf{X} and \mathbf{Y} as

$$\mathbf{X} \circ \mathbf{Y} = [x_{ij}y_{ij}]. \quad (2)$$

A square real matrix is said to be stable if its eigenvalues are all in the interior of the unit disc.

Consider a stable discrete-time closed-loop control system, consisting of a linear time invariant plant $P(z)$ and a digital controller $C(z)$. The plant model $P(z)$ is assumed to be strictly proper with a state-space description

$$\begin{cases} \mathbf{x}_P(k+1) = \mathbf{A}_P \mathbf{x}_P(k) + \mathbf{B}_P \mathbf{u}(k) \\ \mathbf{y}(k) = \mathbf{C}_P \mathbf{x}_P(k) \end{cases} \quad (3)$$

where $\mathbf{A}_P \in \mathcal{R}^{m \times m}$, $\mathbf{B}_P \in \mathcal{R}^{m \times l}$ and $\mathbf{C}_P \in \mathcal{R}^{q \times m}$. The controller $C(z)$ is described by

$$\begin{cases} \mathbf{x}_C(k+1) = \mathbf{A}_C \mathbf{x}_C(k) + \mathbf{B}_C \mathbf{y}(k) \\ \mathbf{u}(k) = \mathbf{C}_C \mathbf{x}_C(k) + \mathbf{D}_C \mathbf{y}(k) \end{cases} \quad (4)$$

where $\mathbf{A}_C \in \mathcal{R}^{n \times n}$, $\mathbf{B}_C \in \mathcal{R}^{n \times q}$, $\mathbf{C}_C \in \mathcal{R}^{l \times n}$ and $\mathbf{D}_C \in \mathcal{R}^{l \times q}$. It can be shown easily that the transition matrix of the closed-loop system is

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_P + \mathbf{B}_P \mathbf{D}_C \mathbf{C}_P & \mathbf{B}_P \mathbf{C}_C \\ \mathbf{B}_C \mathbf{C}_P & \mathbf{A}_C \end{bmatrix} \in \mathcal{R}^{(m+n) \times (m+n)}. \quad (5)$$

It is well-known that a discrete-time closed-loop system is stable if and only if its transition matrix is stable. Since the closed-loop system, consisting of (3) and (4), is designed to be stable, \mathbf{A} is stable. Let

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \in \mathcal{R}^{(m+n) \times (l+n)}, \quad (6)$$

$$\mathbf{C} = \begin{bmatrix} \mathbf{C}_P & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \in \mathcal{R}^{(q+n) \times (m+n)}, \quad (7)$$

$$\mathbf{W} = \begin{bmatrix} \mathbf{D}_C & \mathbf{C}_C \\ \mathbf{B}_C & \mathbf{A}_C \end{bmatrix} \in \mathcal{R}^{(l+n) \times (q+n)}, \quad (8)$$

where $\mathbf{0}$ and \mathbf{I} denote the zero and identity matrices of appropriate dimensions, respectively. Define the set

$$\mathcal{S} = \{\Delta : \Delta \in \mathcal{R}^{(l+n) \times (q+n)}, \mathbf{A} + \mathbf{B}(\mathbf{W} \circ \Delta)\mathbf{C} \text{ is stable}\} \quad (9)$$

and further define

$$v = \inf\{\|\Delta\|_{\max} : \Delta \in \mathcal{R}^{(l+n) \times (q+n)}, \Delta \notin \mathcal{S}\}. \quad (10)$$

The open problem is: calculate the value of v .

2 MOTIVATION OF THE PROBLEM

The classical digital controller design methodology often assumes that the controller is implemented exactly, even though in reality a control law can only be realized with a digital processor of finite word length (FWL). It may seem that the uncertainty resulting from finite-precision computing of the digital controller is so small, compared to the uncertainty within the plant, such that this controller “uncertainty” can simply be ignored. Increasingly, however, researchers have realized that this is not necessarily the case. Due to the FWL effect, a casual controller implementation may degrade the designed closed-loop performance or even destabilize the designed stable closed-loop system, if the controller implementation structure is not carefully chosen [1, 2].

With decreasing in price and increasing in availability, the use of floating-point processors in controller implementations has increased dramatically.

When a real number x is implemented in a floating-point format, it is perturbed to $x(1 + \delta)$ with $|\delta| < \eta$, where η is the maximum round-off error of the floating-point representation [3]. It can be seen that the perturbation resulting from finite-precision floating-point arithmetic is multiplicative.

For the closed-loop system described in section 1, when $C(z)$ is implemented in finite-precision floating-point format, the controller realization \mathbf{W} is perturbed to $\mathbf{W} + \mathbf{W} \circ \Delta$. Each element of Δ is bounded by $\pm\eta$, that is,

$$\|\Delta\|_{\max} < \eta. \quad (11)$$

With the perturbation Δ , the transition matrix of the closed-loop system becomes $\mathbf{A} + \mathbf{B}(\mathbf{W} \circ \Delta)\mathbf{C}$. If an eigenvalue of $\mathbf{A} + \mathbf{B}(\mathbf{W} \circ \Delta)\mathbf{C}$ is outside the open unit disc, the closed-loop system, designed to be stable, becomes unstable with the FWL floating-point implemented \mathbf{W} .

It is therefore critical to know the ability of the closed-loop stability to tolerate the coefficient perturbation Δ in \mathbf{W} resulted from finite-precision implementation. This means that we would like to know the largest “cube” in the perturbation space, within which the closed-loop system remains stable. The measure v defined in (10) gives the exact size of the largest “stable perturbation cube” for \mathbf{W} . If the value of v can be computed, it becomes a simple matter to check whether \mathbf{W} is “robust” to FWL errors, because $\mathbf{A} + \mathbf{B}(\mathbf{W} \circ \Delta)\mathbf{C}$ remains stable when $v > \eta$.

Furthermore, \mathbf{W} or $(\mathbf{A}_C, \mathbf{B}_C, \mathbf{C}_C, \mathbf{D}_C)$ is a realization of the controller $C(z)$. The realizations of $C(z)$ are not unique. Different realizations are all equivalent if they are implemented in infinite precision. In fact, suppose $(\mathbf{A}_C^0, \mathbf{B}_C^0, \mathbf{C}_C^0, \mathbf{D}_C^0)$ is a realization of $C(z)$, then all the realizations of $C(z)$ form a set

$$\mathcal{S}_C = \left\{ \mathbf{W} : \mathbf{W} = \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{T}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{D}_C^0 & \mathbf{C}_C^0 \\ \mathbf{B}_C^0 & \mathbf{A}_C^0 \end{bmatrix} \begin{bmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{T} \end{bmatrix} \right\} \quad (12)$$

where the transformation matrix $\mathbf{T} \in \mathcal{R}^{n \times n}$ is an arbitrary nonsingular matrix. A useful observation is that different \mathbf{W} have different values of v . Provided that the value of v is computationally tractable, an optimal realization of $C(z)$, which has a maximum tolerance to FWL errors, can be obtained via optimization.

The open problem defined in section 1 was first seen in [3]. At present, there exists no available result. An approach to bypass the difficulty in computing v is to define some approximate upper bound of v using a first-order approximation, which is computationally tractable (see [3]).

One of the thorny items in the open problem is the Hadamard product $\mathbf{W} \circ \Delta$. The form of structured perturbation, which was adopted in μ -analysis methods [4], may be used to deal with this Hadamard product: Δ can be transformed into a generalized perturbation $\tilde{\Delta}$ that has certain structure such as block-diagonal. The fixed matrices $\tilde{\mathbf{A}}$, $\tilde{\mathbf{B}}$ and $\tilde{\mathbf{C}}$ may be obtained such that the stability of $\tilde{\mathbf{A}} + \tilde{\mathbf{B}}\tilde{\Delta}\tilde{\mathbf{C}}$ is equivalent to that of $\mathbf{A} + \mathbf{B}(\mathbf{W} \circ \Delta)\mathbf{C}$. Although the stability of $\tilde{\mathbf{A}} + \tilde{\mathbf{B}}\tilde{\Delta}\tilde{\mathbf{C}}$ can be treated

satisfactorily by μ -analysis methods, the open problem cannot be solved successfully by μ -analysis methods. This is because μ -analysis methods are concerned about the maximal singular value $\bar{\sigma}(\tilde{\Delta})$ of $\tilde{\Delta}$. In fact, the distance between $\bar{\sigma}(\tilde{\Delta})$ and $\|\Delta\|_{\max}$ can be quite large, and $\|\Delta\|_{\max}$ is the other thorny item which makes the open problem difficult.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of the United Kingdom Royal Society under a KC Wong fellowship (RL/ART/CN/XFI/KCW/11949). Jun Wu wishes to thank the support of the National Natural Science Foundation of China under Grant 60174026 and the Scientific Research Foundation for the Returned Overseas Chinese Scholars of Zhejiang province under Grant J20020546.

BIBLIOGRAPHY

- [1] M. Gevers and G. Li, *Parameterizations in Control, Estimation and Filtering Problems: Accuracy Aspects*, London, Springer Verlag, 1993.
- [2] R. S. H. Istepanian and J. F. Whidborne, eds., *Digital Controller Implementation and Fragility: A Modern Perspective*, London, Springer Verlag, 2001.
- [3] J. Wu, S. Chen, J. F. Whidborne and J. Chu, "Optimal floating-point realizations of finite-precision digital controllers," In: *Proc. 41st IEEE Conference on Decision and Control*, Las Vegas, USA, Dec. 10-13, 2002, pp. 2570-2575.
- [4] J. Doyle, "Analysis of feedback systems with structured uncertainties," *Proc. IEE*, vol. 129-D, no 6, pp. 242-250, 1982.

PART 9

Identification, Signal Processing

Problem 9.1

A conjecture on Lyapunov equations and principal angles in subspace identification

Katrien De Cock and Bart De Moor¹

Dept. of Electrical Engineering (ESAT-SCD)

K.U.Leuven

Kasteelpark Arenberg 10

B-3001 Leuven

Belgium

<http://www.esat.kuleuven.ac.be/sista-cosic-docarch>

decock@esat.kuleuven.ac.be, demoor@esat.kuleuven.ac.be

1 DESCRIPTION OF THE PROBLEM

The following conjecture relates the eigenvalues of certain matrices that are derived from the solution of a Lyapunov equation that occurred in the analysis of stochastic subspace identification algorithms [3]. First, we formulate the conjecture as a pure matrix algebraic problem. In Section 2, we will describe its system theoretic consequences and interpretation.

Conjecture: Let $A \in \mathbb{R}^{n \times n}$ be a real matrix and $v, w \in \mathbb{R}^n$ be real vectors so that there are no two eigenvalues λ_i and λ_j of $\begin{pmatrix} A & 0 \\ 0 & A + vw^T \end{pmatrix}$ for which $\lambda_i \lambda_j = 1$ ($i, j = 1, \dots, 2n$). If the $n \times n$ matrices P , Q and R satisfy the

¹Katrien De Cock is a research assistant at the K.U.Leuven. Dr. Bart De Moor is a full professor at the K.U.Leuven. Our research is supported by grants from several funding agencies and sources: Research Council KUL: Concerted Research Action GOA-Mefisto 666, IDO, several Ph.D., postdoctoral & fellow grants; Flemish Government: Fund for Scientific Research Flanders (several Ph.D. and postdoctoral grants, projects G.0256.97, G.0115.01, G.0240.99, G.0197.02, G.0407.02, research communities ICCoS, ANMMM), AWI (Bil. Int. Collaboration Hungary/ Poland), IWT (Soft4s, STWW-Genprom, GBOU-McKnow, Eureka-Impact, Eureka-FLiTE, several PhD grants); Belgian Federal Government: DWTC (IUAP IV-02 (1996-2001) and IUAP V-22 (2002-2006)), Program Sustainable Development PODO-II (CP/40); Direct contract research: Verhaert, Electrabel, Elia, Data4s, IPCOS.

Lyapunov equation

$$\begin{pmatrix} P & R \\ R^T & Q \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & (A + vw^T)^T \end{pmatrix} \begin{pmatrix} P & R \\ R^T & Q \end{pmatrix} \begin{pmatrix} A^T & 0 \\ 0 & A + vw^T \end{pmatrix} + \begin{pmatrix} v \\ w \end{pmatrix} (v^T \ w^T) , \quad (1)$$

and P , Q and $(I_n + PQ)$ are nonsingular,² then the matrices $P^{-1}RQ^{-1}R^T$ and $(I_n + PQ)^{-1}$ have the same eigenvalues.

Note that the condition $\lambda_i \lambda_j \neq 1$ ($\forall i, j = 1, \dots, 2n$) ensures that there exists a solution $\begin{pmatrix} P & R \\ R^T & Q \end{pmatrix}$ of the Lyapunov equation (1) and that the solution is unique.

We have checked the similarity of $P^{-1}RQ^{-1}R^T$ and $(I_n + PQ)^{-1}$ for numerous examples (“proof by Matlab”) and it is simple to prove the conjecture for $n = 1$. Furthermore, via a large detour (see [3]) we can also prove it from the system theoretic interpretation, which is given in section 5. However, we have not been able to find a general and elegant proof.

We also remark that the requirement that v and w are vectors is necessary for the conjecture to hold. One can easily find counterexamples for the case $V, W \in \mathbb{R}^{n \times m}$, where $m > 1$. It is consequently clear that this condition on v and w should be used in the proof.

2 BACKGROUND AND MOTIVATION

Although the conjecture is formulated as a pure matrix algebraic problem, its system theoretic interpretation is particularly interesting. In order to explain the consequences, we first have to introduce some concepts: the principal angles between subspaces (section 3) and their statistical counterparts, the canonical correlations of random variables (section 4). Next, in section 5 we will show how the conjecture, when proved correct, would enable us to prove in an elegant way that the nonzero canonical correlations of the past and the future of the output process of a linear stochastic model are equal to the sines of the principal angles between two specific subspaces that are derived from the model. This result, in its turn, is instrumental for further derivations in [3], where a cepstral distance measure is related to canonical correlations and to the mutual information of two processes (see also section 5). Moreover, by this new characterization of the canonical correlations, we gain insight in the geometric properties of subspace based techniques.

²The matrix I_n is the $n \times n$ identity matrix.

3 THE PRINCIPAL ANGLES BETWEEN TWO SUBSPACES

The concept of principal angles between and principal directions in subspaces of a linear vector space is due to Jordan in the nineteenth century [8]. We give the definition and briefly describe how the principal angles can be computed. Let S_1 and S_2 be subspaces of \mathbb{R}^n of dimension p and q , respectively, where $p \leq q$. Then, the p principal angles between S_1 and S_2 , denoted by $\theta_1, \dots, \theta_p$, and the corresponding principal directions $u_i \in S_1$ and $v_i \in S_2$ ($i = 1, \dots, p$) are recursively defined as

$$\begin{aligned}\cos \theta_1 &= \max_{u \in S_1} \max_{v \in S_2} |u^T v| = u_1^T v_1 \\ \cos \theta_k &= \max_{u \in S_1} \max_{v \in S_2} |u^T v| = u_k^T v_k \quad (k = 2, \dots, p)\end{aligned}$$

subject to $\|u\| = \|v\| = 1$, and for $k > 1$: $u^T u_i = 0$ and $v^T v_i = 0$, where $i = 1, \dots, k-1$.

If S_1 and S_2 are the row spaces of the matrices $A \in \mathbb{R}^{l \times n}$ and $B \in \mathbb{R}^{m \times n}$, respectively, then the cosines of the principal angles $\theta_1, \dots, \theta_p$, can be computed as the largest p generalized eigenvalues of the matrix pencil

$$\begin{pmatrix} 0 & AB^T \\ BA^T & 0 \end{pmatrix} - \begin{pmatrix} AA^T & 0 \\ 0 & BB^T \end{pmatrix} \lambda.$$

Furthermore, if A and B are full row rank matrices, i.e., $l = p$ and $m = q$, then the squared cosines of the principal angles between the row space of A and the row space of B are equal to the eigenvalues of

$$(AA^T)^{-1}AB^T(BB^T)^{-1}BA^T.$$

Numerically stable methods to compute the principal angles via the QR and singular value decomposition can be found in [5, pp. 603–604].

4 THE CANONICAL CORRELATIONS OF TWO RANDOM VARIABLES

Canonical correlation analysis, due to Hotelling [6], is the statistical version of the notion of principal angles.

Let $X \in \mathbb{R}^p$ and $Y \in \mathbb{R}^q$, where $p \leq q$, be zero-mean random variables with full rank joint covariance matrix³

$$Q = E \left\{ \begin{pmatrix} X \\ Y \end{pmatrix} \begin{pmatrix} X^T & Y^T \end{pmatrix} \right\} = \begin{pmatrix} Q_x & Q_{xy} \\ Q_{yx} & Q_y \end{pmatrix}.$$

The canonical correlations of X and Y are defined as the largest p eigenvalues of the pencil $\begin{pmatrix} 0 & Q_{xy} \\ Q_{yx} & 0 \end{pmatrix} - \begin{pmatrix} Q_x & 0 \\ 0 & Q_y \end{pmatrix} \lambda$. More information on canonical correlation analysis can be found in [1, 6].

³ $E\{\cdot\}$ is the expected value operator.

5 SYSTEM THEORETIC INTERPRETATION OF CONJECTURE

Let $\{y(k)\}_{k \in \mathbb{Z}}$ be a real, discrete-time, scalar and zero-mean stationary stochastic process that is generated by the following single-input, single-output (SISO), asymptotically stable state space model in forward innovation form:

$$\begin{cases} x(k+1) &= Ax(k) + Ku(k), \\ y(k) &= Cx(k) + u(k), \end{cases} \quad (2)$$

where $\{u(k)\}_{k \in \mathbb{Z}}$ is the innovation process of $\{y(k)\}_{k \in \mathbb{Z}}$, $A \in \mathbb{R}^{n \times n}$, $K \in \mathbb{R}^{n \times 1}$ is the Kalman gain and $C \in \mathbb{R}^{1 \times n}$. The state space matrices of the inverse model (or whitening filter) are $A - KC$, K and $-C$, respectively, as is easily seen by writing $u(k)$ as an output with $y(k)$ as an input.

By substituting the vector v in (1) by K , and w by $-C^T$, the matrices P , Q and R in (1) can be given the following interpretation. The matrix P is the controllability Gramian of the model (2) and Q is the observability Gramian of the inverse model, while R is the cross product of the infinite controllability matrix of (2) and the infinite observability matrix of the inverse model. Otherwise formulated:

$$\begin{pmatrix} P & R \\ R^T & Q \end{pmatrix} = \begin{pmatrix} \mathcal{C}_\infty \\ \Gamma_\infty^T \end{pmatrix} (\mathcal{C}_\infty^T \quad \Gamma_\infty) ,$$

$$\text{where } \mathcal{C}_\infty = (K \quad AK \quad A^2K \quad \dots) \text{ and } \Gamma_\infty = - \begin{pmatrix} C \\ C(A - KC) \\ C(A - KC)^2 \\ \vdots \end{pmatrix} .$$

Due to the stability and the minimum phase property of the forward innovation model (2), these infinite products result in finite matrices and in addition, the condition $\lambda_i \lambda_j \neq 1$ in conjecture 1 is fulfilled. Furthermore, under fairly general conditions, P , Q , and $I_n + PQ$ are nonsingular, which follows from the positive definiteness of P and Q under general conditions. The matrix $P^{-1}RQ^{-1}R^T$ in conjecture 1 is now equal to the product

$$(\mathcal{C}_\infty \mathcal{C}_\infty^T)^{-1} (\mathcal{C}_\infty \Gamma_\infty) (\Gamma_\infty^T \Gamma_\infty)^{-1} (\Gamma_\infty^T \mathcal{C}_\infty^T) .$$

Consequently, its n eigenvalues are the squared cosines of the principal angles between the row space of \mathcal{C}_∞ and the column space of Γ_∞ (see Section 3). The angles will be denoted by $\theta_1, \dots, \theta_n$ (in nondecreasing order).

The eigenvalues of the matrix $(I_n + PQ)^{-1}$, on the other hand, are related to the canonical correlations of the past and the future stochastic processes of $\{y(k)\}_{k \in \mathbb{Z}}$, which are defined as the canonical correlations of the random variables

$$y_p = \begin{pmatrix} y(-1) \\ y(-2) \\ y(-3) \\ \vdots \end{pmatrix} \quad \text{and} \quad y_f = \begin{pmatrix} y(0) \\ y(1) \\ y(2) \\ \vdots \end{pmatrix} ,$$

and denoted by ρ_1, ρ_2, \dots (in nonincreasing order). It can be shown [3] that the largest n canonical correlations of y_p and y_f are equal to the square roots of the eigenvalues of $I_n - (I_n + PQ)^{-1}$. The other canonical correlations are equal to 0.

Conjecture 1 now gives us the following characterization of the canonical correlations of the past and the future of $\{y(k)\}_{k \in \mathbb{Z}}$: the largest n canonical correlations are equal to the sines of the principal angles between the row space of \mathcal{C}_∞ and the column space of Γ_∞ and the other canonical correlations are equal to 0:

$$\rho_1 = \sin \theta_n, \rho_2 = \sin \theta_{n-1}, \dots, \rho_n = \sin \theta_1, \rho_{n+1} = \rho_{n+2} = \dots = 0. \quad (3)$$

This result can be used to prove that a recently defined cepstral norm [9] for a model as in (2) is closely related to the mutual information of the past and the future of its output process. Let the transfer function of the system in (2) be denoted by $H(z)$. Then the complex cepstrum $\{c(k)\}_{k \in \mathbb{Z}}$ of the model is defined as the inverse Z -transform of the complex logarithm of $H(z)$:

$$c(k) = \frac{1}{2\pi i} \oint_C \log(H(z)) z^{k-1} dz,$$

where the complex logarithm of $H(z)$ is appropriately defined (see [10, pp. 495–497]) and the contour C is the unit circle. The cepstral norm that we consider, is defined as

$$\|\log H\|^2 = \sum_{k=0}^{\infty} kc(k)^2.$$

As we have proven in [2], it can be characterized in terms of the principal angles $\theta_1, \dots, \theta_n$ between the row space of \mathcal{C}_∞ and the column space of Γ_∞ as follows:

$$\|\log H\|^2 = -\log \prod_{i=1}^n \cos^2 \theta_i,$$

and from (3) we obtain

$$\|\log H\|^2 = -\log \prod (1 - \rho_i^2).$$

The relation $\sum_{k=0}^{\infty} kc(k)^2 = -\log \prod (1 - \rho_i^2)$ was also reported in [7, proposition 2]. Moreover, if $\{y(k)\}_{k \in \mathbb{Z}}$ is a Gaussian process, then the expression $-\frac{1}{2} \log \prod (1 - \rho_i^2)$ is equal to the mutual information of its past and future (see, e.g., [4]), which is denoted by $I(y_p; y_f)$. Consequently,

$$\|\log H\|^2 = \sum_{k=0}^{\infty} kc(k)^2 = 2I(y_p; y_f).$$

6 CONCLUSIONS

We presented a matrix algebraic conjecture on the eigenvalues of matrices that are derived from the solution of a Lyapunov equation. We showed that

a proof of conjecture 1 would provide yet another elegant geometric result in the subspace based study of linear stochastic systems. Moreover, it can be used to express a cepstral distance measure that was defined in [9] in terms of canonical correlations and also as the mutual information of two processes.

BIBLIOGRAPHY

- [1] T. W. Anderson, *An Introduction to Multivariate Statistical Analysis*, John Wiley & Sons, New York, 1984.
- [2] K. De Cock and B. De Moor, "Subspace angles between ARMA models," Technical Report⁴ 00-44a, ESAT-SCD, K.U.Leuven, Leuven, Belgium, accepted for publication in *Systems & Control Letters*, 2002.
- [3] K. De Cock, *Principal Angles in System Theory, Information Theory and Signal Processing*, Ph.D. thesis,⁵ Faculty of Applied Sciences, K.U.Leuven, Leuven, Belgium, 2002.
- [4] I. M. Gel'fand and A. M. Yaglom, "Calculation of the amount of information about a random function contained in another such function," *American Mathematical Society Translations, Series (2)*, 12, pp. 199–236, 1959.
- [5] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, 1996.
- [6] H. Hotelling, "Relations between two sets of variates," *Biometrika*, 28, pp. 321–372, 1936.
- [7] N. P. Jewell, P. Bloomfield, and F. C. Bartmann, "Canonical correlations of past and future for time series: Bounds and computation," *The Annals of Statistics*, 11, pp. 848–855, 1983.
- [8] C. Jordan, "Essai sur la géométrie à n dimensions," *Bulletin de la Société Mathématique*, 3, pp. 103–174, 1875.
- [9] R. J. Martin, "A metric for ARMA processes," *IEEE Transactions on Signal Processing*, 48, pp. 1164–1170, 2000.
- [10] A. V. Oppenheim and R. W. Schaffer, *Digital Signal Processing*, Prentice/Hall International, 1975.

⁴This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/sista/reports* as file *00-44a.ps.gz*.

⁵This thesis will also be made available by anonymous ftp from *ftp.esat.kuleuven.ac.be* as file *pub/sista/decock/reports/phd.ps.gz*

Problem 9.2

Stability of a nonlinear adaptive system for filtering and parameter estimation

Masoud Karimi-Ghartemani

Department of Electrical and Computer Engineering

University of Toronto

Toronto, Ontario

Canada M5S 3G4

masoud@ele.utoronto.ca

Alireza K. Ziarani

Department of Electrical and Computer Engineering

Clarkson University

Potsdam, NY

USA 13699-5720

aziarani@clarkson.edu

1 DESCRIPTION OF THE PROBLEM

We are concerned about the mathematical properties of the dynamical system presented by the following three differential equations:

$$\begin{cases} \frac{dA}{dt} &= -2\mu_1 A \sin^2 \phi + 2\mu_1 \sin \phi f(t), \\ \frac{d\omega}{dt} &= -\mu_2 A^2 \sin(2\phi) + 2\mu_2 A \cos \phi f(t), \\ \frac{d\phi}{dt} &= \omega + \mu_3 \frac{d\omega}{dt} \end{cases} \quad (1)$$

where parameters μ_i , $i = 1, 2, 3$ are positive real constants and $f(t)$ is a function of time having a general form of

$$f(t) = A_o \sin(\omega_o t + \delta_o) + f_1(t). \quad (2)$$

A_o , ω_o and δ_o are fixed quantities and it is assumed that $f_1(t)$ has no frequency component at ω_o . Variables A and ω are in \mathbb{R}^1 and ϕ varies on the one-dimensional circle \mathbb{S}^1 with radius 2π .

The dynamical system presented by (1) is designed to (i) take the signal $f(t)$ as its input signal and extract the component $f_o(t) = A_o \sin(\omega_o t + \delta_o)$ as its

output signal, and (ii) estimate the basic parameters of the extracted signal $f_o(t)$, namely its amplitude, phase, and frequency. The extracted signal is $y = A \sin \phi$ and the basic parameters are the amplitude A , frequency ω and phase angle $\phi = \omega t + \delta$.

Consider the three variables (A, ω, ϕ) in the three-dimensional space $\mathbb{R}^1 \times \mathbb{R}^1 \times \mathbb{S}^1$. The sinusoidal function $f_o(t) = A_o \sin(\omega_o t + \delta_o)$ corresponds to the T_o -periodic curve

$$\Gamma_o(t) = (A_o, \omega_o, \omega_o t + \delta_o) \quad (3)$$

in this space, with $T_o = \frac{2\pi}{\omega_o}$.

The following theorem, which the authors have proved in [1], presents some of the mathematical properties of the dynamical system presented by 1.

Theorem 1: Consider the dynamical system presented by the set of ordinary differential equations (1) in which the function $f(t)$ is defined in (2) and $f_1(t)$ is a bounded T_1 -periodic function with no frequency component at ω_o . The three variables (A, ω, ϕ) are in $\mathbb{R}^1 \times \mathbb{R}^1 \times \mathbb{S}^1$. The parameters μ_i , $i = 1, 2, 3$ are small positive real numbers. If $T_1 = \frac{T_o}{n}$ for any arbitrary $n \in \mathbb{N}$, the dynamical system of (1) has a stable T_o -periodic orbit in a neighborhood of $\Gamma_o(t)$ as defined in (3).

The behavior of the system, as examined within the simulation environments, has led the authors to the following two conjectures, the proofs of which are desired.

Conjecture 1: With the same assumptions as those presented in Theorem 1, if $T_1 = \frac{p}{q}T_o$ for any arbitrary $(p, q) \in \mathbb{N}^2$ with $(p, q) = 1$, the dynamical system presented by (1) has a stable mT_o -periodic orbit which lies on a torus in a neighborhood of $\Gamma_o(t)$ as defined in (3). The value of $m \in \mathbb{N}$ is determined by the pair (p, q) .

Conjecture 2: With the same assumptions as those presented in Theorem 1, if $T_1 = \alpha T_o$ for irrational α , the dynamical system presented by (1) has an attractor set that is a torus in a neighborhood of $\Gamma_o(t)$ as defined in (3). In other words, the response is a never closing orbit that lies on the torus. Moreover, this orbit is a dense set on the torus.

For both conjectures, the neighborhood in which the torus is located depends on the values of parameters μ_i , $i = 1, 2, 3$ and the function $f_1(t)$. If the function $f_1(t)$ is small in order and the parameters are properly selected, the neighborhood can be made to be very small, meaning that the filtering and estimation processes may be achieved with a high degree of accuracy.

Theorem 1 deals with the local stability analysis of the dynamical system (1). In other words, the existence of an attractor (periodic orbit or torus) and an attraction domain around the attractor is proved. However, this theorem does not deal with this domain of attraction. It is desirable to specify this domain of attraction in terms of the function $f_1(t)$ and parameters μ_i , $i = 1, 2, 3$, hence the following open problem:

Open Problem: Consider the dynamical system presented by the ordinary

differential equations (1) in which the function $f(t)$ is defined in (2) and $f_1(t)$ is a bounded T_1 -periodic function with no frequency component at ω_o . Three variables (A, ω, ϕ) are in $\mathbb{R}^1 \times \mathbb{R}^1 \times \mathbb{S}^1$. Parameters μ_i , $i = 1, 2, 3$ are small positive real numbers. This system has an attractor set that may be either a periodic orbit or a torus based on the value of T_1 . It is desired to specify the extent of the attraction domain associated with the attractor in terms of the function $f_1(t)$ and the parameters μ_i , $i = 1, 2, 3$. In other words, and in a simplified case, for a three-parameter representation of $f_1(t)$ as $f_1(t) = a_1 \sin(2\pi/T_1 t + \delta_1)$, it is desirable to parameterize, in terms of the nine-parameter set of $\{\mu_1, \mu_2, \mu_3, A_o, T_o, \delta_o, a_1, T_1, \delta_1\}$, the attractor set, and also the whole region of points (A, ω, ϕ) in $\mathbb{R}^1 \times \mathbb{R}^1 \times \mathbb{S}^1$ that falls in the attraction domain of the attractor.

2 MOTIVATION AND HISTORY OF THE PROBLEM

The dynamical system presented by (1) was proposed by the authors to devise a system for the extraction of a sinusoidal component with time-varying parameters when it is corrupted by other sinusoids and noise [1, 2]. This is of significant interest in power system applications, for instance [3]. Estimation of the basic parameters of the extracted sinusoid, namely the amplitude, phase, and frequency, was another object of the work. These parameters provide important information useful in electrical engineering applications. Some applications of the system in biomedical engineering are presented in [2, 4]. This dynamical system presents an alternative structure for the well-known phase-locked loop (PLL) system with significantly advantageous features.

3 AVAILABLE RESULTS

Theorem 1, corresponding to the case of $T_1 = \frac{T_o}{n}$, has been proved by the authors in [1] where the existence, local uniqueness and stability of a T_o -periodic orbit are shown using the Poincaré map theorem as stated in [5, page 70]. Extensive computer simulations verified by laboratory experimental results are presented in [1, 2]. Some of the wide-ranging applications of the dynamical system are presented in [2, 3, 4]. The algorithm governed by the proposed dynamical system presents a powerful signal processing method of analysis/synthesis of nonstationary signals. Alternatively, it may be thought of as a nonlinear adaptive notch filter capable of estimation of parameters of the output signal.

BIBLIOGRAPHY

- [1] M. Karimi-Ghartemani and A. K. Ziarani, "Periodic orbit analysis of two dynamical systems for electrical engineering applications," *Journal of Engineering Mathematics*, 45, pp. 135-154, 2003.
- [2] A. K. Ziarani, *Extraction of Nonstationary Sinusoids*, Ph.D. dissertation, University of Toronto, 2002.
- [3] M. Karimi-Ghartemani and M. R. Iravani, "A Nonlinear Adaptive Filter for On-Line Signal Analysis in Power Systems: Applications," *IEEE Transactions on Power Delivery*, 17, pp. 617-622, 2002.
- [4] A. K. Ziarani and A. Konrad, "A Nonlinear Adaptive Method of Elimination of Power Line Interference in ECG Signals," *IEEE Transactions on Biomedical Engineering*, 49, pp. 540-547, 2002.
- [5] S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, New York: Springer-Verlag, 1983.

PART 10

Algorithms, Computation

Problem 10.1

Root-clustering for multivariate polynomials and robust stability analysis

Pierre-Alexandre Bliman

INRIA

Rocquencourt BP 105

78153 Le Chesnay cedex

France

pierre-alexandre.bliman@inria.fr

1 DESCRIPTION OF THE PROBLEM

Given the $(m+1)$ complex matrices A_0, \dots, A_m of size $n \times n$ and denoting $\overline{\mathbb{D}}$ (resp. $\overline{\mathbb{C}^+}$) the closed unit ball in \mathbb{C} (resp. the closed right-half plane), let us consider the following problem: determine whether

$$\forall s \in \overline{\mathbb{C}^+}, \forall z \stackrel{\text{def}}{=} (z_1, \dots, z_m) \in \overline{\mathbb{D}}^m, \det(sI_n - A_0 - z_1 A_1 - \dots - z_m A_m) \neq 0. \quad (1)$$

We have proved in [1] that property (1) is *equivalent* to the existence of $k \in \mathbb{N}$ and $(m+1)$ matrices $P, Q_1 \in \mathcal{H}^{k^m n}$, $Q_2 \in \mathcal{H}^{k^{m-1}(k+1)n}$, \dots , $Q_m \in \mathcal{H}^{k^{m-1}(k+1)^{m-1}n}$, such that

$$P > 0_{k^m n} \text{ and } R(P, Q_1, \dots, Q_m) < 0_{(k+1)^m n}. \quad (2)$$

Here, \mathcal{H}^n represents the space of $n \times n$ hermitian matrices, and R is a linear application taking values in $\mathcal{H}^{(k+1)^m n}$, defined as follows. Let $\hat{J}_k \stackrel{\text{def}}{=} \begin{pmatrix} I_k & 0_{k \times 1} \end{pmatrix}$, $\check{J}_k \stackrel{\text{def}}{=} \begin{pmatrix} 0_{k \times 1} & I_k \end{pmatrix}$, then (using the power of Kronecker product with the natural meaning):

$$\begin{aligned} R &\stackrel{\text{def}}{=} \left(\left(\hat{J}_k^{m \otimes} \otimes A_0 \right) + \sum_{i=1}^m \left(\hat{J}_k^{(m-i) \otimes} \otimes \check{J}_k \otimes \hat{J}_k^{(i-1) \otimes} \otimes A_i \right) \right)^H P \left(\hat{J}_k^{m \otimes} \otimes I_n \right) \\ &+ \left(\hat{J}_k^{m \otimes} \otimes I_n \right)^T P \left(\left(\hat{J}_k^{m \otimes} \otimes A_0 \right) + \sum_{i=1}^m \left(\hat{J}_k^{(m-i) \otimes} \otimes \check{J}_k \otimes \hat{J}_k^{(i-1) \otimes} \otimes A_i \right) \right) \\ &+ \sum_{i=1}^m \left(\hat{J}_k^{(m-i+1) \otimes} \otimes I_{(k+1)^{i-1}n} \right)^T Q_i \left(\hat{J}_k^{(m-i+1) \otimes} \otimes I_{(k+1)^{i-1}n} \right) \end{aligned}$$

$$- \sum_{i=1}^m \left(\hat{j}_k^{(m-i)\otimes} \otimes \check{J}_k \otimes I_{(k+1)^i-1n} \right)^T Q_i \left(\hat{j}_k^{(m-i)\otimes} \otimes \check{J}_k \otimes I_{(k+1)^i-1n} \right) \quad (3)$$

Problem (2,3) is a *linear matrix inequality* in the $m+1$ unknown matrices P, Q_1, \dots, Q_m , a convex optimization problem.

The LMIs (2,3) obtained for increasing values of k constitute indeed a family of weaker and weaker sufficient conditions for (1). Conversely, property (1) necessarily implies solvability of the LMIs for a certain rank k and beyond. See [1] for details.

Numerical experimentations have shown that the precision of the criteria obtained for small values of k (2 or 3) may be remarkably good already, but rational use of this result requires to have a priori information on the size of the least k , if any, for which the LMIs are solvable. Bounds, especially upper bound, on this quantity are thus highly desirable, and they should be computed with low complexity algorithms.

Open Problem 1: Find an integer-valued function $k^*(A_0, A_1, \dots, A_m)$ defined on the product $(\mathbb{C}^{n \times n})^{m+1}$, whose evaluation necessitates polynomial time, and such that property (1) holds if and only if LMI (2,3) is solvable for $k = k^*$.

One may imagine that the previous quantity exists, depending upon n and m only.

Open Problem 2: Determine whether the quantity $k_{n,m}^* \stackrel{\text{def}}{=} \sup \{k^*(A_0, A_1, \dots, A_m) : A_0, A_1, \dots, A_m \in \mathbb{C}^{n \times n}\}$ is finite. In this case, provide an upper bound of its value.

If $k_{n,m}^* < +\infty$, then, for any $A_0, A_1, \dots, A_m \in \mathbb{C}^{n \times n}$, property (1) holds if and only if LMI (2,3) is solvable for $k = k_{n,m}^*$.

2 MOTIVATIONS AND COMMENTS

We expose here some problems related to property (1).

Robust stability

Property (1) is equivalent to asymptotic stability of the uncertain system

$$\dot{x} = (A_0 + z_1 A_1 + \dots + z_m A_m)x, \quad (4)$$

for any value of $z \in \overline{\mathbb{D}}^m$. Usual approaches leading to sufficient LMI conditions for robust stability are based on search for quadratic Lyapunov functions $x(t)^H S x(t)$ with constant S , see related bibliography in [2, p. 72–73], or *parameter-dependent* $S(z)$, namely *affine* [8, 7, 5, 6, 12] and more recently *quadratic* [19, 20]. Methods based on piecewise quadratic Lyapunov functions [21, 13] and LMIs with augmented number of variables [9, 11] also provide sufficient conditions for robust stability.

The approach leading to the result exposed in §1 systematizes the idea of expanding $S(z)$ in powers of the parameters. Indeed, robust stability of (4) guarantees existence of a Lyapunov function $x(t)^H S(z)x(t)$ with $S(z)$ polynomial with respect to z and \bar{z} in \mathbb{D}^m , and the integer k is related to the degree of this polynomial [1].

Computation of structured singular values with repeated scalar blocks

Property (1) is equivalent to $\mu_\Delta(A) < 1$, for a certain matrix A deduced from A_0, A_1, \dots, A_m , and a set Δ of complex uncertainties with $m+1$ repeated scalar blocks. Evaluation of the structured singular values (a standard and powerful tool of robust analysis) has been proved to be a NP-hard problem, see [3, 16]. Hope had dawned that its standard, efficiently computable, upper bound could be a satisfying approximant [17], but the gap between the two measures has latter on been proved infinite [18, 14].

The approach in §1 could offer attractive numerical alternative for the same purpose, as resolution of LMIs is a convex problem. It provides a family of convex relaxations, of arbitrary precision, of a class of NP-hard problems. The complexity results evoked previously imply the existence of $k^*(A_0, A_1, \dots, A_m)$ such that property (1) is equivalent to solvability of LMI (2,3) for $k = k^*$: first, check that $\mu_\Delta(A) < 1$; if this is true, then assess to k^* the value of the smallest k such that LMI (2,3) is solvable, otherwise put $k^* = 1$. This algorithm is, of course, a disaster from the point of view of complexity and computation time, and it does not answer Problem 1. Concerning the value of $k_{n,m}^*$ in Problem 2, its growth at infinity should be faster than any power in n , except if $P=NP$.

Delay-independent stability of delay systems with noncommensurate delays

Property (1) is a strong version of the *delay-independent stability* of the functional differential equation of retarded type $\dot{x} = A_0x(t) + A_1x(t - h_1) + \dots + A_mx(t - h_m)$, that is the asymptotical stability for any value of $h_1, \dots, h_m \geq 0$, see [10, 2, 4]. This problem has been recognized as NP-hard [15]. Solving LMI (2,3) provides explicitly a quadratic Lyapunov-Krasovskii functional independent upon the values of the delays [1].

Robust stability of discrete-time systems and stability of multidimensional (nD) systems

Understanding how to cope with the choice of k to apply LMI (2,3), should also lead to progress in the analysis of the discrete-time analogue of (4), the uncertain system $x_{k+1} = (A_0 + z_1A_1 + \dots + z_mA_m)x_k$. Similarly, stability analysis for multidimensional systems (a discrete-time analogue of the functional differential equations of neutral type) would benefit from such contributions.

BIBLIOGRAPHY

- [1] P.-A. Bliman, "A convex approach to robust stability for linear systems with uncertain scalar parameters," Report research no 4316, INRIA, 2001. Available online at <http://www.inria.fr/rrrt/rr-4316.html>
- [2] S. Boyd, L. El Ghaoui, E. Feron, V. Balakrishnan, "Linear matrix inequalities in system and control theory," *SIAM Studies in Applied Mathematics*, vol. 15, 1994.
- [3] R. P. Braatz, P. M. Young, J. C. Doyle, M. Morari, "Computational complexity of μ calculation," *IEEE Trans. Automat. Control*, 39, no 5, 1000–1002, 1994.
- [4] J. Chen, H.A. Latchman, "Frequency sweeping tests for stability independent of delay," *IEEE Trans. Automat. Control*, 40, no 9, 1640–1645, 1995.
- [5] M. Dettori, C. W. Scherer, "Robust stability analysis for parameter dependent systems using full block S-procedure," *Proc. of 37th IEEE CDC*, Tampa, Florida, 2798–2799, 1998.
- [6] M. Dettori, C. W. Scherer, "New robust stability and performance conditions based on parameter dependent multipliers," *Proc. of 39th IEEE CDC*, Sydney, Australia, 2000.
- [7] E. Feron, P. Apkarian, P. Gahinet, "Analysis and synthesis of robust control systems via parameter-dependent Lyapunov functions," *IEEE Trans. Automat. Control*, 41, no 7, 1041–1046, 1996.
- [8] P. Gahinet, P. Apkarian, M. Chilali, "Affine parameter-dependent Lyapunov functions and real parametric uncertainty," *IEEE Trans. Automat. Control*, 41, no 3, 436–442, 1996.
- [9] J. C. Geromel, M. C. de Oliveira, L. Hsu, "LMI characterization of structural and robust stability," *Linear Algebra Appl.*, 285, no 1-3, 69–80, 1998.
- [10] J. K. Hale, E. F. Infante, F. S. P. Tsen, "Stability in linear delay equations," *J. Math. Anal. Appl.*, 115, 533–555, 1985.
- [11] D. Peaucelle, D. Arzelier, O. Bachelier, J. Bernussou, "A new robust D-stability condition for real convex polytopic uncertainty," *Systems and Control Letters*, 40, no 1, 21–30, 2000.
- [12] D. C. W. Ramos, P. L. D. Peres, "An LMI approach to compute robust stability domains for uncertain linear systems," *Proc. ACC*, Arlington, Virginia, 4073–4078, 2001.

- [13] A. Rantzer, M. Johansson, "Piecewise linear quadratic optimal control," *IEEE Trans. Automat. Control*, 45, no 4, 629–637, 2000.
- [14] M. Sznaier, P. A. Parrilo, "On the gap between μ and its upper bound for systems with repeated uncertainty blocks," *Proc. of 38th IEEE CDC*, Phoenix, Arizona, 4511–4516, 1999.
- [15] O. Toker, H. Özbay, "Complexity issues in robust stability of linear delay-differential systems," *Math. Control Signals Systems*, 9, no 4, 386–400, 1996.
- [16] O. Toker, H. Özbay, "On the complexity of purely complex μ computation and related problems in multidimensional systems," *IEEE Trans. Automat. Control*, 43, no 3, 409–414, 1998.
- [17] O. Toker, B. de Jager, "Conservatism of the standard upper bound test: is $\sup(\bar{\mu}/\mu)$ finite? is it bounded by 2?," *Open problems in mathematical systems and control theory* (V. D. Blondel, E. D. Sontag, M. Vidyasagar, J.C. Willems eds), Springer-Verlag London, 1999.
- [18] S. Treil, "The gap between complex structured singular value μ and its upper bound is infinite," 1999. Available online at <http://www.math.msu.edu/treil/papers/mu/mu-abs.html>
- [19] A. Trofino, "Parameter dependent Lyapunov functions for a class of uncertain linear systems: a LMI approach," *Proc. of 38th IEEE CDC*, Phoenix, Arizona, 2341–2346, 1999.
- [20] A. Trofino, C. E. de Souza, "Bi-quadratic stability of uncertain linear systems," *Proc. of 38th IEEE CDC*, Phoenix, Arizona, 1999.
- [21] L. Xie, S. Shishkin, M. Fu, "Piecewise Lyapunov functions for robust stability of linear time-varying systems," *Syst. Contr. Lett.*, 31, no 3, 165–171, 1997.

Problem 10.2

When is a pair of matrices stable?

Vincent D. Blondel, Jacques Theys

Department of Mathematical Engineering

University of Louvain

Avenue Georges Lematre, 4

B-1348 Louvain-la-Neuve

Belgium

`blondel@inma.ucl.ac.be`, `theys@inma.ucl.ac.be`

John N. Tsitsiklis

Laboratory for Information and Decision Systems

Massachusetts Institute of Technology

Cambridge, MA 02139

USA

`jnt@mit.edu`

1 STABILITY OF ALL PRODUCTS

We consider problems related to the stability of sets of matrices. Let Σ be a finite set of $n \times n$ real matrices. Given a system of the form

$$x_{t+1} = A_t x_t \quad t = 0, 1, \dots$$

suppose that it is known that $A_t \in \Sigma$, for each t , but that the exact value of A_t is not a priori known because of exogenous conditions or changes in the operating point of the system. Such systems can also be thought of as a time-varying systems. We say that such a system is *stable* if

$$\lim_{t \rightarrow \infty} x_t = 0$$

for all initial states x_0 and all sequences of matrix products. This condition is equivalent to the requirement

$$\lim_{t \rightarrow \infty} A_{i_t} \cdots A_{i_1} A_{i_0} = 0$$

for all infinite sequences of indices. Sets of matrices that satisfy this condition are said to be *stable*.

Problem 1. Under what conditions is a given set of matrices stable?

Condition for stability are trivial for matrices of dimension one (all scalar must be of magnitude strictly less than one), and are well-known for sets that contain only one matrix (the eigenvalues of the matrix must be of magnitude strictly less than one). We are asking stability conditions for more general cases.

The matrices in the set must of course have all their eigenvalues of magnitude strictly less than one. This condition does not suffice in general as it is possible to obtain an unstable dynamical system by switching between two stable linear dynamics. Consider, for instance, the matrices

$$A_0 = \alpha \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \text{ and } A_1 = \alpha \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

These matrices are stable iff $|\alpha| < 1$. Consider, then, the product

$$A_0 A_1 = \alpha^2 \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$$

It is immediate to verify that the stability of this matrix is equivalent to the condition $|\alpha| < ((2/(3 + 5^{1/2}))^{1/2} = 0.618$ and so the stability of A_0, A_1 does not imply that of the set $\{A_0, A_1\}$.

Except for elementary cases, no satisfactory conditions are presently available for checking the stability of sets of matrices. In fact the problem is open even in the case of matrices of dimension two. From a set of m matrices of dimension n , it is easy to construct *two* matrices of dimension nm whose stability is equivalent to that of the original set. Indeed, let $\Sigma = \{A_1, \dots, A_m\}$ be a given set and define $B_0 = \text{diag}(A_1, \dots, A_m)$ and $B_1 = T \otimes I$ where T is a $m \times m$ cyclic permutation matrix, \otimes is the Kronecker matrix product, and I the $n \times n$ identity matrix. Then the stability of the pair of matrices $\{B_0, B_1\}$ is easily seen equivalent to that of Σ (see [3] for a more detailed argument). Our question is thus: *When is a pair of matrices stable?*

Several results are available in the literature for this problem, see, e.g., the Lie algebra condition given in [9]. The conditions presently available are only partly satisfactory in that they are either incomplete (they do not cover all cases), they are approximate (see, e.g., [1] and [8]), or they are not effective. We say that a problem is *effectively decidable* (or, *decidable*) if there is an algorithm that, upon input of the data associated with an instance of the problem, provides a yes-no answer after a finite amount of computation. The precise definition of what is meant by an *algorithm* is not critical; most algorithm models proposed so far are known to be equivalent from the point of view of their computing capabilities, and they also coincide with

the intuitive notion of what can be effectively achieved (see [10] for a general description of decidability, and [4] for a survey on decidability in systems and control). Problem 1 can thus be made more explicit by asking for an effective decision algorithm for stability of arbitrary finite sets. Problems similar to this one are known to be undecidable (see, e.g. [2] and [3]); also, attempts (including by the authors of this contribution) of finding such an algorithm have so far failed, we therefore risk the conjecture:

Conjecture 1: The problem of determining if a given pair of matrices is stable is undecidable.

2 STABILITY OF ALL PERIODIC PRODUCTS

Problem 1 is related to the generalized spectral radius of sets of matrices, a notion that generalizes to sets of matrices the usual notion of spectral radius of a single matrix. Let $\rho(A)$ denote the *spectral radius* of a real matrix A ,

$$\rho(A) := \max\{|\lambda| : \lambda \text{ is an eigenvalue of } A\}.$$

The *generalized spectral radius* $\rho(\Sigma)$ of a finite set of matrices Σ is defined in [7] by

$$\rho(\Sigma) = \limsup_{k \rightarrow \infty} \rho_k(\Sigma),$$

where for each $k \geq 1$

$$\rho_k(\Sigma) = \sup\{(\rho(A_1 A_2 \cdots A_k))^{1/k} : \text{each } A_i \in \Sigma\}.$$

When Σ consist of just one single matrix, this quantity is equal to the usual spectral radius. Moreover, it is easy to see that, as for the single matrix case, the stability of the set Σ is equivalent to the condition $\rho(\Sigma) < 1$, and so problem 1 is the problem of finding effective conditions on Σ for $\rho(\Sigma) < 1$.

It is conjectured in [11] that the equality $\rho(\Sigma) = \rho_k(\Sigma)$ always occur for some finite k . This conjecture, known as the *finiteness conjecture*, can be restated by saying that, if a set of matrices Σ is unstable, then there exists a finite unstable product, i.e., if $\rho(\Sigma) \geq 1$, then there exists some $k \geq 1$ and $A_i \in \Sigma$ ($i = 1, \dots, k$) such that

$$\rho(A_1 A_2 \cdots A_k) \geq 1.$$

The existence of a finite unstable product is equivalent to the existence of an infinite *periodic* product that does not converge to zero. We say that a set of matrices is *periodically stable* if all infinite *periodic* products of matrices taken in the set converge to zero. Stability clearly implies periodic stability; according to the finiteness conjecture, the converse is also true. The conjecture has been proved to be false in [6]. A simple counterexample is provided

in [5], where it is shown that there are uncountably many values of the real parameters a and b for which the pair of matrices

$$a \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, b \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$$

is not stable but is periodically stable. Since stability and periodic stability are not equivalent, the following question naturally arises.

Problem 2: Under what conditions is a given finite set of matrices periodically stable?

BIBLIOGRAPHY

- [1] N. E. Barabanov, "A method for calculating the Lyapunov exponent of a differential inclusion," *Avtomat. i Telemekh.*, 4: 53-58, 1989; translation in *Automat. Remote Control* 50:4, part 1, 475-479, 1989.
- [2] V. D. Blondel and J. N. Tsitsiklis, "When is a pair of matrices mortal?", *Information Processing Letters*, 63:5, 283-286, 1997.
- [3] V. D. Blondel and J. N. Tsitsiklis, "The boundedness of all products of a pair of matrices is undecidable," *Systems and Control Letters*, 41(2):135-140, 2000.
- [4] V. D. Blondel and J. N. Tsitsiklis, "A survey of computational complexity results in systems and control," *Automatica*, 36(9), 1249-1274, 2000.
- [5] V. D. Blondel, J. Theys, and A. A. Vladimirov, "An elementary counterexample to the finiteness conjecture," forthcoming in *SIAM Journal of Matrix Analysis*.
- [6] T. Bousch and J. Mairesse, "Asymptotic height optimization for topological IFS, Tetris heaps, and the finiteness conjecture," *J. Amer. Math. Soc.*, 15, 77-111, 2002.
- [7] I. Daubechies and J. C. Lagarias, "Sets of matrices all infinite products of which converge," *Linear Algebra Appl.*, 162, 227-263, 1992.
- [8] G. Gripenberg, "Computing the joint spectral radius," *Linear Algebra Appl.*, 234, 43-60, 1996.
- [9] L. Gurvits, "Stability of discrete linear inclusion," *Linear Algebra Appl.*, 231, 47-85, 1995.
- [10] J. E. Hopcroft and J. D. Ullman, *Introduction to Automata Theory, Languages, and Computation*, Addison-Wesley, Reading, MA, 1979.

- [11] J. C. Lagarias and Y. Wang, "The finiteness conjecture for the generalized spectral radius of a set of matrices," *Linear Algebra Appl.*, 214, 17-42, 1995.
- [12] C. H. Papadimitriou, *Computational Complexity*, Addison-Wesley, Reading, 1994.
- [13] J. N. Tsitsiklis, "The stability of the products of a finite set of matrices," In: *Open Problems in Communication and Computation*, Springer-Verlag, 1987.
- [14] J. N. Tsitsiklis and V. D. Blondel, "Spectral quantities associated with pairs of matrices are hard -when not impossible- to compute and to approximate," *Mathematics of Control, Signals, and Systems*, 10, 31-40, 1997.

Problem 10.3

Freeness of multiplicative matrix semigroups

Vincent D. Blondel

Department of Mathematical Engineering
University of Louvain
Avenue Georges Lemaitre, 4
B-1348 Louvain-la-Neuve
Belgium
`blondel@inma.ucl.ac.be`

Julien Cassaigne

Institut de Mathématiques de Luminy
UPR 9016 - Campus de Luminy, Case 907
13288 Marseille Cedex 9
France
`upr9016@iml.univ-mrs.fr`

Juhani Karhumäki

Dept. of Math. & TUCS
University of Turku
FIN-20014 Turku
Finland
`karhumak@cs.utu.fi`

1 DESCRIPTION OF THE PROBLEM

Matrices play a major role in control theory. In this note, we consider a decidability question for finitely generated multiplicative matrix semigroups. Such semigroups arise, for example, when considering switched linear systems. We consider embeddings of the free semigroup $\Sigma^+ = \{a_0, \dots, a_{k-1}\}^+$ into the multiplicative semigroup of 2×2 matrices over nonnegative integers \mathbf{N} :

$$\varphi : \Sigma^+ \hookrightarrow M_{2 \times 2}(\mathbf{N}).$$

For a two generator semigroup, i.e., $\Sigma^+ = \{a, b\}^+$, such embeddings are defined, for example, by mappings:

$$\varphi_1 : \begin{array}{l} a \mapsto \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \\ b \mapsto \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \end{array} \quad \text{and} \quad \varphi_2 : \begin{array}{l} a \mapsto \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \\ b \mapsto \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \end{array}. \quad (1)$$

Actually, φ_1 provides an embedding of the two generator free group into the multiplicative semigroup of unimodular matrices, e.g., into $SL(2, \mathbf{N})$. The embedding φ_2 , in turn, directly extends to all finitely generated free semigroups. Indeed, the mapping

$$\varphi_i : a_i \mapsto \begin{pmatrix} k & i \\ 0 & 1 \end{pmatrix} \quad \text{for } i = 0, \dots, k-1$$

yields an embedding

$$\{a_0, \dots, a_{k-1}\}^+ \hookrightarrow M_{2 \times 2}(\mathbf{N}). \quad (2)$$

To see this, it is enough to verify that

$$\varphi_i(w) = \begin{pmatrix} k^{|w|} & k(w) \\ 0 & 1 \end{pmatrix},$$

where $k(w)$ denotes the number represented in base k by the word $w \in \{a_0, \dots, a_{k-1}\}^+$ under the identification: a_i corresponds the digit i . Embeddings of countably generated free semigroups are obtained by employing a morphism $\{a_0, a_1, \dots\}^+ \hookrightarrow \{a, b\}^+$, given, for example, by the mapping $\tau : a_i \mapsto a^i b$. Then $\varphi_2 \circ \tau$ yields a required embedding.

In the above examples it is easy to check, as we did for φ_i , $i \geq 2$, that the mappings are indeed embeddings. In general, the situation is strikingly different. In fact, we formulate:

Problem 1: Is it decidable whether a given morphism $\varphi : \Sigma^+ \rightarrow M_{2 \times 2}(\mathbf{N})$ is an embedding, or equivalently, whether a finite set $X = \{A_0, \dots, A_{k-1}\}$ of 2×2 matrices over \mathbf{N} is a free generating set of X^+ ?

Problem 1 deserves two comments. First, we could consider matrices over rational numbers rather than nonnegative integers. This variant is -as it is not too difficult to see- equivalent to the case where matrices are integer matrices. Second, the problem is open even if only two matrices are considered:

Problem 2: Is it decidable whether the multiplicative semigroup generated by two 2×2 matrices over \mathbf{N} is free?

Of course, the nontrivial part of problem 2 is the case when the semigroup is of rank 2. In many concrete examples, the freeness is easy to conclude, as we saw. Amazingly, however, the problem remains even if the matrices are upper-triangular, as is φ_2 above.

2 MOTIVATION AND HISTORY

The importance of problem 1 should be obvious, without any further motivation. Indeed, product of matrices is one of the most fundamental operations in mathematics. In linear algebra it witnesses the composition of linear mappings, and in automata theory it defines the behavior of finite automaton, cf. [7], to mention just two examples. However, the importance of Problem 1 goes far beyond these general reasons.

The existence of embeddings like (2) have been known for a long time. Already in the 1920s J. Neilsen [12] was using these when studying free groups. Such embeddings are extremely useful for both the theories involved. In one hand, these can be used to transfer results on words into those of matrices. The undecidability is an example of a property that is natural and common in the theory of words, and translatable to matrices via these embeddings. This, indeed, is essential in the spirit of this note.

On the other hand, there exist many deep results on matrices that have turned out useful for solving problems on words. A splendid example is Hilbert Bases Theorem, which implies -again via above embeddings- a fundamental compactness property of word equations, so-called *Ehrenfeucht Compactness Property*, cf. [5].

According to the knowledge of the author, the problems mentioned were first discussed in [10], where problem 1 was explicitly stated, and its variant for 3×3 matrices over \mathbf{N} was shown to be undecidable. In [4] the undecidability was extended to upper-triangular 3×3 matrices over \mathbf{N} , and moreover problem 2 was formulated.

Similar problems on matrices have been considered much earlier. Among the oldest results is a remarkable paper by M. Paterson [13], where he shows that it is undecidable whether the multiplicative semigroup generated by a finite set of 3×3 integer matrices contains the zero matrix. In other words, *the mortality problem*, cf. [16], for 3×3 integer matrices is undecidable. According to the current knowledge, it remains undecidable even in the cases when a finite set consists of only seven 3×3 integer matrices or of only two 21×21 integer matrices, cf. [11] and [8, 3, 2]. For 2×2 matrices, the mortality problem is open.

The above undecidability results can be interpreted as follows. First, the existence of the zero element in a two generator (matrix) semigroup is undecidable, cf. [3]. Second, it is also undecidable whether some composition of an effectively given finite set of linear transformation of Euclidian space \mathbf{R}^3 equals to the zero mapping.

The above motivates a related question: is it decidable whether a finitely generated semigroup contains the unit element? In terms of matrices, we state:

Problem 3: Is it decidable whether the multiplicative semigroup S gener-

ated by a finite set of $n \times n$ integer matrices contains the unit matrix?

For $n = 2$ this is shown to be decidable in the case of two matrices in [11], and in the case of an arbitrary number of matrices in [6], but in general the problem is open. A related problem, also open at the moment, asks whether the semigroup S contains a diagonal matrix. In this context, the following example is of interest.

Example 1: For two morphisms $h, g : \{a_0, \dots, a_{k-1}\}^+ \rightarrow \{2, 3\}^+$ define the matrices

$$M(i) = \begin{pmatrix} 10^{|h(a_i)|} & 10^{|h(a_i)|} - 10^{|g(a_i)|} & h(a_i) - g(a_i) \\ 0 & 10^{|g(a_i)|} & g(a_i) \\ 0 & 0 & 1 \end{pmatrix}$$

for $i = 0, \dots, k-1$. A straightforward computation shows that, for any $w = a_{i_1} \dots a_{i_t}$:

$$M(i_1) \dots M(i_t) = \begin{pmatrix} 10^{|h(w)|} & 10^{|h(w)|} - 10^{|g(w)|} & h(w) - g(w) \\ 0 & 10^{|g(w)|} & g(w) \\ 0 & 0 & 1 \end{pmatrix}.$$

Consequently, due to the undecidability of Post Correspondence Problem, cf. [14], it is also undecidable whether the multiplicative semigroup generated by a finite set of 3×3 integer matrices contains a matrix of the form

$$\begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & \delta \\ 0 & 0 & \gamma \end{pmatrix}.$$

We do not know how to get rid of δ .

3 AVAILABLE RESULTS

Due to the embedding $\Sigma^+ \hookrightarrow M_{2 \times 2}(\mathbf{N})$, one way to view Problem 1 is to consider it as an extension of the problem asking to decide whether a finite set of words of Σ^+ generates a free subsemigroup of Σ^+ . This problem is basic in the theory of codes, cf. [1]. It is decidable, even efficiently, as it is not too difficult to see, cf. e.g. [15].

Very little seems to be known about problem 1. As we already said the corresponding problem for 3×3 matrices is undecidable, the proof being a reduction to Post Correspondence Problem, as in example 1. A bit more intriguing reduction techniques were used in [4] in order to show that the undecidability holds even for upper-triangular 3×3 matrices over \mathbf{N} . A fundamental observation in these proofs is that the product monoid $\Sigma^+ \times \Delta^+$ is not embeddable only into the semigroup of matrices of dimension four but also into that of dimension three. In other words, there exists an embedding

$$\varphi : \Sigma^+ \times \Delta^+ \hookrightarrow M_{3 \times 3}(\mathbf{N}).$$

On the other hand, as also shown in [4], there does not exist any such embedding into the semigroup of 2×2 matrices, i.e., into $M_{2 \times 2}(\mathbf{N})$.

Problem 2 was formulated after vain attempts to solve it in [4]. Actually, even the case when both of the matrices are upper-triangular, i.e., of the form

$$\begin{pmatrix} \alpha & \beta \\ 0 & \gamma \end{pmatrix}$$

remained unanswered. Only several sufficient (effective) conditions for the freeness were established. Even for some very particular cases, we do not know if the semigroup is free. In particular, we do not know whether the matrices

$$A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 3 & 5 \\ 0 & 5 \end{pmatrix}$$

generate a free semigroup. We only know that these matrices do not satisfy any equation where both sides are of length at most 20.

As a conclusion, we hope, we have been able to point out a problem that is not only very simply formulated, but also fundamental and challenging.

BIBLIOGRAPHY

- [1] J. Berstel and D. Perrin, *Theory of Codes*, Academic Press, 1986.
- [2] V. Blondel and J. Tsitsiklis, “When is a pair of matrices mortal,” *Inform. Process. Lett.* 63, 283–286, 1997.
- [3] J. Cassaigne and J. Karhumäki, “Examples of undecidable problems for 2-generator matrix semigroups,” *Theoret. Comput. Sci.* 204, 29–34, 1998.
- [4] J. Cassaigne, T. Harju and J. Karhumäki, “On the undecidability of freeness of matrix semigroups,” *Int. J. Algebra Comp.* 9, 295–305, 1999.
- [5] C. Choffrut and J. Karhumäki, “Combinatorics of words,” In: G. Rozenberg and A. Salomaa (eds.), *Handbook of Formal Languages*, vol. I, Springer, 1997, 329–438.
- [6] C. Choffrut and J. Karhumäki, “On decision problems on integer matrices,” manuscript.
- [7] S. Eilenberg, *Automata, Languages and Machines*, vol. A, Academic Press, 1974.
- [8] V. Halava and T. Harju, “Mortality in matrix semigroups,” *Amer. Math. Monthly* 108, 649–653, 2001.

- [9] T. Harju and J. Karhumäki, Morphisms, In: G. Rozenberg and A. Salomaa (eds.), *Handbook of Formal Languages*, vol. I, Springer, 1997, 439–510.
- [10] D. Klarner, J.-C. Birget and W. Satterfield, “On the undecidability of the freeness of integer matrix semigroups,” *Int. J. Algebra Comp.* 1, 223–226, 1991.
- [11] F. Mazoit, “Autour de quelques problèmes de décidabilité sur des semi-groupes de matrices,” manuscript, 1998.
- [12] J. Nielsen, “Die Isomorphismengruppe der freien Gruppen,” *Math. Ann.* 91, 169–209, 1924.
- [13] M. Paterson, “Unsolvability in 3×3 matrices,” *Studies Appl. Math.* 49, 105–107, 1970.
- [14] E. Post, “A variant of recursively unsolvable problem,” *Bull. Amer. Math. Soc.* 52, 264–268, 1949.
- [15] W. Rytter, “The space complexity of the unique decipherability problem,” *Inform. Process. Lett.* 23, 1–3, 1986.
- [16] P. Schultz, “Mortality of 2×2 matrices,” *Amer. Math. Monthly* 84, 463–464, 1977.

Problem 10.4

Vector-valued quadratic forms in control theory

Francesco Bullo

Coordinated Science Laboratory
University of Illinois
Urbana-Champaign, IL 61801
USA
bullo@uiuc.edu

Jorge Cortés

Coordinated Science Laboratory
University of Illinois
Urbana-Champaign, IL 61801
USA
jcortes@uiuc.edu

Andrew D. Lewis

Mathematics & Statistics
Queen's University
Kingston, ON K7L 3N6
Canada
andrew@mast.queensu.ca

Sonia Martínez

Matemática Aplicada IV
Universidad Politécnica de Cataluña
Barcelona, 08800
Spain
soniam@mat.upc.es

1 PROBLEM STATEMENT AND HISTORICAL REMARKS

For finite dimensional \mathbb{R} -vector spaces U and V , we consider a symmetric bilinear map $B: U \times U \rightarrow V$. This then defines a quadratic map $Q_B: U \rightarrow V$ by $Q_B(u) = B(u, u)$. Corresponding to each $\lambda \in V^*$ is a \mathbb{R} -valued quadratic form λQ_B on U defined by $\lambda Q_B(u) = \lambda \cdot Q_B(u)$. B is *definite* if there

exists $\lambda \in V^*$ so that λQ_B is positive-definite. B is *indefinite* if for each $\lambda \in V^* \setminus \text{ann}(\text{image}(Q_B))$, λQ_B is neither positive nor negative-semidefinite, where ann denotes the annihilator.

Given a symmetric bilinear map $B: U \times U \rightarrow V$, the problems we consider are as follows.

- i. Find necessary and sufficient conditions characterizing when Q_B is surjective.*
- ii. If Q_B is surjective and $v \in V$, design an algorithm to find a point $u \in Q_B^{-1}(v)$.*
- iii. Find necessary and sufficient conditions to determine when B is indefinite.*

From the computational point of view, the first two questions are the more interesting ones. Both can be shown to be NP-complete, whereas the third one can be recast as a semidefinite programming problem.¹ Actually, our main interest is in a geometric characterization of these problems. Section 4 below constitutes an initial attempt to unveil the essential geometry behind these questions. By understanding the geometry of the problem properly, one may be lead to simple characterizations like the one presented in Proposition 3, which turn out to be checkable in polynomial time for certain classes of quadratic mappings.

Before we comment on how our problem impinges on control theory, let us provide some historical context for it as a purely mathematical one. The classification of \mathbb{R} -valued quadratic forms is well understood. However, for quadratic maps taking values in vector spaces of dimension two or higher, the classification problem becomes more difficult. The theory can be thought of as beginning with the work of Kronecker, who obtained a finite classification for pairs of symmetric matrices. For three or more symmetric matrices, that the classification problem has an uncountable number of equivalence classes for a given dimension of the domain follows from the work of Kac [12]. For quadratic forms, in a series of papers Dines (see [8] and references cited therein) investigated conditions when a finite collection of \mathbb{R} -valued quadratic maps were simultaneously positive-definite. The study of vector-valued quadratic maps is ongoing. A recent paper is [13], to which we refer for other references.

2 CONTROL THEORETIC MOTIVATION

Interestingly, and perhaps not obviously, vector-valued quadratic forms come up in a variety of places in control theory. We list a few of these here.

¹We thank an anonymous referee for these observations.

Optimal control: Agračhev [2] explicitly realizes second-order conditions for optimality in terms of vector-valued quadratic maps. The geometric approach leads naturally to the consideration of vector-valued quadratic maps, and here the necessary conditions involve definiteness of these maps. Agračhev and Gamkrelidze [1, 3] look at the map $\lambda \mapsto \lambda Q_B$ from V^* into the set of vector-valued quadratic maps. Since λQ_B is a \mathbb{R} -valued quadratic form, one can talk about its *index* and *rank* (the number of -1 's and nonzero terms, respectively, along the diagonal when the form is diagonalized). In [1, 3] the topology of the surfaces of constant index of the map $\lambda \mapsto \lambda Q_B$ is investigated.

Local controllability: The use of vector-valued quadratic forms arises from the attempt to arrive at feedback-invariant conditions for controllability. Basto-Gonçalves [6] gives a second-order sufficient condition for local controllability, one of whose hypotheses is that a certain vector-valued quadratic map be indefinite (although the condition is not stated in this way). This condition is somewhat refined in [11], and a necessary condition for local controllability is also given. Included in the hypotheses of the latter is the condition that a certain vector-valued quadratic map be definite.

Control design via power series methods and singular inversion:

Numerous control design problems can be tackled using power series and inversion methods. The early references [5, 9] show how to solve the optimal regulator problem and the recent work in [7] proposes local steering algorithms. These strong results apply to linearly controllable systems, and no general methods are yet available under only second-order sufficient controllability conditions. While for linearly controllable systems the classic inverse function theorem suffices, the key requirement for second-order controllable systems is the ability to check surjectivity and compute an inverse function for certain vector-valued quadratic forms.

Dynamic feedback linearisation: In [14] Sluis gives a necessary condition for the dynamic feedback linearization of a system

$$\dot{x} = f(x, u), \quad x \in \mathbb{R}^n, \quad u \in \mathbb{R}^m.$$

The condition is that for each $x \in \mathbb{R}^n$, the set $D_x = \{f(x, u) \in T_x \mathbb{R}^n \mid u \in \mathbb{R}^m\}$ admits a *ruling*, that is, a foliation of D_x by lines. Some manipulations with differential forms turns this necessary condition into one involving a symmetric bilinear map B . The condition, it turns out, is that $Q_B^{-1}(0) \neq \{0\}$. This is shown by Agračhev [1] to generically imply that Q_B is surjective.

3 KNOWN RESULTS

Let us state a few results along the lines of our problem statement that are known to the authors. The first is readily shown to be true (see [11] for the proof). If X is a topological space with subsets $A \subset S \subset X$, we denote by $\text{int}_S(A)$ the interior of A relative to the induced topology on S . If $S \subset V$,

$\text{aff}(S)$ and $\text{conv}(S)$ denote, respectively, the affine hull and the convex hull of S .

Proposition 1: *Let $B: U \times U \rightarrow V$ be a symmetric bilinear map with U and V finite-dimensional. The following statements hold:*

- (i) *B is indefinite if and only if $0 \in \text{int}_{\text{aff}(\text{image}(Q_B))}(\text{conv}(\text{image}(Q_B)))$;*
- (ii) *B is definite if and only if there exists a hyperplane $P \subset V$ so that $\text{image}(Q_B) \cap P = \{0\}$ and so that $\text{image}(Q_B)$ lies on one side of P ;*
- (iii) *if Q_B is surjective then B is indefinite.*

The converse of (iii) is false. The quadratic map from \mathbb{R}^3 to \mathbb{R}^3 defined by $Q_B(x, y, z) = (xy, xz, yz)$ may be shown to be indefinite but not surjective. Agračhev and Sarychev [4] prove the following result. We denote by $\text{ind}(Q)$ the index of a quadratic map $Q: U \rightarrow \mathbb{R}$ on a vector space U .

Proposition 2: *Let $B: U \times U \rightarrow V$ be a symmetric bilinear map with V finite-dimensional. If $\text{ind}(\lambda Q_B) \geq \dim(V)$ for any $\lambda \in V^* \setminus \{0\}$ then Q_B is surjective.*

This sufficient condition for surjectivity is not necessary. The quadratic map from \mathbb{R}^2 to \mathbb{R}^2 given by $Q_B(x, y) = (x^2 - y^2, xy)$ is surjective, but does not satisfy the hypotheses of Proposition 2.

4 PROBLEM SIMPLIFICATION

One of the difficulties with studying vector-valued quadratic maps is that they are somewhat difficult to get one's hands on. However, it turns out to be possible to simplify their study by a reduction to a rather concrete problem. Here we describe this process, only sketching the details of how to go from a given symmetric bilinear map $B: U \times U \rightarrow V$ to the reformulated end problem. We first simplify the problem by imposing an inner product on U and choosing an orthonormal basis so that we may take $U = \mathbb{R}^n$.

We let $\text{Sym}_n(\mathbb{R})$ denote the set of symmetric $n \times n$ matrices with entries in \mathbb{R} . On $\text{Sym}_n(\mathbb{R})$ we use the canonical inner product

$$\langle \mathbf{A}, \mathbf{B} \rangle = \text{tr}(\mathbf{AB}).$$

We consider the map $\pi: \mathbb{R}^n \rightarrow \text{Sym}_n(\mathbb{R})$ defined by $\pi(\mathbf{x}) = \mathbf{x}\mathbf{x}^t$, where t denotes transpose. Thus the image of π is the set of positive semidefinite symmetric matrices of rank at most one. If we identify $\text{Sym}_n(\mathbb{R}) \simeq \mathbb{R}^n \otimes \mathbb{R}^n$, then $\pi(\mathbf{x}) = \mathbf{x} \otimes \mathbf{x}$. Let K_n be the image of π and note that it is a cone of dimension n in $\text{Sym}_n(\mathbb{R})$ having a singularity only at its vertex at the origin. Furthermore, K_n may be shown to be a subset of the hypercone in

$\text{Sym}_n(\mathbb{R})$ defined by those matrices A in $\text{Sym}_n(\mathbb{R})$ forming angle $\arccos(\frac{1}{\sqrt{n}})$ with the identity matrix. Thus the ray from the origin in $\text{Sym}_n(\mathbb{R})$ through the identity matrix is an axis for the cone K_N . In algebraic geometry, the image of K_n under the projectivization of $\text{Sym}_n(\mathbb{R})$ is known as the *Veronese surface* [10], and as such is well-studied, although perhaps not along lines that bear directly on the problems of interest in this article.

We now let $B: \mathbb{R}^n \times \mathbb{R}^n \rightarrow V$ be a symmetric bilinear map with V finite-dimensional. Using the universal mapping property of the tensor product, B induces a linear map $\tilde{B}: \text{Sym}_n(\mathbb{R}) \simeq \mathbb{R}^n \otimes \mathbb{R}^n \rightarrow V$ with the property that $\tilde{B} \circ \pi = B$. The dual of this map gives an injective linear map $\tilde{B}^*: V^* \rightarrow \text{Sym}_n(\mathbb{R})$ (here we assume that the image of B spans V). By an appropriate choice of inner product on V , one can render the embedding \tilde{B}^* an isometric embedding of V in $\text{Sym}_n(\mathbb{R})$. Let us denote by L_B the image of V under this isometric embedding. One may then show that with these identifications, the image of Q_B in V is the orthogonal projection of K_n onto the subspace L_B . Thus we reduce the problem to one of orthogonal projection of a canonical object, K_n , onto a subspace in $\text{Sym}_n(\mathbb{R})$! To simplify things further, we decompose L_B into a component along the identity matrix in $\text{Sym}_n(\mathbb{R})$ and a component orthogonal to the identity matrix. However, the matrices orthogonal to the identity are readily seen to simply be the traceless $n \times n$ symmetric matrices. Using our picture of K_n as a subset of a hypercone having as an axis the ray through the identity matrix, we see that questions of surjectivity, indefiniteness, and definiteness of B impact only on the projection of K_n onto that component of L_B orthogonal to the identity matrix.

The following summarizes the above discussion.

The problem of studying the image of a vector-valued quadratic form can be reduced to studying the orthogonal projection of $K_n \subset \text{Sym}_n(\mathbb{R})$, the unprojectivized Veronese surface, onto a subspace of the space of traceless symmetric matrices.

This is, we think, a beautiful interpretation of the study of vector-valued quadratic mappings, and will surely be a useful formulation of the problem. For example, with it one easily proves the following result.

Proposition 3: *If $\dim(U) = \dim(V) = 2$ with $B: U \times U \rightarrow V$ a symmetric bilinear map, then Q_B is surjective if and only if B is indefinite.*

BIBLIOGRAPHY

- [1] A. A. Agrache, "The topology of quadratic mappings and Hessians of smooth mappings," *J. Soviet Math.*, 49(3):990–1013, 1990.

- [2] A. A. Agrachev, "Quadratic mappings in geometric control theory," *J. Soviet Math.*, 51(6):2667–2734, 1990.
- [3] A. A. Agrachev and R. V. Gamkrelidze, "Quadratic mappings and vector functions: Euler characteristics of level sets," *J. Soviet Math.*, 55(4):1892–1928, 1991.
- [4] A. A. Agrachev and A. V. Sarychev, "Abnormal sub-Riemannian geodesics: Morse index and rigidity," *Ann. Inst. H. Poincaré. Anal. Non Linéaire*, 13(6):635–690, 1996.
- [5] È. G. Al'brikht, "On the optimal stabilization of nonlinear systems," *J. Appl. Math. and Mech.*, 25:1254–1266, 1961.
- [6] J. Basto-Gonçalves, "Second-order conditions for local controllability," *Systems Control Lett.*, 35(5):287–290, 1998.
- [7] W. T. Cerven and F. Bullo, "Constructive controllability algorithms for motion planning and optimization," *IEEE Trans Automat Control*, Forthcoming.
- [8] L. L. Dines, "On linear combinations of quadratic forms," *Bull. Amer. Math. Soc. (N.S.)*, 49:388–393, 1943.
- [9] A. Halme, "On the nonlinear regulator problem," *J. Optim. Theory Appl.*, 16(3-4):255–275, 1975.
- [10] J. Harris, *Algebraic Geometry: A First Course*, Number 133 in Graduate Texts in Mathematics, Springer-Verlag, New York Heidelberg Berlin, 1992.
- [11] R. M. Hirschorn and A. D. Lewis, "Geometric local controllability: Second-order conditions," preprint, February 2002.
- [12] V. G. Kac, "Root systems, representations of quivers and invariant theory," In: *Invariant Theory*, number 996 in Lecture Notes in Mathematics, pages 74–108, Springer-Verlag, New York-Heidelberg-Berlin, 1983.
- [13] D. B. Leep and L. M. Schueller, "Classification of pairs of symmetric and alternating bilinear forms," *Exposition. Math.*, 17(5):385–414, 1999.
- [14] W. M. Sluis, "A necessary condition for dynamic feedback linearization," *Systems Control Lett.*, 21(4):277–283, 1993.

Problem 10.5

Nilpotent bases of distributions

Henry G. Hermes

Department of Mathematics
University of Colorado at Boulder,
Boulder, CO 80309
USA
hermes@euclid.colorado.edu

Matthias Kawski¹

Department of Mathematics and Statistics
Arizona State University
Tempe, AZ 85287-1804
USA
kawski@asu.edu

1 DESCRIPTION OF THE PROBLEM

When modeling controlled dynamical systems one commonly chooses individual control variables u_1, \dots, u_m that appear *natural* from a physical, or practical point of view. In the case of nonlinear models evolving on \mathbf{R}^n (or more generally, an analytic manifold M^n) that are affine in the control, such a choice corresponds to selecting vector fields $f_0, f_1, \dots, f_m: M \mapsto TM$, and the system takes the form

$$\dot{x} = f_0(x) + \sum_{k=1}^m u_k f_k(x). \quad (1)$$

From a geometric point of view such a choice appears arbitrary, and the natural objects are not the vector fields themselves but their linear span. Formally, for a set $\mathcal{F} = \{v_1, \dots, v_m\}$ of vector fields define the *distribution spanned by \mathcal{F}* as $\Delta_{\mathcal{F}}: p \mapsto \{c_1 v_1(p) + \dots + c_m v_m(p) : c_1, \dots, c_m \in \mathbf{R}\} \subseteq T_p M$. For *systems with drift* f_0 , the geometric object is the map $\tilde{\Delta}_{\mathcal{F}}(x) = \{f_0(x) + c_1 f_1(x) + \dots + c_m f_m(x) : c_1, \dots, c_m \in \mathbf{R}\}$ whose image at every point x is an affine subspace of $T_x M$. The geometric character of the distribution is captured by its invariance under the *group of feedback transformations*.

¹Supported in part by NSF-grant DMS 00-72369.

In traditional notation (here formulated for systems with drift) these are (analytic) maps (defined on suitable subsets) $\alpha: M^n \times \mathbf{R}^m \mapsto \mathbf{R}^m$ such that for each fixed $x \in M^n$ the map $v \mapsto \alpha(x, v)$ is affine and invertible. Customarily, one identifies $\alpha(x, \cdot)$ with a matrix and writes

$$u_k(x) = \alpha_{0k}(x) + v_1 \alpha_{1k}(x) + \dots v_m \alpha_{mk}(x) \quad \text{for } k = 1, \dots, m. \quad (2)$$

This transformation of the controls induces a corresponding transformation of the vector fields defined by $\dot{x} = f_0(x) + \sum_{k=1}^m u_k f_k(x) \stackrel{!}{=} g_0(x) + \sum_{k=1}^m v_k g_k(x)$

$$\begin{aligned} g_0(x) &= f_0(x) + \alpha_{01}(x)f_1(x) + \dots \alpha_{0m}(x)f_m(x) \\ g_k(x) &= \alpha_{k1}(x)f_1(x) + \dots \alpha_{km}(x)f_m(x), \quad k = 1, \dots, m \end{aligned} \quad (3)$$

Assuming linear independence of the vector fields such feedback transformations amount to *changes of basis* of the associated distributions. One naturally studies the *orbits* of any given system under this group action, i.e., the collection of equivalent systems. Of particular interest are *normal forms*, i.e., natural distinguished representatives for each orbit. Geometrically (i.e., without choosing local coordinates for the state x), these are characterized by properties of the Lie algebra $L(g_0, g_1, \dots, g_m)$ generated by the vector fields g_k (acknowledging the special role of g_0 if present).

Recall that a Lie algebra L is called nilpotent (solvable) if its central descending series $L^{(k)}$ (derived series $L^{<k>}$) is finite, i.e., there exists $r < \infty$ such that $L^{(r)} = \{0\}$ ($L^{<r>} = \{0\}$). Here $L = L^{(1)} = L^{<1>}$ and inductively $L^{(k+1)} = [L^{(k)}, L^{(1)}]$ and $L^{<k+1>} = [L^{<k>}, L^{<k>}]$.

The main questions of practical importance are:

Problem 1: Find necessary and sufficient conditions for a distribution $\Delta_{\mathcal{F}}$ spanned by a set of analytic vector fields $\mathcal{F} = \{f_1, \dots, f_m\}$ to admit a *basis* of analytic vector fields $\mathcal{G} = \{g_1, \dots, g_m\}$ that generate a Lie algebra $L(g_1, \dots, g_m)$ that has a *desirable structure*, i.e., that is **a.** nilpotent, **b.** solvable, or **c.** finite dimensional.

Problem 2: Describe an algorithm that constructs such a basis \mathcal{G} from a given basis \mathcal{F} .

2 MOTIVATION AND HISTORY OF THE PROBLEM

There is an abundance of mathematical problems, which are hard as given, yet are almost trivial when written in the *right* coordinates. Classical examples of finding the *right coordinates* (or, rather, the right bases) are transformations that *diagonalize* operators in linear algebra and functional analysis. Similarly, every system of (ordinary) differential equation is equivalent (via a choice of local coordinates) to the system $\dot{x}_1 = 1, \dot{x}_2 = 0, \dots, \dot{x}_n = 0$ (in the neighborhood of every point that is not an equilibrium). In control, for many purposes the most convenient form is the controller canonical form

(e.g., in the case of $m = 1$) $\dot{x}_1 = u$ and $\dot{x}_k = x_{k-1}$ for $1 < k \leq n$. Every controllable linear system can be brought into this form via feedback and a linear coordinate change. For control systems that are not equivalent to linear systems, the next best choice is a polynomial cascade system $\dot{x}_1 = u$ and $\dot{x}_k = p_k(x_1, \dots, x_{k-1})$ for $1 < k \leq n$. (Both linear and nonlinear cases have natural multi-input versions for $m > 1$.) What makes such linear or polynomial cascade form so attractive for both analysis and design is that trajectories $x(t, u)$ may be computed from controls $u(t)$ by *quadratures* only, obviating the need to solve nonlinear ODEs. Typical examples include pole placement and path planning [11, 16, 19]. In particular, if the Lie algebra is nilpotent (or similarly nice), the general solution formula for $x(\cdot, u)$ as an exponential Lie series [20] (which generalizes *matrix exponentials* to nonlinear systems) collapses and becomes innately manageable.

It is well-known that a system can be brought into such polynomial cascade form via a coordinate change if and only if the Lie algebra $L(f_1, \dots, f_m)$ is nilpotent [9]. Similar results for solvable Lie algebras are available [1]. This leaves open only the geometric question about when does a distribution admit a nilpotent (or solvable) basis.

3 RELATED RESULTS

In [5] it is shown that for every $2 \leq k \leq (n - 1)$ there is a k -distribution Δ on \mathbf{R}^n that does not admit a solvable basis in a neighborhood of zero. This shows the problems of nilpotent and solvable bases are not trivial.

Geometric properties, such as small-time local controllability (STLC) are, by their very nature, unaffected by feedback transformations. Thus conditions for STLC provide valuable information whether any two systems can be feedback equivalent. Typical such information, generalizing the controllability indices of linear systems theory, is contained in the *growth vector*, that is the dimensions of the *derived distributions* that are defined inductively by $\Delta^{(1)} = \Delta$ and $\Delta^{(k+1)} = \Delta^{(k)} + \{[v, w] : v \in \Delta^{(k)}, w \in \Delta^{(1)}\}$.

Of highest practical interest is the case when the system is (locally) equivalent to a linear system $\dot{x} = Ax + Bu$ (for some choice of local coordinates). Necessary and sufficient conditions for such exact *feedback linearization* together with algorithms for constructing the transformation and coordinates were obtained in the 1980s [6, 7]. The geometric criteria are nicely stated in terms of the involutivity (closedness under Lie bracketing) of the distributions spanned by the sets $\{(\text{ad}^j f_0, f_1) : 0 \leq j \leq k\}$ for $0 \leq k \leq m$.

A necessary condition for exact nilpotentization is based on the observation that every nilpotent Lie algebra contains at least one element that commutes with every other element [4].

A well-studied special case is that of nilpotent systems which can be brought into *chained-form*, compare [16]. This is closely related to *differen-*

tially flat systems, compare [2, 8], which have been the focus of much study in the 1990s. The key property is the existence of an *output function* such that all system variables can be expressed in terms of functions of a finite number of derivatives of this output. This work is more naturally performed using a dual description in terms of exterior differential systems and co-distributions $\Delta^\perp = \{\omega: M \mapsto T^*M : \langle \omega, f \rangle = 0 \text{ for all } f \in \Delta\}$. This description is particularly convenient when working with small co-dimension $n - m$, compare [12] for a recent survey. (Special care needs to be taken at singular points where the dimensions of $\Delta^{(k)}$ are nonconstant.) This language allows one to directly employ the machinery of *Cartan's method of equivalence* [3]. However, a *nice* description of a system in terms of differential forms does not necessarily translate in a straightforward manner into a nice description in terms of vector fields (that generate a finite dimensional, or nilpotent Lie algebra).

Some of the most notable recent progress has been made in the general framework of Goursat distributions, see, e.g., [13, 14, 15, 17, 18, 21] for detailed descriptions, the most recent results and further relevant references.

BIBLIOGRAPHY

- [1] P. Crouch, "Solvable approximations to control systems," *SIAM J. Control & Optim.*, 22, pp. 40-45, 1984.
- [2] M. Fliess, J. Levine, P. Martin, and P. Rouchon, "Some open questions related to flat nonlinear systems," *Open problems in Mathematical Systems and Control Theory*, V. Blondel, E. Sontag, M. Vidyasagar, and J. Willems, eds., Springer, 1999.
- [3] R. Gardener, "The method of equivalence and its applications," *CBMS NSF Regional Conference Series in Applied Mathematics*, SIAM, 58, 1989.
- [4] H. Hermes, A. Lundell, and D. Sullivan, "Nilpotent bases for distributions and control systems," *J. of Diff. Equations*, 55, pp. 385-400, 1984.
- [5] H. Hermes, "Distributions and the lie algebras their bases can generate," *Proc. AMS*, 106, pp. 555-565, 1989.
- [6] R. Hunt, R. Su, and G. Meyer, "Design for multi-input nonlinear systems," *Differential Geometric Control Theory*, R. Brockett, R. Millmann, H. Sussmann, eds., Birkhäuser, pp. 268-298, 1982.
- [7] B. Jakubczyk and W. Respondek, "On linearization of control systems," *Bull. Acad. Polon. Sci. Ser. Sci. Math.*, 28, pp. 517-522, 1980.
- [8] F. Jean, "The car with n trailers: Characterization for the singular configurations," *ESAIM Control Optim. Calc. Var.* 1, pp. 241-266, 1996.

- [9] M. Kowski “Nilpotent Lie algebras of vector fields,” *Journal für die Reine und Angewandte Mathematik*, 188, pp. 1-17, 1988.
- [10] M. Kowski and H. J. Sussmann “Noncommutative power series and formal Lie-algebraic techniques in nonlinear control theory,” *Operators, Systems, and Linear Algebra*, U. Helmke, D. Prätzel-Wolters and E. Zerz, eds., Teubner, 111–128, 1997.
- [11] G. Laffariere and H. Sussmann, “A differential geometric approach to motion planning,” *IEEE International Conference on Robotics and Automation*, pages 1148–1153, Sacramento, CA, 1991.
- [12] R. Montgomery, “A Tour of Subriemannian Geometries, Their Geodesics and Applications,” *AMS Mathematical Surveys and Monographs*, 91, 2002.
- [13] P. Mormul, “Goursat flags, classification of co-dimension one singularities,” *J. Dynamical and Control Systems*, 6, 2000, pp. 311–330, 2000.
- [14] P. Mormul, “Multi-dimensional Caratn prolongation and special k -flags,” technical report, University of Warsaw, Poland, 2002.
- [15] P. Mormul, “Goursat distributions not strongly nilpotent in dimensions not exceeding seven,” *Lecture Notes in Control and Inform. Sci.*, 281, pp. 249–261, Springer, Berlin, 2003.
- [16] R. Murray, “Nilpotent bases for a class of non-integrable distributions with applications to trajectory generation for nonholonomic systems,” *Mathematics of Controls, Signals, and Systems*, 7, pp. 58–75, 1994.
- [17] W. Pasillas-Lpine, and W. Respondek, “On the geometry of Goursat structures,” *ESAIM Control Optim. Calc. Var.* 6, pp. 119–181, 2001.
- [18] W. Respondek, and W. Pasillas-Lpine, “Extended Goursat normal form: a geometric characterization,” In: *Lecture Notes in Control and Inform. Sci.*, 259 pp. 323–338, Springer, London, 2001.
- [19] J. Strumper and P. Krishnaprasad, “Approximate tracking for systems on three dimensional Lie matrix groups via feedback nilpotentization,” *IFAC Symposium Robot Control* (1997).
- [20] H. Sussmann, “A product expansion of the Chen series,” *Theory and Applications of Nonlinear Control Systems*, C. Byrnes and A. Lindquist eds., Elsevier, pp. 323–335, 1986.
- [21] M. Zhitomirskii, “Singularities and normal forms of smooth distributions,” *Banach Center Publ.*, 32, pp. 379-409, 1995.

Problem 10.6

What is the characteristic polynomial of a signal flow graph?

Andrew D. Lewis

Department of Mathematics & Statistics

Queen's University

Kingston, ON K7L 3N6

Canada

andrew@mast.queensu.ca

1 PROBLEM STATEMENT

Suppose one is given signal flow graph \mathcal{G} with n nodes whose branches have gains that are real rational functions (the open loop transfer functions). The gain of the branch connecting node i to node j is denoted R_{ji} , and we write $R_{ji} = \frac{N_{ji}}{D_{ji}}$ as a coprime fraction. The closed-loop transfer function from node i to node j for the closed-loop system is denoted T_{ji} .

The problem can then be stated as follows:

Is there an algorithmic procedure that takes a signal flow graph \mathcal{G} and returns a “characteristic polynomial” $P_{\mathcal{G}}$ with the following properties:

- i. $P_{\mathcal{G}}$ is formed by products and sums of the polynomials N_{ji} and D_{ji} , $i, j = 1, \dots, n$;*
- ii. all closed-loop transfer functions T_{ji} , $i, j = 1, \dots, n$, are analytic in the closed right half-plane (CRHP) if and only if $P_{\mathcal{G}}$ is Hurwitz?*

The gist of condition *i* is that the construction of $P_{\mathcal{G}}$ depends only on the topology of the graph, and not on manipulations of the branch gains. That is, the definition of $P_{\mathcal{G}}$ should not depend on the choice of branch gains R_{ji} , $i, j = 1, \dots, n$. For example, one would be prohibited from factoring polynomials or from computing the GCD of polynomials. This excludes unhelpful solutions of the problem of the form, “Let $P_{\mathcal{G}}$ be the product of the characteristic polynomials of the closed-loop transfer functions T_{ji} , $i, j = 1, \dots, n$.”

2 DISCUSSION

Signal flow graphs for modelling system interconnections are due to Mason [3, 4]. Of course, when making such interconnections, the stability of the interconnection is nontrivially related to the open-loop transfer functions that weight the branches of the signal flow graph. There are at least two things to consider in the course of making an interconnection: (1) is the interconnection *BIBO stable* in the sense that all closed-loop transfer functions between nodes have no poles in the CRHP?; and (2) is the interconnection *well-posed* in the sense that all closed-loop transfer functions between nodes are proper? The problem stated above concerns only the first of these matters. Well-posedness when all branch gains R_{ji} , $i, j = 1, \dots, n$, are proper is known to be equivalent to the condition that the determinant of the graph be a biproper rational function. We comment that other forms of stability for signal flow graphs are possible. For example, Wang, Lee, and Ho [5] consider *internal stability*, wherein not the transfer functions between signals are considered, but rather that all signals in the signal flow graph remain bounded when bounded inputs are provided. Internal stability as considered by Wang, Lee, and Ho and BIBO stability as considered here are different. The source of this difference accounts for the source of the open problem of our paper, since Wang, Lee, and Ho show that internal stability can be determined by an algorithmic procedure like that we ask for for BIBO stability. This is discussed a little further in section 3.

As an illustration of what we are after, consider the single-loop configuration of figure 10.6.1. As is well-known, if we write $R_i = \frac{N_i}{D_i}$, $i = 1, 2$, as coprime

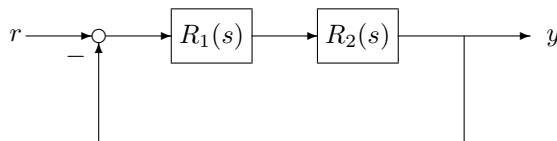


Figure 10.6.1 Single-loop interconnection

fractions, then all closed-loop transfer functions have no poles in the CRHP if and only if the polynomial $P_g = D_1 D_2 + N_1 N_2$ is Hurwitz. Thus P_g serves as the characteristic polynomial in this case. The essential feature of P_g is that one computes it by looking at the topology of the graph, and the exact character of R_1 and R_2 are of no consequence. For example, pole/zero cancellations between R_1 and R_2 are accounted for in P_g .

3 APPROACHES THAT DO NOT SOLVE THE PROBLEM

Let us outline two approaches that yield solutions having one of properties i and ii, but not the other.

The problems of internal stability and well-posedness for signal flow graphs can be handled effectively using the polynomial matrix approach, e.g., [1]. Such an approach will involve the determination of a coprime matrix fractional representation of a matrix of rational functions. This will certainly solve the problem of determining internal stability for any given example. That is, it is possible using matrix polynomial methods to provide an algorithmic construction of a polynomial satisfying property *ii* above. However, the algorithmic procedure will involve computing GCDs of various of the polynomials N_{ji} and D_{ji} , $i, j = 1, \dots, n$. Thus the conditions developed in this manner have to do not only with the topology of the signal flow graph, but also the specific choices for the branch gains, thus violating condition *i* above. The problem we pose demands a simpler, more direct answer to the question of determining when an interconnection is BIBO stable.

Wang, Lee, and He [5] provide a polynomial for a signal flow graph using the determinant of the graph which we denote by Δ_g (see [3, 4]). Specifically, they define a polynomial

$$P = \prod_{(i,j) \in \{1, \dots, n\}^2} D_{ji} \Delta_g. \quad (1)$$

Thus one multiplies the determinant by all denominators, arriving at a polynomial in the process. This polynomial has the property *i* above. However, while it is true that if this polynomial is Hurwitz then the system is BIBO stable, the converse is generally false. Thus property *ii* is not satisfied by P . What is shown in [5] is that all *signals* in the graph are bounded for bounded inputs if and only if P is Hurwitz. This is different from what we are asking here, i.e., that all closed-loop transfer functions have no CRHP poles. It is true that the polynomial P in (1) gives the desired characteristic polynomial for the interconnection of Figure 10.6.1. It is also true that if the signal flow graph has no loops (in this case $\Delta_g = 1$) then the polynomial P of (1) satisfies condition *ii*. We comment that the condition of Wang, Lee, and Ho is the same condition one would obtain by converting (in a specific way) the signal flow graph to a polynomial matrix system, and then ascertaining when the resulting polynomial matrix system is internally stable.

The problem stated is very basic, one for which an inquisitive undergraduate would demand a solution. It was with some surprise that the author was unable to find its solution in the literature, and hopefully one of the readers of this article will be able to provide a solution, or point out an existing one.

BIBLIOGRAPHY

- [1] F. M. Callier and C. A. Desoer, *Multivariable Feedback Systems*, Springer-Verlag, New York-Heidelberg-Berlin, 1982.
- [2] A. D. Lewis, "On the stability of interconnections of SISO, time-invariant, finite-dimensional, linear systems," preprint, July 2001.

- [3] S. J. Mason, "Feedback theory: Some properties of signal flow graphs," *Proc. IRE*, 41:1144–1156, 1953.
- [4] S. J. Mason, "Feedback theory: Further properties of signal flow graphs," *Proc. IRE*, 44:920–926, 1956.
- [5] Q.-G. Wang, T.-H. Lee, and J.-B. He, "Internal stability of interconnected systems," *IEEE Trans. Automat. Control*, 44(3):593–596, 1999.

Problem 10.7

Open problems in randomized μ analysis

Onur Toker

College of Computer Sciences and Engineering

KFUPM, P.O. Box 14

Dhahran 31261

Saudi Arabia

onur@kfupm.edu.sa

1 INTRODUCTION

In this chapter, we review the current status of the problem reported in [5], and discuss some open problems related to randomized μ analysis. Basically, what remains still unknown after Treil's result [6] are the growth rate of the $\bar{\mu}/\mu$ ratio, and how likely it is to observe a high conservatism. In the context of randomized μ analysis, we discuss two open problems (i) Existence of polynomial time Las Vegas type randomized algorithms for robust stability against structured LTI uncertainties, and (ii) The minimum sample size to guarantee that $\mu/\hat{\mu}$ conservatism will be bounded by g , with a confidence level of $1 - \epsilon$.

2 DESCRIPTION AND HISTORY OF THE PROBLEM

The structured singular value [1] is a quite general framework for analysis/design against component level LTI uncertainties. Although for small number of uncertain blocks, the problem is of reasonable difficulty, all initial studies implied that the same is not likely to be true for the general case. Under these observations, convex upper bound tests became popular alternatives for the structured singular value. Later, it has been proved that these upper bound tests are indeed nonconservative robustness measures for different classes of component level uncertainties, and the structured singular value analysis problem is NP-hard. See the paper [5] and references therein for further details on the history of the problem.

What remains still open after Treil's result? An important open problem was the conservativeness of the standard upper bound test for the complex μ [5]. Recently, Treil showed that the gap between μ and its upper bound $\bar{\mu}$ can be arbitrarily large [6]. Despite this negative result, computational experiments show that most of the time the gap is very close to one for matrices of reasonable size. The following are still open problems:

- What is the growth rate of the gap? In other words, what is the growth rate of

$$\sup_{\mu(M) \neq 0} \frac{\bar{\mu}(M)}{\mu(M)}$$

as a function of $n = \dim(M)$. It is suspected that the growth rate is $O(\log(n))$ [6].

- How likely it is to observe low conservatism? In other words, what is the relative volume of the set

$$\{M : (1 + \epsilon)\mu(M) \geq \bar{\mu}(M) \geq \mu(M)\}$$

in the set of all $n \times n$ matrices with all entries having absolute value at most 1.

Randomized algorithms and some open problems Randomized algorithms are known to be quite powerful tools for dealing with difficult problems. A recent paper of Vidyasagar and Blondel [8] has both a nice summary of earlier results in this area, and provides a strong justification for the importance of randomized algorithms for tackling difficult control related problems. Randomized structured singular value analysis is studied in detail in the recent paper [3], which also has many references to related work in this area.

Las Vegas type algorithm for μ analysis

There are two possible ways of utilizing the results of randomized algorithms, in particular the randomized μ analysis. Let us assume that several random uncertainty matrices with norm bounded by 1, Δ_k 's, $k = 1, \dots, S$, are generated according to some probability distribution, and $\hat{\mu}(M)$ is set to

$$\hat{\mu}(M) = \max_{1 \leq k \leq S} \rho(M\Delta_k).$$

The first interpretation is the following: with a high probability, the inequality $\rho(M\Delta) \leq \hat{\mu}(M)$ is satisfied for all Δ 's except for a set of small relative volume [7]. The second interpretation is to consider the whole process of generating random Δ samples and checking the condition $\rho(M\Delta) < 1$, as a Monte Carlo type algorithm for the complement of robust stability [2]. Indeed, after generating several Δ samples, if $\hat{\mu}(M) \geq 1$, then the (M, Δ) system is not robustly stable, otherwise one can either say the test is inconclusive or conclude that the (M, Δ) system is robustly stable (which

sometimes can be the wrong conclusion). This unpleasant phenomena is a standard characteristic of Monte Carlo algorithms. What is not known is whether there is also a polynomial time Monte Carlo type randomized algorithm for the robust stability condition itself. Combining these two Monte Carlo algorithms will result an algorithm that never gives a false answer, and the probability of getting inconclusive answers goes to zero as we generate more and more random samples.

Problem 1 (Las Vegas type algorithm for μ analysis): Is there a polynomial time Las Vegas type randomized algorithm for testing robust stability against structured LTI uncertainties?

Why this problem is important?

An algorithm like this can be used to check whether the (M, Δ) system is robustly stable or not by generating random Δ matrices: There is no possibility of getting false answers, and probability of getting inconclusive answers goes to zero as the sample size goes to infinity. However, the rate of convergence of the probability of getting inconclusive answers to zero, is also an important factor for the algorithm to be practical.

Relationship between the conservatism of $\mu/\hat{\mu}$, sample size, and confidence levels

Conservativeness of the randomized lower bound $\hat{\mu}$ is also an open problem. More precisely, we have very little knowledge about the relationship between the conservatism ratio $\mu/\hat{\mu}$, the sample size S , and the conservatism bound g . For simplicity, let n denotes the dimension of the matrix M for the rest of this section. The following is a major open problem:

Problem 2: Find the best lower found, $S(n, g, \epsilon)$, such that generating $S \geq S(n, g, \epsilon)$ random samples is enough to claim that, for all M ,

$$\mu(M) \geq \max_{1 \leq k \leq S(n, g, \epsilon)} \rho(M\Delta_k) \geq g^{-1}\mu(M), \quad \text{with confidence level } \geq 1 - \epsilon.$$

In other words, the probability inequality

$$\begin{aligned} \text{Prob} \left\{ \Delta_1, \dots, \Delta_{S(n, g, \epsilon)} : \mu(M) \geq \max_{1 \leq k \leq S(n, g, \epsilon)} \rho(M\Delta_k) \geq g^{-1}\mu(M) \right\} \geq \\ \geq 1 - \epsilon, \end{aligned}$$

is satisfied for all M matrices.

Why this is an important problem?

In a robust stability analysis, one can set a confidence level very close one, say $1 - 10^{-6}$, generate many random Δ samples, and compute the randomized

lower bound $\hat{\mu}$. Ideally, a control engineer would like know how conservative is the obtained $\hat{\mu}$ compared to the actual μ , in order to have a better feeling of the system at hand.

There is very little known about this problem, and some partial results are reported in [4]. The following are simple corollaries:

Result 1 (Polynomial number of samples): For any positive universal constants $C, \alpha \in \mathbb{Z}$, generating $S_n = Cn^\alpha$ random samples *is not enough* to claim that, for all M ,

$$\mu(M) \geq \max_{1 \leq k \leq S_n} \rho(M\Delta_k) \geq 0.99\mu(M), \quad \text{with confidence level} \geq 1 - 10^{-6}.$$

Result 2 (Exponential number of samples): There is a universal constant C such that, generating $S_n = Ce^{n^{2.01}}$ random samples *is enough* to claim that, for all M ,

$$\mu(M) \geq \max_{1 \leq k \leq S_n} \rho(M\Delta_k) \geq 0.99\mu(M), \quad \text{with confidence level} \geq 1 - 10^{-6}.$$

Alternatively, one can fix a confidence level, say $1 - 10^{-6}$, and study the relationship between the sample size S , and the best conservatism bound, $g(S)$, that can be guaranteed with this confidence level. Again not much is known about how fast/slow the best conservatism bound $g(S)$ converges to 1 as the sample size S goes to infinity.

BIBLIOGRAPHY

- [1] A. Packard and J. C. Doyle, "The complex structured singular value," *Automatica*, 29, pp. 71–109, 1993.
- [2] Papadimitriou, C. H., *Computational Complexity*, Addison-Wesley, 1994.
- [3] G. C. Calafiore, F. Dabbene, and R. Tempo, "Randomized algorithms for probabilistic robustness with real and complex structured uncertainty," *IEEE Transactions on Automatic Control*, vol. 45, pp. 2218–2235, 2000.
- [4] O. Toker, "Conservatism of Randomized Structured Singular Value," accepted for publication in *IEEE Transactions on Automatic Control*.
- [5] O. Toker, and B. Jager, "Conservatism of the Standard Upper Bound Test: Is $\sup(\bar{\mu}/\mu)$ finite? Is it bounded by 2?," In: *Open Problems in Mathematical Systems and Control Theory*, V. Blondel, et al. eds., chapter 43, pp. 225–228.
- [6] S. Treil, "The gap between complex structured singular value μ and its upper bound is infinite," *Forthcoming In: IEEE Transactions on Automatic Control*.

- [7] M. Vidyasagar, *A Theory of Learning and Generalization*, Springer-Verlag, London, 1997.
- [8] M. Vidyasagar and V. Blondel, “Probabilistic solutions to some NP-hard matrix problems,” *Automatica*, vol. 37, pp. 1397–1405, 2001.