2002 MTNS Problem Book

Open Problems on the Mathematical Theory of Systems

August 12-16, 2002

Editors

Vincent D. Blondel
Alexander Megretski

Associate Editors

Roger Brockett
Jean-Michel Coron
Miroslav Krstic
Anders Rantzer
Joachim Rosenthal
Eduardo D. Sontag
M. Vidyasagar
Jan C. Willems

Some of the problems appearing in this booklet will appear in a more extensive forthcoming book on open problems in systems theory. For more information about this future book, please consult the website

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

We wish the reader much enjoyment and stimulation in reading the problems in this booklet.

The editors

Program and table of contents

Monday August 12.

Co-chairs: Vincent Blondel, Roger Brockett.

8:00 PM Linear Systems

8:48 PM Stochastic Systems

9:00 PM Nonlinear Systems

9:24 PM Hybrid Systems

9:48 PM Infinite dimensional systems

9:54 PM Identification, Signal Processing

9:24 PM Controllability, Observability

Problem 75

State and first order representations

Jan C. Willems
Department of Electrical Engineering - SCD (SISTA)
University of Leuven
Kasteelpark Arenberg 10
B-3001 Leuven-Heverlee
Belgium
Jan.Willems@esat.kuleuven.ac.be

75.1 Abstract

We conjecture that the solution set of a system of linear constant coefficient PDE's is Markovian if and only if it is the solution set of a system of first order PDE's. An analogous conjecture regarding state systems is also made.
Keywords: Linear differential systems, Markovian systems, state systems, kernel representations.

75.2 Description of the problem

75.2.1 Notation

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

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

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

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

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

Finally, we define concatenation of maps on $\mathbb{R}^{\mathtt{n}}$. Let $f_-, f_+ : \mathbb{R}^{\mathtt{n}} \to \mathfrak{F}$, and let $\pi = (S_-, S_0, S_+) \in \mathfrak{P}$.

Define the map $f_- \wedge_\pi f_+ : \mathbb{R}^{\mathtt{n}} \to \mathfrak{F}$, called the *concatenation* of (f_-, f_+) along π, by

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

75.2.2 Markovian systems

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

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

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:

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

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

$$R_0 w + R_1 \frac{\partial}{\partial x_1} w + R_2 \frac{\partial}{\partial x_2} w + \cdots R_{\mathtt{n}} \frac{\partial}{\partial x_{\mathtt{n}}} w = 0.$$

Oberst [2] has proven that there is a one-to-one relation between $\mathfrak{L}_{\mathtt{n}}^{\mathtt{w}}$ and the submodules of $\mathbb{R}^{\mathtt{w}}[\xi_1, \xi_2, \ldots, \xi_{\mathtt{n}}]$, each $\mathfrak{B} \in \mathfrak{L}_{\mathtt{n}}^{\mathtt{w}}$ being identifiable with its set of *annihilators*

$$\mathfrak{N}_{\mathfrak{B}} := \{n \in \mathbb{R}^{\mathtt{w}}[\xi_1, \xi_2, \ldots, \xi_{\mathtt{n}}] \mid n^\top(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \ldots, \frac{\partial}{\partial x_{\mathtt{n}}})\mathfrak{B} = 0\}.$$

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

75.2.3 State systems

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

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

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

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

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:

$\mathfrak{B} \in \mathfrak{L}_{\mathtt{n}}^{\mathtt{w}+\mathtt{z}}$ is a state system
if and only if
it has a kernel representation
that is first order in the state variables **z**
and zero-th order in the manifest variables **w**.

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

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

75.3 Motivation and history of the problem

These open problems aim at understanding state and state construction for $\mathtt{n}$-D systems.

Maxwell's equations constitute an example of a Markovian system. The diffusion equation and the wave equation are non-examples.

75.4 Available results

It is straightforward to prove the 'if'-part of both conjectures. The conjectures are true for $\mathtt{n} = 1$, 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*, volume 37, pages 388-408, 1999.

[2] U. Oberst, Multidimensional constant linear systems, *Acta Applicandae Mathematicae*, volume 20, pages 1-175, 1990.

[3] P. Rapisarda and J.C. Willems, State maps for linear systems, *SIAM Journal on Control and Optimization*, volume 35, pages 1053-1091, 1997.

Problem 59

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

59.1 Description of the problem

Problem: Let $R \in \mathbb{C}[\eta_1, \dots, \eta_{\mathtt{m}}, \xi]^{\mathtt{g} \times \mathtt{w}}$ and let $\mathfrak{B}$ be the behaviour 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,

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

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

$$D_R \begin{bmatrix} w_1 \\ \vdots \\ w_{\mathtt{w}} \end{bmatrix} = \begin{bmatrix} \sum_{\mathtt{k}=1}^{\mathtt{w}} r_{1\mathtt{k}} \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_{\mathtt{m}}}, \frac{\partial}{\partial t} \right) w_{\mathtt{k}} \\ \vdots \\ \sum_{\mathtt{k}=1}^{\mathtt{w}} r_{\mathtt{gk}} \left(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_{\mathtt{m}}}, \frac{\partial}{\partial t} \right) w_{\mathtt{k}} \end{bmatrix}.$$

Time-controllability is a property of the behaviour, defined as follows. The behaviour $\mathfrak{B}$ 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

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

59.2 Motivation and history of the problem

The behavioural 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 behavioural 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 behaviour, 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 which 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.

59.3 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 $\mathtt{r}_0 \in \mathbb{N} \cup \{0\}$ such that for all $\lambda \in \mathbb{C}$, $\operatorname{rank}(R(\lambda)) = \mathtt{r}_0$. This condition is in turn equivalent with the torsion freeness of the $\mathbb{C}[\xi]$-module $\mathbb{C}[\xi]^{\mathtt{w}}/\mathbb{C}[\xi]^{\mathtt{g}}R$.

Let us consider the following statements

*A*1. The $\mathbb{C}(\eta_1, \dots, \eta_{\mathtt{m}})[\xi]$-module $\mathbb{C}(\eta_1, \dots, \eta_{\mathtt{m}})[\xi]^{\mathtt{w}}/\mathbb{C}(\eta_1, \dots, \eta_{\mathtt{m}})[\xi]^{\mathtt{g}}R$ is torsion free.

*A*2. There exists a $\chi \in \mathbb{C}[\eta_1, \dots, \eta_{\mathtt{m}}, \xi]^{\mathtt{w}} \setminus \mathbb{C}(\eta_1, \dots, \eta_{\mathtt{m}})[\xi]^{\mathtt{g}}R$ and there exists a nonzero $p \in \mathbb{C}[\eta_1, \dots, \eta_{\mathtt{m}}, \xi]$ such that $p \cdot \chi \in \mathbb{C}(\eta_1, \dots, \eta_{\mathtt{m}})[\xi]^{\mathtt{g}}R$, and $\deg(p) = \deg(\jmath(p))$, where $\jmath$ denotes the homomorphism $p(\xi, \eta_1, \dots, \eta_{\mathtt{m}}) \mapsto p(\xi, 0, \dots, 0) : \mathbb{C}[\xi, \eta_1, \dots, \eta_{\mathtt{m}}] \to \mathbb{C}[\xi]$.

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

$$\begin{array}{ccc} \mathfrak{B} & \text{is} & \text{time-controllable} \\ \Downarrow \not\Uparrow & & \Uparrow \\ \neg A2 & \begin{array}{c}\Leftarrow \\ \not\Rightarrow\end{array} & A1 \end{array}$$

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 to[1] 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 $\mathtt{k} \in \mathbb{N}$, $|\varphi^{(\mathtt{k})}(t)| \leq C_\epsilon \epsilon^{\mathtt{k}} (\mathtt{k}!)^2$ 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 behaviour that is not in the class $\mathcal{W}$. Also whether the whole behaviour (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 behaviour corresponding to $p \in \mathbb{C}[\eta_1, \dots, \eta_{\mathtt{m}}, \xi]$ is time-controllable iff $p \in \mathbb{C}[\eta_1, \dots, \eta_{\mathtt{m}}]$, which would eventually help in settling the question of the equivalence of $A1$ and time-controllability.

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

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).

[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 58

Schur Extremal Problems

Large Lev Sakhnovich
Courant Institute of Mathematical Science
New York, NY,11223
USA
Lev.Sakhnovich@verizon.net

58.1 Description of the problem

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

$$w^{\star}(z)w(z)\leq\rho_{min}^2, |z| < 1 \tag{58.1}$$

where $w(z)$ and ρ_{min} are mxm 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 doesn't 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 $a_0, a_1, ..., a_n$ are given.To describe the set of $m\times m$ matrix functions $w(z)$ holomorphic in the circle $|z| < 1$ and satisfying the relation

$$w(z) = a_0 + a_1 z + ... + a_n z^n + ... \tag{58.2}$$

and the inequality (1.1)
A necessary condition of the solvability of the Schur extremal problem is the inequality

$$R_{min}^2 - S\geq 0 \tag{58.3}$$

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

$$S = C_n C_n^{\star}, R_{min} = diag[\rho_{min}, \rho_{min}, ..., \rho_{min}] \tag{58.4}$$

$$C_n = \begin{pmatrix} a_0 & 0 & ... & 0 \\ a_1 & a_0 & ... & 0 \\ ... & ... & ... & ... \\ a_n & a_{n-1} & ... & a_0 \end{pmatrix} \tag{58.5}$$

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

$$R_{min}^2 - S\geq 0 \tag{58.6}$$

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

$$R^2 - S\geq 0 \tag{58.7}$$

then

$$rank(R^2_{min} - S) \leq rank(R^2 - S) \tag{58.8}$$

where $R = diag[\rho, \rho, ..., \rho]$
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 $\rho^2_{min} = \lambda_{max}$, 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 $S = C_n C_n^\star$ in the following block form

$$\begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \tag{58.9}$$

where S_{22} is the $m \times m$ matrix.
Proposition 1 [1]If $\rho = q > 0$ satisfies inequality (1.7) and the relation

$$q^2 = S_{22} + S^\star_{12}(Q^2 - S_{11})^{-1} S_{12} \tag{58.10}$$

where $Q = diag[q, q, ..., q]$ then $\rho_{min} = q$.
We shall apply the method of successive approximation when studying equation (1.10).We put

$q^2_0 = S_{22}, q^2_{k+1} = S_{22} + S^\star_{12}(Q^2_k - S_{11})^{-1} S_{12}$ where $k \geq 0, Q_k = diag[q_k, q_k, ..., q_k]$.We suppose that

$$Q^2_0 - S_{11} > 0 \tag{58.11}$$

Theorem 1[1].The sequence $q^2_0, q^2_2, q^2_4, ...$ monotonically increases and has the limit m_1.The sequence $q^2_1, q^2_3, q^2_5, ...$ monotonically decreases and has the limit m_2.The inequality $m_1 \leq m_2$ is true. If $m_1 = m_2$ then $\rho^2_{min} = q^2$
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]).
Remark4.In book [1] we give an example in which ρ_{min} is constructed in explicit form.

Bibliography

[1] Lev Sakhnovich"Interpolation Theory and Applications",it Kluwer Acad.Publ.(1997)
[2] J.Helton and L.Sakhnovich "Extremal Problems of Interpolation Theory"(to appear)
[3] N.J.Young"The Nevanlinna-Pick Problem for matrix-valued functions",it J. Operator Theory,15 pp239-265,(1986).
[4] V.V.Peller and N.J.Young "Superoptimal analytic approximations of matrix functions",it Journal of Functional Analysis,120,pp300-343(1994)
[5] M.Green M. and D.Limebeer "Linear Robust Control"it Prentice -Hall(1995)
[6] "Lecture Notes in Control and Information"it Sciences,Springer Verlag, 135,(1989)
[7] N.I.Akhiezer "On a Minimum in Function Theory",it Operator Theory,Adv. and Appl.95 pp19-35(1997)
[8] A.Ferrante and B.C.Levy "Hermitian Solutions of the Equations $X = Q + NX^{-1}N^\star$", Linear Algebra and Applications 247 pp359-373(1996)

Problem 69

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

69.1 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 $\{\mathbb{R}, \mathbb{R}^{\mathtt{w}}, \mathcal{B}\}$, 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 $w : \mathbb{R} \to \mathbb{R}^{\mathtt{w}}$ of a set of higher order, linear, constant coefficient, differential equations. More precisely,

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

for some polynomial matrix $R \in \mathbb{R}^{\bullet \times \mathtt{w}}[\xi]$. The set of all such systems Σ is denoted by $\mathcal{L}^{\mathtt{w}}$. Often, we simply refer to a system by talking about its behavior, and we write $\mathcal{B} \in \mathcal{L}^{\mathtt{w}}$ instead of $\Sigma \in \mathcal{L}^{\mathtt{w}}$. Behaviors $\mathcal{B} \in \mathcal{L}^{\mathtt{w}}$ 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 $w = (w_1, w_2, \ldots, w_{\mathtt{w}})$ 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 $\mathtt{m}(\mathcal{B})$, its number of free, 'input', variables. The remaining number of variables, $\mathtt{w} - \mathtt{m}(\mathcal{B})$, is called the *output cardinality* of $\mathcal{B}$ and is denoted by $\mathtt{p}(\mathcal{B})$. Finally, a third integer invariant associated with a system behavior $\mathcal{B} \in \mathcal{L}^{\mathtt{w}}$ is its McMillan degree. It can be shown that (modulo permutation of the components of the external variable w) any $\mathcal{B} \in \mathcal{L}^{\mathtt{w}}$ 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 $\mathtt{n}(\mathcal{B})$.

Suppose now $\Sigma_1 = \{\mathbb{R}, \mathbb{R}^{\mathtt{w}_1} \times \mathbb{R}^{\mathtt{w}_2}, \mathcal{B}_1\} \in \mathcal{L}^{\mathtt{w}_1 + \mathtt{w}_2}$ and $\Sigma_2 = \{\mathbb{R}, \mathbb{R}^{\mathtt{w}_2} \times \mathbb{R}^{\mathtt{w}_3}, \mathcal{B}_2\} \in \mathcal{L}^{\mathtt{w}_2 + \mathtt{w}_3}$ are linear differential systems with common factor $\mathbb{R}^{\mathtt{w}_2}$ 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

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

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 $\Sigma_1 \wedge_{w_2} \Sigma_2$:

$$\mathtt{p}(\mathcal{B}_1 \wedge_{\mathtt{w}_2} \mathcal{B}_2) = \mathtt{p}(\mathcal{B}_1) + \mathtt{p}(\mathcal{B}_2).$$

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 McMilan degree of the interconnection:

$$\mathtt{n}(\mathcal{B}_1 \wedge_{\mathtt{w}_2} \mathcal{B}_2) = \mathtt{n}(\mathcal{B}_1) + \mathtt{n}(\mathcal{B}_2).$$

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 which 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].

69.2 Statement of the problem

Suppose $\mathcal{P}_{\text{full}} \in \mathcal{L}^{\mathtt{w}+\mathtt{c}}$ 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 $\mathcal{C} \in \mathcal{L}^{\mathtt{c}}$ (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* $\mathcal{P}_{\text{full}} \wedge_c \mathcal{C} \in \mathcal{L}^{\mathtt{w}+\mathtt{c}}$. The *manifest controlled behavior* $\mathcal{K} \in \mathcal{L}^{\mathtt{w}}$ is obtained by projecting all trajectories $(w, c) \in \mathcal{P}_{\text{full}} \wedge_c \mathcal{C}$ on their first coordinate:

$$\mathcal{K} := \{w \mid \text{ there exists } c \text{ such that } (w, c) \in \mathcal{P}_{\text{full}} \wedge_c \mathcal{C}\}. \tag{69.1}$$

If this holds, then we say that $\mathcal{C}$ *implements* $\mathcal{K}$. If, for a given $\mathcal{K} \in \mathcal{L}^{\mathtt{w}}$ there exists $\mathcal{C} \in \mathcal{L}^{\mathtt{c}}$ 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 $\mathcal{P}_{\text{full}} \in \mathcal{L}^{\mathtt{w}+\mathtt{c}}$, the set of all behaviors $\mathcal{K} \in \mathcal{L}^{\mathtt{w}}$ that are implementable by regular feedback. In other words:

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

1. $\mathcal{C}$ implements $\mathcal{K}$ (meaning that (69.1) holds),
2. $\mathtt{p}(\mathcal{P}_{\text{full}} \wedge_{\mathtt{c}} \mathcal{C}) = \mathtt{p}(\mathcal{P}_{\text{full}}) + \mathtt{p}(\mathcal{C})$,
3. $\mathtt{n}(\mathcal{P}_{\text{full}} \wedge_{\mathtt{c}} \mathcal{C}) = \mathtt{n}(\mathcal{P}_{\text{full}}) + \mathtt{n}(\mathcal{C})$.

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

69.3 Background

Our open problem is to find conditions for a given $\mathcal{K} \in \mathcal{L}^{\mathtt{w}}$ 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 $\mathcal{L}^{\mathtt{w}}$, 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 $\mathcal{K} \in \mathcal{L}^{\mathtt{w}}$ is implementable if and only if

$$\mathcal{N} \subseteq \mathcal{K} \subseteq \mathcal{P}, \tag{69.2}$$

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

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

with $R(\xi_1, \xi_2, \ldots, \xi_n)$ a polynomial matrix in n indeterminates. Recently, in [6] a variation of condition (69.2) was shown to be sufficient for implementability of system behaviors in a more general (including non-linear) context.

For a system behavior $\mathcal{K} \in \mathcal{L}^{\mathtt{w}}$ 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 $\mathcal{K} \in \mathcal{L}^{\mathtt{w}}$ is regularly implementable if and only if, in addition to condition (69.2), the following condition holds:

$$\mathcal{K} + \mathcal{P}_{\text{cont}} = \mathcal{P}. \tag{69.3}$$

Condition (69.3) 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 (69.3) 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 (69.2) and (69.3) for regular implementability in the context of nD system behaviors can also be found in [19]. 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, to appear in *IEEE Transactions on Automatic Control*, May 2002.

[2] M. Kuijper, Why do stabilizing controllers stabilize, *Automatica*, volume 31, pages 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, Number 1, pages 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 2002, submitted for publication.

[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, pages 53-69, 2002.

[8] J.C. Willems, On Interconnections, Control and Feedback, *IEEE Transactions on Automatic Control*, volume 42, pages 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, Number 4, pages 1072-1086, 2001.

Problem 36

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

36.1 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:

1. *the construction of $P_{\mathcal{G}}$ depends only on the topology of the graph, and not on manipulations of the branch gains;*
2. *all closed-loop transfer functions T_{ji}, $i, j = 1, \ldots, n$, are analytic in the closed right half plane (CRHP) if and only if $P_{\mathcal{G}}$ is Hurwitz?*

The condition 1 is a little vague. It may be thought of as a sort of "meta-condition," and perhaps can be replaced with something like

1'. *$P_{\mathcal{G}}$ is formed by products and sums of the polynomials N_{ji} and D_{ji}, $i, j = 1, \ldots, n$.*

The idea is that the definition of $P_{\mathcal{G}}$ should not depend on the choice of branch gains R_{ji}, $i, j = 1, \ldots, n$. 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}, $i, j = 1, \ldots, n$."

36.2 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 two things to consider in the course of making an interconnection: (1) is the interconnection *internally 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}, $i, j = 1, \ldots, n$, are proper is known to be equivalent to the condition that the determinant of the graph be a biproper rational function.

As an illustration of what we are after, consider the single-loop configuration of Figure 36.1. As is

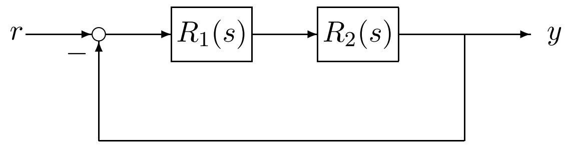


Figure 36.1: Single-loop interconnection

well-known, if we write $R_i = \frac{N_i}{D_i}$, $i = 1, 2$, as coprime fractions, then all closed-loop transfer functions have no poles in the CRHP if and only if the polynomial $P_{\mathcal{G}} = D_1 D_2 + N_1 N_2$ is Hurwitz. Thus $P_{\mathcal{G}}$ serves as the characteristic polynomial in this case. The essential feature of $P_{\mathcal{G}}$ 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 $P_{\mathcal{G}}$.

36.3 Approaches that do not solve the problem

Let us outline two approaches that yield solutions having one of properties 1 and 2, 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 2 above. However, the algorithmic procedure will involve computing GCD's of various of the polynomials N_{ji} and D_{ji}, $i, j = 1, \ldots, n$. 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 1 (and 1') above. The problem we pose demands a simpler, more direct answer to the question of determining when an interconnection is internally stable.

Wang, Lee, and He [5] propose a characteristic polynomial for a signal flow graph using the determinant of the graph which we denote by $\Delta_{\mathcal{G}}$ (see [3, 4]). Specifically, they suggest using

$$P = \prod_{(i,j) \in \{1,\ldots,n\}^2} D_{ji} \Delta_{\mathcal{G}} \tag{36.1}$$

as the characteristic polynomial. Thus one multiplies the determinant by all denominators, arriving at a polynomial in the process. This polynomial has the property 1 (or 1') above. However, while it is true that if this polynomial is Hurwitz then the system is internally stable, the converse is false (see [2] for a simple counterexample). Thus property 2 is not satisfied by P. It is true that the polynomial P in (36.1) gives the correct characteristic polynomial for the interconnection of Figure 36.1. It is also true that if the signal flow graph has no loops (in this case $\Delta_{\mathcal{G}} = 1$) then the polynomial P of (36.1) satisfies condition 2.

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 35

Bases for Lie algebras and a continuous CBH formula

Matthias Kawski[1]
Department of Mathematics and Statistics
Arizona State University,
Tempe, Arizona 85287-1804
USA
kawski@asu.edu

35.1 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

$$\dot{x} = f_0(x) + \sum_{k=1}^{m} u_k \, f_k(x). \tag{35.1}$$

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 $X_1, \ldots X_m$, Sussmann [16] gave a general formula as a directed infinite product of exponentials

$$x(T, u) = \prod_{H \in \mathcal{H}}^{\rightarrow} \exp(\xi_H(T, u) \cdot f_H). \tag{35.2}$$

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

$$\xi_{HK} = \xi_H * \xi_K \tag{35.3}$$

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

if H, K, HK are Hall words and the *left factor* of K is not equal to H [9, 16]. (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* 2[nd] *kind"*) is a single exponential of an infinite Lie series (*"coordinates of the* 1[st] *kind"*).

$$x(T, u) = \exp\Big(\sum_{B \in \mathcal{B}} \zeta_B(T, u) \cdot f_B\Big) \tag{35.4}$$

It is straightforward to obtain explicit formulas for ζ_B for some spanning sets $\mathcal{B}$ of the free Lie algebra [16], but much preferable are series that use *bases* $\mathcal{B}$, and which, in addition, yield as simple formulas for ζ_B as (35.3) 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 $X_1, \ldots X_m$ such that the corresponding iterated integral functionals ζ_B defined by (35.4) have *simple formulas* (similar to (35.3)), suitable for control applications (both analysis and design).

The formulae (35.4) and (35.2) 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 (27.1) 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 $\mathcal{F} = \{f_1, \ldots, f_m\}$ on M^n a basis for $L(f_1, \ldots, f_m)$ 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. $L(f_1, \ldots, f_m)(p) = T_pM$ at a specified initial point p. It makes sense to first consider special classes of systems $\mathcal{F}$, e.g. which are such that $L(f_1, \ldots, f_m)$ 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 (35.3), which has proven convenient for theoretical studies, numerical computation, and practical implementations.

35.2 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 $\{X_1, \ldots, X_m\}$ yields the Chen Fliess series [2, 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 (35.2)) 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 [10, 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 $H(a, b) = a + b + \frac{1}{2}[a, b] + \frac{1}{12}[a, [a, b]] - \frac{1}{12}[b, [a, b] + \ldots$. 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 unwieldily expressions. The expansion (35.2) is the natural *continuous* analogue of the CBH formula – and the problem is to find the most useful form.

The uses of these expansions (35.2) and (35.4) 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 (35.4) 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 (35.2) and (35.4) is the Chen series [2], which is well known to be an exponential Lie series, compare [18], thus guaranteeing the existence of the alternative expansions

$$\sum_{w\in Z^*} w\otimes w \overset{!}{=} \exp\Big(\sum_{B\in\mathcal{B}} \zeta_B\otimes B\Big) \overset{!}{=} \overrightarrow{\prod_{B\in\mathcal{B}}} \exp\left(\xi_B\otimes B\right) \tag{35.5}$$

The first bases for free Lie algebras build on Hall's work in the 1930s on commutator groups. While several *other* bases (Lyndon, Sirsov) 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 Poincare-Birckhoff-Witt bases (of the universal enveloping algebra of a Lie algebra) and Lazard elimination. Formulas for the dual PBW bases ξ_B have been given by Schützenberger, Sussmann[16], 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.

35.3 Available related results

The direct expansion of the logarithm into a formal power series may be simplified using symmetrization [18, 16], but this still does not yield well-defined *"coordinates"* with respect to a basis.

Explicit, but quite *unattractive* formulas forthe 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 don't 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 [16], but are not bases. These formulas can be refined to apply to Hall-bases, but at the cost of loosing 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 (35.3). 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) (to appear).
[5] M. Fliess, "Fonctionelles causales nonlinéaires et indeterminées 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., pp. 111–128, Teubner (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", (1993), Oxford (Clarendon Press).
[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 (1986), 323–335.
[23] G. Viennot, "Algèbres de Lie Libres et Monoïdes Libres", *Lecture Notes Math.*, 692, Springer, Berlin, (1978).

Problem 10

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
Systems, Signals and Control
University of Twente
Enschede, 7500 AE
The Netherlands
j.cortesmonforte@utwente.nl

Andrew D. Lewis
Mathematics & Statistics
Queen's University
Kingston, ON K7L 3N6
Canada
andrew@mast.queensu.ca

Sonia Martínez
Institute of Mathematics & Physics
High Council of Scientific Research
Madrid, 28006
Spain
s.martinez@imaff.cfmac.csic.es

10.1 Problem statement and historical remarks

For finite dimensional $\mathbb{R}$-vector spaces U and V we consider a symmetric bilinear map $B: U \times U \to V$. This then defines a quadratic map $Q_B: U \to V$ 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 $\lambda \in V^* \setminus \operatorname{ann}(\operatorname{image}(Q_B))$, λQ_B is neither positive nor negative-semidefinite, where ann denotes the annihilator.

Given a symmetric bilinear map $B: U \times U \to V$, the problems we consider are as follows.

1. *Find necessary and sufficient conditions characterizing when Q_B is surjective.*

2. *If Q_B is surjective and $v \in V$, design an algorithm to find a point $u \in Q_B^{-1}(v)$.*

3. 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 10.4 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 clases 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.

10.2 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.

Optimal control: Agračhev [2] explicitly realises 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 diagonalised). 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 linearisation of a system

$$\dot{x} = f(x, u), \quad x \in \mathbb{R}^n, \ u \in \mathbb{R}^m.$$

The condition is that for each $x \in \mathbb{R}^n$, the set $D_x = \{f(x, u) \in T_x\mathbb{R}^n | \ u \in \mathbb{R}^m\}$ admits a *ruling*, that is, a foliation of D_x by lines. Some manipulations with differential forms turns this necessary condition

[1] We thank an anonymous referee for these observations.

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.

10.3 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 $\mathrm{int}_S(A)$ the interior of A relative to the induced topology on S. If $S \subset V$, $\mathrm{aff}(S)$ and $\mathrm{conv}(S)$ denote, respectively, the affine hull and the convex hull of S.

Proposition 1 *Let $B\colon U \times U \to V$ be a symmetric bilinear map with U and V finite-dimensional. The following statements hold:*

(i) B is indefinite if and only if $0 \in \mathrm{int}_{\mathrm{aff}(\mathrm{image}(Q_B))}(\mathrm{conv}(\mathrm{image}(Q_B)))$;

(ii) B is definite if and only if there exists a hyperplane $P \subset V$ so that $\mathrm{image}(Q_B) \cap P = \{0\}$ and so that $\mathrm{image}(Q_B)$ 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 $\mathrm{ind}(Q)$ the index of a quadratic map $Q\colon U \to \mathbb{R}$ on a vector space U.

Proposition 2 *Let $B\colon U \times U \to V$ be a symmetric bilinear map with V finite-dimensional. If $\mathrm{ind}(\lambda Q_B) \geq \dim(V)$ 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.

10.4 Problem simplification

One of the difficulties with studying vector-valued quadratic maps is that they are somewhat difficult to get ones 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 $B\colon U \times U \to V$ 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 $\mathrm{Sym}_n(\mathbb{R})$ denote the set of symmetric $n \times n$ matrices with entries in $\mathbb{R}$. On $\mathrm{Sym}_n(\mathbb{R})$ we use the canonical inner product

$$\langle \boldsymbol{A}, \boldsymbol{B} \rangle = \mathrm{tr}(\boldsymbol{A}\boldsymbol{B}).$$

We consider the map $\pi\colon \mathbb{R}^n \to \mathrm{Sym}_n(\mathbb{R})$ defined by $\pi(\boldsymbol{x}) = \boldsymbol{x}\boldsymbol{x}^t$, where ^{t} denotes transpose. Thus the image of π is the set of symmetric matrices of rank at most one. If we identify $\mathrm{Sym}_n(\mathbb{R}) \simeq \mathbb{R}^n \otimes \mathbb{R}^n$, then $\pi(\boldsymbol{x}) = \boldsymbol{x} \otimes \boldsymbol{x}$. Let K_n be the image of π and note that it is a cone of dimension n in $\mathrm{Sym}_n(\mathbb{R})$ having a singularity only at its vertex at the origin. Furthermore, K_n may be shown to be a subset of the hypercone in $\mathrm{Sym}_n(\mathbb{R})$ defined by those matrices $\boldsymbol{A}$ in $\mathrm{Sym}_n(\mathbb{R})$ forming angle $\arccos(\frac{1}{n})$ with the identity matrix. Thus the ray from the origin in $\mathrm{Sym}_n(\mathbb{R})$ through the identity matrix is an axis for the cone K_N. In algebraic geometry, the image of K_n under the projectivisation of $\mathrm{Sym}_n(\mathbb{R})$ 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 $B\colon \mathbb{R}^n \times \mathbb{R}^n \to V$ be a symmetric bilinear map with V finite-dimensional. Using the universal mapping property of the tensor product, B induces a *linear* map $\tilde{B}\colon \mathrm{Sym}_n(\mathbb{R}) \simeq \mathbb{R}^n \otimes \mathbb{R}^n \to V$ with the property that $\tilde{B} \circ \pi = B$. The dual of this map gives an injective linear map $\tilde{B}^*\colon V^* \to$

$\mathrm{Sym}_n(\mathbb{R})$ (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 $\mathrm{Sym}_n(\mathbb{R})$. 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 $\mathrm{Sym}_n(\mathbb{R})$! To simplify things further, we decompose L_B into a component along the identity matrix in $\mathrm{Sym}_n(\mathbb{R})$ 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 summarises the above discussion.

The problem of studying the image of a vector-valued quadratic form can be reduced to studying the orthogonal projection of $K_n \subset \mathrm{Sym}_n(\mathbb{R})$*, the unprojectivised 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* $B: U \times U \to V$ *a symmetric bilinear map, then* Q_B *is surjective if and only if* B *is indefinite.*

Bibliography

[1] A. A. Agračhev. The topology of quadratic mappings and Hessians of smooth mappings. *J. Soviet Math.*, 49(3):990–1013, 1990.

[2] A. A. Agračhev. Quadratic mappings in geometric control theory. *J. Soviet Math.*, 51(6):2667–2734, 1990.

[3] A. A. Agračhev 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. Agračhev 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'brekht. 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. Preprint, November 2001.

[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 4

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

4.1 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.

The 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?

4.2 Motivation and history of the problem

The modern stochastic optimal control and filtering theories use white noise driven systems. The 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 which describe real noise more adequately. Such 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 the 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. 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].

4.3 Available results and discussion

The random process φ with the property $\operatorname{cov}(\varphi(t+s),\varphi(t)) = \lambda(t,s)$ if $0 \le s < \varepsilon$ and $\operatorname{cov}(\varphi(t+s),\varphi(t)) = 0$ if $s \ge \varepsilon$, 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 *autocovatiance function* of φ) depends only on s (see Fleming and Rishel [1]).

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

$$\varphi(t) = \int_{-\min(t,\varepsilon)}^{0} \phi(\theta) w(t+\theta)\, d\theta,\ t \ge 0, \tag{4.1}$$

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

$$\int_{-\varepsilon}^{-s} \phi(\theta)\phi(\theta+s)\, d\theta = \lambda(s),\ 0 \le s \le \varepsilon. \tag{4.2}$$

The solution φ of (4.2) is called a *relaxing function.* Since in (4.2) ϕ has only one variable the process φ from (4.1) 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 (4.2) in the space $L_2(-\varepsilon,0)$ if λ is a nonzero function a.e. on $[-\varepsilon,0]$.*

The nonuniqueness of solution of the equation (4.2) 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 (4.2) generates the random process φ in the form (4.1) that could be considered as a less or more adequate model of real noise. In order to make a reasonable decision about 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 Theorem **??** from [8, 9] can be useful for construction different solutions of the equation (4.2).

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*, 7p. (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).

Problem 25

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

25.1 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.

25.2 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.

25.3 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, Sect. 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 which 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 (1985). *Differential-Geometrical Methods in Statistics.* Lecture Notes in Statistics 28. Springer Verlag, Berlin.

[2] S.-I. Amari (1987). Differential Geometry of a Parametric Family of Invertible Linear Systems — Riemannian Metric, Dual Affine Connections, and Divergence. *Mathematical Systems Theory*, **20**, 53–82.

[3] C. Atkinson and A. F. S. Mitchell (1981). Rao's Distance Measure. *Sankhyā: The Indian Journal of Statistics*, Series A, **43** (3), 345–365.

[4] R. W. Brockett and P. S. Krishnaprasad (1980). A scaling theory for linear systems. *IEEE Trans. Aut. Contr.*, **AC-25**, 197–206.

[5] J. M. C. Clark (1976). The consistent selection of parametrizations in system identification. *Proc. Joint Automatic Control Conference*, 576–580. Purdue University, Lafayette, Ind.

[6] B. Friedlander (1984). 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 (1983). Evaluation of the covariance matrix for the maximum likelihood estimator of a Gaussian autoregressive-moving average process. *Biometrika*, **70** (1), 279–284.

[8] E. J. Hannan and M. Deistler (1988). *The Statistical Theory of Linear Systems*. John Wiley & Sons, New York.

[9] B. Hanzon (1989). *Identifiability, Recursive Identification and Spaces of Linear Dynamical Systems*, CWI Tracts 63 and 64. Centrum voor Wiskunde en Informatica (CWI), Amsterdam.

[10] B. Hanzon and R. L. M. Peeters (1993). 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.

[11] A. Klein and G. Mélard (1989). On Algorithms for Computing the Covariance Matrix of Estimates in Autoregressive Moving Average Processes. *Computational Statistics Quarterly*, **1**, 1–9.

[12] J. Milnor (1969). A problem in cartography. *American Math. Monthly*, **76**, 1101–1112.

[13] R. L. M. Peeters (1994). *System Identification Based on Riemannian Geometry: Theory and Algorithms*. Tinbergen Institute Research Series, Vol. **64**. Thesis Publishers, Amsterdam.

[14] R. L. M. Peeters and B. Hanzon (1999). Symbolic computation of Fisher information matrices for parametrized state-space systems. *Automatica*, **35**, 1059–1071.

[15] C. R. Rao (1945). Information and accuracy attainable in the estimation of statistical parameters. *Bull. Calcutta Math. Soc.*, **37**, 81–91.

[16] N. Ravishanker, E. L. Melnick and C.-L. Tsai (1990). Differential geometry of ARMA models. *Journal of Time Series Analysis*, **11**, 259–274.

[17] A. L. Rijkeboer (1993). 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.

[18] T. Söderström and P. Stoica (1989). *System Identification*. Prentice-Hall, New York.

Problem 81

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 analysed by studying the associated matrix differential equations. Such gradient-like nonlinear differential equations have an intricate convergence behaviour 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.

81.1 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 ODE's 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 1 (PSA/MSA Flow) *A normalised 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 orthornormal for all t.*
3. *$\lim_{t\to\infty} X(t) = X_\infty$ exists for all full rank initial conditions X_0.*
4. *X_∞ is an orthornormal 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 $t \to \infty$ 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]

$$\dot{X} = (I - XX')CX. \tag{81.1}$$

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]:

$$\dot{X} = CXN - XNX'CX \tag{81.2}$$

Here $N = N' > 0$ denotes an arbitrary diagonal $k \times k$ matrix with distinct eigenvalues. This system is actually a joint generalisation of Oja's flow (81.1) and of Brockett's [1] gradient flow on orthogonal matrices

$$\dot{X} = [C, XNX']X \tag{81.3}$$

for symmetric matrix diagonalisation; see also [6]. In [19], Oja's flow was re-derived by first proposing the gradient flow

$$\dot{X} = (C(I - XX') + (I - XX')C)X \tag{81.4}$$

and then omitting the first term $C(I - XX')X$ because $C(I - XX')X = CX(I - X'X) \to 0$, a consequence of both terms in (81.4) forcing X to the invariant manifold $\{X : X'X = I\}$. Interesting, it has recently been realised [8] that (81.4) has certain computational advantages compared with (81.1), 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 (81.1), (81.2) and (81.4). 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 (81.1), (81.2) and (81.4), 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 discretisations are to be found. Since we are working in unconstrained Euclidean matrix space $\mathbb{R}^{n\times p}$, we consider Euler step discretisations. Thus, e.g., for system (81.1) consider

$$X_{t+1} = X_t - s_t(I - X_tX_t')CX_t, \tag{81.5}$$

[1] It 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 $Q \to 0$. However, counter-examples are known [15] where $G(t)$ is strictly negative definite (with constant eigenvalues) yet Q diverges.

with suitably small step sizes. Such Euler discretisation 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).

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.*

81.2 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

$$X_{t+1} = X_t - s_t(I - X_t X_t')u_t u_t' X_t. \tag{81.6}$$

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 (81.1). Here, $C = E(u_t u_t')$ is the covariance matrix. Similarly, the other flows, (81.2) and (81.4), have analogous interpretations.

In [9, 11] it is shown for $p = 1$ that (81.1) is a PSA flow, i.e. it converges for generic initial conditions to a normalised dominant eigenvector of C. In [11], the system (81.1) was studied for arbitrary values of p and it was conjectured that (81.1) 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 inital 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 (81.2) 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 (81.1) was given, but without recognising that (81.1) is actually a gradient flow. There have been confusing remarks in the literature claiming that (81.1) cannot be a gradient system as the linearisation is not a symmetric matrix. However, this is due to a misunderstanding of the concept of a gradient. In [20] it is shown that (81.2), and in particular (81.1), is actually a gradient flow for the cost function $f(X) = 1/4tr(AXNX')^2 - 1/2tr(A^2XD^2X')$ 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 towards a phase portrait analysis of (81.4) are made in [8].

81.3 Available Results

In [12, 13, 17], the equilibrium points of (81.2) 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 (81.2), and hence (81.1), is a PSA. An analogous result for (81.4) is forthcoming; see [8] for first steps in this direction. Sufficient conditions for initial matrices in the Oja flow (81.1) to converge to a dominant subspace basis are given in [18, 4], but not for the other, unstable equilibria, nor for system (81.2). A complete characterisation 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.

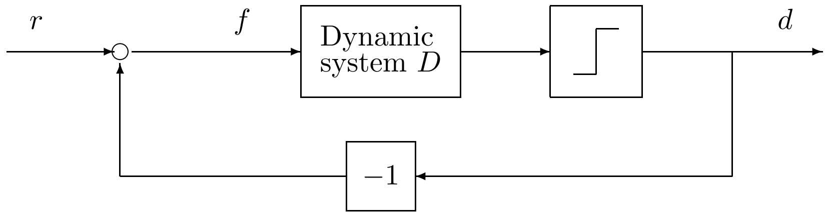
Problem 55

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

55.1 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 $[\omega_1, \omega_2]$.



To make a more 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 $k = 0, 1, 2, \ldots$. A map $D : \ell[0, \infty] \to \ell[0, \infty]$ 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 static operator

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

Problem

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

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

The problem is to find a 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

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

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.

55.2 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

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

it holds that

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

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 44

Locomotive Drive Control by Using Haar Wavelet Packets

P. Mercorelli
Institut für Automatisierung und Informatik GmbH
Zentrum für Industrielle Forschung und Entwicklung
Schlachthof 4 Wernigerode
Germany
pmercorelli@hs-harz.de
mercorelli@iai-wr.de

P. Terwiesch
ABB Corporate Research
Segelhof, CH-5405 Baden
Switzerland
Peter.Terwiesch@ch.abb.com

D. Prattichizzo
Dip. di Ingegneria dell' Informazione – Università di Siena
via Roma, 56
53100 Siena
Italy
prattichizzo@ing.unisi.it

44.1 Description of the problem

Preliminary simulations and roughly algorithms show how the use of wavelet packets are very promising in itermodulation disturbance rejection for converter bridge control. The presented contribution wants to pay attention on an original way to states a possible converter bridge control by using Haar wavelet packets. The presented attack to the problem could represent a possible alternative to the traditional PWM control for the converter bridge. The idea is to find an algorithm which guaranties the needed power supply and in the meantime a high level of the $\frac{S}{N}$ ratio. One proposes a roughly algorithm which find the length of the active switching function looking for suitable scaling function. The open problem consists of checking the suitability of its formulation in order to build a rigorous algorithm which can find a switching control law.

44.2 Motivation and history of the problem

The modern vehicles produced during the past decade is radically different from previous generations of vehicles. Power semiconductors, such as GTOs (gate turn-off thyristors) and IGBTs (insulated gate bipolar transistors) combined with computer control systems have greatly enhanced the functionality of rail vehicles. New vehicles are now able to operate under different AC and DC supply networks with only moderate additional equipment effort, which is essential when traveling over country borders. It was found on the overline of the rail system of several countries that particular disturbances, generated by drive control, produce dangerous oscillation problems with resonance and sometimes instability phenomena, thus to recognize these particular frequencies and to actively suppress them is a necessary and ambitious task. Classical PWM control is very sensitive to the level of the dc link voltage but it is not sensitive to the frequency disturbance in superposition.

Theoretical aspect on wavelets are known already from many years [2] or [**?**] and in particular wavelet packets as in [1]. Nevertheless they industrial application have been recently presented. Recent technical publications see for instance [6], [10] pointed out how wavelet tools are particular suitable in order to describe power control problems and they gave already good results. There are in [5] and [4] significant contributions where the Haar functions are used in order to analyze electrical systems.

More, very recent patents, see [7] and for instance [8], make clear the interest of the industry community to invest on this area and indicate already a real possibility to apply with success wavelet packets. Preliminary simulations and theoretic considerations show a good chances to achieve several interesting results connected to the applications as well as to theoretical aspects. The presented contribution is organized as follows. In section 44.3 one states the problem in a mathematical form. In section 44.4 preliminary simulation of a roughly idea of solution is shown and in the meantime one discusses some arising problems. At the end one comes to the conclusions.

44.3 General Problem Formulation

Given the following non linear system,

$$\dot{x}(t) = f(x,t) + b(x,t)u_{sw}(t). \tag{44.1}$$

Let the tracking error be $\tilde{\mathrm{x}}(t) \in \Re$ with

$$e(t) = \tilde{\mathrm{x}}(t) = x_w(t) - x(t),$$

where the $x_w(t)$ is the wanted function.

Find a suitable $u_{sw}(t)$, switching input function, such that $\forall\ t\ e(t)$ is as smaller as possible under the following constraint condition on the system

$$\mathcal{C}\big(x(t), \dot{x}(t)\big) = 0$$

44.3.1 A Conjecture in Converter Bridge Control

One can start with a particular linear time varying system characterized as follows:

$$\dot{x}(t) = \mathbf{A}x(t) + \mathbf{B}\Big(U_n(t) - U_d(t)u_{sw}(t)\Big), \tag{44.2}$$

where the $\mathbf{A}$ and the $\mathbf{B}$ matrices are the maps of the linear system and $U_d(t)$ could be interpreted like an external disturbances, $u_{sw}(t)$ is a switching function. According to the preliminary model discussed in [9], the mathematical model becomes:

$$\dot{I}_n(t) + \frac{R}{L}I_n(t) = \frac{1}{L}[U_n(t) - U_d(t)u_{sw}(t)], \tag{44.3}$$

where $U_n(t)$ is a sinusoidal feeding voltage and $U_d(t)$ is the dc-link voltage, the disturbance is due to the charge of the motor.

In other words this system could be seen like a time-varying system. In order to 'localize' the disturbance, the problem is, roughly speaking, to find a switching function $u_{sw}(t)$ such that the secondary current of the transformer has sinousoidal shape depending only on the fundamental frequency and with the same phase as the feeding voltage.

One of the possible idea of the solution is the following.

If

$$U_d(t) < V - 0.05V, \tag{44.4}$$

where V is the required voltage level, then
find a suitable level "d" of the tree Haar wavelet functions such that

$$\min_d(\mathcal{D}(I_w(t), I_n(t)))^1 \tag{44.5}$$

and able to obtain

$$\int I_n(t)U_d(t)u_{sw}(t)d(t) \geq \frac{1}{2}CV^2. \tag{44.6}$$

Where $I_w(t)$ and $I_n(t)$ are the secondary wanted and nominal current of the transformer, $U_d(t)$ is the dc link voltage and the product CV is the required charge condition from the capacitor battery at the wanted voltage V. One can remark:

The condition (44.6) could be again easily interpreted as a distance condition to satisfy on the packet tree. In other words, the problem is to be able to find a compatible solution.

Another important aspect to remark is how, thanks to the time-frequency localization of the wavelet function, the problem of the phase shift of the signals is easily resolved by using the idea of the functional distance D.

About the constraint (44.6), one doesn't risk a dangerous over voltage because the system is technically protected from over tension.

44.4 Several Preliminary Simulations and Considerations

Preliminary simulations show how it exists an $u_{ws}(t)$ control low such that the amplitude of the disturbances of intermodulation are often considerably reduced, even though the noise level is increased. The simulations have been performed with realistic values given from ADtranz, see [9]. One considered for the simulation

$$I_w(t) = x_w(t) = 300cos(2\pi * 50 * t). \tag{44.7}$$

The disturbance with frequency of 25 Hz and amplitude of around 100 A, this condition simulates a possible start-in of the train. One can observe that the sincronismus is kept, in fact the main disturbances, intermodulation disturbance, has frequencies equal to 25 Hz and 75 Hz, 50 ± 25.
By using, a window with 8 samples then the levels of the packets are 3, a sampling time equal to 2 ms then tha analysing signal has length equal to 16 ms. Looking on the wavelet packet tree one can find a suitable level in order to find the length of the modulating input. In other words, the length of the scaling function corresponding to the selected level is the active wanted time of the switching wanted function. From the preliminary simulations one can see the two disturbances due to the modulation and how they are reduced from the proposed approach.
Arising problems:

- Does exist, for every charge condition, in the set of "d" levels, under the condition (44.6), a solution which guaranties the condition (44.5) ?
- If one considers more samples with the same sampling time then the number "d" of the levels is increased. Then, the set of the possible candidate functions is bigger, this yields more possibilities to find a suitable solution. But, a longer window increases the information on the past real time axis, this yields not necessarely a suitable solution.
- The level of the white noise is sensitively higher than with PWM pulse pattern.

[1] Where with $\mathcal{D}$ one indicates one functional distance.

44.5 Conclusion

The contribution presents an original way to perform a converter bridge control by using wavelet packets and proposes a line of research to follow. The idea could be a possible alternative to the PWM control in locomotive drive where one needs to have a high level of the $\frac{S}{N}$ ratio. This preliminary idea seems quite challenging and promising in order to guarantee the required power supply and a good power factor. One proposes a roughly algorithm which finds the length of the active switching function looking for the suitable scaling function. The open problem consists of building a rigorous algorithm which can find a switching control law.

Bibliography

[1] R. R Coifman and M. V. Wickerhauser, "Entropy based algorithm for best basis selection", *IEEE Trans. Inform. Theory*,32,712-718,(1992).

[2] R. R. Coifman, Y. Meyer and M.V. Wickerhauser, "Wavelet Analysis and Signal Processing", Publisher Jones and Bartlett,Boston (USA), 153-178,(1992).

[3] I. Daubechies,"Ten Lectures on Wavelets", Publisher Society for Industrial and Applied Mathematics, Philadelphia (Pennsylvania), (1995).

[4] A. Gandelli and A. Monti and F. Ponci and E. Santi, "Power converter control design based on haar analysis", *PESC. 2001 IEEE 32nd Annual Power Electronics Specialists Conference*,3,1742-1746,(2001).

[5] , A. Gandelli and S. Leva,"Haar state equations for power electronics system modeling", The 2001 IEEE International Symposium on Circuits and Systems 2001 ISCAS 2001,2, 4525-528,(2001).

[6] P. Mercorelli and P. Terwiesch, "A Black Box Identification in Harmonic Domain", **VDE** *European Transactions on Electrical Power*, to appear 2002.

[7] P. Mercorelli and A. Frick, "System and Methodology for Noise Level Detection by Using Wavelet Basis Functions in Wavelet Packet Trees", In Patentamt ABB Corporate Research Ladenburg (Germany),2002.

[8] , P. Mercorelli and A. Frick, "System and Methodology for Gross Error Detection by Using Wavelet Basis Functions in Wavelet Packet Trees", In Patentamt ABB Corporate Research Ladenburg (Germany),(2002).

[9] , A. Petersen, "Calculation Tools-Mathematical Background", Thecnical Report, 3EHA610201 by ADtranz, (1997).

[10] , P. Terwiesch, S. Menth and S. Schmidt, " Analysis of Transients in Electrical Railway Networks Using Wavelets", IEEE Trans. on Industrial Electronics, 45,6,955-959,(1998).

Problem 61

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
¡mailto:sklar@sus.univ.szczecin.pl¿sklar@sus.univ.szczecin.pl, ¡mailto:sklyar@univer.kharkov.ua ¿sklyar@univer.kharkov.ua

S. Yu. Ignatovich
Kharkov National University
Svoboda sqr. 4, 61077 Kharkov, Ukraine
¡mailto:bob@online.kharkiv.com¿bob@online.kharkiv.com

61.1 Motivation and history of the problem

The time-optimal control problem is one of the most natural and at the same time hard 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

$$\dot{x}_i = -t^{q_i}u,\ i = 1, \ldots, n,\ q_1 < \cdots < q_n,\ x(0) = x^0, x(\theta) = 0, |u| \le 1, \theta \to \min\,. \tag{1}$$

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

$$\dot{x} = a(t, x) + ub(t, x),\ a(t, 0) \equiv 0,\ x(0) = x^0,\ x(\theta) = 0,\ |u| \le 1,\ \theta \to \min, \tag{2}$$

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 $\mathrm{ad}\,^{m+1}_{R_a} R_b = [R_a, \mathrm{ad}^m_{R_a} R_b]$, $m \ge 0$; $\mathrm{ad}\,^0_{R_a} R_b = R_b$, where $[\cdot, \cdot]$ is the operator commutator. Denote $E(x) \equiv x$.

THEOREM 1. The conditions $\mathrm{rank}\{\mathrm{ad}^j_{R_a} R_b E(x)\big|_{\substack{t=0\\x=0}}\}_{j\ge 0} = n$ and

$$[\mathrm{ad}^{m_1}_{R_a} R_b, \cdots [\mathrm{ad}^{m_{k-1}}_{R_a} R_b, \mathrm{ad}^{m_k}_{R_a} R_b] \cdots] E(x)\big|_{\substack{t=0\\x=0}} \in \mathrm{Lin}\left\{\mathrm{ad}^j_{R_a} R_b E(x)\big|_{\substack{t=0\\x=0}}\right\}_{j=0}^{m-2} \tag{3}$$

for any $k \ge 2$ and $m_1, \ldots, m_k \ge 0$, where $m = m_1 + \cdots + m_k + k$, 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

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

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

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

$$x^0 = \sum_{m=1}^{\infty} \sum_{m_1+\cdots+m_k+k=m} v_{m_1\ldots m_k} \xi_{m_1\ldots m_k}(\theta, u), \tag{4}$$

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

$$v_{m_1\ldots m_k} = \frac{(-1)^k}{m_1! \cdots m_k!} \mathrm{ad}^{m_1}_{R_a} R_b\ \mathrm{ad}^{m_2}_{R_a} R_b \cdots \mathrm{ad}^{m_k}_{R_a} R_b E(x)\big|_{\substack{t=0\\x=0}}.$$

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

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

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

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

$$x^0_i = \xi_{q_i}(\theta, u), \qquad i = 1, \ldots, n.$$

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

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 $\xi_{m_1\ldots m_k}$ over $\mathbb{R}$ as *a free algebra* with the basis $\{\xi_{m_1\ldots m_k} : k \ge 1, m_1, \ldots, m_k \ge 0\}$ and the product $\xi_{m_1\ldots m_k}\xi_{n_1\ldots n_s} = \xi_{m_1\ldots m_k n_1 \ldots n_s}$. Introduce the inner product in $\mathcal{A}$ assuming the basis $\{\xi_{m_1\ldots m_k}\}$ 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 $\mathcal{A}_m = \mathrm{Lin}\{\xi_{m_1\ldots m_k} : \mathrm{ord}(\xi_{m_1\ldots m_k}) = m\}$.

Consider now a system of the form (2). The series in (4) naturally defines the linear mapping $v : \mathcal{A} \to \mathbb{R}^n$ by the rule $v(\xi_{m_1\ldots m_k}) = v_{m_1\ldots m_k}$. Further we assume the system (2) to be n-dimensional, i.e. $\dim v(L) = n$. Note that the form of coefficients $v_{m_1\ldots m_k}$ 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 $D_r = v(L\cap(\mathcal{A}_1+\cdots+\mathcal{A}_r)) \subset \mathbb{R}^n$, and put $r_0 = \min\{r : \dim D_r = n\}$. For any $r \le r_0$ 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, $\widetilde{\ell}_1,\dots,\widetilde{\ell}_n$ be a basis of $\sum_{r=1}^{r_0}(L_{J^\perp}\cap\mathcal{A}_r)$ such that $\mathrm{ord}(\widetilde{\ell}_i) \le \mathrm{ord}(\widetilde{\ell}_j)$ as $i<j$. Then there exists a nonsingular analytic transformation Φ of a neighborhood of the origin which reduces (4) to the following form

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

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

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

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

$$x_i^0 = \widetilde{\ell}_i, \qquad i = 1,\dots,n. \tag{6}$$

(B) Suppose there exists an open domain $\Omega \subset \mathbb{R}^n\backslash\{0\}$, $0 \in \overline{\Omega}$, 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 $(\widetilde{\theta}_{x^0}, \widetilde{u}_{x^0}(t))$, $x^0 \in \Omega$, such that the control $\widetilde{u}_{x^0}(t)$ steers the point $\Phi(x^0)$ to the origin in the time $\widetilde{\theta}_{x^0}$ w.r.t. system (2) and

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

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

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 $r_0 \ge 0$ 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 $\mathfrak{J}$ the set of all such ideals. For any $J \in \mathfrak{J}$ one can construct a control system of the form (5) such that its representation (4) is of the form (6).

61.2 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 the 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 $\mathfrak{J}$. 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 estremals in the time minimal control problem in $\mathbb{R}^3$", *SIAM J. Control Optimiz.,* 23, pp. 794-802 (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", *submitted to 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 28

$\mathcal{L}_2$-Induced Gains of Switched Linear Systems

João P. Hespanha[1]
Dept. Electrical and Computer Engineering,
University of California, Santa Barbara
USA
hespanha@ece.ucsb.edu

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.

28.1 Switched Linear Systems

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* $\mathcal{S} := \{\sigma : [0, \infty) \to \mathcal{P}\}$. 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

$$\dot{x} = A_\sigma x + B_\sigma u, \qquad y = C_\sigma x + D_\sigma u, \qquad \sigma \in \mathcal{S} \tag{28.1}$$

and by a *solution to* (28.1), we mean a pair (x, σ) for which $\sigma \in \mathcal{S}$ and x is a solution to the time-varying system

$$\dot{x} = A_{\sigma(t)} x + B_{\sigma(t)} u, \qquad y = C_{\sigma(t)} x + D_{\sigma(t)} u, \qquad t \geq 0. \tag{28.2}$$

Given a set of switching signals $\mathcal{S}$, we define the $\mathcal{L}_2$-*induced gain of* (28.1) 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 (28.1). The $\mathcal{L}_2$-induced gain of (28.1) 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.

28.2 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

[1]This material is based upon work supported by the National Science Foundation under Grant No. ECS-0093762.

$\mathcal{S}[\tau_D]$ 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 $\mathfrak{g} : [0, \infty) \to [0, \infty]$ that maps each *dwell-time* τ_D with the $\mathcal{L}_2$-induced gain of (28.1), for the set of dwell-time switching signals $\mathcal{S} := \mathcal{S}[\tau_D]$. Until recently little more was known about $\mathfrak{g}$ other than the following:

1. $\mathfrak{g}$ is monotone decreasing
2. $\mathfrak{g}$ is bounded below by

$$\mathfrak{g}_{\text{static}} := \sup_{p \in \mathcal{P}} \|C_p(sI - A_p)^{-1}B_p + D_p\|_\infty,$$

   where $\|T\|_\infty := \sup_{\Re[s] \geq 0} \|T(s)\|$ denotes the $\mathcal{H}_\infty$-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 $\mathfrak{g}_{\text{static}}$ is strict and in general there is a gap between $\mathfrak{g}_{\text{static}}$ and

$$\mathfrak{g}_{\text{slow}} := \lim_{\tau_D \to \infty} \mathfrak{g}[\tau_D].$$

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 $\mathfrak{g}_{\text{slow}}$. Opposite to what had been conjectured, $\mathfrak{g}_{\text{slow}}$ is realization dependent and cannot be determined just from the transfer functions of the systems being switched.

The function $\mathfrak{g}$ thus looks roughly like the ones shown in Figure 28.1, 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 $\tau_{\min}$ denotes the smallest dwell-time for which instability can occur for some switching signal in $\mathcal{S}[\tau_{\min}]$.

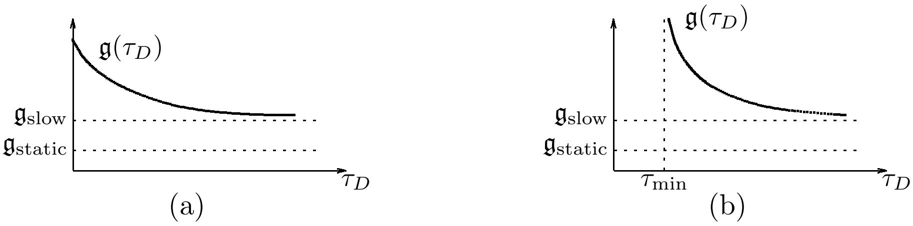


Figure 28.1: $\mathcal{L}_2$-induced gain versus the dwell-time.

Several important basic question remain open:

1. Under what conditions is $\mathfrak{g}$ 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 $\mathfrak{g}$ is unbounded (case (b) in Figure 28.1), how to compute the position of the vertical asymptote? Or equivalently, what is the smallest dwell-time $\tau_{\min}$ for which one can have instability.
3. Is $\mathfrak{g}$ a convex function? Is it smooth (or even continuous)?

Even if direct computation of $\mathfrak{g}$ 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 known, currently only very course upper-bounds for $\mathfrak{g}$ are available. These are obtained by computing a conservative upper-bound τ_{upper} for $\tau_{\min}$ and then an upper-bound for $\mathfrak{g}$ 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." To be 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 48

Is Monopoli's Model Reference Adaptive Controller Correct?

A. S. Morse [1]
Center for Computational Vision and Control
Department of Electrical Engineering
Yale University, New Haven, CT 06520 USA.

48.1 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.

48.2 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 modelled siso process $\mathbb{P}$ to approach and track the output y_{ref} of a pre-specified reference model $\mathbb{M}_{\text{ref}}$ 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 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 pre-specified 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 $\mathbb{M}_{\text{ref}}$ is a system of the form

$$\dot{y}_{\text{ref}} = -\lambda y_{\text{ref}} + \bar{c}x_{\text{ref}} + \bar{d}r \qquad \dot{x}_{\text{ref}} = \bar{A}x_{\text{ref}} + \bar{b}r \tag{48.1}$$

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

48.3 Monopoli's Proposed Solution

Monopoli's proposed solution is based on a special representation of $\mathbb{P}$ which 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 $\mathbb{M}_{\text{ref}}$ to obtain the *direct control model reference parameterization*

$$\dot{z} = \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} z + \begin{bmatrix} b \\ 0 \end{bmatrix} y + \begin{bmatrix} 0 \\ b \end{bmatrix} u \tag{48.2}$$

$$\dot{x} = \bar{A}x + \bar{b}(u - z'p^* - r) \tag{48.3}$$

$$\dot{\mathbf{e}}_{\mathbf{T}} = -\lambda \mathbf{e}_{\mathbf{T}} + \bar{c}x + \bar{d}(u - z'p^* - r) \tag{48.4}$$

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

$$\mathbf{e}_{\mathbf{T}} \stackrel{\Delta}{=} y - y_{\text{ref}} \tag{48.5}$$

and $x \stackrel{\Delta}{=} \bar{x} - x_{\text{ref}}$. Note that it is possible to generate an asymptotically correct estimate $\widehat{z}$ of z using a copy of (48.2) with $\widehat{z}$ replacing z. To keep the exposition simple we're going to ignore the exponentially decaying estimation error $\widehat{z} - z$ and assume that z can be measured directly.

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

$$u = z'\widehat{p} + r \tag{48.6}$$

where $\widehat{p}$ 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 (48.6), then this would cause $\mathbf{e}_{\mathbf{T}}$ to tend to zero exponentially fast and tracking would therefore be achieved.

Monopoli proposed to generate $\widehat{p}$ using two sub-systems which we will refer to here as a "multi-estimator" and a " tuner" respectively. A *multi-estimator* $\mathbb{E}(\widehat{p})$ is a parameter-varying linear system with parameter $\widehat{p}$, whose inputs are u, y and r and whose output is an estimate $\widehat{e}$ of $\mathbf{e}_{\mathbf{T}}$ which would be asymptotically correct were $\widehat{p}$ held fixed at p^*. It turns out that there are two different but very similar types of multi-estimators which 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 (48.2) as a sub-system.

48.3.1 Version 1

There are two versions of the adaptive controller which a relevant to the problem at hand. In this section we describe the multi-estimator and tuner which together with reference model (48.1) and control law (48.6), comprise the first version.

Multi-Estimator 1

The form of the first multi-estimator $\mathbb{E}_1(\widehat{p})$ 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

$$\dot{H}_1 = \bar{A}H_1 + \bar{b}z' \qquad \text{and} \qquad \dot{w}_1 = \bar{A}w_1 + \bar{b}(u - r) \tag{48.7}$$

respectively, then $w_1 - H_1 p^*$ is a solution to (48.3). In other words $x = w_1 - H_1 p^* + \epsilon$ 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 (48.4) can be re-written as

$$\dot{\mathbf{e}}_{\mathbf{T}} = -\lambda \mathbf{e}_{\mathbf{T}} - (\bar{c}H_1 + \bar{d}z')p^* + \bar{c}w_1 + \bar{d}(u - r)$$

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

$$\dot{\widehat{e}}_1 = -\lambda \widehat{e}_1 - (\bar{c}H_1 + \bar{d}z')\widehat{p} + \bar{c}w_1 + \bar{d}(u - r) \tag{48.8}$$

From this it clearly follows that the multi-estimator $\mathbb{E}_1(\widehat{p})$ defined by (48.2), (48.7) and (48.8) has the required property of delivering an asymptotically correct estimate $\widehat{e}_1$ of $\mathbf{e_T}$ if $\widehat{p}$ is fixed at p^*.

Tuner 1

From (48.8) and the differential equation for $\mathbf{e_T}$ directly above it, it can be seen that the estimation error[2]

$$e_1 \stackrel{\Delta}{=} \widehat{e}_1 - \mathbf{e_T} \tag{48.9}$$

satisfies the error equation

$$\dot{e}_1 = -\lambda e_1 + \phi_1'(\widehat{p} - p^*) \tag{48.10}$$

where

$$\phi_1' = -(\bar{c}H_1 + \bar{d}z') \tag{48.11}$$

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

$$\dot{\widehat{p}}_1 = -\phi_1 e_1 \tag{48.12}$$

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

What's Known About Version 1?

The overall model reference adaptive controller proposed by Monopoli thus consists of the reference model (48.1), the control law (48.6), the multi-estimator (48.2), (48.7), (48.8), the output estimation error (48.9) and the tuner (48.11), (48.12). *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 which is relevant to the problem. In the first place, note that (48.1), (48.2) together with (48.5) - (48.11) define a parameter varying linear system $\Sigma_1(\widehat{p})$ with input r, state $\{y_{\text{ref}}, x_{\text{ref}}, z, H_1, w_1, \widehat{e}_1, e_1\}$ and output e_1. The consequence of the assumption that every system in $\mathcal{M}$ is minimum phase is that $\Sigma_1(\widehat{p})$ is detectable through e_1 for every fixed value of $\widehat{p}$ [5]. Meanwhile the form of (48.10) enables one to show by direct calculation, that the rate of change of the partial Lyapunov function $V \stackrel{\Delta}{=} e_1^2 + ||\widehat{p} - p^*||^2$ along a solution to (48.12) and the equations defining $\Sigma_1(\widehat{p})$, satisfies

$$\dot{V} = -2\lambda e_1^2 \leq 0 \tag{48.13}$$

From this it is evident that V is a bounded monotone nonincreasing function and consequently that e_1 and $\widehat{p}$ are bounded wherever they exist. Using and the fact that $\Sigma_1(\widehat{p})$ is a *linear* parameter-varying system, it can be concluded that solutions exist globally and that e_1 and $\widehat{p}$ are bounded on $[0, \infty)$. By integrating (48.13) it can also be concluded that e_1 has a finite $\mathcal{L}^2[0, \infty)$-norm and that $||e_1||^2 + ||\widehat{p} - p^*||^2$ tends to a finite limit as $t \to \infty$. Were it possible to deduce from these properties that $\widehat{p}$ tended to a limit $\bar{p}$, then it would possible to establish correctness of the overall adaptive controller using the detectability of $\Sigma_1(\bar{p})$.

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 $\widehat{p} \stackrel{\Delta}{=} \widehat{q}k$ where $\widehat{q}$ is a scalar parameter tuned by the equation $\dot{\widehat{q}} = -k'\phi_1 e_1$ [6].

[2] Monopoli called e_1 an *augmented error.*

48.3.2 Version 2

In the sequel we describe the multi-estimator and tuner which together with reference model (48.1) and control law (48.6), comprise the second version of them adaptive controller relevant to the problem at hand.

Multi-Estimator 2

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

$$\bar{A}_2 \stackrel{\Delta}{=} \begin{bmatrix} \bar{A} & 0 \\ \bar{c} & -\lambda \end{bmatrix} \qquad \text{and} \qquad \bar{b}_2 \stackrel{\Delta}{=} \begin{bmatrix} \bar{b} \\ \bar{d} \end{bmatrix}$$

The form of $\mathbb{E}_2(\widehat{p})$ 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

$$\dot{H}_2 = \bar{A}_2 H_2 + \bar{b}_2 z' \qquad \text{and} \qquad \dot{w}_2 = \bar{A}_2 w_2 + \bar{b}_2 (u - r) \tag{48.14}$$

then $w_2 - H_2 p^*$ is a solution to (48.3) - (48.4). In other words, $[\ x' \quad \mathbf{e_T}\ ]' = w_2 - H_2 p^* + \epsilon$ 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 $\bar{c}_2 = [\ 0 \ \ \cdots \ \ 0 \ \ 1\ ]$. Thus in this case a natural way to generate an estimate $\widehat{e}_2$ of $\mathbf{e_T}$ is by means of the equation

$$\widehat{e}_2 = \bar{c}_2 w_2 - \bar{c}_2 H_2 \widehat{p} \tag{48.15}$$

It is clear that the multi-estimator $\mathbb{E}_2(\widehat{p})$ defined by (48.2), (48.14) and (48.15) has the required property of delivering an asymptotically correct estimate $\widehat{e}_2$ of $\mathbf{e_T}$ if $\widehat{p}$ is fixed at p^*.

Tuner 2

Note that in this case the estimation error

$$e_2 \stackrel{\Delta}{=} \widehat{e}_2 - \mathbf{e_T} \tag{48.16}$$

satisfies the error equation

$$e_2 = \phi_2'(\widehat{p}_2 - p^*) \tag{48.17}$$

where

$$\phi_2' = -\bar{c}_2 H_2 \tag{48.18}$$

Equation (48.17) suggests that one consider a pseudo-gradient tuner of the form

$$\dot{\widehat{p}} = -\phi_2 e_2 \tag{48.19}$$

What's Known About Version 2?

The overall model reference adaptive controller in this case, thus consists of the reference model (48.1), the control law (48.6), the multi-estimator (48.2), (48.14), (48.15), the output estimation error (48.16) and the tuner (48.18), (48.19). *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, (48.1), (48.2) together with (48.5), (48.6) (48.14) - (48.18) define a parameter varying linear system $\Sigma_2(\widehat{p})$ 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 $\Sigma_2(\widehat{p})$ is detectable through e_2 for every fixed value of $\widehat{p}$ [5]. Meanwhile the form of (48.17) enables one to show by direct calculation, that the rate of change of the partial Lyapunov function $V \stackrel{\Delta}{=} ||\widehat{p} - p^*||^2$ along a solution to (48.19) and the equations defining $\Sigma_2(\widehat{p})$, satisfies

$$\dot{V} = -2\lambda e_2^2 \leq 0 \tag{48.20}$$

It is evident that V is a bounded monotone nonincreasing function and consequently that $\widehat{p}$ is bounded wherever they exist. From this and the fact that $\Sigma_2(\widehat{p})$ is a linear parameter-varying system, it can be concluded that solutions exist globally and that $\widehat{p}$ is bounded on $[0, \infty)$. By integrating (48.20) it can also be concluded that e_2 has a finite $\mathcal{L}^2[0, \infty)$-norm and that $||\widehat{p} - p^*||^2$ tends to a finite limit as $t \to \infty$. Were it possible to deduce from these properties that $\widehat{p}$ tended to a limit $\bar{p}$, then it would to 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 $\widehat{p} \stackrel{\Delta}{=} \widehat{q}k$ where $\widehat{q}$ is a scalar parameter tuned by the equation $\dot{\widehat{q}} = -k'\phi_2 e_2$. The underlying reason why things go through is because in this special case, the fact that $||\widehat{p} - p^*||^2$ and consequently $||\widehat{q} - q^*||$ tend to a finite limits, means that $\widehat{q}$ tends to a finite limit as well.

48.4 The Essence of the Problem

In this section we write down a stripped down version of the problem which retains all the essential feature which 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 $\bar{A}_2 = -\lambda$ 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

$$\dot{z} = \begin{bmatrix} A & 0 \\ 0 & A \end{bmatrix} z + \begin{bmatrix} b \\ 0 \end{bmatrix} (w + \phi' p^*) + \begin{bmatrix} 0 \\ b \end{bmatrix} \widehat{p}' z \tag{48.21}$$

$$\dot{\phi} = -\lambda\phi - z \tag{48.22}$$

$$\dot{w} = -\lambda w + \widehat{p}' z \tag{48.23}$$

$$e = \phi'(\widehat{p} - p^*) \tag{48.24}$$

$$\dot{\widehat{p}} = -\phi e \tag{48.25}$$

To recap, p^* is unknown and constant but is such that the linear parameter-varying system $\Sigma(\widehat{p})$ defined by (48.21) to (48.24) is detectable through e for each fixed value of $\widehat{p}$. Solutions to the system (48.21) - (48.25) exist globally. The parameter vector $\widehat{p}$ and integral square of e are bounded on $[0, \infty)$ and $||\widehat{p} - p^*||$ tends to a finite limit as $t \to \infty$. *The open problem here is to show for every initialization of (48.21)-(48.25), that the state of $\Sigma(\widehat{p})$ 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*, pages 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.* PhD thesis, Yale University, 1978.

Problem 82

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
ziarani@ieee.org

82.1 Description of the problem

We are concerned about the mathematical properties of the dynamical system presented by the following three differential equations:

$$\begin{cases} \frac{dA}{dt} &= -2\mu_1 A \sin^2\phi + 2\mu_1 \sin\phi\ f(t), \\ \frac{d\omega}{dt} &= -\mu_2 A^2 \sin(2\phi) + 2\mu_2 A \cos\phi\ f(t), \\ \frac{d\phi}{dt} &= \omega + \mu_3 \frac{d\omega}{dt} \end{cases} \tag{82.1}$$

where parameters $\mu_i,\ i = 1, 2, 3$ are positive real constants and $f(t)$ is a function of time having a general form of

$$f(t) = A_o \sin(\omega_o t + \delta_o) + f_1(t). \tag{82.2}$$

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 (82.1) 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

$$\Gamma_o(t) = (A_o, \omega_o, \omega_o t + \delta_o) \tag{82.3}$$

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 82.1.

Theorem 1: Consider the dynamical system presented by the set of ordinary differential equations (82.1) in which the function $f(t)$ is defined in (82.2) 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 $\mu_i,\ i = 1, 2, 3$ are small positive real numbers. If $T_1 = \frac{T_o}{n}$ for any arbitrary $n \in \mathbb{N}$, the dynamical system of (82.1) has a stable T_o-periodic orbit in a neighborhood of $\Gamma_o(t)$ as defined in (82.3).

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 (82.1) has a stable mT_o-periodic orbit which lies on a torus in a neighborhood of $\Gamma_o(t)$ as defined in (82.3). 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 (82.1) has an attractor set which is a torus in a neighborhood of $\Gamma_o(t)$ as defined in (82.3). In other words, the response is a never closing orbit which 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 $\mu_i,\ i = 1, 2, 3$ 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 (82.1). 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 $\mu_i,\ i = 1, 2, 3$, hence the following open problem:

Open Problem: Consider the dynamical system presented by the ordinary differential equations (82.1) in which the function $f(t)$ is defined in (82.2) 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 $\mu_i,\ i = 1, 2, 3$ are small positive real numbers. This system has an attractor set which 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 $\mu_i,\ i = 1, 2, 3$. 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 determine the whole region of points (A, ω, ϕ) in $\mathbb{R}^1 \times \mathbb{R}^1 \times \mathbb{S}^1$ which falls in the attraction domain of the attractor.

82.2 Motivation and history of the problem

The dynamical system presented by (82.1) was proposed by the authors for the first time [1, 2]. The primary motivation was to devise a system for the extraction of a sinusoidal component with time-varying parameters when it is corrupted by other sinusoids and noise. 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 [5].

82.3 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 [6, page 70]. Extensive computer simulations verified by laboratory experimental results are presented in [1, 2, 5]. 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", accepted for publication in *International Journal of Engineering Mathematics.*

[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] M. Karimi-Ghartemani and M.R. Iravani, "An Alternative Structure for Phase-Locked Loop (PLL) System", submitted to *IEEE Transactions on Circuits and Systems I: Fundamental Theory and Applications.*

[6] S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, New York: Springer-Verlag (1983).

Problem 70

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

70.1 Description of the problem

At the request of Vincent Blondel, the following control problem has been formulated.

Problem - 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 on-line information on their partial observations, state estimates, or input values but their 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 - Decentralized control of a discrete-event system with communication between supervisors Consider a discrete-event system in the form of a generator and r supervisors,

$$\begin{aligned} G &= (Q, E, f, q_0),\ L(G) = \{s \in E^* | f(q_0, s) \text{ is defined}\}, \\ & \forall k \in \mathbb{Z}_r,\ \text{a partition},\ E = E_{c,k} \cup E_{uc,k},\ E_{cp,k} = \{E_s \subseteq E | E_s \subseteq E_{cp,k}\}, \\ & \forall k \in \mathbb{Z}_r,\ \text{a partition},\ E = E_{o,k} \cup E_{uo,k},\ p_k : E \to E_{o,k},\ \forall k \in \mathbb{Z}_r, \\ & \text{an event is enabled only if it is enabled by all supervisors}, \\ & \{v_k : p_k(L(G)) \to E_{cp,k}, \forall k \in \mathbb{Z}_r\}, \\ & \text{the set of supervisors based on partial observations}, \\ & L_r, L_a \subseteq L(G),\ \text{required and admissable language, respectively.} \end{aligned}$$

A variant of the problem is to determine a set of subsets of the event set which represent the events to be communicated by each supervisor to the other supervisors and a set of supervisors such that the closed-loop system satisfies,

$$\begin{aligned} & \forall (i,j) \in \mathbb{Z}_r \times \mathbb{Z}_r,\ E_{o,i,j} \subseteq E_{o,i},\ p_{i,j} : E \to E_{o,i,j}, \\ & \{v_k(p_k(s), p_{1,k}(s), \ldots, p_{r,k}(s)) \mapsto E_{cp,k}, \forall k \in \mathbb{Z}_r\}, \\ & \text{the set of supervisors based on partial observations and on communications}, \\ & L(v_1 \wedge \ldots \wedge v_r/G) \subseteq L_a, \text{`the closed-loop language, such that}, \\ L_r &\subseteq L(v_1 \wedge \ldots \wedge v_r/G) \subseteq L_a,\ \text{and the controlled system is nonblocking.} \end{aligned}$$

Problem - Decentralized control of a finite-dimensional linear system with communication between controllers . Consider a finite-dimensional linear system with r input signals and r output signals,

$$\begin{aligned} \dot{x}(t) &= Ax(t) + \sum_{k=1}^{r} B_k u_k(t),\ x(t_0) = x_0, \\ y_j(t) &= C_j x(t) + \sum_{k=1}^{r} D_{j,k} u_k(t),\ \ \forall\ j \in \mathbb{Z}_r = \{1, 2, \ldots, r\}, \end{aligned}$$

where the dimensions of the state, the input signals, the output signals, and of the matrices have been omitted. The ith controller observes output y_i and provides to the system input u_i. Suppose that Controller 2 communicates 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 minization? The period will have to take account of the feedback achievable time constants of the system. A further restriction on the communication channel ist 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].

70.2 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 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 view point 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 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 relative cheap and therefore the trade-off has shifted towards the use of more communciation.

70.3 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 necesary 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, 1, 12, 13, 11, 14, 15, 16, 19]. Besides the control problem, the corresponding problem for failure diagnosis has also been analyzed, see [6, 7, 9, 8]. 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.

70.4 Approach

Suggestions for the solution of the problem follow. 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 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 fix point 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 is then 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. *TITLE.* PhD 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, 1998. IEEE Press.

[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*, pages 3281–3286, New York, 1998. IEEE Press.

[4] G. Barrett and S. Lafortune. Some issues concerning decentralized supervisory control with communication. In *Proceedings 38th IEEE Conference on Decision and Control*, pages 2230–2236, New York, 1999. IEEE Press.

[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 Jan H. van Schuppen. Decentralized failure diagnosis for discrete-event systems with constrained communication between diagnosers. In *Proceedings Workshop on Discrete-Event Systems*, 2002.

[7] R. Debouk. *Failure diagnosis of decentralized discrete-event systems.* PhD 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] Rami Debouk, Stéphane Lafortune, and Demosthesis Teneketzis. Coordinated decentralized protocols for failure diagnosis of discrete event systems. Report CGR-97-17, College of Engineering, University of Michcigan, 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, editors, *Decision and organization*, pages 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.* PhD 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)*, pages 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*, pages 1326–1332, New York, 1999. IEEE Press.

[15] S.L. Ricker and J.H. van Schuppen. Asynchronous communication in timed discrete event systems. In *Proceedings of the American Control Conference (ACC2001)*, pages 305–306, 2001.

[16] S.L. Ricker and J.H. van Schuppen. Decentralized failure diagnosis with asynchronuous communication between supervisors. In *Proceedings of the European Control Conference (ECC2001)*, pages 1002–1006, 2001.

[17] S.C. Tatikonda. *Control under communication constraints.* PhD 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)*, pages 36–41, London, 1998. IEE.

[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)*, pages 284–289, London, 1996. IEE.

Problem 80

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

80.1 Introduction

In recent years, as a consequence of the dramatic increases in computing power and of the continuing refinement of the numerical algorithms available, it has become possible the numerical treatment of control problems for systems governed by partial differential equations 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 nonlinear 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. That is 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 the special features present in each system governed by partial differential equations. So that, at least in some special cases, should be possible to improve on it.

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 section 80.2 we summarize the results obtained in [6], [2], and in section 80.3 we present two problems that we belive can be approached in a way similar to the one described in [6], [2].

80.2 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.

80.3 Two control problems

Let $\mathbf{R}$ be the set of real numbers, $\boldsymbol{x} = (x_1, x_2, x_3)^T \in \mathbf{R}^3$ (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 $\boldsymbol{n}(\boldsymbol{x}) = (n_1(\boldsymbol{x}), n_2(\boldsymbol{x}), n_3(\boldsymbol{x}))^T \in \mathbf{R}^3$, $\boldsymbol{x} \in \partial\Omega$ be the outward unit normal vector in $\boldsymbol{x}$ for $\boldsymbol{x} \in \partial\Omega$. Note that $\boldsymbol{n}(\boldsymbol{x})$ exists almost everywhere in $\boldsymbol{x}$ for $\boldsymbol{x} \in \partial\Omega$. 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 $(\boldsymbol{\mathcal{E}}^i(\boldsymbol{x},t), \boldsymbol{\mathcal{B}}^i(\boldsymbol{x},t)), (\boldsymbol{x},t) \in \mathbf{R}^3 \times \mathbf{R}$ (where $\boldsymbol{\mathcal{E}}^i$ is the electric field and $\boldsymbol{\mathcal{B}}^i$ is the magnetic induction field) be the incoming electomagnetic field propagating in the medium filling $\mathbf{R}^3 \setminus \Omega$ and satisfying the Maxwell equations (80.1)-(80.3) in $\mathbf{R}^3 \times \mathbf{R}$. Let $(\boldsymbol{\mathcal{E}}^s(\boldsymbol{x},t), \boldsymbol{\mathcal{B}}^s(\boldsymbol{x},t)), (\boldsymbol{x},t) \in (\mathbf{R}^3 \setminus \overline{\Omega}) \times \mathbf{R}$ be the electromagnetic field scattered by the obstacle Ω when hit by the incoming field $(\boldsymbol{\mathcal{E}}^i(\boldsymbol{x},t), \boldsymbol{\mathcal{B}}^i(\boldsymbol{x},t)), (\boldsymbol{x},t) \in \mathbf{R}^3 \times \mathbf{R}$. The scattered electric field $\boldsymbol{\mathcal{E}}^s$ and the scattered magnetic induction field $\boldsymbol{\mathcal{B}}^s$ satisfy the following equations:

$$\left(curl \boldsymbol{\mathcal{E}}^s + \frac{\partial \boldsymbol{\mathcal{B}}^s}{\partial t} \right)(\boldsymbol{x},t) = \mathbf{0},\ (\boldsymbol{x},t) \in (\mathbf{R}^3 \setminus \overline{\Omega}) \times \mathbf{R}, \tag{80.1}$$

$$\left(curl \boldsymbol{\mathcal{B}}^s - \frac{1}{c^2} \frac{\partial \boldsymbol{\mathcal{E}}^s}{\partial t} \right)(\boldsymbol{x},t) = \mathbf{0},\ (\boldsymbol{x},t) \in (\mathbf{R}^3 \setminus \overline{\Omega}) \times \mathbf{R}, \tag{80.2}$$

$$div\,\mathcal{B}^s(\boldsymbol{x},t)=0,\quad div\,\mathcal{E}^s(\boldsymbol{x},t)=0,\quad (\boldsymbol{x},t)\in(\mathbf{R}^3\setminus\overline{\Omega})\times\mathbf{R}, \tag{80.3}$$

$$\begin{array}{l}[\boldsymbol{n}(\boldsymbol{x}),\mathcal{E}^s(\boldsymbol{x},t)]-c\chi[\boldsymbol{n}(\boldsymbol{x}),[\boldsymbol{n}(\boldsymbol{x}),\mathcal{B}^s(\boldsymbol{x},t)]]=\\ -[\boldsymbol{n}(\boldsymbol{x}),\mathcal{E}^i(\boldsymbol{x},t)]+c\chi[\boldsymbol{n}(\boldsymbol{x}),[\boldsymbol{n}(\boldsymbol{x}),\mathcal{B}^i(\boldsymbol{x},t)]],\ (\boldsymbol{x},t)\in\partial\Omega\times\mathbf{R},\end{array} \tag{80.4}$$

$$\mathcal{E}^s(\boldsymbol{x},t)=O\left(\frac{1}{r}\right),\ [\mathcal{B}^s(\boldsymbol{x},t),\hat{\boldsymbol{x}}]-\frac{1}{c}\mathcal{E}^s(\boldsymbol{x},t)=o\left(\frac{1}{r}\right),\ r\to+\infty,\ t\in\mathbf{R}, \tag{80.5}$$

where $\mathbf{0}=(0,0,0)^T$, $c=1/\sqrt{\epsilon\,\nu}$, $r=\|\boldsymbol{x}\|$, $\boldsymbol{x}\in\mathbf{R}^3$, $\hat{\boldsymbol{x}}=\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$, $\boldsymbol{x}\neq\mathbf{0},\boldsymbol{x}\in\mathbf{R}^3$, $O(\cdot)$ and $o(\cdot)$ are the Landau symbols, and $curl\cdot$ and $div\cdot$ denote the curl and the divergence operator of $\cdot$ with respect to the $\boldsymbol{x}$ 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 $(\mathcal{E}^s(\boldsymbol{x},t),\mathcal{B}^s(\boldsymbol{x},t))$, $(\boldsymbol{x},t)\in(\mathbf{R}^3\setminus\overline{\Omega})\times\mathbf{R}$ solution of (80.1)-(80.5). 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'. More in details 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* $\overline{D}\subset\Omega$ *with electromagnetic boundary impedance* χ'*, choose a vector control function* $\boldsymbol{\psi}$ *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.* $(\mathcal{E}^s_D,\mathcal{B}^s_D)$*, 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 $\boldsymbol{\psi}$ 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 $\tilde{\boldsymbol{v}}=\tilde{\boldsymbol{v}}(\boldsymbol{x},t)$, $(\boldsymbol{x},t)\in\Omega_t\times\mathbf{R}$. 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 $\boldsymbol{w}=\boldsymbol{w}(\boldsymbol{\xi},t), t\in\mathbf{R}$ of the center of mass of the obstacle (i.e. $\boldsymbol{w}=\dfrac{d\boldsymbol{\xi}}{dt}, t\in\mathbf{R}$), 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 $\tilde{\boldsymbol{v}}(\boldsymbol{x},t)$, $(\boldsymbol{x},t)\in\Omega_t\times\mathbf{R}$ can be expressed in terms of $\boldsymbol{w}(\boldsymbol{\xi},t),\boldsymbol{\omega}(\boldsymbol{\xi},t), t\in\mathbf{R}$. 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 $\boldsymbol{v}=(v_1,v_2,v_3)^T$ and p be the velocity field and the pressure field of the fluid respectively, $\boldsymbol{f}$ be the density of the external forces per mass unit acting on the fluid, and $\boldsymbol{v}_{-\infty}$ be an assigned solenoidal vector field. We assume that in the limit $t\to-\infty$ the body Ω_t is at rest in the position $\Omega_{-\infty}$. Under these assumptions we have that in the reference frame given by $\boldsymbol{x}=(x_1,x_2,x_3)^T$ the following system of Navier-Stokes equations holds:

$$\begin{array}{r}\dfrac{\partial\boldsymbol{v}}{\partial t}(\boldsymbol{x},t)+(\boldsymbol{v}(\boldsymbol{x},t),\nabla)\,\boldsymbol{v}(\boldsymbol{x},t)-\eta\Delta\boldsymbol{v}(\boldsymbol{x},t)+\nabla p(\boldsymbol{x},t)=\boldsymbol{f}(\boldsymbol{x},t),\\ (\boldsymbol{x},t)\in(\mathbf{R}^3\setminus\overline{\Omega}_t)\times\mathbf{R}\,,\end{array} \tag{80.6}$$

$$div\,\boldsymbol{v}(\boldsymbol{x},t)=0,\ (\boldsymbol{x},t)\in(\mathbf{R}^3\setminus\overline{\Omega}_t)\times\mathbf{R}\,, \tag{80.7}$$

$$\lim_{t\to-\infty}\boldsymbol{v}(\boldsymbol{x},t)=\boldsymbol{v}_{-\infty}(\boldsymbol{x}),\ \boldsymbol{x}\in\mathbf{R}^3\setminus\overline{\Omega}_{-\infty}\,,\boldsymbol{v}(\boldsymbol{x},t)=\tilde{\boldsymbol{v}}(\boldsymbol{x},t),\ (\boldsymbol{x},t)\in\partial\Omega_t\times\mathbf{R}. \tag{80.8}$$

In (80.6) we have $\nabla = \left(\frac{\partial}{\partial x_1}, \frac{\partial}{\partial x_2}, \frac{\partial}{\partial x_3}\right)^T$ and $(\boldsymbol{v}, \nabla)\boldsymbol{v} = \left(\sum_{j=1}^3 v_j \frac{\partial v_1}{\partial x_j}, \sum_{j=1}^3 v_j \frac{\partial v_2}{\partial x_j}, \ \sum_{j=1}^3 v_j \frac{\partial v_3}{\partial x_j}\right)^T$. The boundary condition in (80.8) requires that the fluid velocity $\boldsymbol{v}$ and the velocity of the obstacle $\tilde{\boldsymbol{v}}$ 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 manoeuvre $\boldsymbol{w}(\boldsymbol{\xi}, t)$, $\boldsymbol{\omega}(\boldsymbol{\xi}, t), t \in \mathbf{R}$ connecting two given states that minimizes the work done by the obstacle $\Omega_t, t \in \mathbf{R}$ against the fluid going from the initial state to the final state, and the "size" of the manoeuvre employed. Note that in this context a manoeuvre connecting two given states is made of two functions $\boldsymbol{w}(\boldsymbol{\xi}, t), \boldsymbol{\omega}(\boldsymbol{\xi}, t), t \in \mathbf{R}$ such that $\lim_{t\to\pm\infty} \boldsymbol{w}(\boldsymbol{\xi}, t) = \boldsymbol{w}^{\pm}$ and $\lim_{t\to\pm\infty} \boldsymbol{\omega}(\boldsymbol{\xi}, t) = \boldsymbol{\omega}^{\pm}$, where $\boldsymbol{w}^{\pm}$ and $\boldsymbol{\omega}^{\pm}$ are preassigned. The couple $(\boldsymbol{w}^-, \boldsymbol{\omega}^-)$ is the initial state and the couple $(\boldsymbol{w}^+, \boldsymbol{\omega}^+)$ is the final state. For simplicity we have assumed $(\boldsymbol{w}^-, \boldsymbol{\omega}^-) = (\mathbf{0}, \mathbf{0})$. We formulate the following problem:

Problem 2 "Drag" Optimization Problem: *Given a rigid obstacle $\Omega_t, t \in \mathbf{R}$ 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 $(\boldsymbol{w}^+, \boldsymbol{\omega}^+)$, choose a manoeuvre connecting these two states in order to minimize a cost functional that measures the work that the obstacle $\Omega_t, t \in \mathbf{R}$ must exert on the fluid to make the manoeuvre, and the "size" of the manoeuvre 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. 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 $(\boldsymbol{\mathcal{E}}^s_D(\boldsymbol{x}, t), \boldsymbol{\mathcal{B}}^s_D(\boldsymbol{x}, t)), (\boldsymbol{x}, t) \in (\mathbf{R}^3 \setminus \overline{D}) \times \mathbf{R}$ is the solution of problem (80.1)-(80.5) 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 involving different 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", submitted to *Third ISAAC Congress Proceedings* (2002).

[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. Gamkrélidzé and F. Micktckenko, "Théorie Mathématique des Processus Optimaux", Editions Mir, Moscou (1974).

[8] S. S. Sritharan, "An introduction to deterministic and stochastic control of viscous flow", in *Optimal Control of Viscous Flow*, S. S. Sritharan Editor, SIAM, Philadelphia, pp. 1-42 (1998).

Problem 19

A conjecture on Lyapunov equations and principal angles in subspace identification

Katrien De Cock and Bart De Moor[1]
Dept. of Electrical Engineering (ESAT–SCD)
K.U.Leuven
Kasteelpark Arenberg 10
B-3001 Leuven
Belgium
Tel. 32/16/32 17 09, Fax 32/16/32 19 70
http://www.esat.kuleuven.ac.be/sista-cosic-docarch
decock@esat.kuleuven.ac.be, demoor@esat.kuleuven.ac.be

19.1 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 19.2, we will describe its system theoretic consequences and interpretation.

Conjecture 2 *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 $\lambda_i \lambda_j = 1$ $(i, j = 1, \ldots, 2n)$. If the $n \times n$ matrices P, Q and R satisfy the Lyapunov equation*

$$\begin{pmatrix} P & R \\ R^T & Q \end{pmatrix} = \begin{pmatrix} A & 0 \\ 0 & (A + vw^T)^T \end{pmatrix} \begin{pmatrix} P & R \\ R^T & Q \end{pmatrix} \begin{pmatrix} A^T & 0 \\ 0 & A + vw^T \end{pmatrix} + \begin{pmatrix} v \\ w \end{pmatrix} \begin{pmatrix} v^T & w^T \end{pmatrix} , \quad (19.1)$$

and P, Q and $(I_n + PQ)$ are non-singular[2], then the matrices $P^{-1}RQ^{-1}R^T$ and $(I_n + PQ)^{-1}$ have the same eigenvalues.

[1]Katrien 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 PhD, postdoc & fellow grants; Flemish Government: Fund for Scientific Research Flanders (several PhD and postdoc 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.

[2]The matrix I_n is the $n \times n$ identity matrix.

Note that the condition $\lambda_i \lambda_j \neq 1$ ($\forall i, j = 1, \ldots, 2n$) ensures that there exists a solution $\begin{pmatrix} P & R \\ R^T & Q \end{pmatrix}$ of the Lyapunov equation (19.1) 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 19.2.3. 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.

19.2 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 19.2.1) and their statistical counterparts, the canonical correlations of random variables (Section 19.2.2). Next, in Section 19.2.3 we will show how the conjecture – when proved correct – would enable us to prove in an elegant way that the non-zero 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 19.2.3). Moreover, by this new characterization of the canonical correlations we gain insight in the geometric properties of subspace based techniques.

19.2.1 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 $\theta_1, \ldots, \theta_p$, and the corresponding principal directions $u_i \in S_1$ and $v_i \in S_2$ ($i = 1, \ldots, p$) are recursively defined as

$$\cos\theta_1 = \max_{u \in S_1} \max_{v \in S_2} |u^T v| = u_1^T v_1$$
$$\cos\theta_k = \max_{u \in S_1} \max_{v \in S_2} |u^T v| = u_k^T v_k \quad (k = 2, \ldots, p)$$

subject to $\|u\| = \|v\| = 1$, and for $k > 1$: $u^T u_i = 0$ and $v^T v_i = 0$, where $i = 1, \ldots, k-1$.
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 $\theta_1, \ldots, \theta_p$, 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].

19.2.2 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 matrix[3]

$$Q = E\left\{\begin{pmatrix} X \\ Y \end{pmatrix} \begin{pmatrix} X^T & Y^T \end{pmatrix}\right\} = \begin{pmatrix} Q_x & Q_{xy} \\ Q_{yx} & Q_y \end{pmatrix} .$$

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].

19.2.3 System theoretic interpretation of Conjecture 2

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:

$$\left\{ \begin{array}{rcl} x(k+1) & = & Ax(k) + Ku(k) \ , \\ y(k) & = & Cx(k) + u(k) \ , \end{array} \right. \tag{19.2}$$

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 (19.1) by K, and w by $-C^T$, the matrices P, Q and R in (19.1) can be given the following interpretation. The matrix P is the controllability Gramian of the model (19.2) and Q is the observability Gramian of the inverse model, while R is the cross product of the infinite controllability matrix of (19.2) and the infinite observability matrix of the inverse model. Otherwise formulated:

$$\begin{pmatrix} P & R \\ R^T & Q \end{pmatrix} = \begin{pmatrix} \mathcal{C}_\infty \\ \Gamma_\infty^T \end{pmatrix} \begin{pmatrix} \mathcal{C}_\infty^T & \Gamma_\infty \end{pmatrix} ,$$

where $\mathcal{C}_\infty = \begin{pmatrix} K & AK & A^2K & \cdots \end{pmatrix}$ 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 (19.2), these infinite products result in finite matrices and in addition, the condition $\lambda_i\lambda_j \neq 1$ in Conjecture 2 is fulfilled. Furthermore, under fairly general conditions, P, Q and $I_n + PQ$ are non-singular, which follows from the positive definiteness of P and Q under general conditions.
The matrix $P^{-1}RQ^{-1}R^T$ in Conjecture 2 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 $\mathcal{C}_\infty$ and the column space of Γ_∞ (see Section 19.2.1). The angles will be denoted by $\theta_1, \ldots, \theta_n$ (in non-decreasing 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 $\rho_1, \rho_2, \ldots$ (in non-increasing 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.

[3] $E\{\cdot\}$ is the expected value operator.

Conjecture 2 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 $\mathcal{C}_\infty$ and the column space of Γ_∞ and the other canonical correlations are equal to 0:

$$\rho_1 = \sin\theta_n,\ \rho_2 = \sin\theta_{n-1},\ \ldots,\ \rho_n = \sin\theta_1,\ \rho_{n+1} = \rho_{n+2} = \cdots = 0\ . \qquad (19.3)$$

This result can be used to prove that a recently defined cepstral norm [9] for a model as in (19.2) 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 (19.2) 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 $\theta_1,\ldots,\theta_n$ between the row space of $\mathcal{C}_\infty$ and the column space of Γ_∞ as follows:

$$\|\log H\|^2 = -\log\prod_{i=1}^{n}\cos^2\theta_i\ ,$$

and from (19.3) 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,

$$\|\log H\|^2 = \sum_{k=0}^{\infty} kc(k)^2 = 2I(y_p;y_f)\ .$$

19.3 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 2 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 Report[4] 00-44a, ESAT-SCD, K.U.Leuven (Leuven, Belgium), accepted for publication in *Systems & Control Letters* (2002).

[4]This report is available by anonymous ftp from *ftp.esat.kuleuven.ac.be* in the directory *pub/sista/reports* as file *00-44a.ps.gz*.

[3] K. De Cock, "Principal Angles in System Theory, Information Theory and Signal Processing", PhD 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. Schafer, *Digital Signal Processing*, Prentice/Hall International (1975).

[5]This 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 84

On the Computational Complexity of Non-Linear Projections and its Relevance to Signal Processing

Jonathan H. Manton
Department of Electrical and Electronic Engineering
The University of Melbourne, Victoria 3010, Australia
jon@ee.mu.oz.au

Robert Mahony
Department of Engineering
Australian National University, ACT 0200, Australia
mahony@ieee.org

84.1 Description of the Problem

Recent advances in finding the roots of a non-linear equation to within a specified level of accuracy has lead to a novel approach to studying computational complexity of real valued problems [1]. This chapter considers a similar problem, that of globally solving non-linear projection problems to within a specified level of accuracy. A host of important problems derived from signal processing and telecommunications engineering may be written in this form. Optimal solutions are usually intractable and it is important to understand the computational complexity of solving such problems approximately in order that effective sub-optimal algorithms may be developed.

Consider a non-linear (non-convex) minimisation problem of the form

$$\arg\min_{x\in\mathbb{R}^n} \|y - f(x)\|^2 \tag{84.1}$$

for a given $y \in \mathbb{R}^m$ and function $f : \mathbb{R}^n \to \mathbb{R}^m$. Although (84.1) has several interpretations, including a least-squares interpretation and a non-linear regression interpretation, the interpretation favoured here is that if x minimises (84.1) then $f(x)$ is the point in the image of f closest to y. Indeed, (84.1) is solved conceptually by first projecting y onto the set

$$\mathcal{Y} = \{y \in \mathbb{R}^m : y = f(x),\ x \in \mathbb{R}^n\} \tag{84.2}$$

and then computing the pre-image. Sometimes only the projection $f(x)$ and not the pre-image x is of interest, while other times, only certain elements of x are of interest. This motivates studying the more general problem:

$$\text{Calculate } g(\widehat{x}) \text{ where } \widehat{x} = \arg\min_{x\in\mathbb{R}^n} \|y - f(x)\|^2 \tag{84.3}$$

for some function $g : \mathbb{R}^n \to \mathbb{R}^p$. Clearly, if g is taken to be $g(\widehat{x}) = f(\widehat{x})$ then (84.3) is precisely the projection of y onto the image of f.

The practical goal of an engineer in solving a problem of the form (84.3) is to develop a tractable numerical algorithm that, for any input y, outputs an approximate solution that is sufficiently accurate for the application considered. Inherent in the formulation is the trade-off between computational complexity and accuracy; the more accurate the solution is required to be, the harder it is to obtain. To evaluate a potential algorithm one wishes to compare its complexity with the *intrinsic complexity* of the problem (84.3). The following two definitions provide a framework in which this comparison may be made. The definitions here are motivated by those proposed recently by Smale *et al.* [1].

Definition 1 (ϵ-accurate solution) *For a given $\epsilon > 0$, an ϵ-accurate solution of (84.3) is a vector $s \in \mathbb{R}^p$ satisfying the condition*

$$\|s - g(\widehat{x})\| < \epsilon, \quad \widehat{x} = \arg\min_{x\in\mathbb{R}^n} \|y - f(x)\|^2. \tag{84.4}$$

(In the event that $\|y - f(x)\|^2$ does not have a unique minimum, an ϵ-accurate solution is one satisfying (84.4) for some $\widehat{x}$ in the set of all x minimising $\|y - f(x)\|^2$.)

Definition 2 (Computational complexity) *Only algorithms of the following form are considered for solving (84.3). Let R be an integer and let $r_i : \mathbb{R}^m \to \mathbb{R}$ for $i = 1, \cdots, 2^R - 1$ and $h_j : \mathbb{R}^m \to \mathbb{R}^p$ for $j = 1, \cdots, 2^R$ be sequences of rational[1] functions. The r_i are used to divide the space y lies in into 2^R regions as follows. Initially, set $i := 1$. While $i < 2^R$, evaluate $r_i(y)$ and set $i := 2i$ if $r_i(y) < 0$ and set $i := 2i + 1$ otherwise. Finally, set $j := i - 2^R + 1$. (This results in a j lying between 1 and 2^R.) Based on this region j, the algorithm outputs $s := h_j(y)$. Each of these operations involves a certain number of multiplications and additions; for simplicity, the complexity of evaluating a rational function is taken to be the degree of its numerator plus the degree of its denominator. (The h_j are treated as comprising p rational functions from $\mathbb{R}^m$ to $\mathbb{R}$.) For a given y, the complexity of the above algorithm is deduced by summing the complexity of the functions encountered. The overall complexity of the algorithm is defined to be the largest complexity over all possible y.*
Given a function f and an accuracy ϵ, the computational complexity of implementing the non-linear projection operator (84.3) is said to be upper bounded by $N(\epsilon; f)$ if there exists an algorithm of the above form with complexity at most N and which is guaranteed to find an ϵ-accurate solution of (84.3) for all y.

These definitions provide a formulation in which one may evaluate the relative difficulty of a number of interesting problems.

Problem 1 *Let $g : \mathbb{R}^n \to \mathbb{R}^p$ be a known function[2]. Consider three classes of function f*

1. *Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a bilinear function.*
2. *Let D be a positive integer. Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a generic polynomial of degree D.*
3. *Let $\rho > 0$ be a positive constant. Let f be a smooth function with $||D^2 f|| < 1/\rho$. That is that the curvature of the injective image $f : \mathbb{R}^n \to \mathbb{R}^m$ is locally bounded above by ρ.*

For each of the given classes of functions f compute a bound on the complexity $N(\epsilon; f)$ (cf. Def. 2) for any function f in the class considered.

Of the three classes of functions proposed the first is strongly motivated by the joint filtering/system identification problem for a linear system. The second class of functions considered (generic polynomials) are chosen to link directly to the structure of Definition 2 and earlier work on the algebraic complexity of finding the roots of polynomials [2, 1, 3]. Finally, applying a condition on the curvature of the image of $f(x)$ should reduce the complexity of the local projection problem. The global complexity may still lead to an intractable problem in the general formulation, however, in practice this

[1] A rational function is a ratio of two polynomial functions.

[2] The three typical choices for the function g are i) $g(x) = x$, ii) $g(x)$ is a linear projection of x, and iii) $g(x) = f(x)$.

difficulty may not be encountered when real engineering problems are considered (cf. Prob. 2). The classes of functions considered in Problem 1 are clearly a drop in the ocean of possibilities. An equally important question would be to identify the key classes of functions f that relate closely to practical engineering problems and lead to a tractable computational complexity analysis.

Requiring the algorithm in Definition 2 to find an ϵ-accurate solution for all y is a strong global condition. In practical applications, extra knowledge about y is often available. An example is when $y = f(x)+n$ where n is a noise vector. If the noise variance is sufficiently small then y will generally lie close to the image of f. Exploiting this fact may lead to significantly better performance of a filter [4]. Indeed, most signal processing algorithms perform well when the noise variance is sufficiently small. This motivates the following problem:

Problem 2 *Let $g : \mathbb{R}^n \to \mathbb{R}^p$ be a known function (cf. Prob. 1)). For a given $\delta > 0$, define the set*

$$\mathcal{Y}_\delta = \{y \in \mathbb{R}^m : \exists x \in \mathbb{R}^n,\ \|y - f(x)\| \leq \delta\}. \tag{84.5}$$

Define $N(\epsilon, \delta; f)$ as in Definition 2 but with the weaker requirement that the algorithm find an ϵ-accurate solution for all $y \in \mathcal{Y}_\delta$.

Consider the three classes of functions f described in Problem 1 parts 1)-3). Compute a bound on the complexity $N(\epsilon, \delta; f)$ for any function f in the class considered.

Remarks: i) A feature of Definition 2 is that it ignores round-off error. For example, according to Definition 2, $y = Ax$ can be inverted exactly in a finite number of operations. Although Definition 2 can be modified to take round-off error into account, it is preferable to ignore round-off error and focus instead on the inherent difficulty in computing *algebraically* a non-linear projection.

ii) If f is a polynomial function then the closure of $\mathcal{Y}$ is a variety (i.e., an irreducible algebraic set) [3] and can thus be made into a manifold by removing a set of measure zero. Therefore, the essential problem considered here is the complexity of computing the projection of a point onto a manifold.

84.2 Motivation and History

The motivation for studying the complexity of non-linear projection operators is twofold; from a mathematical perspective, it is a natural generalisation of Smale's recent work [1]. From a signal processing perspective, because the main computational difficulty in many applications is that of computing a non-linear projection, understanding the intrinsic complexity of a non-linear projection operator is a significant step towards understanding the intrinsic complexity of signal processing problems.

84.2.1 Intrinsic Complexity of Signal Processing Algorithms

Some signal processing problems have a reputation for being harder to solve than others. A fundamental question then is whether or not this notion of difficulty can be made mathematically precise. Motivating examples of signal processing problems reducing to a problem of the form (84.3) are given below. This shows their inherent complexity can be defined as in Section 1.

The harmonic retrieval problem, in its simplest form, is to determine the frequency f of a noise corrupted sinusoid $y_i = a\cos(2\pi f i + \theta) + n_i$ where n_i is noise. Let $x = (a, f, \theta)$ be a three-dimensional vector and define $f : \mathbb{R}^3 \to \mathbb{R}^m$ component-wise by the rule $f_i(a, f, \theta) = a\cos(2\pi f i + \theta)$. Define $g : \mathbb{R}^3 \to \mathbb{R}$ by $g(a, f, \theta) = f$. Then the solution of (84.3) is the least-squares estimate of the frequency f.

An important class of problems of significant interest is the dual identification and and filtering problem for any system. Let $x = (\theta, s)$ where θ parameterises a class of systems functions f_θ and s denotes the input sequence. The minimization problem

$$\hat{s} = \arg\min_s \left\{ \inf_\theta \|y - f_\theta(s)\|^2 \right\},$$

solves for an estimate of the input signal based on minimizing a least squares cost in the observation space. If the system class f_θ is linearly parameterised and consists of linear systems then $f(\theta, s) := f_\theta(s)$ is a bilinear function. This formulation leads directly to Problem 1-1).

84.2.2 Real Complexity Theory

Rigorously defining the complexity of an engineering problem is challenging. Traditional computational complexity theory, which studies the number of operations a Turing machine requires to solve a particular problem, is not suitable. Indeed, a Turing machine cannot process real valued numbers as it has no way of storing or manipulating irrational numbers in general.

In a series of recent papers (see [1] for references), Steve Smale and his coworkers proposed an alternative definition of computational complexity which can handle problems involving real valued numbers. Central to Smale's work is a measure of how computationally complicated it is to find all the roots of the equation $f(x) = 0$ to within an accuracy ϵ. Also of interest is the fact that, for a given f and accuracy ϵ, Smale proposed a homotopy-like iterative algorithm which is guaranteed to find all the roots of f to within ϵ, and moreover, requires at most N iterations to do so, where N depends on f and on ϵ; see [1] for details.

Although a root finding algorithm can be used to solve (84.1), determining the complexity of (84.1) is a more general problem than determining the complexity of root finding. For a given y, define $g(x; y)$ to be the derivative of $\|y - f(x)\|^2$. Then (84.1) can be solved by finding all the roots of $g(x; y) = 0$ and determining which roots globally minimise $\|y - f(x)\|^2$. However, it has not been established that root finding is the best way of solving (84.1) or (84.3); that is, the intrinsic computational complexity of (84.1) conceivably may be considerably lower than the computational complexity of the associated root finding problem.

Therefore, determining the complexity of non-linear projections is an extension of the complexity results for root finding in [1].

84.3 Available Results

Real computational complexity theory, as opposed to traditional computational complexity theory, is a new subject [1]. Nevertheless, some advanced results appear in [1] which can be applied to the non-linear projection problem. For instance, the homotopy-like method in [1] for finding all the roots of $g(x) = 0$ can be used to solve (84.1) for a given y by setting $g(x)$ to the derivative of $\|y - f(x)\|^2$. It is likely though that better techniques exist, and it is hoped this article will stimulate research in this area.

Acknowledgments: The authors would like to thank Gunnar Rätsch for a number of interesting and useful discussions.

Bibliography

[1] L. Blum, F. Cucker, M. Shub, and S. Smale. *Complexity and Real Computation*. Springer, 1997.
[2] D. A. Cox, J. B. Little, and D. O'Shea. *Using Algebraic Geometry*. Springer-Verlag, 1998.
[3] R. Hartshorne. *Algebraic Geometry*. Springer-Verlag, 1977.
[4] J. H. Manton and B.-N. Vo. Fisher information decision directed quantisation: A fast sub-optimal solution to discrete optimisation problems with applications in wireless communications. In *International Conference on Optimisation Theory and Applications*, volume 1, pages 316–323, Hong Kong, December 2001.

Problem 2

H_∞-norm approximation

A.C. Antoulas
Department of Electrical and Computer Engineering
Rice University
6100 South Main Street
Houston, Texas 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@ic.ac.uk

2.1 Description of the problem

Let $\mathcal{RH}_\infty^m$ be the (Hardy) space of real-rational scalar[1] 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] $A^\star(s) \in \mathcal{RH}_\infty^n$ such that

$$A^\star(s) = \arg\min_{A(s)\in\mathcal{RH}_\infty^n} \|G(s) - A(s)\|_\infty. \tag{2.1}$$

For such a problem, let

$$\gamma_n^\star = \min_{A(s)\in\mathcal{RH}_\infty^n} \|G(s) - A(s)\|_\infty,$$

then two further problems can be posed.

(D) (Optimal Distance problem in the H_∞ norm)
Given $G(s) \in \mathcal{RH}_\infty^N$ and an integer $n < N$ find $\gamma_n^\star$.

(A) (Sub-optimal Approximation in the H_∞ norm)
Given $G(s) \in \mathcal{RH}_\infty^N$, an integer $n < N$ and $\gamma > \gamma_n^\star$ find $\tilde{A}(s) \in \mathcal{RH}_\infty^n$ such that

$$\gamma_n^\star \leq \|G(s) - \tilde{A}(s)\|_\infty \leq \gamma.$$

[1] Similar considerations can be done for the non-scalar case.
[2] By *find* we mean *find an exact solution* or an algorithm converging to the exact solution.

The optimal H_∞ approximation problem can be formally posed as a constrained min-max problem. For, note that any function in $\mathcal{RH}_\infty^n$ 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 $\gamma_n^\star$ can be posed as

$$\gamma_n^\star = \min_{\theta \in \Omega} \max_{\omega \in R} \|G(j\omega) - A(j\omega)\|, \tag{2.2}$$

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 (2.2) 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 $A^\star(s) \in \mathcal{RH}_\infty^n$ be such that equation (2.1) holds. Suppose

$$|W(j\omega^\star) - A^\star(j\omega^\star)| = \gamma_n^\star, \tag{2.3}$$

and

$$A(j\omega^\star) \neq 0 \tag{2.4}$$

for $\omega^\star = 0$. Then there exists a constant $\tilde{\omega} \neq \omega^\star$ such that

$$|W(j\tilde{\omega}) - A^\star(j\tilde{\omega})| = \gamma_n^\star,$$

i.e. if the value $\gamma_n^\star$ is attained by the function $|W(j\omega) - A^\star(j\omega)|$ at $\omega = 0$ it is also attained at some $\omega \neq 0$.

Proof. We prove the statement by contradiction. Suppose

$$|W(j\omega) - A^\star(j\omega)| < \gamma_n^\star, \tag{2.5}$$

for all $\omega \neq \omega^\star$ and consider the approximant $\tilde{A}(s) = (1 + \lambda)A^\star(s)$, with $\lambda \in I\!R$. By equation (2.5), condition (2.4) and by continuity with respect to λ and ω of

$$|W(j\omega) - \tilde{A}(j\omega)|,$$

there is a $\lambda^\star$ (sufficiently small) such that

$$\max_{\omega} |W(j\omega) - (1 + \lambda^\star)A^\star(j\omega)| < \gamma_n^\star,$$

or, what is the same, it is possible to obtain an approximant which is better then $A^\star(s)$, hence a contradiction. $\triangleleft$

It would be interesting to show that the above fact holds (or it does not hold) when $\omega^\star \neq 0$.

2.2 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 the book [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 $\gamma_n^\star$ and a solution of the sub-optimal 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

$$\lambda_{\min}(PQ) = (\gamma_n^\star)^2 \ \text{ with multiplicity } \ N - n. \tag{2.6}$$

However, the above procedure, as also noted by the authors of [5], is not computationally viable, and presupposes the knowledge of $\gamma_n^\star$. 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 (27.1).

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 which 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 which 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 $\epsilon \to 0$, 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, to appear (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. Kavranoglu 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 14

Exact 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

14.1 Description of the problem

We consider a generalized linear continuous-time system Σ characterized by

$$\Sigma \ : \ \begin{cases} \dot{x} = A\ x + \ B\ \ u + \ E\ \ w \\ y = C_1\ x + D_{11}\ u + \ D_1\ w \\ h = C_2\ x + \ D_2\ \ u + D_{22}\ w \end{cases} \tag{14.1}$$

where $x \in \mathbb{R}^n$ 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 $\Sigma_{\rm P}$ be the subsystem characterized by the matrix quadruple (A, B, C_2, D_2) and $\Sigma_{\rm Q}$ 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,

$$\Sigma_{\rm cmp} \ : \ \begin{cases} \dot{v} = A_{\rm cmp}\ v + B_{\rm cmp}\ y \\ u = C_{\rm cmp}\ v + D_{\rm cmp}\ y \end{cases} \tag{14.2}$$

such that when it is applied to the given plant (**??**), 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

$$\|T_{hw}\|_\infty := \sup_{\omega\in[0,\infty)} \sigma_{\max}[T_{hw}(j\omega)] = \sup_{\|w\|_2=1} \frac{\|h\|_2}{\|w\|_2}, \tag{14.3}$$

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

$$\gamma^* := \inf \Big\{ \ \|T_{hw}(\Sigma\times\Sigma_{\rm cmp})\|_\infty \ | \ \ \Sigma_{\rm cmp} \text{ internally stabilizes } \Sigma \ \Big\}. \tag{14.4}$$

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 (**??**), the designer should know before hand the infimum γ^*, which represents the best achievable level of disturbance attenuation. Unfortunately, the problem of the exact computation of this γ^* for general systems still remains unsolved in the open literature.

14.2 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 focussed 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.*, [2, 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.

14.3 Available results

There are quite a few researchers who have attempted to develop procedures for finding the exact 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 [1]) have developed a non-iterative procedures for computing the exact 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 a linear system Σ_* characterized by a matrix quadruple (A_*, B_*, C_*, D_*), we define

1. $\mathcal{V}^-(\Sigma_*)$, a weakly unobservable subspace, is the maximal subspace of $\mathbb{R}^n$ which is $(A_*+B_*F_*)$-invariant and contained in $\mathrm{Ker}\,(C_* + D_*F_*)$ 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

2. $\mathcal{S}^-(\Sigma_*)$, a strongly controllable subspace, is the minimal $(A_* + K_*C_*)$-invariant subspace of $\mathbb{R}^n$ containing $\mathrm{Im}\,(B_*+K_*D_*)$ 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 exact computation of γ^* has been solved by Chen and his co-workers [3, 4] (see also [1]) for a class of systems that satisfy the following conditions:

1. $\mathrm{Im}(E) \subset \mathcal{V}^-(\Sigma_\mathrm{P}) + \mathcal{S}^-(\Sigma_\mathrm{P})$; and

2. $\mathrm{Ker}(C_2) \supset \mathcal{V}^-(\Sigma_\mathrm{Q}) \cap \mathcal{S}^-(\Sigma_\mathrm{Q})$,

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 [1] that the exact 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 Edition, Birkhäuser, Boston, Massachusetts, 1995.

[2] B. M. Chen, *H_∞ Control and Its Applications*, Springer, London, 1998.

[3] B. M. Chen, Y. Guo and Z. L. Lin, "Non-iterative 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 $\mathcal{L}_\infty$ 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, Massachusetts, 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 40

The Computational Complexity of Markov Decision Processes

Omid Madani
Department of Computing Science University Of Alberta
Edmonton, AB T6G 2E8
Canada
madani@cs.ualberta.ca

40.1 Description of the problem

A Markov decision process (MDP) model consists of a set Q of n states, and a finite set of actions Σ. Time is discretized and at each time point t, $t = 0, 1, \cdots$, the system occupies a single state $s^{(t)}$ in Q, which is called the (current) state of the system at time t. The means of change of time and system state is action execution, and there is uncertainty about the outcome of actions. Consider the states indexed: $s_i \in Q, 1 \leq i \leq n$. Associated with each action $a \in \Sigma$ is an $n \times n$ stochastic matrix $\mathbf{M_a}$ which specifies the state transition probabilities for action a. $\mathbf{M_a}[i, j]$ has the semantics that if the state of the system is s_i at a given time point t and action a is executed, with probability $\mathbf{M_a}[i, j]$ the state of the system at time point $t + 1$ is s_j. We refer to $\mathbf{M_a}$ as the transition matrix of action a. Also associated with each action a is an $n \times 1$ vector of rewards $\mathbf{R_a}$, with the semantics that the decision maker gains $\mathbf{R_a}[i]$, if the state of the system is s_i and action a is executed.

The states and actions (basically the transition matrices and the reward vectors) are completely specified as part of the problem instance. Additionally, a discount factor $0 < \beta < 1$ is given, and the problem is to compute an *optimal policy*, as described next. A *policy* $\mathcal{P}$ is a mapping assigning to each state s a single action $\mathcal{P}(s) \in \Sigma$. The value (vector) of a policy, denoted $\mathcal{V}_{\mathcal{P}}$ is a vector of n values, where $\mathcal{V}_{\mathcal{P}}[i]$ is the *value* of state s_i under policy $\mathcal{P}$, defined as the expectation of total reward $\sum_{t=0}^{\infty} \beta^t r(\mathcal{P}(s^{(t)}))$, when $s^{(0)} = s_i$, and $r(\mathcal{P}(s^{(t)}))$ denotes the reward obtained from executing action $\mathcal{P}(s^{(t)})$ at state $s^{(t)}$. The value of a policy can be computed in polynomial time by essentially a matrix inversion [1, 2]. An MDP model enjoys the property that a policy $\mathcal{P}^*$ exists under which all state values are maximized: $\exists$ policy $\mathcal{P}^*, \forall s_i \in Q, \mathcal{V}_{\mathcal{P}^*}[i] = \max_{\mathcal{P}} \mathcal{V}_{\mathcal{P}}[i]$, where the maximum is taken over all possible policies. Note that there are only a finite number of possible policies. Such a policy is called an optimal policy. Thus the MDP problem is to compute an optimal policy, given the states and actions and a discount factor.

MDP problems can be formulated as linear programs and thus solved in polynomial time by polynomial algorithms for linear programs [1, 2]. An important open problem is to give polynomial algorithms other than the more general linear programming algorithms [3]. In particular it is open whether the widely used *policy iteration* algorithm (described below) is polynomial.

40.2 Motivation and history of the problem

The reader is referred to books on MDPs, such as [1, 2], for a comprehensive introduction to MDPs. The problem we described is more accurately identified as the infinite-horizon fully observable MDP problem under the discounted total reward objective function[1]. MDP problems remain polynomial time equivalent under several common variations in the objective function, such as dropping the discount factor, or changing the objective to maximum average reward per action execution or maximizing the probability of reaching a certain state.

The MDP problem is a classic optimization problem and several well-studied combinatorial optimization problems such as shortest-paths problems are special subproblems. On the other hand, there are game generalizations of MDPs (e.g. simple stochastic games [5]) for which no polynomial time algorithm is known. The complexity of MDPs were first investigated in [3]. Later work in this direction include [4, 5, 6, 8, 7]. See for example the survey [9].

Simple and elegant dynamic programming algorithms called policy iteration are the preferred method for solving MDPs. These algorithms tend to converge quickly in practice, but it is not known whether they have polynomial complexity, i.e. whether they converge in a polynomial number of iterations (polynomial in the number of states, actions, and the binary representation of the numbers). We conjecture that a very common variant, that we shall refer to as "parallel" policy iteration, is a polynomial time algorithm. An additional motivation for studying these algorithms is the fact that they are similar to variants of the simplex algorithm for solving general linear programs, and it is a major open problem whether these simplex methods are polynomial in solving linear programs [12]. Analysis techniques that would establish (parallel) policy iteration polynomial on MDPs may suggest ways of establishing the simplex algorithm polynomial on more special linear programs such as MDPs and perhaps ultimately on general linear programs. We describe policy iteration and in particular parallel policy iteration next.

Policy iteration algorithms begin with an arbitrary policy and iteratively change and improve the policy, by changing action assignments for a subset of the states, until no more improvement is possible. A policy is improved as follows. Under any given policy, each state s_i has a value $\mathcal{V}_{\mathcal{P}}[i]$ which is obtained when the policy is evaluated as described in the problem definition. Let us define the value of action a (in policy $\mathcal{P}$) for state i as:

$$\mathbf{R_a}[i] + \beta \sum_{s_j \in Q} M_a[i,j]\mathcal{V}_{\mathcal{P}}[j].$$

Therefore, the value of an action is its immediate reward at state s_i plus the discounted expected reward if policy $\mathcal{P}$ is used thereafter. It can be shown that for each state s_i, the difference between the value of a highest valued action a^* for state s_i and the state value $\mathcal{V}_{\mathcal{P}}[i]$ under the policy is nonnegative, i.e., $\left(\max_{a\in\Sigma} \mathbf{R_a}[i] + \beta \sum_{s_j\in Q} M_a[i,j]\mathcal{V}_{\mathcal{P}}[j]\right) - \mathcal{V}_{\mathcal{P}}[i]) \geq 0$, and policy $\mathcal{P}$ is suboptimal if and only if for some state the difference is positive. Let us call the states for which the difference is positive the *improvable* states. The improvable states are exactly the states for which the highest valued action is different from the action assigned by the current policy $\mathcal{P}$.

In the common parallel variant of policy iteration, each state is assigned its highest valued action after policy evaluation. It is the case that unless policy $\mathcal{P}$ is already optimal, actions assigned to improvable states, change from the assignment prescribed by policy $\mathcal{P}$, leading to a new strictly *improved* policy $\mathcal{P}'$, where policy improvement is in the following sense: the value of no state under $\mathcal{P}'$ is lower than its value under $\mathcal{P}$ ($\forall s_i \in Q, \mathcal{V}_{\mathcal{P}'}[i] \geq \mathcal{V}_{\mathcal{P}}[i]$), and for some state, its strictly higher ($\exists s_i \in Q, \mathcal{V}_{\mathcal{P}'}[i] > \mathcal{V}_{\mathcal{P}}[i]$). Under other policy iteration methods actions assigned to a subset of improvable states, for example only for a single improvable state chosen by some heuristic, are changed, but the same improvement properties hold.

In any policy iteration algorithm the cycle of policy evaluation and improvement is repeated until no more improvable states exist. The evaluation and improvement steps are polynomial time computations, and as there are a finite number of policies, these algorithms converge to an optimal policy in finite time. Therefore, the main question is whether the number of iterations to convergence is polynomial.

[1] There are still other distinctions. For example, our problem involves a finite number of states and actions with stationary transitions and rewards.

40.3 Available results

It is known that parallel policy iteration has pseudo-polynomial run time, i.e., it has polynomial run time if the numbers in the problem instance are represented in unary representation instead of the more compact and standard binary representation [4, 6]. On the other hand, several variants of policy iteration, in which only the action assigned to a single state is changed to improve the policy, have been shown to take exponential time irrespective of the number representation (i.e., they do not run in pseudo-polynomial time) [11]. We have shown that parallel policy iteration has polynomial complexity on MDPs with special state transition structure: For any pair of states s_i and s_j, if an action a takes state s_i to s_j with positive probability (i.e., $M_a[i,j] > 0$), then there is no single action execution that takes s_j to s_i with positive probability, with the exception that there is a special 'bottleneck' state b that has no such restrictions, i.e., any state can have positive transition probability to state b under some actions, and so can state b have a transition to any state under some actions. This restriction effectively forces all the possible cycles in the graph structure of the MDP to go through state b [7, 10]. Viewing policy iteration as a Newton's method for finding the zero of a function and utilizing the geometric constraints that apply are the key insights in establishing the algorithm polynomial in this case. We expect that analyses along similar lines will establish parallel policy iteration and similar variants polynomial on the more general problems.

Bibliography

[1] M. L. Puterman. *Markov Decision Processes*. Wiley Inter-science. 1994.

[2] D. P. Bertsekas. *Dynamic Programming and Optimal Control*. Athena Scientific. 1995.

[3] C. H. Papadimitriou and J. N. Tsitsiklis, "The complexity of Markov decision processes", *Mathematics of operations research*, 12,3, pp. 441-450 (1987).

[4] P. Tseng,"Solving H-horizon stationary Markov decision process in time proportional to $\log(H)$", *Operations Research Letters*, 9,5, pp. 287-297 (1990).

[5] A. Condon, "The Complexity of Simple Stochastic Games", *Information and Computation*, 96,2, pp. 203-224 (1992).

[6] M. Littman. *Algorithms for Sequential Decision Making*. Brown. 1996.

[7] O. Madani. *Complexity Results for Infinite-Horizon Markov Decision Processes*. University of Washington. 2000.

[8] Y. Mansour and S. Singh, "On the Complexity of Policy Iteration", in proceedings of *Uncertainty in AI*, 96,2, pp. 401-408 (1999).

[9] V. D. Blondel and J. N. Tsitsiklis, "A Survey of Computational Complexity Results in Systems and Control", *Automatica*, 36,9, pp. 1249-1274 (2000).

[10] O. Madani, "On Policy Iteration as a Newton's Method and Polynomial Policy Iteration Algorithms", To appear in the proceedings of *AAAI02* (2002).

[11] M. Melekopoglou and A. Condon, "On the Complexity of the Policy Improvement Algorithm for Markov Decision Processes", *ORSA Journal on Computing*, 6, 2 (1994).

[12] V. Klee and P. Kleinschmidt, "The d-step conjecture and its relatives", *Mathematics of Operations Research*, 12,4, pp. 718-755 (1987).

Problem 7

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

7.1 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

$$\lim_{t \to \infty} x_t = 0$$

for all initial states x_0 and all sequences of matrix products. This condition is equivalent to the requirement

$$\lim_{t \to \infty} A_{i_t} \cdots A_{i_1} A_{i_0} = 0$$

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 $\Sigma = \{A_1, \ldots, A_m\}$ be a given set and define $B_0 = \text{diag}(A_1, \ldots, A_m)$ 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 [2] for a more detailled 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 [7]. The conditions presently available are only partly satisfactory in that they are either incomplete (they do not cover all cases), or they are complete but do allow to *effectively decide* if a given pair of matrices is stable. We say that a problem is (effectively) *decidable* if there is an algorithm which, 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 [8] for a general description of decidability, and [3] 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. [1] and [2]); 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 with rational entries is stable is undecidable.

7.2 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 [6] by

$$\rho(\Sigma) = \limsup_{k \to \infty} \rho_k(\Sigma),$$

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 [9] 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 $A_i \in \Sigma$ $(i = 1, \ldots, k)$ 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 doesn't 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 [5]. A simple counterexample is provided in [4] 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] Blondel V. D. and J. N. Tsitsiklis, "When is a pair of matrices mortal?", *Information Processing Letters*, **63**:5, 283-286, 1997.

[2] Blondel V. D. 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.

[3] Blondel V. D. and J. N. Tsitsiklis, "A survey of computational complexity results in systems and control, *Automatica*, 36(9), 1249–1274, 2000.

[4] Blondel V. D., J. Theys and A. A. Vladimirov, "An elementary counterexample to the finiteness conjecture", submitted.

[5] Bousch T. and Jean Mairesse, "Asymptotic height optimization for topical IFS, Tetris heaps, and the finiteness conjecture", J. Amer. Math. Soc., **15**, 77-111, 2002.

[6] I. Daubechies and J. C. Lagarias (1992), Sets of matrices all infinite products of which converge, Linear Algebra Appl., **162** 227-263.

[7] L. Gurvits. Stability of discrete linear inclusion. *Linear Algebra Appl.*, 231:47–85, 1995.

[8] Hopcroft J. E. and J. D. Ullman, "Introduction to automata theory, languages, and computation", Addison-Wesley, Reading, MA, 1979.

[9] J. C. Lagarias and Y. Wang (1995), *The finiteness conjecture for the generalized spectral radius of a set of matrices*, Linear Algebra Appl., **214** 17-42.

[10] Papadimitriou, C. H. (1994). *Computational complexity*, Addison-Wesley, Reading.

[11] Tsitsiklis, J. N (1987). The stability of the products of a finite set of matrices, in *Open problems in communication and computation*, Springer-Verlag.

[12] Tsitsiklis, J. N. and V. D. Blondel (1997). *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.

Problem 12

On the Stability of Random Matrices

Giuseppe Calafiore*, Fabrizio Dabbene**
*Dip. di Automatica e Informatica
** IRITI-CNR
Politecnico di Torino
C.so Duca degli Abruzzi, 24
Torino
Italy
{calafiore, dabbene}@polito.it

12.1 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 $\mathrm{Re}\lambda_i < 0$, $i = 1, \ldots, n$, 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 $\mathbf{\Delta}$ and a stable matrix $A \in \mathbb{R}^{n,n}$, state whether $A_\Delta = A + \Delta$ is stable for all $\Delta \in \mathbf{\Delta}$. Since the above deterministic problem may be computationally hard in some cases, a recent line of study proposes to introduce a probability distribution over $\mathbf{\Delta}$, 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 12.3 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.

12.2 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 $\|X\|$ 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. A straightforward conclusion is that for an orthogonal invariant random matrix $\mathbf{X}$, its pdf is a function of only the eigenvalues $\Lambda \doteq \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$ of $\mathbf{X}$, i.e.

$$f_{\mathbf{X}}(X) = g_{\mathbf{X}}(\Lambda). \tag{12.1}$$

This class of 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: Whishart 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 (12.1). 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 $\{X \in \mathbb{R}^{n,n} : \rho(X) \le 1\}$.
6. US_n^F: Symmetric $n \times n$ random matrices from S_n, which are uniform over the set $\{X \in \mathbb{R}^{n,n} : \|X\| \le 1\}$.

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 $\lambda_1 \ge \lambda_2 \ge \ldots \ge \lambda_n$ are the eigenvalues of a random matrix $\mathbf{X}$ belonging to S_n, then their pdf $f_{\mathbf{\Lambda}}(\Lambda)$ is

$$f_{\mathbf{\Lambda}}(\Lambda) = \frac{\pi^{n^2/2}}{\Gamma_n(n/2)} g_{\mathbf{X}}(\Lambda) \prod_{1 \le i < j \le n} (\lambda_i - \lambda_j). \tag{12.2}$$

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

$$W_n : \quad \frac{\pi^{n^2}}{\Gamma_n^2(n/2)} \exp(-\tfrac{1}{2} \textstyle\sum_i \lambda_i) \prod_i \lambda_i^{-1/2} \prod_{1 \le i < j \le n} (\lambda_i - \lambda_j) \tag{12.3}$$

$$\mathrm{GOE}_n : \quad \frac{1}{2^{n/2} \prod_i \Gamma(i/2)} \exp(-\tfrac{1}{2} \textstyle\sum_i \lambda_i^2) \prod_{1 \le i < j \le n} (\lambda_i - \lambda_j) \tag{12.4}$$

$$\mathrm{US}_n^\rho : \quad K_u \prod_{1 \le i < j \le n} (\lambda_i - \lambda_j), \; 1 \ge \lambda_1 \ge \ldots \ge \lambda_n \ge -1. \tag{12.5}$$

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]

$$K_u = n! 2^{\frac{n}{2}(n+1)} \prod_{j=0}^{n-1} \frac{\Gamma(3/2 + j/2)\Gamma^2(1 + j/2)}{\Gamma(3/2)\Gamma((n + j + 3)/2)}. \tag{12.6}$$

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 $n \to \infty$) 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.

12.2.1 Probability of stability for the US_n^ρ ensemble

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

$$P \doteq \int \cdots \int_{\mathrm{Re}\Lambda < 0} f_{\mathbf{\Lambda}}(\Lambda) \mathrm{d}\Lambda. \tag{12.7}$$

We now compute this probability for matrices in the US_n^ρ ensemble, whose pdf is given in (12.5). To this end, we first remove the ordering of the eigenvalues, and therefore divide by $n!$ the pdf (12.5). Then, the probability of stability is

$$P_{US} = \frac{K_u}{n!} \int_{-1}^{0} \cdots \int_{-1}^{0} \prod_{1 \le i < j \le n} |\lambda_i - \lambda_j| \, \mathrm{d}\lambda_1 \cdots \mathrm{d}\lambda_n. \tag{12.8}$$

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

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

12.3 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):

P 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:

P 2. *Determine the probability of stability for the G_n ensemble.*

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:

P 31. *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 12.2.*

12.4 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, editor, *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, editor, *Statistical Theories of Spectra: Fluctuations.* Academic, New York, 1965.

Problem 37

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

37.1 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*

$$\dot{x} = f_\sigma(x), \qquad x \in R^n \tag{37.1}$$

where $\sigma : [0, \infty) \to P$ 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 (37.1) 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 a survey). 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.

37.2 Available results: linear systems

In this section we concentrate on the case when the subsystems are linear. This results in the *switched linear system*

$$\dot{x} = A_\sigma x, \qquad x \in R^n. \tag{37.2}$$

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 (37.2), 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 (37.2) 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 which 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 (37.2) is stable; this was proved in [4] for the discrete-time case. The system (37.2) 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 (37.2) 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 representation 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).

37.3 Open problem: nonlinear systems

Let us now turn to the general nonlinear situation described by equation (37.1). 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 (37.1). 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 (37.1). 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

$$[f_p, f_q](x) := \frac{\partial f_q(x)}{\partial x} f_p(x) - \frac{\partial f_p(x)}{\partial x} f_q(x) = 0 \qquad \forall x \in R^n, \quad \forall p, q \in P$$

then the switched system (37.1) 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?*

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 are 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 of this topic).

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. Kawski. 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 and A. S. Morse. Basic problems in stability and design of switched systems. *IEEE Control Systems Magazine*, pages 59–70, October 1999.

[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 21

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

21.1 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 which 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 $l^2[0,\infty) = \{< x_0, x_1, \cdots >: x_i \in \mathbb{C}^n, \sum \|x_i\|^2 < \infty\}$ and denote by P_n the truncation projection onto the subspace generated by the first n vectors $\{e_0, \cdots, e_n\}$ 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 revelant 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 characterises 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) $G(L) = Ran \begin{bmatrix} M \\ N \end{bmatrix} = Ker[\ -\hat{N} \quad \hat{M}\ ]$, (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 $G(C) = Ran \begin{bmatrix} \hat{Y} - NQ \\ \hat{X} + MQ \end{bmatrix} = Ker[\ -(X + Q\hat{M}) \quad Y - Q\hat{N}\ ]$, *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?

21.2 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 existance of strongly stabilizing controllers were formulated from the point of view of H^∞ control problems. The latest such effort is [7].
In 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 Banach 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 existance 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. Saeks, *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 86

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

86.1 Description of the problem

Consider the system

$$\dot{x}(t) = f(x(t), u(t)) \tag{86.1}$$

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 $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ is locally Lipschitz, and that the system is forward complete. Assume that the output maps $h : \mathbb{R}^n \to \mathbb{R}^{p_y}$ and $g : \mathbb{R}^n \to \mathbb{R}^{p_w}$ 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 $\|z\|_{[0,t]} =$ ess sup $\{|z(t)| : t \in [0, t]\}$.

A function $\gamma : \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ is *of class* $\mathcal{K}$ (denoted $\gamma \in \mathcal{K}$) if it is continuous, positive definite, and strictly increasing; and is of class $\mathcal{K}_\infty$ if in addition it is unbounded. A function $\beta : \mathbb{R}_{\geq 0} \times \mathbb{R}_{\geq 0} \to \mathbb{R}_{\geq 0}$ 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 $t \to \infty$.

The following definitions are given for a forward complete system with two output channels as in (86.1). The outputs y and w are considered as error and measurement signals, respectively.

Definition We say that the system (86.1) is *input-measurement to error stable* (IMES) if there exist $\beta \in \mathcal{KL}$ and γ_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 (86.1), and all $t \geq 0$.

Open Problem

Find a (if possible, smooth) Lyapunov characterization of the IMES property.

86.2 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].

86.3 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 (86.1) 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 (86.1), and all $t \geq 0$.

Definition Let $\rho \in \mathcal{K}$. We say that the system (86.1) satisfies the *stability in three measures* (SIT) property (with gain ρ) if there exists $\beta \in \mathcal{KL}$ so that for any solution of (86.1), if there exists $t_1 > 0$ so that $|y(t)| > \rho(|w(t)|)$ for all $t \in [0, t_1]$, then

$$|y(t)| \leq \beta(|x(0)|, t) \qquad \forall t \in [0, t_1].$$

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 $V : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$ is a *lower semicontinuous SIT-Lyapunov function* for system (86.1) with gain ρ if

- there exist $\alpha_1, \alpha_2 \in \mathcal{K}_\infty$ so that

$$\alpha_1(|h(\xi)|) \le V(\xi) \le \alpha_2(|\xi|), \qquad \forall \xi \text{ so that } |h(\xi)| > \rho(|g(\xi)|),$$

- there exists $\alpha_3 : \mathbb{R}_{\ge 0} \to \mathbb{R}_{\ge 0}$ continuous positive definite so that for each ξ so that $|h(\xi)| > \rho(|g(\xi)|)$,

$$\zeta \cdot v \le -\alpha_3(V(\xi)) \qquad \forall \zeta \in \partial_D V(\xi),\ \forall v \in F(\xi). \tag{86.2}$$

(Here ∂_D denotes a viscosity subgradient.)

Theorem Let a system of the form (86.1) and a function $\rho \in \mathcal{K}$ be given. The following are equivalent.

1. The system satisfies the SIT property with gain ρ.
2. The system admits a lower semicontinuous SIT-Lyapunov function with gain ρ.
3. 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. Kokotovi'c, *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 90's," 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.

[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] "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-Lagarrigue, 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** (1995), pp. 351–359.

[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** (1997), pp. 279–290.

[16] E. D. Sontag and Y. Wang, "Notions of input to output stability", *Systems & Control Letters* **38** (1999), pp. 351–359.

[17] E. D. Sontag and Y. Wang, "Lyapunov characterizations of input to output stability", *SIAM Journal on Control and Optimization* **39** (2001), pp. 226–249.

Problem 13

Copositive Lyapunov functions

M.K. Çamlıbel
Department of Mathematics
University of Groningen
P.O. Box 800, 9700 AV Groningen
The Netherlands
e-mail: `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
e-mail: `j.m.schumacher@kub.nl`

13.1 Preliminaries

The following notational conventions and terminology will be in force. Inequalities for vectors are understood componentwise. Given two matrices M and N with the same number of columns, the notation $\mathrm{col}(M, N)$ 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 *nonnegative (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 (nonnegative 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 $\mathrm{col}(C, CA, \cdots, CA^{n-1})$ is of rank n where A is of order n.

13.2 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 3 [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 $A^TP + PA = Q$ 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

$$\dot{x} = A_\sigma x \tag{13.1}$$

where the *switching signal* $\sigma : [0, \infty) \to \{1, 2\}$ 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 (13.1) is the requirement of stability for arbitrary switching signals.

The dynamics of (13.1) 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 (13.1). 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 4 *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 30.7 and 4 for a class of piecewise linear systems. More precisely, we will consider systems of the form

$$\dot{x} = A_i x \quad \text{for } C_i x \geq 0 \quad i = 1, 2. \tag{13.2}$$

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 (13.2) 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 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 (13.2) with $A_1 = A$, $C_1 = C$, $A_2 = A - BD^{-1}C$, and $C_2 = -C$. Equivalently, it can be described by

$$\dot{x} = f(x) \tag{13.3}$$

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 (13.3) admits a unique (continuously differentiable) solution x for all initial states x_0.

One way of studying the stability of the system (13.2) is simply to utilize Theorem 4. 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 (13.2).

ii. It considers *any* switching signal whereas the initial state determines *the* switching signal in (13.2).

In the next section, we focus on ways of eliminating the conservatism mentioned in i.

13.3 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 $M \overset{\mathcal{C}}{\succcurlyeq} 0$ and $M \overset{\mathcal{C}}{\succ} 0$ respectively for copositivity and strict copositivity. When the cone $\mathcal{C}$ is clear from the context we just write $\succcurlyeq$ or $\succ$.

The first problem that we propose calls for an extension of Theorem 30.7 for linear dynamics restricted to a cone.

Problem 3 *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 $A^T P + PA \prec 0$.*

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 4 to the systems (13.2). In other words, it would be natural to attack the following problem.

Problem 4 *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 $P \overset{\mathcal{C}_i}{\succ} 0$ and $A_i^T P + PA_i \overset{\mathcal{C}_i}{\prec} 0$ for $i = 1, 2$.*

13.4 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 5 *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 which are equal to the sum of a nonnegative definite matrix and a nonnegative 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, CA, 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 56

Delay independent and delay dependent Aizerman problem

Vladimir Răsvan
Department of Automatic Control
University of Craiova
13 A.I.Cuza Street
1100 Craiova ROMANIA
vrasvan@automation.ucv.ro

56.1 Abstract

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.

Keywords: Sector nonlinearity, Time delay, Absolute stability.

56.2 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 equilibrium of a feedback control system composed of a linear dynamic part and a nonlinear static element

$$\dot{x} = Ax - b\varphi(c^\star x) \tag{56.1}$$

where x, b, c are n-dimensional vectors, A is a $n \times n$ matrix and $\varphi : \mathbb{R} \to \mathbb{R}$ is a continuous function. The only additional assumption about φ was its location in some sector

$$\underline{\varphi}\sigma^2 < \varphi(\sigma)\sigma < \overline{\varphi}\sigma^2 \tag{56.2}$$

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, much famous paper of Lurie and Postnikov [17] global asymptotic stability conditions were obtained for a system (56.1) of 3d order with $\varphi(\sigma)$ satisfying $\varphi(\sigma)\sigma > 0$ i.e. satisfying (56.2) with $\underline{\varphi} = 0, \overline{\varphi} = +\infty$. 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 (56.2). 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 (56.1) and its linearized version

$$\dot{x} = Ax - bhc^{\star}x \tag{56.3}$$

i.e. system (56.1) with $\varphi(\sigma) = h\sigma$. It is known that the necessary and sufficient conditions of asymptotic stability for (56.3) will require h to be restricted to some interval $\left(\underline{h}, \overline{h}\right)$ called the *Hurwitz sector*. On the other hand for system (56.1) the absolute stability problem is stated: find conditions of global asymptotic stability of the zero equilibrium for all functions satisfying (56.2). All functions include the linear ones hence the class of systems defined by (56.1) is larger than the class of systems defined by (56.3). Consequently the sector $\left(\underline{\varphi}, \overline{\varphi}\right)$ from (56.2) may be at most as large as the Hurwitz sector $\left(\underline{h}, \overline{h}\right)$. *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 which 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 (56.2) by the class of differentiable functions with slope restrictions

$$\underline{\gamma} < \varphi'(\sigma) < \overline{\gamma} \tag{56.4}$$

The *Kalman problem asks*: do coincide the intervals $\left(\underline{\gamma}, \overline{\gamma}\right)$ and $\left(\underline{h}, \overline{h}\right)$ - 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 counter-example known up to that paper had been published by Fitts [10] 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 sub-system 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 (56.4) instead of (56.2);

- the counter-example 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 counter-examples 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.

56.3 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

$$\dot{x} = A_0 x(t) + A_1 x(t-\tau), \tau > 0 \tag{56.5}$$

Exponential stability of this system is ensured by the location in the LHP(left-hand plane) of the roots of the characteristic equation

$$det\left(\lambda I - A_0 - A_1 e^{-\lambda\tau}\right) = 0 \tag{56.6}$$

where the LHS(left-hand side) is a quasi-polynomial. We have here the *Routh-Hurwitz problem for quasi-polynomials*. This problem has been studied since the first applications of (56.5); 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 quasi-polynomials 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 (56.5) holds. This 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 [12] 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 [13]. There are by now dozens of references concerning *delay-dependent and delay-independent Routh-Hurwitz problem* for (56.5); we send the reader to the books of S.I.Niculescu [19, 20] with their rich reference lists.

To illustrate the difference between the two notions we shall consider the scalar version of (56.5):

$$\dot{x} + a_0 x(t) + a_1 x(t-\tau) = 0, \tau > 0 \tag{56.7}$$

for which the exponential stability is ensured provided the following inequalities hold:

$$1 + a_0\tau > 0, \qquad -a_0\tau < a_1\tau < \psi(a_0\tau) \tag{56.8}$$

where $\psi(\xi)$ is obtained by eliminating the parameter λ between the two equalities below

$$\xi = -\frac{\lambda}{tg\lambda}, \qquad \psi = \frac{\lambda}{sin\lambda} \tag{56.9}$$

The delay-independent stability is ensured provided the simple inequalities

$$a_0 > 0, \qquad |a_1| < a_0 \tag{56.10}$$

are fulfilled. It can be shown [10] that $\psi(\xi) > \xi$ for $\xi > 0$ hence the fulfilment of (56.10) implies the fulfilment of (56.8).

Consider now a special case of (56.8) that is in fact the underlying topic of most references cited in [19, 20] - stability for small delays.

As shown in [10] the stability inequalities are given by

$$a_1 + a_0 > 0, \quad 0 \le \tau < \frac{arccos\left(-\frac{a_0}{a_1}\right)}{\sqrt{a_1^2 - a_0^2}} \tag{56.11}$$

provided $a_1 > |a_0|$ (otherwise (56.10) 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 (56.2) is only natural. The most suitable references on this problem are the monographs of A. Halanay [11] 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

$$\dot{x} = A_0x(t) + A_1x(t-\tau) - b\varphi\left(c_0^\star x(t) + c_1^\star x(t-\tau)\right) \tag{56.12}$$

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 (56.2).

Following author's book [23] we shall consider a scalar version of (56.12):

$$\dot{x} + a_0x(t) + \varphi\left(x(t) + c_1x(t-\tau)\right) = 0 \tag{56.13}$$

where $\varphi(\sigma)\sigma > 0$. Assume that $a_0 > 0$ and apply the frequency domain inequality of Popov for $\overline{\varphi} = +\infty$:

$$Re(1 + j\omega\beta)H(j\omega) > 0, \quad \forall\omega \geq 0 \tag{56.14}$$

Since

$$H(s) = \frac{1 + c_1e^{-\tau s}}{s + a_0}$$

the frequency domain inequality reads

$$\frac{\left(a_0^2 + \omega^2\beta\right)\left(1 + c_1cos\omega\tau\right) + \omega\left(a_0\beta - 1\right)sin\omega\tau}{a_0^2 + \omega^2} > 0$$

By choosing the Popov parameter $\beta = a_0^{-1}$ the above inequality becomes

$$1 + c_1cos\omega\tau > 0, \quad \forall\omega \geq 0 \tag{56.15}$$

which cannot hold for $\forall\omega$ but only with $|c_1| < 1$. The frequency domain inequality of Popov prescribes in this case only a *delay independent absolute stability.*

56.4 The problems of Aizerman for time delay systems

Let us follow the way of Barbashin [6] to introduce the Aizerman problem in the time delay case: given system (56.7) for $a_0 > 0$, if we replace a_0x by $\varphi(x)$ where $\varphi(x)x > 0$, the nonlinear time delay system should be globally asymptotically stable provided

$$\frac{\varphi(\sigma)}{\sigma} > |a_1| \tag{56.16}$$

for the delay-independent stability, or provided

$$\frac{\varphi(\sigma)}{\sigma} > max\left\{-a_1, \frac{1}{\tau}\psi^{-1}(a_1\tau)\right\} \tag{56.17}$$

in the delay-dependent case. The meaning of an Aizerman problem is quite clear and it could be stated as follows:

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 is considered instead or have they to be strengthened?

It is clear that we have gathered here both the delay-independent and delay-dependent cases. In fact we could speak about a *delay-independent Aizerman problem* and also about a *delay-dependent Aizerman problem.* Judging after what is known, the delay-independent Aizerman problem should be easier to analyze.

Consider, for instance, the delay independent Aizerman problem defined above, for system (56.7) replaced by

$$\dot{x} + a_1 x(t-\tau) + \varphi(x(t)) = 0 \tag{56.18}$$

where $\varphi(\sigma)\sigma > 0$. Taking into account that (56.10) 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):

$$\dot{x} + |a_1|x(t) + a_1 x(t-\tau) + f(x(t)) = 0 \tag{56.19}$$

For this system we apply the frequency domain inequality of Popov for $\overline{\varphi} = +\infty$ i.e. inequality (56.14); here

$$H(s) = \frac{1}{s + |a_1| + a_1 e^{-s\tau}} \tag{56.20}$$

and the frequency domain inequality reduces to

$$\beta\omega^2 - (\beta a_1 sin\omega\tau)\omega + |a_1| + a_1 cos\omega\tau \geq 0 \tag{56.21}$$

which is fulfilled provided the free Popov parameter β is chosen from

$$0 < \beta|a_1| < 2 \tag{56.22}$$

(more details concerning manipulation of the frequency domain inequality for time delay systems may be found in author's book [23]).

It follows that (56.19) 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 (56.7) and (56.18) has been answered positively.

56.5 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 which 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 lineaire" dans la theorie 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.*, (2002) - in print.
[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, contrrollability, observability" (in Russian), *Nauka Publ.House*, Moscow, (1979).

Problem 6

Root location of multivariable polynomials and stability analysis

Pierre-Alexandre Bliman
INRIA
Rocquencourt BP 105
78153 Le Chesnay cedex
France
pierre-alexandre.bliman@inria.fr

6.1 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

$$\forall s \in \overline{\mathbb{C}^+},\ \forall z \stackrel{\text{def}}{=} (z_1, \dots, z_m) \in \overline{\mathbb{D}}^m,\ \det(sI_n - A_0 - z_1 A_1 - \cdots - z_m A_m) \neq 0\ . \tag{6.1}$$

We have proved in [1] that property (6.1) is *equivalent* to the existence of a positive integer k and $(m+1)$ hermitian matrices $P, Q_1 \in \mathbb{C}^{k^m n \times k^m n}$, $Q_2 \in \mathbb{C}^{k^{m-1}(k+1)n \times k^{m-1}(k+1)n}, \dots, Q_m \in \mathbb{C}^{k(k+1)^{m-1} n \times k(k+1)^{m-1} n}$, such that

$$P > 0 \text{ and } R(P, Q_1, \dots, Q_m) < 0\ . \tag{6.2}$$

Here, R is a linear application taking values in the set of hermitian matrices in $\mathbb{C}^{(k+1)^m n \times (k+1)^m n}$, and which is defined as follows. Let $\hat{J}_k \stackrel{\text{def}}{=} (I_k \quad 0_{k\times 1})$, $\check{J}_k \stackrel{\text{def}}{=} (0_{k\times 1} \quad I_k)$, and define the matrices $J_k \in \mathbb{R}^{(m+1)k^m \times (k+1)^m}$ and $\hat{J}_{k,i}, \check{J}_{k,i} \in \mathbb{R}^{k^{m-i+1}(k+1)^{i-1} \times (k+1)^m}$, $1 \leq i \leq m$ by ($\otimes$ denotes the Kronecker product)

$$J_k \stackrel{\text{def}}{=} \begin{pmatrix} \hat{J}_k^{m\otimes} \\ \hat{J}_k^{(m-1)\otimes} \otimes \check{J}_k \\ \hat{J}_k^{(m-2)\otimes} \otimes \check{J}_k \otimes \hat{J}_k \\ \vdots \\ \check{J}_k \otimes \hat{J}_k^{(m-1)\otimes} \end{pmatrix},$$

$$\hat{J}_{k,i} \stackrel{\text{def}}{=} \hat{J}_k^{(m-i+1)\otimes} \otimes I_{(k+1)^{i-1}},\ \check{J}_{k,i} \stackrel{\text{def}}{=} \hat{J}_k^{(m-i)\otimes} \otimes \check{J}_k \otimes I_{(k+1)^{i-1}}\ .$$

The matrices $J_k, \hat{J}_k, \check{J}_k$ have 0 or 1 entries. Then,

$$R \stackrel{\text{def}}{=} (J_k \otimes I_n)^T \left(\begin{array}{c|c} (I_{k^m} \otimes A_0)^H P + P(I_{k^m} \otimes A_0) & P(I_{k^m} \otimes A_1) \quad \dots \quad P(I_{k^m} \otimes A_m) \\ \hline (I_{k^m} \otimes A_1)^H P & \\ \vdots & 0_{mk^m n} \\ (I_{k^m} \otimes A_m)^H P & \end{array} \right) (J_k \otimes I_n)$$
$$+ \sum_{i=1}^{m} \left((\hat{J}_{k,i} \otimes I_n)^T Q_i (\hat{J}_{k,i} \otimes I_n) - (\check{J}_{k,i} \otimes I_n)^T Q_i (\check{J}_{k,i} \otimes I_n) \right) . \tag{6.3}$$

Problem (6.2,6.3) is a *linear matrix inequality* in the $m+1$ matrix unknowns $P, Q_1, \dots, Q_m$, easily solvable numerically as a convex optimization problem.

The LMIs (6.2,6.3) obtained for increasing values of k constitute indeed a family of weaker and weaker sufficient conditions for (6.1). Conversely, property (6.1) 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* (6.1) *holds* if and only if *LMI* (6.2,6.3) *is solvable for $k = k^*$.*

One may imagine that such a 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 (6.1) holds *if and only if* LMI (6.2,6.3) is solvable for $k = k^*_{n,m}$.

6.2 Motivations and comments

Robust stability Property (6.1) is equivalent to asymptotic stability of the uncertain system

$$\dot{x} = (A_0 + z_1 A_1 + \dots + z_m A_m)x , \tag{6.4}$$

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 §6.1 systematizes the idea of expanding $S(z)$ in powers of the parameters. Indeed, robust stability of (6.4) guarantees existence of a Lyapunov function $x(t)^H S(z) x(t)$ with $S(z)$ *polynomial wrt z and $\bar{z}$* in $\overline{\mathbb{D}}^m$, and the integer k is related to the *degree* of this polynomial [1].

Computation of structured singular values (ssv) with repeated scalar blocks Property (6.1) 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 ssv 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 §6.1 offers attractive numerical alternative for the same purpose, as resolution of LMIs is a convex problem: it provides a family of simple P problems for approximation, with 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 (6.1) is equivalent to solvability of LMI (6.2,6.3) 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 (6.2,6.3) 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 (6.1) is a strong version of the delay-independent stability of system $\dot{x} = A_0x(t) + A_1x(t-h_1) + \dots + A_mx(t-h_m)$, $h_1, \dots, h_m \geq 0$, see [10, 2, 4]. This problem has been recognized as NP-hard [15]. Solving LMI (6.2,6.3) provides explicitly a quadratic Lyapunov-Krasovskii functional [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 (6.2,6.3), should also lead to progress in the analysis of the discrete-time analogue of (6.4), the uncertain system $x_{k+1} = (A_0 + z_1A_1 + \dots + z_mA_m)x_k$. Similarly, stability analysis for multidimensional systems would benefit from such contributions.

Bibliography

[1] P.-A. Bliman (2001). A convex approach to robust stability for linear systems with uncertain scalar parameters, Report research no 4316, INRIA. Available online at `http://www.inria.fr/rrrt/rr-4316.html`

[2] S. Boyd, L. El Ghaoui, E. Feron, V. Balakrishnan (1994). *Linear matrix inequalities in system and control theory*, SIAM Studies in Applied Mathematics vol. 15

[3] R.P. Braatz, P.M. Young, J.C. Doyle, M. Morari (1994). Computational complexity of μ calculation, *IEEE Trans. Automat. Control* **39** no 5, 1000–1002

[4] J. Chen, H.A. Latchman (1995). Frequency sweeping tests for stability independent of delay, *IEEE Trans. Automat. Control* **40** no 9, 1640–1645

[5] M. Dettori, C.W. Scherer (1998). Robust stability analysis for parameter dependent systems using full block S-procedure, *Proc. of 37th IEEE CDC*, Tampa (Florida), 2798–2799

[6] M. Dettori, C.W. Scherer (2000). New robust stability and performance conditions based on parameter dependent multipliers, *Proc. of 39th IEEE CDC*, Sydney (Australia)

[7] E. Feron, P. Apkarian, P. Gahinet (1996). Analysis and synthesis of robust control systems via parameter-dependent Lyapunov functions, *IEEE Trans. Automat. Control* **41** no 7, 1041–1046

[8] P. Gahinet, P. Apkarian, M. Chilali (1996). Affine parameter-dependent Lyapunov functions and real parametric uncertainty, *IEEE Trans. Automat. Control* **41** no 3, 436–442

[9] J.C. Geromel, M.C. de Oliveira, L. Hsu (1998). LMI characterization of structural and robust stability, *Linear Algebra Appl.* **285** no 1-3, 69–80

[10] J.K. Hale, E.F. Infante, F.S.P. Tsen (1985). Stability in linear delay equations, *J. Math. Anal. Appl.* **115**, 533–555

[11] D. Peaucelle, D. Arzelier, O. Bachelier, J. Bernussou (2000). A new robust D-stability condition for real convex polytopic uncertainty, *Systems and Control Letters* **40** no 1, 21–30

[12] D.C.W. Ramos, P.L.D. Peres (2001). An LMI approach to compute robust stability domains for uncertain linear systems, *Proc. ACC*, Arlington (Virginia), 4073–4078

[13] A. Rantzer, M. Johansson (2000). Piecewise linear quadratic optimal control, *IEEE Trans. Automat. Control* **45** no 4, 629–637

[14] M. Sznaier, P.A. Parrilo (1999). On the gap between μ and its upper bound for systems with repeated uncertainty blocks, *Proc. of 38th IEEE CDC*, Phoenix (Arizona), 4511–4516

[15] O. Toker, H. Özbay (1996). Complexity issues in robust stability of linear delay-differential systems, *Math. Control Signals Systems* **9** no 4, 386–400

[16] O. Toker, H. Özbay (1998). On the complexity of purely complex μ computation and related problems in multidimensional systems, *IEEE Trans. Automat. Control* **43** no 3, 409–414
[17] O. Toker, B. de Jager (1999). 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
[18] S. Treil ($\leq$ 1999). The gap between complex structured singular value μ and its upper bound is infinite. Available online at `http://www.math.msu.edu/treil/papers/mu/mu-abs.html`
[19] A. Trofino (1999). Parameter dependent Lyapunov functions for a class of uncertain linear systems: a LMI approach, *Proc. of 38th IEEE CDC*, Phoenix (Arizona), 2341–2346
[20] A. Trofino, C.E. de Souza (1999). Bi-quadratic stability of uncertain linear systems, *Proc. of 38th IEEE CDC*, Phoenix (Arizona)
[21] L. Xie, S. Shishkin, M. Fu (1997). Piecewise Lyapunov functions for robust stability of linear time-varying systems, *Syst. Contr. Lett.* **31** no 3, 165–171

Problem 18

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.ca and miller@hobbes.uwaterloo.ca

18.1 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:

$$DM(P_0, C) := \sup\{\overline{\tau} \;:\; \text{for all } \tau \in [0, \overline{\tau}], \text{ the feedback control system with controller } C(s) \text{ and plant } P(s) = P_0(s)exp(-s\tau) \text{ is internally stable}\}.$$

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 cross-over 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 pre-specified 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 pre-specified 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

$$DM_{sup}(P_0) := \sup\{DM(P_0, C) \ : \ \text{the feedback control system with controller } C(s) \text{ and plant } P_0(s) \text{ is internally stable}\}$$

is finite?

Problem 2: If the answer to Problem 1 is affirmative, devise a computationally feasible algorithm which, 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 which, given $P_0(s)$ and a value T in the range $0 < T < DM_{sup}(P_0)$, constructs a $C(s)$ that satisfies $DM(P_0, C) \geq T$.

18.2 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 53

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 Extramadura
Avda. De Elvas s/n, 06071-Badajoz,
Spain
bvinagre@unex.es

53.1 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 non-commensurate real orders [7] of the following form:

$$G(s) = \frac{b_m s^{\beta_m} + \ldots + b_1 s^{\beta_1} + b_0 s^{\beta_0}}{a_n s^{\alpha_n} + \ldots + a_1 s^{\alpha_1} + a_0 s^{\alpha_0}} = \frac{Q(s^{\beta_k})}{P(s^{\alpha_k})}, \tag{53.1}$$

where α_k, β_k $(k = 0, 1, 2, \ldots)$ are real numbers and without loss of generality they can be arranged as $\alpha_n > \ldots > \alpha_1 > \alpha_0$, $\beta_m > \ldots > \beta_1 > \beta_0$. The coefficients a_k, b_k $(k = 0, 1, 2, \ldots)$ 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 non-commensurate order systems (or, in general, with fractional order systems) it is important to bear in mind that $P(s^\alpha)$, $\alpha \in \mathrm{R}$ 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 (53.1) is still open.

53.2 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 the complex variable 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, $\alpha_k = \alpha k, \beta_k = \alpha k, (0 < \alpha < 1), \forall k \in \mathrm{Z}$, and the transfer function has the following form:

$$G(s) = K_0 \frac{\sum_{k=0}^{M} b_k (s^\alpha)^k}{\sum_{k=0}^{N} a_k (s^\alpha)^k} = K_0 \frac{Q(s^\alpha)}{P(s^\alpha)} \tag{53.2}$$

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:

$$G(s) = K_0 \left[\sum_{i=1}^{N} \frac{A_i}{s^\alpha + \lambda_i} \right] \tag{53.3}$$

where $\lambda_i \, (i = 1, 2, .., N)$ are the roots of the polynomial $P(s^\alpha)$ or the system poles which are assumed to be simple. Stability condition can then be stated that [2, 3]:

> A commensurate order system described by a rational transfer function (53.2) 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 $\sigma = s^\alpha$, $\alpha \in \mathrm{R}^+, (0 < \alpha < 1)$, 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:

- **step1:** Rewrite the LTI FOS $G(s)$ of the commensurate order α, to the equivalence system $H(\sigma)$, where transformation is: $s^\alpha \to \sigma$, $\alpha \in \mathrm{R}^+$;
- **step2:** Write the interval polynomial $P(\sigma, q)$ of the equivalence system $H(\sigma)$, where interval polynomial is defined as
$$P(\sigma, q) = \sum_{i=0}^{n} [q^-, q^+] \sigma^i;$$
- **step3:** For interval polynomial $P(\sigma, q)$, construct four Kharitonov's polynomials:
$$p^{--}(\sigma), p^{-+}(\sigma), p^{+-}(\sigma), p^{++}(\sigma);$$

- **step4:** Test the four Kharitonov's polynomials whether they satisfy the stability condition: $|\arg(\sigma_i)| > \alpha\frac{\pi}{2}$, $\forall\sigma \in \mathrm{C}$, 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.

53.3 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 non-commensurate orders with known parameters.
- A *robust stability* test procedure of LTI FOS of non-commensurate 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, 1996, Lille, France, pp. 963-968.
[3] D. Matignon: *Stability properties for generalized fractional differential systems.* In: Proceeding of Fractional Differential Systems: Models, Methods and Applications, vol. 5, 1998, pp. 145-158.
[4] D. A. Pierre: *Minimum Mean-Square-Error Design of Distributed Parameter Control Systems*, ISA Transactions, vol. 5, 1966, pp. 263-271.
[5] I. Petráš and Ľ. Dorčák: *The Frequency Method for Stability Investigation of Fractional Control Systems*, J. of SACTA, vol. 2, no. 1-2, 1999, pp. 75-85.
[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 ICCC2002, May 27-30, Beskydy, 2002, (accepted).
[7] I. Podlubny: *Fractional Differential Equations*, Academic Press, San Diego, 1999.

Problem 73

Stability of Fractional-order Model Reference Adaptive Control

BLAS M. VINAGRE
Dept. of Electronic and Electromechanical Engineering
Industrial Engineering School, University of Extramadura
Avda. De Elvas s/n, 06071-Badajoz
Spain
bvinagre@unex.es

IVO PETRÁŠ, IGOR PODLUBNY
Department of Informatics and Process Control
BERG Faculty, Technical University of Košice
B. Němcovej 3, 042 00 Košice
Slovak Republic
ivo.petras; igor.podlubny@tuke.sk

YANGQUAN CHEN
Center for Self-Organizing and Intelligent System (CSOIS)
Dept. of Electrical and Computer Engineering
Utah State University, Logan, UT-84322-4160
USA
yqchen@ieee.org

73.1 Description of the problem

MRAS or MRAC: The model reference adaptive system (MRAS) [1] is one of the main approaches to adaptive control, in which the desired performance is expressed in terms of a reference model (a model that describes the desired input-output properties of the closed-loop system) and the parameters of the controller are adjusted based on the error between the reference model output and the system output.

The gradient approach to model-reference adaptive control [1] is based on the assumption that the parameters change more slowly than the other variables in the system. This assumption, which admits a quasi-stationary treatment, is essential for the computation of the sensitivity derivatives that are needed in the adaptation.

Let e denote the error between the system output, y, and the reference output, y_m. Let θ denote the parameters to be updated. By using the criterion

$$J(\theta) = \frac{1}{2}e^2, \tag{73.1}$$

the adjustment rule for changing the parameters in the direction of the negative gradient of J is that

$$\frac{\mathrm{d}\theta}{\mathrm{d}t} = -\gamma \frac{\partial J}{\partial \theta} = -\gamma e \frac{\partial e}{\partial \theta} \tag{73.2}$$

If it is assumed that the parameters change much more slowly than the other variables in the system, the derivative $\frac{\partial e}{\partial \theta}$, that is, the sensitivity derivative of the system, can be evaluated under the assumption that θ is constant.

There are many variants about the MIT rules for the parameter adjustment. For example, the *sign-sign* algorithm is widely used in communication systems [1]; the PI-adjustment rule is used in [2].

Fractional-order MRAC (FO-MRAC): Here, we will introduce a new variant of the MIT rules for the parameter adjustment by using the fractional order calculus. As can be observed in equation (73.2), the rate of change of the parameters depends solely on the adaptation gain, γ. Taking into account the properties of the fractional differential operator, it is possible to make the rate of change depending on both the adaptation gain, γ, and the derivative order, α, by using the adjustment rule

$$\frac{\mathrm{d}^\alpha \theta}{\mathrm{d}^\alpha t} = -\gamma \frac{\partial J}{\partial \theta} = -\gamma e \frac{\partial e}{\partial \theta} \tag{73.3}$$

where α is a positive real number and $\frac{\mathrm{d}^\alpha}{\mathrm{d}^\alpha t}$ denotes the fractional order derivative.

Fractional Derivative Definition: For self-containing purpose, we introduce in the following the definition of fractional order derivative. Fractional calculus is a generalization of integration and differentiation to non-integer (fractional) order fundamental operator ${}_a\mathrm{D}_t^\alpha$, where a and t are the limits and α, $(\alpha \in \Re)$ the order of the operation. The two definitions used for the general fractional integro-differential are the Grünwald-Letnikov (GL) definition and the Riemann-Liouville (RL) definition [3, 4]. The GL definition is that

$${}_a\mathrm{D}_t^\alpha f(t) = \lim_{h \to 0} h^{-\alpha} \sum_{j=0}^{[\frac{t-a}{h}]} (-1)^j \binom{\alpha}{j} f(t - jh) \tag{73.4}$$

where $[\cdot]$ means the integer part while the RL definition

$${}_a\mathrm{D}_t^\alpha f(t) = \frac{1}{\Gamma(n-\alpha)} \frac{\mathrm{d}^n}{\mathrm{d}t^n} \int_a^t \frac{f(\tau)}{(t-\tau)^{\alpha-n+1}} \mathrm{d}\tau \tag{73.5}$$

for $(n - 1 < r < n)$ and where $\Gamma(\cdot)$ is the Euler's *gamma* function.

For convenience, Laplace domain notion is usually used to describe the fractional integro-differential operation [4]. The Laplace transform of the RL fractional derivative/integral (73.5) under zero initial conditions for order α, $(0 < \alpha < 1)$ is given by [3]:

$$\pounds\{{}_a\mathrm{D}_t^{\pm\alpha} f(t); s\} = s^{\pm\alpha} F(s). \tag{73.6}$$

A Simple Case: Based on the introduced fractional-order calculus notations, the above parameter updating rule (73.3) can be expressed as follows:

$$\theta = -\gamma I^\alpha \left[\frac{\partial J}{\partial \theta}\right] = -\gamma I^\alpha \left[e \frac{\partial e}{\partial \theta}\right]; \qquad I^\alpha \equiv D^{-\alpha} \tag{73.7}$$

For example, consider the first-order SISO system to be controlled:

$$\frac{dy}{dt} + ay = bu \tag{73.8}$$

where y is the output, u is the input and the system parameters a and b are unknown constants or unknown slowly time-varying. Assume that the corresponding reference model is given by

$$\frac{dy_m}{dt} + a_m y_m = b_m u_c \tag{73.9}$$

where u_c is the reference input signal for the reference model, y_m is the output of the reference model and a_m and b_m are known constants. Perfect model-following can be achieved with the controller defined by

$$u(t) = \theta_1 u_c(t) - \theta_2 y(t) \tag{73.10}$$

where

$$\theta_1 = \frac{b_m}{b}; \qquad \theta_2 = \frac{a_m - a}{b} \tag{73.11}$$

¿From Eqs. (73.8) and (73.10), assuming that $a + b\theta_1 \approx a_m$, and taking into account that b can be absorbed in γ, the equations for updating the controller parameters can be designed as (see, e.g., [1]),

$$\frac{\mathrm{d}^\alpha \theta_1}{\mathrm{d}t^\alpha} = -\gamma \left(\frac{1}{p + a_m} \right) u_c e \tag{73.12}$$

$$\frac{\mathrm{d}^\alpha \theta_2}{\mathrm{d}t^\alpha} = \gamma \left(\frac{1}{p + a_m} \right) y e \tag{73.13}$$

where $p = \frac{d}{dt}$, and γ is the adaptation gain, a small positive real number. Equivalently, in frequency domain, (73.12) and (73.13) can be written as

$$\theta_1 = -\frac{\gamma}{s^\alpha} \left(\frac{1}{s + a_m} \right) u_c e \tag{73.14}$$

$$\theta_2 = \frac{\gamma}{s^\alpha} \left(\frac{1}{s + a_m} \right) y e. \tag{73.15}$$

Clearly, the conventional MRAC [1] is the case when $\alpha = 1$.

The Open Problem: Even for linear systems, the stability analysis for the fractional order MRAC is open; how to design γ the adaptation gain and the order α is also open. We conjecture that, the analysis should be easier in the frequency domain.

73.2 Motivation and history of the problem

Fractional calculus is a 300-years-old topic. The theory of fractional-order derivative was developed mainly in the 19-th century. Recent books [3, 5, 4] provide a good source of references on fractional calculus. However, applying fractional-order calculus to dynamic systems control is just a recent focus of interest [6, 7, 8, 9]. For pioneering work on this regard, we cite [10, 11].

The model reference approach was developed by Whitaker and his colleagues around 1960 [12]. MRAC (Model Reference Adaptive Control) has become a standard part in textbooks on adaptive control [1, 13]. The well known MIT rule for MRAC is to adjust or update the unknown parameter using gradient information.

73.3 Available results

So far, there is no stability analysis result for the FO-MRAC scheme. However, some experimental results using numerical simulation is reported in [14] with some illustrated benefits. To simulate the FO-MRAC, one can use the approximate discretization scheme for s^α (e.g. [15] and the references therein).

How to determine the fractional order α is an interesting and open problem.

Bibliography

[1] K. J. Astrom and B. Wittenmark, *Adaptive Control*, Addison-Wesley Publishing Company Inc., Second Edition. Reading, 1995.

[2] C. C. Hang and P. C. Parks, "Comparative studies of model reference adaptive control systems," *IEEE Trans. Automatic Control*, vol. 18, pp. 419–428, 1973.

[3] K. B. Oldham and J. Spanier, *The Fractional Calculus*, Academic Press, New York, 1974.

[4] I. Podlubny, *Fractional Differential Equations*, Academic Press, San Diego, 1999.

[5] K. S. Miller and B. Ross, *An Introduction to the Fractional Calculus and Fractional Differential Equations*, Wiley, New York, 1993.

[6] Boris J. Lurie, "Three-parameter tunable tilt-integral-derivative (TID) controller," *US Patent US5371670*, 1994.

[7] Igor Podlubny, "Fractional-order systems and $PI^{\lambda}D^{\mu}$-controllers," *IEEE Trans. Automatic Control*, vol. 44, no. 1, pp. 208–214, 1999.

[8] A. Oustaloup, B. Mathieu, and P. Lanusse, "The CRONE control of resonant plants: application to a flexible transmission," *European Journal of Control*, vol. 1, no. 2, 1995.

[9] H.-F. Raynaud and A. ZergaInoh, "State-space representation for fractional order controllers," *Automatica*, vol. 36, pp. 1017–1021, 2000.

[10] S. Manabe, "The non-integer integral and its application to control systems," *JIEE (Japanese Institute of Electrical Engineers) Journal*, vol. 80, no. 860, pp. 589–597, 1960.

[11] A. Oustaloup, "Linear feedback control systems of fractional order between 1 and 2," in *Proc. of the IEEE Symposium on Circuit and Systems*, Chicago, USA, 4 1981.

[12] P. V. Osburn, H. P. Whitaker, and A. Kezer, "Comparative studies of model reference adaptive control systems," *Institute of Aeronautical Sciences*, pp. Paper no. 61–39, 1961.

[13] Y. D. Landau, *Adaptive Control: the Model Reference Approach*, Marcel Dekker, New York, 1979.

[14] B. M. Vinagre, I. Petras, I. Podlubny, and Y. Q. Chen, "Using fractional-order adjustment rules and fractional order reference models in model reference adaptive control," *Nonlinear Dynamics: An Int. J. of Nonlinear Dynamics and Chaos in Engineering Systems*, vol. 27, no. 6, pp. (in press), 2002.

[15] YangQuan Chen and Kevin L. Moore, "Discretization schemes for fractional order differentiators and integrators," *IEEE Trans. on Circuits and Systems I: Fundamental Theory and Applications*, vol. 49, no. 3, pp. 363–367, 2002.

Problem 29

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

29.1 Description of the problem

By definition, a system of the form

$$\dot{x}(t) = Ax(t), \quad t \geq 0 \tag{29.1}$$

($A \in \mathbb{K}^{n\times n}$, $\mathbb{K} = \mathbb{R}, \mathbb{C}$) is exponentially stable if and only if there are constants $M \geq 1$, $\beta < 0$ such that

$$\|e^{At}\| \leq Me^{\beta t}, \quad t \geq 0. \tag{29.2}$$

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 (29.2) 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 exponentially stability. More precisely, for every $\beta > \gamma(A)$ there *exists* a constant $M \geq 1$ such that (29.2) is satisfied. However it is unknown how to determine the minimal value of M such that (29.2) 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

$$\dot{x} = Ax + Bu, \tag{29.3}$$

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 $(M,\beta)-$*stability radius* defined by

$$r_{(M,\beta)}(A;D,E) = \inf\left\{\|\Delta\| \in \mathbb{K}^{\ell\times q},\ \exists \tau > 0 :\ \|e^{(A+D\Delta E)\tau}\| \geq Me^{\beta\tau}\right\}. \tag{29.4}$$

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

$$r_{(M,\beta)}(A - BF;D,E) \geq \gamma\,. \tag{29.5}$$

b) Provide a method for the computation of feedback matrices F such that (29.5) is satisfied.

29.2 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 which decide on the quality of a controller in applications. As such, the notion of $(M,\beta)-$stability is related to such classical design criteria as "overshoot" of system responses. The question 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 29.3.

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.

29.3 Available results

For Problem 1 a number of results are available. 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 $(M, \beta)-$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 (29.1) 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

$$A^*P + PA \leq 2\beta P\,,$$

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

$$P(A - BF) + (A - BF)^*P \leq 2\beta P \quad \text{and} \quad \kappa(P) = \|P\| \|P^{-1}\| \leq M^2\,, \tag{29.6}$$

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 (29.6) is far from being equivalent to Problem 2.

Concerning Problem 3 differential Riccati equations are used to derive bounds for the $(M, \beta)-$ stability radius in [6]. Suppose there exist positive definite Hermitian matrices P^0, Q, R of suitable dimensions such that the differential Riccati equation

$$\dot{P} - (A - \beta I)P - P(A - \beta I)^* - E^*QE - PDRD^*P = 0 \tag{29.7}$$

$$P(0) = P^0 \tag{29.8}$$

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 $(M,\beta)-$stability radius is at least

$$r_{(M,\beta)}(A;D,E) \geq \sqrt{\underline{\sigma}(Q)\underline{\sigma}(R)}\,, \tag{29.9}$$

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 (29.9) 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. Bristeau, M.-O. (ed.) et al., 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. To appear 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. preprint Fernuniversität Hagen, Hagen, Germany, 2001.

Problem 85

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

85.1 Description of the problem

The open problem presented here is the Generalized Lyapunov Theory and its Ω-transformable regions. First, the definition of the Ω-transformable regions and its degrees is described. Then, an open problem of the generalized Lyapunov theory regarding the Ω-transformable regions (with the degree above two) is presented.

Definition 1. (Gutman & Jury 1981) A region

$$\Omega_v = \{(x, y) \mid f(\lambda, \lambda^*) = f(x + jy, x - jy) = f_{xy}(x, y) < 0\} \tag{85.1}$$

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 both are Ω-transformable regions. However, some regions are non-Ω-transformable regions.

Open Problem. (Generalized Lyapunov Theory) Let $A \in C^{n\times n}$ and consider 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 is the degree of the boundary function (any positive integer number) and

$$f(\lambda, \lambda^*) = \sum_{p+q\leq v,\, p,\, q=1}^{v} c_{pq}\lambda^p\lambda^{*q},\ \lambda = x + jy \tag{85.2}$$

λ is a point on the complex plane. For the eigenvalues of 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)

$$\sum_{p+q\leq v,\, p,q=1}^{v} c_{pq}A^pPA^{*q} = -Q \tag{85.3}$$

Strictly say, the above open problem is for Ω -transformable regions with degree v greater than two. In order to let the problem be a more general picture, we present it as above for any positive integer v.

85.2 Motivation and history of the problem

The Lyapunov theory is well known for Hurwitz stability and Schur stability, i.e., continuous-time system and discrete-time system, respectively. Now, the above described generalized Lyapunov theory (GLT) is a general theory that takes both continuous-time and discrete-time Lyapunov theories as its special cases. Furthermore, it is well known that the system closed-loop 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, especially 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 Lyapunov theory to the stability problems.

Let us briefly review the history or 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 if the Lyapunov equation is satisfied.

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)

$$AP + PA^* = -Q \tag{85.4}$$

For discrete-time systems, the interest is to check if the system matrix eigenvalues lie within the unit disk. The corresponding Lyapunov theory for discrete-time systems is as follows.

Lyapunov Theory (discrete-time). For the eigenvalues of 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

$$APA^* - P = -Q \tag{85.5}$$

Thus, it is clear that the Lyapunov theory for Hurwitz stability and Schur stability is a special case of Generalized Lyapunov Theory described in the above-presented open problem with the specific Ω-transformable regions, the left half-plane and the unit disk of the degrees one and two, respectively.

When robust control is considered, we need robust performance in additional robust stability. Thus, a robust pole clustering, or robust root clustering or robust Gamma stability is needed and called in the literature (Ackermann, Kaesbauer & Muench 1991, Barmish 1994, Wang & Shieh 1994a,b, Yedavalli 1993, among others). First, define the region boundary function. Then, the GLT is very useful for us to determine/guarantee the system performance, similar to the system stability. It will also be used to determine robust pole clustering in general Ω-transformable regions, i.e., for robust performance. Also, these kinds of general regions considered may be very interesting in the study of discrete-time systems where the transient behavior is hard to specify in terms of common pole-clustering regions. On other areas, such as two and multidimensional digital filters and multidimensional systems, the Ω-transformable regions and its related GLT are and will be further useful, as well as non-Ω-transformable regions which tell us the GLT is not valid there. This is a motivation for investigating the open problem GLT and its Ω-transformable regions.

85.3 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 (1) with its boundary function f, where $v = 1, 2$ and

$$f(\lambda, \lambda^*) = \sum_{p+q\leq v,\, p,q=1}^{v} c_{pq}\lambda^p \lambda^{*q} \tag{85.6}$$

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

$$\sum_{p+q\leq v,\, p,q=1}^{v} c_{pq} A^p P A^{*q} = -Q \tag{85.7}$$

Notice that the GLT is proved/valid only 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. It is well-known that if a region is Ω-transformable, its complement is not Ω-transformable.

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

$$\Gamma_v = \{(x,\, y) \mid |\psi(s)|^2 - |\phi(s)|^2 < 0,\; s = x + jy\} \tag{85.8}$$

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 (8). The complement of a Γ-transformable region is also a Γ-transformable region.

It is also noticed that Horng, Horng and Chou (1993) and Yedavalli (1993) discussed robust pole clustering in Ω -transformable regions with degrees one and two by applications of the GLT. On the other hand, Wang and Shieh (1994a,b) used a Rayleigh principle approach for analysis of robust pole clustering in general Ω -regions, described as

$$\Omega_v = \{(x,\, y) \mid f(\lambda,\, \lambda^*) = \sum_{p+q \leq v,\, p,q=1}^{v} c_{pq} \lambda^p \lambda^{*q} < 0,\; \lambda = x + jy\} \tag{85.9}$$

which they called Hermitian regions or general Ω regions, including both Ω -transformable and non-Ω -transformable regions, as well as Γ regions. Notice that the general Ω regions 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 and 2002) 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 address any solution about the above specific open problem. Thus, the above open problem is still an open problem so 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, "Algebriac 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). *Automatica*, (to appear).

[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 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 (1994).
[11] 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).
[12] 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).
[13] R. K. Yedavalli, "Robust root clustering for linear uncertain system using generalized Lyapunov theory", *Automatica*, 29(1), pp.237-240 (1993).

Problem 20

Linearization of linearly controllable systems

R.Devanathan
School of Electrical and Electronic Engineering (Block S1)
Nanyang Technological University
Singapore 639798
E-mail:edevan@ntu.edu.sg

20.1 Description of the problem

We consider a class of systems of the form

$$\dot{\xi} = f(\xi) + g(\xi)\zeta \tag{20.1}$$

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 (20.1) be linearly controllable (Krener [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 problem is to develop a systematic explicit methodology to linearize the above system upto an arbitrary order.

20.2 Motivation and history of the problem

Linearization of a nonlinear dynamic system (without a control input) was originally investigated by Poincare [2, 3]. It is shown that, around an equilibrium point, a near identity (normalizing) transformation of a nonlinear system 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 at the order of these residual nonlinearities.

The residual nonlinearities belong to the null space of the mapping characterized by what is known as the homological equation whose solution corresponds to the normalizing transformation.The solution for the normalizing transformation is in the form of an infinite series whose convergence has been proved under certain assumptions [4, 5]. 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 non-resonant) through such a transformation thus providing an approximate linearization of the dynamic system.

20.3 Available results

Krener et. al. [6] have considered a nonlinear system with a control input, such as, (20.1) and showed that a generalized form of the homological equation can be formulated in this case. Devanathan [7] has shown that, with an appropriate choice of state feedback, the system matrix can be made non-resonant. This concept is further exploited in [8] to find a solution to the second order linearization.

For an arbitrary order linearization, one can proceed as follows:
The power series expansion of (20.1) about the origin can be written as

$$\dot{x} = Fx + G\phi + O_1(x)^{(2)} + \gamma_1(x, \phi)^{(1)} \tag{20.1}$$

where, without loss of generality, F and G can be considered to be in Brunovsky form [9], 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 now introduce state feedback as in [7]

$$\phi = -Kx + u \tag{20.2}$$

where

$$K = [k_n, k_{n-1}, \cdots, k_2, k_1]^t \tag{20.3}$$

(20.1) then becomes

$$\dot{x} = Ax + G\,u + O(x)^{(2)} + \gamma(x, u)^{(1)} \tag{20.4}$$

where

$$A = F - GK \tag{20.5}$$

We can choose the eigenvalues of matrix A in (20.5), without loss of generality, to be real, distinct and non-resonant by the proper choice of the matrix K [7]. Putting (20.4) into the form

$$\dot{x} = Ax + Gu + f_2(x) + f_3(x) + \cdots + g_1(x)u + g_2(x)u + \cdots \tag{20.6}$$

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

$$x^m \ = x_1^{m_1} x_2^{m_2} \cdots x_n^{m_n}, \sum_{i=1}^{n} m_i = m, m \geq 2.$$

Consider a change of variable as in

$$x = y + h(y) \tag{20.7}$$

and a change of input as in

$$v = (1 + \beta(x))u + \alpha(x)\,, \;\; 1 + \beta(x) \neq 0 \tag{20.8}$$

and put

$$f_2(x) + f_3(x) + f_4(x) + \cdots = f_2'(y) + f_3'(y) + f_4'(y) + \cdots \tag{20.9}$$

$$g_1(x)u + g_2(x)u + g_3(x)u + \cdots = g_1'(y)u + g_2'(y)u + g_3'(y)u + \cdots \tag{20.10}$$

$$\alpha(x) = \alpha_2(x) + \alpha_3(x) + \alpha_4(x) + \cdots = \alpha_2'(y) + \alpha_3'(y) + \alpha_4'(y) + \cdots \tag{20.11}$$

$$\beta(x) = \beta_1(x) + \beta_2(x) + \beta_3(x) + \cdots = \beta_1'(y) + \beta_2'(y) + \beta_3'(y) + \cdots \tag{20.12}$$

for some appropriate polynomials $f_m'(.), g_{m-1}'(.), \alpha_m'(.)\, and\, \beta_{m-1}'(.), m \geq 2.$

Substituting (20.7) and (20.8) into (20.6) and using (20.9)-(20.12), consider the polynomials of the form $y^m\, and\, y^{m-1}u$, $m = 2, 3, \cdots$. Then, the terms "$f_m(x)$" and "$g_{m-1}(x)u$" of arbitrary order can be removed from (20.6) progressively , m = 2,3, etc. provided the following generalized homological equations, are satisfied.

$$\frac{\partial h_m(y)}{\partial y}(Ay) - Ah_m(y) + G\alpha_m'(y) = f_m''(y), m \geq 2 \tag{20.13}$$

$$\frac{\partial h_m(y)}{\partial y}(Gu) + G\beta_{m-1}'(y)u = g_{m-1}''(y)u, \;\; \forall\, u, m \geq 2 \tag{20.14}$$

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

Assume that $h_{m-j}(y), \alpha_{m-j}(y), \beta_{m-j-1}(y), j = 1, 2, \cdots, (m-2)$, $m > 2$, are known. Then, since the matrix A is non-resonant, $h_m(y)$ can be explicitly solved for in an unique way from (20.13) in terms of $\alpha_m(y)$. Substitution of the solution of $h_m(y)$ into (20.14) results in a linear system of $n \times \sharp S_{m-1}$ equations in $\sharp S_m + \sharp S_{m-1}$ variables, $m > 2$, where $S_m = [(m_1, m_2, \cdots, m_n) \mid \sum_{i=1}^{n} m_i = m]$ and $\sharp S_m$ corresponds to the number of elements in the finite set S_m.The rank of the matrix corresponding to such a linear system of equations needs to be established.

For $m = 2$, the corresponding rank is shown in [8] to be $(\frac{n(n+1)}{2} + n - 1)$. Further, in the case of m=2, the system of linear equations can be reduced to a system of $(\frac{n(n-1)}{2})$ equations in n variables whose rank is $(n-1)$.

It is conjectured that a corresponding reduction of the linear system of equations in the arbitrary order m case should also be possible. Formulation of the properties of the solution of these equations for the coefficients of the polynomials $\alpha_m(.)$ and $\beta_{m-1}(.)$ together with $h_m(.)$ 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] V.I.Arnold, "Geometrical Methods in the Theory of Ordinary Differential Equations", *Springer-Verlag, New York*, pp.177-188 (1983).

[3] J. Guckenheimer and P.Holmes, "Nonlinear Oscillations, Dynamical Systems and Bifurcation of Vector Fields",*Springer-Verlag, New York* (1983).

[4] Bruno A.D. "Local Methods in Nonlinear Differential Equations",*Springer- Verlag, Berlin* (1989).

[5] Giampaolo Cicogna, "On the Convergence of Normalizing Transformations in the Presence of Symmetries",*Departmento di Fisica, Universita di Pisa, P.za Torricelli 2, I-56126, Pisa,Italy.*

[6] A.J.Krener, S.Karahan, and M.Hubbard, " Approximate Normal Forms of Nonlinear Systems", *Proc. 27th Conf. Decision and Control*, pp.1223-1229 (1988).

[7] R. Devanathan,"Linearization Condition Through State Feedback", *IEEE Transactions on Automatic Control*, Vol.46, No.8, pp.1257-1260 (2001).

[8] 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).

[9] Brunovsky, P.," A Classification of Linear Controllable Systems", *Kybernetica cislo*, Vol. 3, pp.173-188 (1970).

Problem 5

The dynamical Lame system with the boundary control: on the structure of the reachable sets

M.I. Belishev
Dept. of the Steklov Mathematical Institute (POMI)
Fontanka 27
St. Petersburg 191011
Russia
belishev@pdmi.ras.ru [1]

5.1 Motivation

The origin of the questions posed below is the 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 the inverse problems based on their relations with the boundary control theory [1], [2].

5.2 Geometry

Let $\Omega \subset \mathbf{R}^3$ be a bounded domain with the ($C^\infty-$) smooth boundary Γ; λ, μ, ρ smooth functions (Lame parameters) satisfying $\rho > 0$, $\mu > 0$, $3\lambda + 2\mu > 0$ in $\bar{\Omega}$.
The parameters determine two metrics in $\bar{\Omega}$

$$dl^2 = \frac{|dx|^2}{c_\alpha^2}, \qquad \alpha = p, s$$

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.
The distant functions (eikonals)

$$\tau_\alpha(x) := \mathrm{dist}_\alpha(x, \Gamma), \qquad x \in \bar{\Omega}$$

determine the subdomains

$$\Omega_\alpha^T := \{x \in \Omega \,|\, \tau_\alpha(x) < T\}, \qquad T > 0$$

and the values $T_\alpha := \inf\{T > 0 \,|\, \Omega_\alpha^T = \Omega\}$ ($\alpha = p, s$). 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

$$u(x) = u(x)_p + u(x)_s \,, \quad x \in \Omega_p^T \qquad (*)$$

with $u(x)_p \parallel \nu_p(x)$ and $u(x)_s \perp \nu_s(x)$.

5.3 The 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 Lame 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 $\rho\, dx$); $\mathcal{H}_\alpha^T := \{\, y \in \mathcal{H} \,|\, \operatorname{supp} y \subset \bar{\Omega}_\alpha^T\}$. As shown in [3], the correspondence $f \mapsto u^f$ 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

$$\mathcal{U}^T := \{\, u^f(\cdot, T) \,|\, f \in L_2(\Gamma \times [0, T]; \mathbf{R}^3)\}$$

is embedded into $\mathcal{H}_p^T$. As proved in the same paper, the relation

$$\operatorname{clos} \mathcal{U}^T \supset \mathcal{H}_s^T$$

is valid for *any* $T > 0$, i.e. an approximate controllability ever holds in the subdomain Ω_s^T filled with the shear waves, whereas the elements of the defect subspace

$$\mathcal{N}^T := \mathcal{H}_p^T \ominus \operatorname{clos}_{\mathcal{H}} \mathcal{U}^T$$

('unreachable states') can be supported only in $\Delta \Omega^T$. On the other hand, it is not difficult to show the examples with $\mathcal{N}^T \neq \{0\}, \ T < T_s$.

5.4 Problems and hypotheses

The open problem is to characterize the defect subspace $\mathcal{N}^T$. The following is the reasonable hypotheses.

- The defect space is ever nontrivial: $\mathcal{N}^T \neq \{0\}$ for $T < T_s$ in the general case (not only in examples). Note that due to the standart 'controllability – observability' duality this property would lead to existence of *the slow waves* which forward front propagates with the velocity c_s in *inhomogeneous* isotropic elastic medium.

- 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 specify the operator K^T.

- The decomposition $(*)$ diagonalizes the principal part of the Lame system.

The progress in these questions would be of great importance for the inverse problems of the elasticity theory which is now 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 Lame system: regularity of solutions, boundary controllability and boundary data continuation. *ESAIM COCV, 7, 2002 (to appear).*

Problem 30

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
Faculty of Mathematical Sciences
University of Twente
P.O. Box 217
7500 AE Enschede
The Netherlands
h.j.zwart@math.utwente.nl

30.1 Description of the problem

We consider the abstract system

$$\dot{x}(t) = Ax(t), \ x(0) = x_0, \quad t \geq 0 \tag{30.1}$$
$$y(t) = Cx(t), \quad t \geq 0, \tag{30.2}$$

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 (30.1) 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 (30.2) is well-defined, and continuous. However, in many PDE's, rewritten in the form (30.1)-(30.2), 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

$$\int_0^\infty \|CT(t)x\|_Y^2 \, dt \leq L\|x\|_H^2, \qquad x \in D(A). \tag{30.3}$$

Assuming that the observation operator C is admissible, system (30.1)-(30.2) 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

$$\int_0^\infty \|CT(t)x\|_Y^2 \, dt \geq l\|x\|_H^2, \qquad x \in D(A). \tag{30.4}$$

Often, the emphasis is on exact observability on a finite interval, which means that the integral in (30.4) is over $[0, t_0]$ for some $t_0 > 0$. However, for exponentially stable semigroups both notions are equivalent, i.e., if (30.4) 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 (30.1)-(30.2) 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

$$\|(sI - A)x\|_H^2 + |Re(s)|\|Cx\|_Y^2 \geq m|Re(s)|^2\|x\|_H^2 \tag{30.5}$$

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.

30.2 Motivation and history of the conjecture

System (30.1)-(30.2) with $A \in \mathbb{C}^{n\times n}$ and $C \in \mathbb{C}^{p\times n}$ is observable if and only if

$$rank \begin{matrix} sI - A \\ C \end{matrix} = n \qquad \text{for all } s \in \mathbb{C}. \tag{30.6}$$

This is known as the Hautus test, due to Hautus [2] and Popov [8]. If A is a stable matrix, then (30.6) 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 (30.1)-(30.2) 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, $\mathrm{Re}(\lambda_n) = -\rho_1$, $|\lambda_{n+1} - \lambda_n| > \rho_2$, where λ_n are the eigenvalues of A, and $\rho_1, \rho_2 > 0$, [9]; or if
- m in equation (30.5) 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 $\mathbb{C}$.

30.3 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 (30.1)-(30.2) is exactly observable if and only if there exists a bounded operator L which is positive and boundedly invertible and satisfies the Lyapunov equation

$$\langle Ax_1, Lx_2\rangle_H + \langle Lx_1, Ax_2\rangle_H = \langle Cx_1, Cx_2\rangle_Y, \quad \text{for all } x_1, x_2 \in D(A). \tag{30.7}$$

From the admissibility of C and the exponential stability of the semigroup, one easily obtains that equation (30.7) 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 (30.7) 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 transl. publ. 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 34

On the convergence of normal forms for analytic control systems

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

Arthur J. Krener
Department of Mathematics, University of California
Davis, CA 95616
ajkrener@ucdavis.edu

34.1 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 dynamics 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 some systems that are not linearizable, the Poincaré-Dulac Theorem provides the normal form.

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

$$\dot{x} = f(x) = Fx + f^{[2]}(x) + f^{[3]}(x) + \cdots \tag{34.1}$$

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

Definition 1.1 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 1.2 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 in the Siegel Domain, there are no resonances and as $|m| \to \infty$ the rate at which resonances are approached is controlled.
A formal change of coordinates is a formal power series

$$z = Tx + \theta^{[2]}(x) + \theta^{[3]}(x) + \cdots \tag{34.2}$$

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 1.1 *(Poincaré-Dulac) If the system (34.1) is C^∞ then there exists a formal change of coordinates (34.2) 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 (34.1) is real analytic and its eigenvalues lie in the Poincaré Domain (34.2), then $w(z)$ is a polynomial vector field and the change of coordinates(34.2) is real analytic.

Theorem 1.2 (Siegel) *If the system (34.1) 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(34.2) 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.

34.2 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

$$\dot{x} = f(x, u) = Fx + Gu + f^{[2]}(x, u) + f^{[3]}(x, u) + \cdots \tag{34.1}$$

where $x \in \Re^n$ is the state variable, $u \in \Re$ 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 Poincar'e's theory, a homogeneous transformation for (34.1) consists of both change of coordinates and invertible state feedback,

$$z = x + \theta^{[d]}(x), \quad v = u + \kappa^{[d]}(x, u) \tag{34.2}$$

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

$$z = Tx + \textstyle\sum_{d=2}^{\infty} \theta^{[d]}(x), \quad v = Ku + \textstyle\sum_{d=2}^{\infty} \kappa^{[d]}(x, u) \tag{34.3}$$

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].

Theorme 2.1 *Suppose (F, G) in (34.1) is a controllable pair. Under a suitable transformation (34.3), (34.1) can be transformed into the following normal form*

$$\begin{array}{rcl} \dot{z}_i & = & z_{i+1} + \sum_{j=i+2}^{n+1} p_{i,j}(\bar{z}_j) z_j^2 \quad 1 \le i \le n-1 \\ \dot{z}_n & = & v \end{array} \tag{34.4}$$

where $z_{n+1} = v$, $\bar{z}_j = (z_1, z_2, \cdots, z_j)$, *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 (34.4) 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 (34.1) is real analytic and* F, G *is a controllable pair. Under what conditions does there exist a real analytic transformation (34.3) that transforms the system to the normal form (34.4)?*

Normal forms of control systems have proven to be a powerful tool in the analysis of the local bifurcation and the local qualitative performance of a control system. 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 the analytic normal forms.

34.3 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, than 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 th convergence results of Poincaré-Dulac. For the latter the location of the eigenvalues are crucial and these are invariant under change of state 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 is 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, to appear.

[9] A. J. Krener and L. Li, Normal Forms and Bifurcations of Discrete Time Nonlinear Control Systems, *Proc. of* 5^{th} *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.

Problem 27

Nilpotent bases of distributions

Henry G. Hermes
Department of Mathematics
University of Colorado at Boulder,
Boulder, Colorado 80309
USA
hermes@euclid.colorado.edu

Matthias Kawski[1]
Department of Mathematics and Statistics
Arizona State University,
Tempe, Arizona 85287-1804
USA
kawski@asu.edu

27.1 Description of the problem

When modelling controlled dynamical systems one commonly chooses individual control variables $u_1, \ldots u_m$ which 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 $f_0, f_1, \ldots f_m \colon M \mapsto TM$, and the system takes the form

$$\dot{x} = f_0(x) + \sum_{k=1}^{m} u_k\, f_k(x). \tag{27.1}$$

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 $\mathcal{F} = \{v_1, \ldots v_m\}$ of vector fields define the *distribution spanned by* $\mathcal{F}$ as $\Delta_{\mathcal{F}} \colon p \mapsto \{c_1 v_1(p) + \ldots + c_m v_m(p) \colon c_1, \ldots c_m \in \mathbf{R}\} \subseteq T_pM$. For *systems with drift* f_0, the geometric object is the map $\tilde{\Delta}_{\mathcal{F}}(x) = \{f_0(x) + c_1 f_1(x) + \ldots + c_m f_m(x) \colon c_1, \ldots c_m \in \mathbf{R}\}$ whose image at every point x is an affine subspace of T_xM. The geometric character of the distribution is captured by its invariance under the *group of feedback transformations*. In traditional notation (here formulated for systems with drift) these are (analytic) maps (defined on suitable subsets) $\alpha \colon M^n \times \mathbf{R}^m \mapsto \mathbf{R}^m$ 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

$$u_k(x) = \alpha_{0k}(x) + v_1 \alpha_{1k}(x) + \ldots v_m \alpha_{mk}(x) \quad \text{for } k = 1, \ldots m. \tag{27.2}$$

This transformation of the controls induces a corresponding transformation of the vector fields defined by $\dot{x} = f_0(x) + \sum_{k=1}^{m} u_k\, f_k(x) \stackrel{!}{=} g_0(x) + \sum_{k=1}^{m} v_k\, g_k(x)$

$$\begin{array}{rclll} g_0(x) & = & f_0(x) + & \alpha_{01}(x) f_1(x) + \ldots \alpha_{0m}(x) f_m(x) & \\ g_k(x) & = & & \alpha_{k1}(x) f_1(x) + \ldots \alpha_{km}(x) f_m(x), & k = 1, \ldots m \end{array} \tag{27.3}$$

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

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 $L(g_0, g_1, \ldots g_m)$ 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 $\mathcal{F} = \{f_1, \ldots f_m\}$ to admit a *basis* of analytic vector fields $\mathcal{G} = \{g_1, \ldots g_m\}$ which generate a Lie algebra $L(g_1, \ldots g_m)$ 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 $\mathcal{F}$.

27.2 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 $\dot{x}_1 = 1, \dot{x}_2 = 0, \ldots \dot{x}_n = 0$ (in the neighbourhood 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 $\dot{x}_k = p_k(x_1, \ldots, x_{k-1})$ 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 [10, 14, 15]. In particular, if the Lie algebra is nilpotent (or similarly nice), the general solution formula for $x(\cdot, u)$ as an exponential Lie series [16] (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 $L(f_1, \ldots f_m)$ is nilpotent [8]. 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.

27.3 Related results

In [5] it is shown that for every $2 \leq k \leq (n-1)$ there is a k-distribution Δ on $\mathbf{R}^n$ which 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* which are defined inductively by $\Delta^{(1)} = \Delta$ and $\Delta^{(k+1)} = \Delta^{(k)} + \{[v, w] \colon v \in \Delta^{(k)}, w \in \Delta^{(1)}\}$.

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 $\{(\mathrm{ad}^j f_0, f_1) : 0 \leq j \leq k\}$ 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].

Another well-studied special case of nilpotent systems are those which can be brought into *chained-form*, compare [14]. This work builds on the Goursat normal form, and is a natural intermediary towards the dual description in terms of exterior differential systems. This dual description of systems in terms of co-distributions $\Delta^\perp = \{\omega : M \mapsto T^*M \ : \ \langle \omega, f \rangle = 0 \text{ for all } f \in \Delta\}$. is particularly convenient when working with small co-dimension $n - m$. (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], and many more recent theoretical tools and results, see e.g. [11] for a recent survey, and [12, 13] for recent results and further relevant references, Also see [17] for many related results and further references. In the 1990, much work has focused *differentially flat* systems, compare [2]. 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.

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 e.g. generate a finite dimensional, or nilpotent Lie algebra).

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. Jakubvzyk and W. Respondek, "On linearization of control systems", Bull. Acad. Polon. Sci. Ser. Sci. Math., 28, pp. 517–522 (1980).

[8] M. Kawski "Nilpotent Lie algebras of vector fields", *Journal für die Riene und Angewandte Mathematik*, 188, pp. 1-17 (1988).

[9] M. Kawski 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).

[10] G. Laffariere and H. Sussmann, "A differential geometric approach to motion planning", G. Lafferiere and H. Sussmann. Motion planning for controllable systems without drift. In IEEE International Conference on Robotics and Automation, pages 1148–1153, Sacramento, CA, 1991.

[11] R. Montgomery, "A Tour of Subriemannian Geometries, Their Geodesics and Applications", AMS Mathematical Surveys and Monographs, 91, (2002).

[12] P. Mormul, "Goursat flags, classification of co-dimension one singularities", *J. Dynamical and Control Systems*,6, 2000, pp. 311–330 (2000).

[13] P. Mormul, "Minimal nilpotent bases for goursat distributions in dimensions not exceeding seven", Proc. NOLCOS 01 (2001).

[14] 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).

[15] J. Strumper and P. Krishnaprasad, "Approximate tracking for systems on three dimensional Lie matrix groups via feedback nilpotentization", *IFAC Symposium Robot Control* (1997).

[16] 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).

[17] M. Zhitomirskii, "Singularities and normal forms of smooth distributions", Banach Center Publ., 32, pp 379-409 (1995).

Problem 9

Optimal Synthesis for the Dubins' Car with Angular Acceleration Control

Ugo Boscain
SISSA-ISAS
Via Beirut 2-4
34014 Trieste
Italy
boscain@sissa.it

Benedetto Piccoli
IAC – CNR
Viale Policlinico 137
00161 Roma
Italy
piccoli@iac.rm.cnr.it

9.1 Description of the problem

A modified version of the Dubins' car, in which the control is on the angular acceleration of the steering, is given by:

$$\begin{cases} \dot{x} &= \cos(\theta) \\ \dot{y} &= \sin(\theta) \\ \dot{\theta} &= \omega \\ \dot{\omega} &= u. \end{cases} \tag{9.1}$$

with $|u| \leq 1$, $(x, y) \in \mathbf{R}^2$, $\theta \in S^1$ and $\omega \in \mathbf{R}$. We use the notation $\mathbf{x} = ((x, y), \theta, \omega)$ and $M = \mathbf{R}^2 \times S^1 \times \mathbf{R}$.

Problem. Construct a time optimal synthesis to a fixed point $\bar{q} \in M$, i.e. a family of trajectory-control pairs $\{(\mathbf{x}_q(.), u_q(.)) : q \in M\}$ such that $\mathbf{x}_q(.)$ steers q to $\bar{q}$ in minimum time.

9.2 Motivation and history of the problem

One of the simplest model for a car–like robot is the one known as Dubins' car. In this model, the system state is represented by a pair $((x, y), \theta)$ where $(x, y) \in \mathbf{R}^2$ is the position of the center of the car and $\theta \in S^1$ is the angle between its axis and the positive x-axis. The car is assumed to travel with constant (in norm) speed, the control is on the angular velocity of the steering and is assumed to be bounded, thus we obtain the system:

$$\begin{cases} \dot{x} &= \cos(\theta) \\ \dot{y} &= \sin(\theta) \\ \dot{\theta} &= u \end{cases}$$

where $|u| \leq C$ (usually for simplicity one assumes $C = 1$).

This problem was originally introduced by Markov in [8] and studied by Dubins in [3]. In particular Dubins proved that every minimum time trajectory is concatenation of at most three arcs, each of which is either an arc of circle or a straight line. If we consider the possibility of non constant speed and admit also backward motion, then we obtain the model proposed by Reed and Shepp [10]. A family of time optimal trajectories, that are sufficiently reach to join optimally any two points, was given in [13]. Now the situation is more complicated since there are 46 possible combination of straight lines and arcs of circles. Then a time optimal synthesis was built by Soueres and Laumond in [11].

Time optimal trajectories for the system (9.1) were studied mainly by Boissonnat, Cerezo, Kostov, Kostova, Leblond and Sussmann, see [1, 2, 5, 6, 12].

9.3 Available results

As usual the set of candidates optimal trajectories is restricted by means of Pontryagin Maximum Principle (briefly PMP). In this case PMP can be formulated as follows:

Theorem (PMP) *Let T^*M be the cotangent bundle to the state space M. For every $(\mathbf{x},\mathbf{p},\lambda_0,u) \in T^*M \times \mathbf{R} \times [-1,1]$ define (here $\mathbf{p}= (p_1,...,p_4)$):*

$$\mathcal{H}(\mathbf{x},\mathbf{p},\lambda_0,u) = p_1\cos(\theta) + p_2\sin(\theta) + p_3\omega + p_4 u$$
$$H(\mathbf{x},\mathbf{p},\lambda_0) = \max\{\mathcal{H}(\mathbf{x},\mathbf{p},\lambda_0,u) : u \in [-1,1]\}.$$

A couple trajectory-control $(\mathbf{x}(.),u(.)) : [0,T] \to M \times [-1,1]$ is said to be extremal if there exist an absolutely continuous map $\mathbf{p}(.) : t \in [0,T] \mapsto \mathbf{p}(t) \in T^*_{\mathbf{x}(t)}M$ and a constant $\lambda_0 \leq 0$, with $(\mathbf{p}(t),\lambda_0) \neq (0,0)$, that satisfy for a.e. $t \in [0,T]$:

(PMP1) $\dot{p}_1 = 0,\ \dot{p}_2 = 0,\ \dot{p}_3 = p_1\sin(\theta) - p_2\cos(\theta),\ \dot{p}_4 = -p_3,$

(PMP2) $H(\mathbf{x}(t),\mathbf{p}(t),\lambda_0) = \mathcal{H}(\mathbf{x}(t),\mathbf{p}(t),\lambda_0,u(t)) = 0$, *that is* $p_4(t)u(t) = |p_4(t)|$.

We have the following: if a couple trajectory-control $(\mathbf{x}(.),u(.)) : [0,T] \to M \times [-1,1]$ is optimal then it is extremal.

The function $p_4(.)$ is the so called switching function. In fact from **(PMP2)** it follows that:

- if $p_4(t) > 0$ (resp < 0) for every $t \in [a,b]$, then $u \equiv 1$ (resp. $u \equiv -1$) on $[a,b]$. In this case the corresponding trajectory $\mathbf{x}(.)|_{[a,b]}$ is called a *bang* arc and it is an arc of clothoid in the (x,y) space.
- if $p_4(t) \equiv 0$ for every $t \in [a,b]$, then $u \equiv 0$ in $[a,b]$. In this case the trajectory $\mathbf{x}(.)|_{[a,b]}$ is called a *singular* arc and it is a straight line in the (x,y) space.

The main feature of this highly nongeneric problem is that *an optimal trajectory cannot contain points where the control jumps from* ±1 *to* 0 *or from* 0 *to* ±1. In other words a singular arc cannot be preceded or followed by a bang arc.

In [12] it is proved the existence of extremals presenting chattering. More precisely, there exist extremal trajectories $\mathbf{x}(.)$ defined on some interval $[a,b]$, which are singular on some interval $[c,b]$, where $a < c < b$, and such that $p_4(.)$ does not vanish identically on any subinterval of $[a,c]$. Moreover the set of zeros of $p_4(.)$ consists of c together with an increasing sequence $\{t_j\}_{j=1}^{\infty}$ of points converging to c. At each t_j, $\mathbf{x}(.)$ is switches the control from $+1$ to -1 or viceversa. Chattering phenomena where studied for instance in [4, 7, 9, 14], in particular a time optimal synthesis with chattering trajectories was first discovered in [4].

An optimal path can thus have at most a finite number of switchings only if it is a finite concatenation of clothoids, with no singular arc. Existence of optimal trajectories (not only extremal) presenting chattering was proved by Kostov and Kostova in [6]. More precisely if the distance between the initial and final point is big enough, then the shortest path can not be a finite concatenation of clothoids.

Bibliography

[1] J.D. Boissonnat, A. Cerezo and J. Leblond, *A note on shortest path in the plane subject to a constraint on the derivative of the curvature*, Research Report 2160, INRIA, 1998.

[2] J.D. Boissonnat, A. Cerezo, E. Degtiariova-Kostova, V. Kostov, and J. Leblond *Shortest plane paths with bounded derivative of the curvature.* C. R. Acad. Sci., 329-I, pp.613–618, 1999.

[3] Dubins L.E. *On curves of minimal length with a constraint on average curvature and with prescribed initial and terminal position and tangents*, Amer. J. Math, V.79 (1957), pp.497-516.

[4] A.T. Fuller, "Study of an optimal nonlinear system", *J. Electronics Control*, V.15 (1963), pp.63-71.

[5] V. Kostov and E. Degtiariova-Kostova *Suboptimal paths in the problem of a planar motion with bounded derivative of the curvature.* Comptes rendus de l'acadmie des sciences 321, 1995, pp. 1441–1447.

[6] V. Kostov and E. Degtiariova-Kostova, *Irregularity of Optimal Trajectories in a Control Problem for a Car-like Robot.* Research Report 3411, INRIA, 1998.

[7] I. Kupka, *The ubiquity of Fuller's phenomenon,* in Nonlinear Controllability and Optimal Control, H. J. Sussmann Ed., M. Dekker, Inc., New York (1990), pp. 313-350.

[8] A.A. Markov, *Some examples of the solution of a special kind of problem in greatest and least quantities*, (in Russian) Soobshch. Karkovsk. Mat. Obshch. 1, 1887, pp. 250–276.

[9] Marchal, C., *Chattering arcs and chattering controls,* J. Optim Theory Appl. 11 (1973), pp. 441-468.

[10] J.A. Reeds and L.A. Shepp, *Optimal Path for a car that goes both forwards and backwards*, Pacific J. Math, V.145 (1990), pp. 367-393.

[11] P.Soueres , J.P.Laumond, *Shortest paths synthesis for a car-like robot* , IEEE Transactions on Automatic Control, Vol.41, N5, pp.672-688.

[12] H.J. Sussmann, *The Markov-Dubins Problem with Angular Acceleration Control*, Proceedings of the 36th IEEE, Conference on Decision and Control, San Diego, CA, Dec. 1997. IEEE Publications, New York, 1997, pp. 2639-2643.

[13] H.J. Sussmann, G. Q. Tang, *Shortest Paths for the Reeds-Shepp Car: A Worked Out Example of the Use of Geometric Techniques in Nonlinear Optimal Control* Rutgers Center for Systems and Control Technical Report 91-10, September 1991, to appear in SIAM J. Control.

[14] M.I. Zelikin and V.F. Borisov, *Theory of chattering control. With applications to astronautics, robotics, economics, and engineering*, Systems & Control: Foundations & Applications, Birkhäuser, Boston, 1994.