ADVANCES IN AUTOMATIC CONTROL

edited by

Mihail Voicu

KLUWER ACADEMIC PUBLISHERS

ADVANCES IN AUTOMATIC CONTROL

edited by

Mihail Voicu
Technical University "Gh. Asachi" of Iasi, Romania

KLUWER ACADEMIC PUBLISHERS
Boston / Dordrecht / New York / London

Distributors for North, Central and South America:
Kluwer Academic Publishers
101 Philip Drive
Assinippi Park
Norwell, Massachusetts 02061 USA
Telephone (781) 871-6600
Fax (781) 681-9045
E-Mail < kluwer@wkap.com>

Distributors for all other countries:
Kluwer Academic Publishers Group
Post Office Box 322
3300 AH Dordrecht, THE NETHERLANDS
Telephone 31 78 6576 000
Fax 31 78 6576 254
E-Mail < services@wkap.nl>

Electronic Services < http://www.wkap.nl>

Library of Congress Cataloging-in-Publication Data

Advances in Automatic Control / edited by Mihail Voicu.
p.cm.—(Kluwer international series in engineering and computer science; SECS 754)
Includes bibliographical references and index.
ISBN 1-4020-7607-X (alk. paper)
1. Control. 2. Systems & Control Theory. 3. Control & Optimization. I. Voicu, Mihail. II. Series.

Printed on acid-free paper.

Printed in the United States of America.

Contents

Preface

During the academic year 2002-2003, the Faculty of Automatic Control and Computer Engineering of Iaşi (Romania), and its Departments of Automatic Control and Industrial Informatics and of Computer Engineering respectively, celebrated 25 years from the establishment of the specialization named Automatic Control and Computer Engineering within the framework of the former Faculty of Electrical Engineering of Iaşi, and, at the same time, 40 years since the first courses on Automatic Control and Computers respectively, were introduced in the curricula of the former specializations of Electromechanical Engineering and Electrical Power Engineering at the already mentioned Faculty of Electrical Engineering. The reader interested to know some important moments of our evolution during the last five decades is invited to see the Addendum of this volume, where a short history is presented. And, to highlight once more the nice coincidences, it must be noted here that in 2003 our Technical University "Gheorghe Asachi" of Iaşi celebrated 190 years from the emergence of the first cadastral engineering degree course in Iaşi (thanks to the endeavor of Gheorghe Asachi), which is today considered to be the beginning of the engineering higher education in Romania.

Generally speaking, an anniversary is a celebration meant to mark special events of the past, with festivities to be performed solemnly and publicly according to a specific ritual. And, if a deeper insight into the human nature and the social relations and their symbolism is considered, we must recognize in such a celebration an *a posteriori* constitution of an *ad hoc* rite of passage, which periodically actualize founding events marking the advance of some people's life, of some social groups, of some organizations and, corresponding to their emerging viability, of concrete and adequate institutions which must fulfill some well defined and / or recursively definable intellectual, social and economic tasks.

People celebrate fundamental events in many different ways. Taking into consideration that our celebration is the first one of this kind, the Faculty of Automatic Control and Computer Engineering and its two departments decided to mark their beginning moments by publishing two special books respectively. As a part of this double anniversary and to honor its founding events, the Department of Automatic Control Industrial Informatics decided to publish the present volume, titled ***Advances in Automatic Control***, meant to comprise also invited papers authored by well-known scientists who, in various forms, developed collaborative works with this department.

As it can be seen from the contents, the themes dealt with in the papers of this volume cover a large variety of topics which correspondingly reflects the very different research interests of the authors: stabilization of distributed parameter systems, disturbance attenuation in stochastic systems, analysis and simulation of discrete event systems, fault detection, characterization of linear periodic Hamiltonian systems, stability of time delay systems, flow invariance and componentwise asymptotic stability, distributed control, parametrization of stabilizing controller, vibration control, predictive control, fuzzy control, intelligent decision and control, optimal control, computer aided control, robot and CIM control, DVD player control. Nevertheless, throughout this variety of interests we can distinguish two unifying features: the novelty of the approaches and / or results, which can be explicitly perceived by reading the book, and the other one, having for us the same importance but acting rather implicitly from the first conceptual idea about this book, which mirrors the high quality of the human and collaborative relations previously established between the invited authors and the members of our department.

Finally, we wish to thank all the authors for their contributions and for their cooperation in making this book a successful part of the celebration marking the founding moments and the evolution of the Department of Automatic Control and Industrial Informatics. At the same time, we express our gratitude to AUTEC GmbH and especially to Dr. h. c. Hartmut Stärke, who financially contributed to the dissemination of this book and, during the last five years, partially supported the mobility of our students and researchers. Our thanks also go to Dana Serbeniuc, who electronically prepared the camera-ready manuscript, and, in this respect, to Mitică Craus and Laurențiu Marinovici for their benevolent and valuable counseling. At last, but not at least, we express our gratitude to Kluwer, especially to Jennifer Evans and Anne Murray for their efficient and kind cooperation during the entire process the result of which is the present volume.

It is not only a nice duty but also a great pleasure to acknowledge all of these contributions.

The Editor

INTERNAL STABILIZATION OF THE PHASE FIELD SYSTEM

Viorel Barbu
Department of Mathematics
"Al.I. Cuza" University, 6600 Iaşi, Romania
e-mail: vb41@uaic.ro

Abstract The phase-field system is locally exponentially stabilizable by a finite dimensional internal controller acting on a component of the system only.

Keywords: internal stabilization, phase filed system

1. Introduction

Consider the controlled phase field system

$$\begin{aligned} & y_t + \ell\varphi_t - k\Delta y = mu && \text{in } Q = \Omega \times (0,\infty) \\ & \varphi_t - a\Delta\varphi - b(\varphi - \varphi^3) + dy = 0 && \text{in } Q \\ & y = 0,\ \varphi = 0 && \text{on } \partial\Omega \times (0,\infty) \\ & y(x,0) = y_0(x),\ \varphi(x,0) = \varphi_0(x) && \text{in } \Omega, \end{aligned} \tag{1.1}$$

where $\Omega \in R^n$, $n = 1,2,3$ is an open and bounded domain with smooth boundary $\partial\Omega$ and a, b, ℓ, k, d are positive constants. Finally, m is the characteristic function of an open subset $\omega \subset \Omega$ and u is the internal control input.

This system models the phase transition of physical processes and in particular the melting or solidification processes. In this latter case y is temperature and φ is the phase function. The Stefan free boundary problem is a limiting case of problem (1.1).

The local controllability of system (1.1) where internal control inputs arise in both equations was proved in [1] via Carleman's inequality for linear parabolic equations (see [4]).

In [2] it was established the stabilization of null solution to (1.1) via a Riccati equation approach. The main result obtained here, Theorem 1 below is a sharpening of the results obtained in [2] on the lines of [3].

2. The main result

We set $A = -\Delta$ with $D(A) = H^2(\Omega) \cap H_0^1(\Omega)$. Let $\{\psi_i\}_{i=1}^{\infty}$ be an orthonormal basis of eigenfunctions for the operator a, i.e.,

$$A\psi_i = \lambda_i \psi_i,\ i = 1, \dots.$$

(Here each eigenvalue λ_i is repeated according to its multiplicity.)

Denote by A^s, $0 < s < 1$, fractional powers of Q and set $H = L^2(\Omega)$, $W = D\left(A^{\frac{1}{4}}\right)$, $V = D\left(A^{\frac{1}{2}}\right)$, with the usual norms.

For each $\rho > 0$ denote by $\mathcal{W}_\rho$ the open ball in $W \times W$

$$W_\rho = \left\{(y_0, \varphi_0) \in W \times W;\ \left|A^{\frac{1}{4}} y_0\right|^2 + \left|A^{\frac{1}{4}} \varphi_0\right|^2 < \rho^2\right\}.$$

Now we are ready to formulate the main result of this paper.

Theorem 1. *There are N and $R_N : D(R_N) \subset H \times H \to H \times H$, linear, self–adjoint satisfying*

$$C_2\left(\left|A^{\frac{1}{4}} y\right|^2 + \left|A^{\frac{1}{4}} \varphi\right|^2\right) \leq\ < R_N(y,\varphi),(y,\varphi) > \leq C_1\left(\left|A^{\frac{1}{4}} y\right|^2 + \left|A^{\frac{1}{4}} \varphi\right|^2\right) \tag{2.1}$$

$$\|R_N(y,\varphi)\|_{H\times H}^2 \leq C_3\left(\left|A^{\frac{1}{2}} y\right|^2 + \left|A^{\frac{1}{2}} \varphi\right|^2\right) \tag{2.2}$$

and such that the feedback controller

$$u = -\sum_{i=1}^{N} (R_{11} y(t) + R_{12}\varphi(t), \psi_j)_\omega \psi_j \tag{2.3}$$

exponentially stabilizes (1.1) *on W_ρ. More precisely, for all $(y_0, \varphi_0) \in \mathcal{W}_\rho$ we have*

$$|y(t)| + |\varphi(t)| \leq C_4 e^{-\gamma t}(\|y_0\|_W + \|\varphi_0\|_W) \tag{2.4}$$

$$\int_0^\infty \left(\left|A^{\frac{1}{4}} y(t)\right|^2 + \left|A^{\frac{1}{4}} \varphi(t)\right|^2\right) dt \leq C_5(\|y_0\|_W^2 + \|\varphi_0\|_W^2). \tag{2.5}$$

Here $R_N = \left\| \begin{matrix} R_{11} & R_{12} \\ R_{12} & R_{22} \end{matrix} \right\|$, $(\cdot,\cdot)_\omega$ is the scalar product in $L^2(\omega)$ and $\|y\|_W = \left|A^{\frac{1}{4}}\right|$, $|\cdot|$ is the norm of H. Finally, $< \cdot, \cdot >$ is the scalar product of $H \times H$.

The idea of the proof already used in [2], [3] is in few words the following. One proves first that the linearization system associated with (1.1) is exponentially stabilizable (Lemma 1) and use this fact to construct a feedback controller R_N satisfying (2.1), (2.2) (Lemma 2). Finally, one proves that controller (2.3) exponentially stabilizes system (1.1).

3. Stabilization of the linear systems

We shall rewrite system (1.1) as

$$\begin{aligned}&y' + kAy - a\ell A\varphi - \ell dy + \ell b\varphi - \ell b\varphi^3 = mu\\&\varphi' + aA\varphi - b\varphi + dy + b\varphi^3 = 0 \quad \text{in } Q\\&y(0) = y_0,\ \varphi(0) = \varphi_0 \quad \text{in } \Omega.\end{aligned} \tag{3.1}$$

Equivalently,

$$\begin{aligned}&\frac{d}{dt}\begin{pmatrix} y \\ \varphi \end{pmatrix} + \mathcal{A}\begin{pmatrix} y \\ \varphi \end{pmatrix} + F\begin{pmatrix} y \\ \varphi \end{pmatrix} = \begin{pmatrix} mu \\ 0 \end{pmatrix}\\&\begin{pmatrix} y \\ \varphi \end{pmatrix}(0) = \begin{pmatrix} y_0 \\ \varphi_0 \end{pmatrix},\end{aligned} \tag{3.2}$$

where

$$\mathcal{A} = \left\| \begin{matrix} kA - \ell d & -a\ell A + \ell b \\ d & aA - b \end{matrix} \right\| \tag{3.3}$$

$$F\begin{pmatrix} y \\ \varphi \end{pmatrix} = \begin{pmatrix} -\ell b\varphi^3 \\ b\varphi^3 \end{pmatrix}. \tag{3.4}$$

Consider the linear control system

$$\begin{aligned}&\frac{d}{dt}\begin{pmatrix} y \\ \varphi \end{pmatrix} + \mathcal{A}\begin{pmatrix} y \\ \varphi \end{pmatrix} = \begin{pmatrix} mu \\ 0 \end{pmatrix}\\&\begin{pmatrix} y \\ \varphi \end{pmatrix}(0) = \begin{pmatrix} y_0 \\ \varphi_0 \end{pmatrix}\end{aligned} \tag{3.5}$$

i.e.,

$$\begin{aligned}&y' + kAy - a\ell A\varphi - \ell dy + \ell b\varphi = mu\\&\varphi' + aA\varphi - b\varphi + dy = 0\\&y(0) = y_0,\ \varphi(0) = \varphi_0.\end{aligned} \tag{3.6}$$

We set $X_N = \text{span}\{\psi_i\}_{i=1}^N$ and denote by P_N the projector on X_N. We set

$$\begin{aligned}&y = y_N + z_N,\ \varphi = \varphi_N + \zeta_N,\\&y_N = P_N y,\ \varphi_N = P_N\varphi,\ z_N = (I - P_N)y,\ \zeta_N = (I - P_N)\varphi\end{aligned}$$

and rewrite system (3.6) as

$$\begin{aligned}
&\dot{y}_N^j + (k\lambda_j - \ell d)y_N^j - (a\ell\lambda_j + \ell b)\varphi_N^j = (P_N(mu), \psi_j),\ j = 1, ..., N,\\
&\dot{\varphi}_N^j + dy_N^j + (a\lambda_j - b)\varphi_N^j = 0,\ j = 1, ..., N,
\end{aligned} \tag{3.7}$$

$$\begin{aligned}
&\dot{z}_N^j + (k\lambda_j - \ell d)z_N^j - (a\ell\lambda_j + \ell b)\zeta_N^j = ((I - P_N(mu), \psi_j),\ j = N+1, ...,\\
&\dot{\zeta}_N^j + dz_N^j + (a\lambda_j - b)\zeta_N^j = 0,\ j = N+1, ...,\\
&\zeta_N(0) = P_N y_0,\ \varphi_N(0) = P_N \varphi_0,\\
&z_N(0) = (I - P_N)y_0, \zeta_N(0) = (I - P_N)\varphi_0.
\end{aligned} \tag{3.8}$$

Here

$$y_N = \sum_{j=1}^{N} y_N^j \psi_j, \qquad \varphi_n = \sum_{j=1}^{N} \varphi_N^j \psi_j,$$

$$\zeta_N = \sum_{j=N+1}^{N} z_N^j \psi_j, \quad \zeta_N = \sum_{j=N+1}^{N} \zeta_N^j \psi_j.$$

Lemma 1. *There are* $y_j \in L^2(0, \infty)$*,* $j = 1, ..., N$*, such that for* N *large enough the controller*

$$u(x,t) = \sum_{j=1}^{N} u_j(t)\psi_j(x) \tag{3.9}$$

stabilizes exponentially system (3.6)*, i.e.,*

$$|y(t)| + |\varphi(t)| + |u_j(t)| \le Ce^{-\gamma t}(|y_0| + |\varphi_0|),\ \forall t \ge 0. \tag{3.10}$$

for some $\gamma > 0$.

Here $| \cdot |$ denotes the norm in $L^2(\Omega)$.

Proof. To prove Lemma 1 which is the main ingredient of the proof of Theorem 1 we shall prove first the exact null controllability of the finite dimensional system (3.7) for N large enough.

For u given by (3.9) system (3.7) becomes

$$\begin{aligned}
&\dot{y}_N^j + (k\lambda_j - \ell d)y_N^j - (a\ell\lambda_j + \ell b)\varphi_N^j = \sum_{i=1}^{N} u_i(t)(\psi_j, \psi_i)_\omega,\\
&\dot{\varphi}_N^j + dy_N^j + (a\lambda_j - b)\varphi_N^j = 0,\ j = 1, ..., N.
\end{aligned} \tag{3.11}$$

The dual system of (3.7) is the following

$$\begin{aligned}
&\dot{p}_N^j - (k\lambda_j - \ell d)p_N^j - dq_N^j = 0,\ j = 1, ..., N,\\
&\dot{q}_N^j + (a\ell\lambda_j + \ell b)p_N^j - (a\lambda_j - b)q_N^j = 0.
\end{aligned} \tag{3.12}$$

We set

$$B_N = \|(\psi_j, \psi_i)_\omega\|_{i,j=1}^N,$$

where $(\cdot,\cdot)_\omega$, is the scalar product in $L^2(\omega)$. Recall that (3.11) is null controllable on $[0,T]$ if and only if

$$B_N^*(p_N(t)) = 0,\ \forall t \in (0,T) \tag{3.13}$$

implies that

$$p_N(t) \equiv 0,\ q_N(t) \equiv 0. \tag{3.14}$$

By (3.13) we have

$$\sum_{j=1}^N (\psi_j, \psi_i)_\omega p_N^j(t) \equiv 0,\ i = 1, ..., N. \tag{3.15}$$

On the other hand, det $\|(\psi_j, \psi_i)_\omega\| = 0$ because otherwise system $\{\psi_j\}_{j=1}^N$ is dependent on ω and this implies by unique continuation that $\{\psi_j\}_{j=1}^N$ is linearly dependent on Ω. Hence $p_N \equiv 0$ and by (3.12) it follows that $q_N \equiv 0$. Hence system (3.11) is null controllable and this implies that there are $\{u_j\}_{j=1}^N$ (given in feedback form) such that system (3.11) is exponentially stable with arbitrary exponent γ, i.e.,

$$|y_N^j(t)| + |\varphi_N^j(t)| + |u_j(t)| \le Ce^{-\gamma t}|y_N^j(0)| + |\varphi_N^j(0)|,\ \forall t \ge 0. \tag{3.16}$$

Substituting (3.9) into (3.8) and taking in account (3.16) we get

$$\frac{1}{2}\frac{d}{dt}(|z_N^j|^2 + \alpha|\zeta_N^j|^2) + (k\lambda_j - \ell d)|z_N^j|^2 - (a\ell\lambda_j + \ell b)z_N^j \cdot \zeta_N^j +$$
$$+\alpha d z_N^j \zeta_N^j + \alpha(a\lambda_j - b)|\zeta_N^j|^2 = \alpha(I - P_N)(mu, \psi_j)\zeta_N^j,$$

where $\alpha > 0$ is arbitrary.

For α suitable chosen (for instance for $\alpha \ge 2a\ell$) and N large enough we see that

$$|z_N^j(t)|^2 + \alpha|\zeta_N^j(t)|^2 \le e^{-\gamma_N t}(|z_N^j(0)|^2 + |\zeta_N^j(0)|^2) +$$
$$+ \int_0^t e^{-\gamma_N(t-s)} \sum_{j=1}^N |u_j(s)|^2 ds \le Ce^{-\gamma_N^1 t}(|z_N^j(0)|^2 + |\zeta_N^j(0)|^2),\ \forall t \ge 0,$$

where $\gamma_N^1 > 0$.

This completes the proof. ■

Next consider the optimal control problem

$$\left.\begin{aligned} \text{Min}\left\{\frac{1}{2}\int_0^\infty\left(\left|A^{\frac{3}{4}}y(t)\right|^2+\left|A^{\frac{3}{4}}\varphi(t)\right|^2+|u(t)|^2\right)dt\right. \\ \text{subject to (3.7)},\ u=\sum_{j=1}^N u_j(t)\psi_j \end{aligned}\right\}=\Phi(y_0,\varphi_0). \tag{3.17}$$

It is readily seen that

$$\Phi(y_0,\varphi_0)\le C\left(\left|A^{\frac{3}{4}}y_0\right|^2+\left|A^{\frac{3}{4}}\varphi_0\right|^2\right)\ \forall y_0\in D\left(A^{\frac{1}{4}}\right),\ \varphi_0\in D\left(A^{\frac{1}{4}}\right). \tag{3.18}$$

Indeed, multiplying (3.6) by $A^{\frac{1}{2}}y$ and $\alpha A^{\frac{1}{2}}\varphi$, respectively, we get

$$\begin{aligned} \frac{1}{2}\ \frac{d}{dt}\left(\left|A^{\frac{1}{4}}y(t)\right|^2+\alpha\left|A^{\frac{1}{4}}\varphi(t)\right|^2\right)+k\left|A^{\frac{3}{4}}y(t)\right|^2-a\ell\left(A^{\frac{1}{2}}y(t),A\varphi(t)\right)- \\ -\ell d\left(y(t),A^{\frac{1}{2}}y(t)\right)+\ell b\left(\varphi(t),A^{\frac{1}{2}}y(t)\right)+ \\ +\alpha a\left|A^{\frac{3}{4}}\varphi(t)\right|^2+\alpha\left(A^{\frac{1}{2}}y(t),dy(t)-b\varphi(t)\right)=\left(mu(t),A^{\frac{1}{2}}y(t)\right). \end{aligned}$$

For α sufficiently large we get

$$\begin{aligned} \frac{d}{dt}\left(\left|A^{\frac{1}{4}}y(t)\right|^2+\alpha\left|A^{\frac{1}{4}}\varphi(t)\right|^2\right)+\delta\left(\left|A^{\frac{3}{4}}y(t)\right|^2+\left|A^{\frac{3}{4}}\varphi(t)\right|^2\right)\le \\ \le C\left(|y(t)|^2+|\varphi(t)|^2+|u(t)|^2\right),\ t>0. \end{aligned}$$

Integrating on $(0,\infty)$ and using Lemma 1 we get (3.18) as claimed.

Hence there is a symmetric continuous operator $R_N : W\times W\to W'\times W$ such that

$$\Phi(y_0,\varphi_0)=\frac{1}{2}\ < R_N(t_0,\varphi_0),(y_0,\varphi_0)>\ \forall(y_0,\varphi_0)\in W\times W. \tag{3.19}$$

We set

$$R_N=\left\|\begin{matrix} R_{11} & R_{12} \\ R_{12} & R_{22}\end{matrix}\right\|.$$

We have also

Lemma 2. *Let (y^*,φ^*,u^*) be optimal in (3.17). We have*

$$u_j^*(t)=-(R_{11}y(t)+R_{12}\varphi(t),\psi_j)_\omega,\ \forall t\ge 0,\ j=1,...,N. \tag{3.20}$$

Moreover,

$$|R_N(y,\varphi)|_{H\times H} \le C(\|y\| + \|\varphi\|),\ \forall (y,\varphi)\in V \tag{3.21}$$

and

$$< R_N(y,\varphi)(y,\varphi) > \ge \omega(|y|_W^2 + |\varphi|_W^2)\ \forall (y,\varphi)\in W\times W. \tag{3.22}$$

Finally, R_N is the solution to the Riccati equation

$$\begin{aligned} &< kAy - \ell aA\varphi - \ell dy + \ell b\varphi, R_{11}y + R_{12}\varphi > + \\ &+ < aA\varphi + dy - b\varphi, R_{12}y + R_{22}\varphi > + \\ &+\frac{1}{2}\left|\sum_{j=1}^{N}(R_{11}y + R_{12}\varphi, \psi_j)_\omega^2\right| = \\ &= \frac{1}{2}\left(\left|A^{\frac{3}{4}}y(t)\right|^2 + \left|A^{\frac{3}{4}}\varphi(t)\right|^2\right),\ \forall (y,\varphi)\in D(A)\times D(A). \end{aligned} \tag{3.23}$$

The proof of Lemma 2 is exactly the same as that given in [2], [3] and so it will be omitted.

4. Proof of Theorem 1

Consider the closed loop system

$$\begin{aligned} & y_t + kAy - \ell aA\varphi - \ell dy + \ell b\varphi - \ell b\varphi^3 + \\ & \qquad +m\sum_{i=1}^{N}(R_{11}y + R_{12}\varphi, \psi_j)_\omega \psi_j = 0 \\ & \varphi_t + aA\varphi - b\varphi + dy = 0, \quad t\ge 0, \\ & y(0) = y_0, \quad \varphi(0) = \varphi_0. \end{aligned} \tag{4.1}$$

It is easily seen that for each $(y_0,\varphi_0)\in H\times H$ this system has a unique solution $(y,\varphi)\in L^2(0;T;V)\times L^2(0,T;V)$. Multiplying first equation (4.1) by $R_{11}y + R_{12}\varphi$ the second by $(R_{12}y + R_{22}\varphi)$ and using (3.23) we obtain after some calculation

$$\frac{d}{dt} < R(y,\varphi),(y,\varphi) > + \left|A^{\frac{3}{4}}y(t)\right|^2 + \left|A^{\frac{3}{4}}\varphi(t)\right|^2 \le C|(R_{11}y + R_{12}\varphi,\varphi^3)|.$$

On the other hand, we have

$$\begin{aligned} |(R_{11}y + R_{12}\varphi,\varphi^3)| &\le C|R_{11}y + R_{12}\varphi|\ |\varphi|^3_{L^6(\Omega)} \le \\ &\le C(\|y\| + \|\varphi\|)|\varphi|^3_{L^6(\Omega)} \le \\ \le C\left(\left|A^{\frac{1}{4}}y\right|^{\frac{1}{2}} + \left|A^{\frac{3}{4}}y\right|^{\frac{1}{2}}|\varphi|^3_{L^6} + \left|A^{\frac{1}{4}}\varphi\right|^{\frac{1}{2}}\left|A^{\frac{3}{4}}\varphi\right|^{\frac{1}{2}}|\varphi|^3_{L^6}\right) &\le \end{aligned}$$

$$\leq C\left(\left|A^{\frac{1}{4}}y\right|^{\frac{1}{2}}\left|A^{\frac{3}{4}}y\right|^{\frac{1}{2}}\left|A^{\frac{1}{4}}\varphi\right|^{3}+\left|A^{\frac{1}{4}}\varphi\right|^{\frac{1}{2}}\left|A^{\frac{3}{4}}\varphi\right|^{\frac{1}{2}}\left|A^{\frac{1}{4}}\varphi\right|^{3}\right)\leq$$

$$\leq C\left(\left|A^{\frac{1}{4}}\varphi\right|^{3}\left|A^{\frac{3}{4}}y\right|+\left|A^{\frac{1}{4}}\varphi\right|\left|A^{\frac{1}{4}}\varphi\right|^{3}\right)\leq C\left|A^{\frac{1}{4}}\varphi\right|^{2}\left(\left|A^{\frac{3}{4}}y\right|^{2}+\left|A^{\frac{3}{4}}\varphi\right|^{2}\right)+$$

$$+C\left(\left|A^{\frac{3}{4}}\varphi\right|^{2}\left|A^{\frac{1}{4}}\varphi\right|^{2}\right)\leq C\Phi(y,\varphi)\left(\left|A^{\frac{1}{4}}y\right|^{2}+\left|A^{\frac{3}{4}}\varphi\right|^{2}\right).$$

Hence for $\Phi(y,\varphi)\leq\rho$ sufficiently small,

$$\frac{d}{dt} < R(y,\varphi),(y,\varphi)> +\left|A^{\frac{3}{4}}y\right|^{2}+\left|A^{\frac{3}{4}}\varphi\right|^{2}\leq 0.$$

Finally, if $< R(y_0,\varphi_0),(y_0,\varphi_0) > \leq\rho$ small enough we arrive to conclusion.

References

1 V. Barbu, Local controllability of the phase field system, *Nonlinear Analysis*, 50 (2002), 363–372.

2 V. Barbu, G. Wang, Internal stabilization of semilinear parabolic systems (to appear).

3 V. Barbu, R. Triggiani, Internal stabilization of Navier–Stokes with finite dimensional controllers (to appear).

4 A.V.Fursikov, O.Yu.Imanuilov, *Controllability of Evolution Equations*, Lectures Notes, 34(1996), Seoul University Press.

A SOLUTION TO THE FIXED END-POINT LINEAR QUADRATIC OPTIMAL PROBLEM

Corneliu Botan, Florin Ostafi and Alexandru Onea
"Gh. Asachi" Technical University of Iasi
Dept. of Automatic Control and Industrial Informatics
Email: { cbotan, fostafi, aonea}@ac.tuiasi.ro

Abstract A linear quadratic optimal problem with fixed end-point is studied for continuous and discrete case. The proposed solution is convenient for control law implementation. Some remarks referring to the existence of the solution are indicated.

Keywords: optimal control, linear quadratic, fixed end-point, continuous-time, discrete-time

1. Introduction

A completely controllable linear time invariant system is considered

$$\dot{x}(t) = Ax(t) + Bu(t)\,, \tag{1}$$

where $x \in R^n$ is the state vector, $u \in R^m$ is the control vector and A and B are matrices of the appropriate dimensions.

The optimal control problem refers to the criterion

$$J = \frac{1}{2}\int_{t_0}^{t_f} [x^T(t)Qx(t) + u^T(t)Pu(t)]dt\,, \tag{2}$$

$Q = Q^T \geq 0$, $P = P^T > 0$ (the symbol T denotes the transposition).

The problem is to find the control u(t) that transfer the system (1) from the initial state $x(t_0) = x^0$ to a given final state $x(t_f) = x^f$ (Anderson and Moore, 1991; Athans and Falb, 1966). The usual case that will be considered is $x(t_f) = 0$. The approach for a more general case, when the target set is $Cx(t_f) = 0$, $C \in R^{p\times n}$ is similar.

A similar problem can be formulated for the discrete-time case (Kuo, 1992), referring to the system

$$x(k+1) = Ax(k) + Bu(k) \tag{3}$$

and to the criterion

$$J = \frac{\tau}{2}\sum_{k=k_0}^{k_f-1} x^T(k)Qx(k) + u^T(k)Pu(k)\,. \tag{4}$$

In (3) and (4), x(k) and u(k) denote the vectors x and u at the discrete moment $k\tau$, $k \in Z$, and τ is the sampling period (we shall consider τ=1). Equation (3) can be obtained via discretization of the equation (1). It is preferred the same notations for matrices although they have, of course, different values. Since the system is time invariant, we may consider t_0=0 and k_0=0.

The argument *t* or *k* will be omitted in the following relations if they are similar for both continuous and discrete case.

From the Hamiltonian conditions one obtains

$$u(t) = -P^{-1}B^T\lambda(t)\,,\ \lambda \in R^n\,, \tag{5}$$

$$\begin{aligned} \dot{x}(t) &= Ax(t) - N\lambda(t), \qquad N = BP^{-1}B^T \\ \dot{\lambda}(t) &= -Qx(t) - A^T\lambda(t) \end{aligned} \tag{6}$$

for the continuous case and

$$u(k) = -P^{-1}B^T\lambda(k+1) \tag{7}$$

$$\begin{aligned} x(k+1) &= Ax(k) - N\lambda(k+1) \\ \lambda(k) &= Qx(k) + A^T\lambda(k+1) \end{aligned} \tag{8}$$

for the discrete case.

If the 2n-order vector $\gamma = \begin{bmatrix} x \\ \lambda \end{bmatrix}$ is introduced, the equations (6) and (8) can be written in the form

$$\dot{\gamma}(t) = G_c\gamma(t) \tag{9}$$

and

$$\gamma(k+1) = G_d\gamma(k)\,, \tag{10}$$

respectively. In the above relations

$$G_c = \begin{bmatrix} A & -N \\ -Q & -A^T \end{bmatrix},\ G_d = \begin{bmatrix} A + NA^{-T}Q & -NA^{-T} \\ -A^{-T}Q & A^{-T} \end{bmatrix}, \tag{11}$$

where $A^{-T} = (A^{-1})^T$.

The solution to (9)/(10) can be expressed as

$$\gamma(.) = \Gamma(.)\gamma^0 , \tag{12}$$

where γ^0 is the vector γ at the initial instant, and

$$\Gamma(.) = \begin{bmatrix} \Gamma_{11}(.) & \Gamma_{12}(.) \\ \Gamma_{21}(.) & \Gamma_{22}(.) \end{bmatrix} \in R^{2nx2n}, \quad \Gamma_{ij} \in R^{nxn}, \quad i,j = 1,2, \tag{13}$$

is the transition matrix for G.

It is possible to compute $\gamma(.)$ from (12) only if the initial value of the co-state vector λ is established. For this purpose, we explicit the vector x from (12)

$$x(.) = \Gamma_{11}(.)x^0 + \Gamma_{12}(.)\lambda^0 . \tag{14}$$

Since $x^f = 0$, one obtains

$$\lambda^0 = -\Gamma_{12f}^{-1}\Gamma_{11f}x^0 , \tag{15}$$

where $\Gamma_{11f} = \Gamma_{11}(t_f,0)$, $\Gamma_{12f} = \Gamma_{12}(t_f,0)$ for the continuous case and $\Gamma_{11f} = \Gamma_{11}(k_f)$, $\Gamma_{12f} = \Gamma_{12}(k_f)$ for the discrete case.

This solution implies that Γ_{12f} is a nonsingular matrix. The conditions for non-singularity of this matrix will be discussed below.

Now, the system (9)/(10) can be solved and the optimal trajectory (14) is in the two cases:

$$x(t) = [\Gamma_{11}(t,0) - \Gamma_{12}(t,0)\Gamma_{12f}^{-1}\Gamma_{11f}]x^0 , \tag{16}$$

$$x(k) = [\Gamma_{11}(k) - \Gamma_{12}(k)\Gamma_{12f}^{-1}\Gamma_{11f}]x^0 . \tag{17}$$

From (12) and (15) it also follows

$$\lambda(t) = [\Gamma_{21}(t,0) - \Gamma_{22}(t,0)\Gamma_{12f}^{-1}\Gamma_{11f}]x^0 , \tag{18}$$

$$\lambda(k) = [\Gamma_{21}(k) - \Gamma_{22}(k)\Gamma_{12f}^{-1}\Gamma_{11f}]x^0 . \tag{19}$$

The optimal control for continuous case is obtained replacing (18) in (5). For the discrete case, it has to express $\lambda(k+1)$ from (8) and then use (19) and (7):

$$u(k) = P^{-1}B^TA^{-T}Qx(k) - P^{-1}B^TA^{-T}[\Gamma_{21}(k) - \Gamma_{22}(k)\Gamma_{21f}^{-1}\Gamma_{11f}]x^0 . \tag{20}$$

The control vector u(k) can be computed with (20) or replacing x(k) in terms of x^0 from (17). In the both cases only the open loop control is obtained. In order to obtain the closed loop control, the vector x^0 is replaced in (20) from (17). But this approach implies a considerable increase of the computing difficulties, because it has to compute in real time the inverse of a time variant matrix. A similar situation also appears in the continuous case.

The method presented below avoids these difficulties.

In the sequel, it will be presented only the proof for the discrete time case. For the continuous time case only the final results will be indicated. Some of these results for continuous case are indicated in (Botan and Onea, 1999).

2. Main results

2.1. Basic relations

The main idea of the method is to perform a change of variable, so that one of the *nxn* matriceal blocks of the system matrix to be a null matrix:

$$\gamma(.) = U\rho(.), \quad \rho(.) = \begin{bmatrix} x(.) \\ v(.) \end{bmatrix} \tag{21}$$

with

$$U = \begin{bmatrix} I & 0 \\ R & I \end{bmatrix} \in R^{2nx2n}, \quad U^{-1} = \begin{bmatrix} I & 0 \\ -R & I \end{bmatrix}. \tag{22}$$

I is the *n* x *n* identity matrix and R is a *n* x *n* constant positive defined matrix.

From (21) and (22) results

$$\lambda(.) = Rx(.) + v(.). \tag{23}$$

The equation for the new variables in the discrete time case is

$$\rho(k+1) = H\rho(k), \tag{24}$$

where

$$H = U^{-1}G_d U = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix} \in R^{2nx2n}, \qquad H_{ij} \in R^{nxn}, \qquad i,j = 1,2, \tag{25}$$

Matrix H_{21} is

$$H_{21} = -RG_{11} - RG_{12}R + G_{22}R + G_{21},$$

where $G_{ij}, i = 1,2$ are the *n* x *n* matriceal blocks of G_d. Using (11), one obtains

$$H_{21} = (I + RN)A^{-T}R - (I + RN)A^{-T}Q - RA$$

or

$$H_{21} = (I + RN)A^{-T}[R - Q - A^{T}(I + RN)^{-1}RA].$$

If we impose

$$R = Q + A^T(I + RN)^{-1}RA \tag{26}$$

then

$$H_{21} = 0\,. \tag{27}$$

Note that the matrix I+RN is non-singular and also that all the inverse matrices which appear in the following relations exist.

One remark that (26) is the discrete Riccati equation for the LQ problem with infinite final time.

Using (26), from (11) and (25) it also follows

$$H_{11} = (I + NR)^{-1}A;\;\; H_{12} = G_{12} = -NA^{-T};\;\; H_{22} = H_{11}^{-T} = (I + RN)A^{-T}\,. \tag{28}$$

The transition matrix for H is $\Omega(k) = H^k \in R^{2n\mathrm{x}2n}$. One obtains for $\Omega(k)$ a similar form as for H

$$\Omega(k) = \begin{bmatrix} \Omega_{11}(k) & \Omega_{12}(k) \\ 0 & \Omega_{22}(k) \end{bmatrix} \tag{29}$$

and

$$\Omega_{11}(k) = H_{11}^k, \quad \Omega_{22}(k) = H_{22}^k, \quad \Omega_{12}(k) = H_{12k} = \sum_{i=0}^{k-1} H_{11}^i H_{12} H_{22}^{k-i-1}\,. \tag{30}$$

The transition matrix $\Gamma(k)$ can be expressed in terms of the transition matrix $\Omega(k)$ taking into account (22), (25) and (29) and has the form:

$$\Gamma(k) = U\Omega(k)U^{-1} = \begin{bmatrix} \Omega_{11}(k) - \Omega_{12}(k)R & \Omega_{12}(k) \\ R\Omega_{11}(k) - R\Omega_{12}(k)R - \Omega_{22}(k)R & R\Omega_{12}(k) + \Omega_{22}(k) \end{bmatrix}. \tag{31}$$

The solution of system (24) is

$$\begin{aligned} x(k) &= \Omega_{11}(k)x^0 + \Omega_{12}(k)v^0 \\ v(k) &= \Omega_{22}(k)v^0. \end{aligned} \tag{32}$$

The initial vector v^0=v(0) results from (15) and (23) for k=0 and it is

$$v^0 = -(\Gamma_{12f}^{-1}\Gamma_{11f} + R)x^0\,. \tag{33}$$

Substituting (30), (31) into (33) it results

$$v^0 = -\Omega_{12f}^{-1}\Omega_{11f}x^0 = -H_{12k_f}^{-1}H_{11}^{k_f}x^0\,. \tag{34}$$

The optimal control is obtained from (7), (8) and (23) as

$$u(k) = -P^{-1}B^TA^{-T}(R - Q)x(k) - P^{-1}B^TA^{-T}v(k)\,.$$

Replacing (32) and (34) it results

$$u(k) = -P^{-1}B^T A^{-T}(R - Q)x(k) + P^{-1}B^T A^{-T} H_{22}^{k} H_{12k_f}^{-1} H_{11}^{k_f} x^0 . \quad (35)$$

One can remark that the optimal control can be expressed as

$$u(k) = u_f(k) + u_s(k) , \quad (36)$$

where

$$u_f(k) = -P^{-1}B^T A^{-T}(R - Q)x(k) \quad (37)$$

is the feedback component and

$$u_s(k) = P^{-1}B^T A^{-T} H_{22}^{k} H_{12k_f}^{-1} H_{11}^{k_f} x^0 \quad (38)$$

is a supplementary component that depends on the initial state x^0=x(0).

Note that in (37) only the term $u(k) = -P^{-1}B^T A^{-T}Rx(k)$ is a proper feedback component and it is identical with the control vector obtained in the LQ problem with infinite final time.

Therefore the real time computing of the optimal control u(k) implies to establish a usual state feedback component $u_f(k)$ and a supplementary component $u_s(k)$. The last one contains only one time variant element: the transition matrix H_{22}^{k}, which evidently can be recursively computed.

For the continuous time case, the transformed system can be written as

$$\dot{\rho}(t) = H\rho(t) , \quad (39)$$

with

$$H = U^{-1}G_c U = \begin{bmatrix} F & -N \\ 0 & -F^T \end{bmatrix} , \quad (40)$$

where

$$F = A - NR . \quad (41)$$

The form (40) is obtained if we impose

$$RNR - RA - A^T R - Q = 0 , \quad (42)$$

that is R satisfies the Riccati algebraic equation that appears in the continuous LQ problem with infinite final time.

The transition matrix corresponding to H is

$$\Omega(t,\tau) = \begin{bmatrix} \Psi(t,\tau) & \Omega_{12}(t,\tau) \\ 0 & \phi(t,\tau) \end{bmatrix} , \quad (43)$$

where $\Psi(.)$ and $\phi(.)$ are the transition matrices for F and $-F^T$, respectively and

$$\Omega_{12}(t,\tau) = \int_t^{\tau} \Psi(t,\theta) N \phi(\theta,\tau) d\theta . \quad (44)$$

Since $\Gamma(.) = U\Omega(.)U^{-1}$, one obtains

$$\Gamma(.) = \begin{bmatrix} \Psi(.) - \Omega_{12}(.)R & \Omega_{12}(.) \\ R\Psi(.) - R\Omega_{12}(.)R - \phi(.)R & \phi(.) + R\Omega_{12}(.) \end{bmatrix}. \tag{45}$$

The solution to the system (39) is

$$\begin{bmatrix} x(t) \\ v(t) \end{bmatrix} = \Omega(t,0) \begin{bmatrix} x^0 \\ v^0 \end{bmatrix}, \tag{46}$$

where

$$v^0 = -\Omega_{12f}^{-1}\Omega_{11f}x^0 = -\Omega_{12f}^{-1}\Psi_f x^0, \quad \Psi_f = \Psi(t_f,0) \tag{47}$$

with

$$\Omega_{12f} = \Omega_{12}(t_f,0) \text{ and } \Omega_{11f} = \Omega_{11}(t_f,0) = \Psi(t_f,0).$$

From (46) and (47) it follows

$$\begin{aligned} x(t) &= [\Psi(t,0) + \Omega_{12}(t,0)\Omega_{12f}^{-1}\Psi_f]x^0 \\ v(t) &= -\phi(t,0)\Omega_{12f}^{-1}\Psi_f x^0. \end{aligned} \tag{48}$$

The optimal control u(t) is

$$u(t) = u_f(t) + u_s(t), \tag{49}$$

where the feedback component is

$$u_f(t) = -P^{-1}B^T R x(t) \tag{50}$$

and the supplementary component is

$$u_s(t) = -P^{-1}B^T\phi(t,0)\Omega_{12f}^{-1}\Psi_f x^0 \tag{51}$$

and depends on the initial state x^0.

As in the discrete case, the last component contains only one time variant element, namely the transition matrix $\phi(t,0)$; this matrix can be recursively computed.

Remark 1. It is usually desired to maintain x(t) = 0 for t > t_f. For this purpose it is necessary to adopt u(t) = 0 for t > t_f.□

Remark 2. The performed simulation tests have indicated that a significant increase of the sampling period only for the supplementary component leads to a small difference in the system behavior. This aspect is important because offers the possibility of the decrease of the real time computing volume. □

2.2 The existence of the solution

As results from the previous relations, the problem has solution if the matrix $\Gamma_{12f} = \Omega_{12f}$ is non-singular. Indeed, one remarks from (15) that there is a unique initial vector $\lambda^0 = \lambda(0)$ for a given $x^0 = x(0)$ if the matrix $\Gamma_{12f} = \Gamma_{12}(t_f, 0)$ is non-singular and then the formulated problem has solution. The condition for non-singularity of the matrix Γ_{12f} is given by the following

Theorem *If the pair* (A,B) *is completely controllable, then matrix* Γ_{12f} *is non-singular.*

Proof. For the continuous time case, from (31) and (44) one obtains

$$\Gamma_{12}(t_f, 0) = \Omega_{12}(t_f, 0) = -\int_0^{t_f} e^{F(t_f-\theta)} BP^{-1}B^T e^{F^T(t_f-\theta)} d\theta \phi(t_f, 0) . \quad (52)$$

Since the transition matrix $\phi(t_f, 0)$ is non-singular, $\Gamma_{12}(t_f, 0)$ is non-singular if the matrix $\Pi(t_f, 0) = \int_0^{t_f} e^{F\sigma} BP^{-1}B^T e^{F^T\sigma} d\sigma$ is non-singular. One can prove (Botan, 1991) that $\Pi(t_f, 0) > 0$ if (A,B) is completely controllable and thus the theorem is proved.□

For the discrete time case, the proof is similar, but some supplementary transformations are necessary because the matriceal block $H_{12} = -NA^{-T} \in R^n$ of the matrix $H \in R^{2nx2n}$ contains the factor A^{-T}.

Firstly, we will establish another expression for the matriceal blocks H_{ij}, i,j=1,2, given by (28). From (28) one obtains

$$H_{11}^{-1} = A^{-1}(R^{-1} + N)R . \quad (53)$$

Let us denote the matrix

$$X^{-1} = R^{-1} + N = R^{-1} + BP^{-1}B^T . \quad (54)$$

Multiplying (54) with X and then with R, it results

$$R = X + XBP^{-1}B^T R . \quad (55)$$

Multiplying (55) with B, one obtains

$$RB = XBP^{-1}(P + B^T RB) .$$

This relation is multiplied with $(P + B^T RB)^{-1}$, then with $B^T R$ and then we subtract from R:

$$R - RB(P + B^T RB)^{-1} B^T R = R - XBP^{-1}B^T R . \quad (56)$$

Let us denote

$$\bar{P} = P + B^T RB$$
$$\bar{N} = B\bar{P}^{-1}B^T . \tag{57}$$

Using these notations and (55), relation (56) becomes

$$X = R - R\bar{N}R . \tag{58}$$

From (54) and (58) one obtains

$$(R^{-1} + N)^{-1} = R - R\bar{N}R \tag{59}$$

and then

$$H_{11} = R^{-1}(R^{-1} + N)^{-1} A = (I - \bar{N}R)A \tag{60}$$

or

$$H_{11} = A + BK = F , \; K = -\bar{P}^{-1}B^T RA . \tag{61}$$

From (28) one obtains

$$H_{22} = H_{11}^{-T} = F^{-T} , \tag{62}$$

$$H_{12} = -NA^{-T} = -NA^{-T}F^T F^{-T} = -NA^{-T}A^T(I - R\bar{N})F^{-T} = -(N - NR\bar{N})F^{-T} .$$

But from (59) it follows $N - NR\bar{N} = \bar{N}$, so that

$$H_{12} = -\bar{N}F^{-T} . \tag{63}$$

Having in view (61), (62) and (63), the matrix H can be written in the form

$$H = \begin{bmatrix} F & -\bar{N}F^{-T} \\ 0 & F^{-T} \end{bmatrix} .$$

Now we introduce the nonsingular matrix

$$\chi = \begin{bmatrix} I & 0 \\ 0 & F^{-T} \end{bmatrix}$$

and carry out the transformation

$$D = \chi H \chi^{-1} = \begin{bmatrix} F & -\bar{N} \\ 0 & F^{-T} \end{bmatrix}$$

and analogous for the corresponding transition matrix

$$\Delta(k) = \chi \Omega(k) \chi^{-1} = \begin{bmatrix} \Delta_{11k} & \Delta_{12k} \\ 0 & \Delta_{2k} \end{bmatrix} , \tag{64}$$

$$\Delta_{11k} = F^k,\ \Delta_{22k} = (F^{-T})^k,\ \Delta_{12k} = \sum_{i=0}^{k-1} D_{11}^i D_{12} D_{22}^{k-i-1}.$$

For $k=k_f$

$$\Delta_{12f} = \Delta_{12}(k_f) = \sum_{i=0}^{k_f-1} D_{11}^i D_{12} D_{22}^{-i} D_{22}^{k_f-1} =$$

$$= -\sum_{i=0}^{k_f-1} F^i \overline{N}(F^T)^i (F^{-T})^{k_f-1} = -\Pi(F^{-T})^{k_f-1}, \qquad (65)$$

where

$$\Pi = \sum_{i=0}^{k_f-1} F^i B \overline{P}^{-1} B^T (F^i)^T.$$

We shall prove now that Π is a positive defined matrix ($\Pi>0$) if the pair (A,B) is completely controllable. Indeed, if (A,B) is controllable, (F,B) is controllable, with F=A+BK.

Since P>0 and R>0, $\overline{P}$ given by (57) is also positive defined, and also $\overline{P}^{-1} > 0$; in this case, there is a unique positive defined matrix V such as $VV^T = \overline{P}^{-1}$.

We can express

$$\Pi = \sum_{i=0}^{k_f-1} F^i BV(BV)^T (F^i)^T$$

$$[BV \quad FBV \quad \ldots \quad F^{n-1}BV] = [B \quad FB \quad \ldots \quad F^{n-1}B] \begin{bmatrix} V & & \\ & \ldots & \\ & & V \end{bmatrix}.$$

Since the last matrix is nonsingular and $[B \quad FB \quad \ldots \quad F^{n-1}B]$ is of rank n, the matrix $[BV \quad FBV \quad \ldots \quad F^{n-1}BV]$ is of the same rank, thus the pair (F,BV) is completely controllable and the matrix Π is positive defined. Since F^{-T} is nonsingular, from (65) it follows that Δ_{12f} is nonsingular.

From the transformation (64) we obtain $\Delta_{12f} = \Omega_{12f} F^T$. Since Δ_{12f} and F^T are nonsingular, Ω_{12f} is nonsingular and $\Gamma_{12f} = \Omega_{12f}$ is nonsingular and the theorem is proved also for the discrete case. ■

3. Simulation results

Some simulation tests were performed for different conditions and different weight matrices in the criterion for both continuous and discrete case. The results presented bellow refer to a continuous system (1) with

$$A = \begin{bmatrix} 0 & 20 \\ -3.5 & -19 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 6 \end{bmatrix}$$

and to the criterion (2) with Q = diag (1, 3) and P=p=0.7.

The corresponding matrices obtained via discretization were adopted for the discrete case. The sampling period is τ=0.002 s. The terminal moments are t_0=0 and t_f=0.3 s (and corresponding k_f=150). The initial state is $x(t_0) = [50 \;\; 0]^T$.

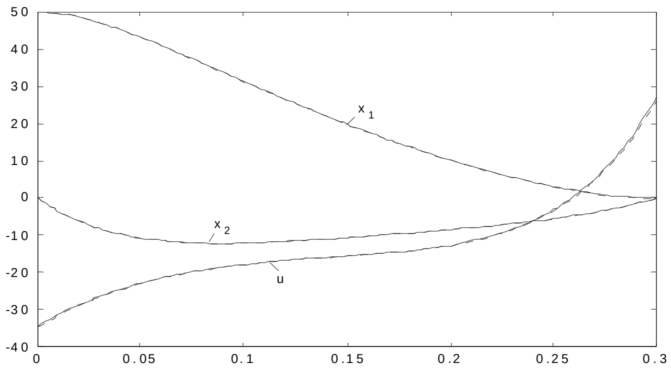


Figure 1. The behavior of the system for continuous and discrete cases

Figure 1 presents the behavior of the system for both continuous and discrete cases; one can remark that the curves are practically the same in the two cases. The more significant differences between continuous and discrete case appears if the sampling period is increased. Figure 2 presents the same situation for the discrete case, but indicates in addition the behavior for $k>k_f$, when the control u(k)=0 is adopted (see Remark 1).

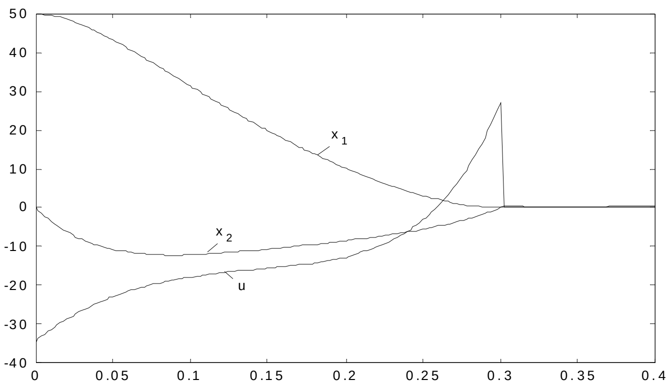


Figure 2. The behavior of a system before and after the final time

Figure 3 presents a comparison between the basic discrete case and the case when the sampling period τ was increased by ten times only for the supplementary component (see Remark 2); one can remark that the differences are not significant, especially referring to the states variables.

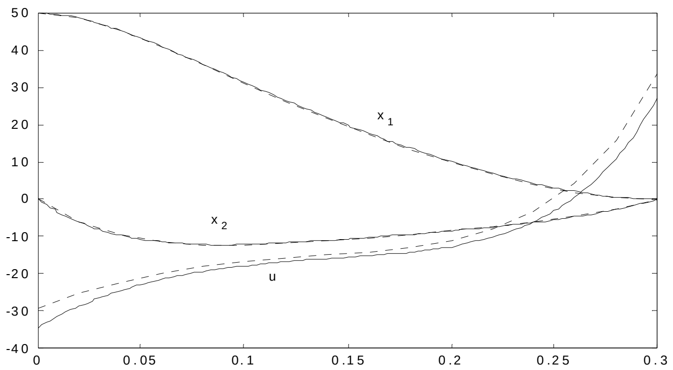


Figure 3. The effect of the sampling period increase for the supplementary component

It was performed a comparison with the similar linear quadratic problem but with free end-point. The difference is that in the last case the state vector does not arrive in zero at the final time, but the control variable is significantly smaller.

4. Conclusions

The linear quadratic optimal problem for continuous and for discrete case is studied; the results are presented especially for the discrete case.

A new method is presented and it is obtained a very convenient form for the feedback optimal control law.

Some considerations about the existence of the solution are presented.

The theoretical and simulation results indicate a similarity between the continuous and discrete time cases.

References

Anderson, B.D.O., and J.B. Moore (1990). *Optimal Control*, Prentice-Hall, New Jersey

Athans, M, and P.L. Falb (1966). *Optimal Control*, McGraw Hill, New York.

Botan, C. (1991) On the Fixed End-Point Linear Quadratic Optimal Problem. In: *Proceedings of the 8th International Conf. on Control Systems and Computer Science.* Polytechnic Inst. of Bucharest, pp 100-105, Bucharest.

Botan, C. and A. Onea (1999) Fixed End-Point Problem for an Electrical Drive System. In: *Proceedings of the IEEE International Symposium on Industrial Electronics, ISIE'99, Bled,* Slovenia, Vol. 3, pp. 1345-1349.

Kuo, B.C. (1992) *Digital Control System*, Saunders College Publishing, Philadelphia.

PATTERN RECOGNITION CONTROL SYSTEMS – A DISTINCT DIRECTION IN INTELLIGENT CONTROL

Emil Ceangă and Laurenţiu Frangu

University "Dunărea de Jos" of Galaţi

Str. Domnească 111, 2200 Galaţi, Romania

E-mail: Emil.Ceanga@ugal.ro

Abstract This paper presents an alternative approach of intelligent control: the pattern recognizing systems. The main idea is to endow the control system with the ability to learn from the experience generated by the interaction with the environment (the control object and the external world). Learning implies generalization and abstraction, through recognition of synthetic entities, which concentrate the essence of the past experience. The notions allowing learning, used in this approach, are called "control situations". The historical evolution of these notions is briefly exposed and they are analytically defined. Some results of the authors, presented in the paper, concern: the usefulness of each control situation in the hierarchical structure of intelligent control, the properties of the clusters, the learning automaton and the connections with other control techniques. For illustrating the approach, some applications of the pattern recognition control systems are presented.

Keywords: intelligent control, pattern recognition, control situations, hierarchy, adaptive control, strategy

1. Introduction

A fundamental property of the intelligent control systems is their ability to extract, through learning, the relevant information from the environment. This action implies generalization and abstraction, which are frequently performed through pattern recognition (PR).

The use of the PR methods, in the space of the observations on the controlled process, started at the end of the '50s [Widrow, 1962, 1964]. A remarkable application of this period was the pattern recognition control system (PRCS) for an inverted pendulum; the neural learning automaton was built using the technology of those years (memistors). PR techniques for control purposes were also used in [Taylor, 1963], [Waltz and Fu, 1966]. In the paper [Nikolic and Fu, 1966], a milestone in the field, the control of a dynamic system with unknown properties is performed by a PRCS. Learning is driven by an uncertain teacher, who learns simultaneously with the PR controller. The paper presents the first theoretical and qualitative results, regarding the convergence of the learning processes.

In some papers of the '60s, the generic term of "situation" was used instead of "pattern", having the meaning of an abstract entity, relevant for the control and diagnosis of the systems. The term was mainly adopted by the researchers in the field of automatic control, such as Aizerman, in the papers that theoretically founded the method of potential functions [Aizerman et al., 1965, 1966]. A refinement of this concept appears in the papers [Drăgan and Ceangă, 1968], [Ceangă, 1969a, 1969b], [Ceangă et al., 1971a, 1971b].

A paper that had an important influence on the evolution of the intelligent control is [Saridis, 1979]. Saridis revealed that intelligent systems are able to perform behavioral learning, i.e. they classify the information and take decisions through PR. This means that the intelligent systems perform generalization, through learning, in order to recognize some synthetic concepts, concerning the environment they are interacting with. The detailed description of such synthetic concepts, referred to with the generic term of "control situations", was approached in [Ceangă et al., 1981, 1984, 1985a, 1985b, 1991]. Recent papers, like [Seem and Nesler, 1996], [Ronco and Gawthrop, 1997a, 1997b], [Grigore, 2000], [Frangu, 2001], make use of the PRCS, in neural implementation. In general, the present approaches in intelligent control (including the PRCS approach) aim at the analogy with the human mind: [Frangu, 2001], [Truță, 2002]. In the above mentioned papers, the PR techniques are used to form abstract concepts, hierarchically structured, according to the principle "Increasing Precision on Decreasing Intelligence" (IPDI) [Saridis, 1988, 1989].

The purpose of this paper is to present some recent results of the authors in the field of PRCS. The connections with previous results are also mentioned, in order to highlight algorithms and techniques that maintain their up-to-dateness. By structuring in a hierarchy the concepts of control situations, it will become clearer how PRCS perform the essential functions of the intelligent systems. According to Albus ([Albus, 1991]), these are: perception, model of the world, value judgment and behaviour generation. Some unsolved aspects will also be presented; in the opinion of the authors, they are important for the evolution of the field.

2. Control situation approaches

Let us consider a control structure, such as that presented in figure 1, where the controlled process is a sampled dynamic system, with unknown or partially known properties. The particular variants of the generic concept of "Control Situation" depend on the function performed by the PRCS. They are described in the sequel.

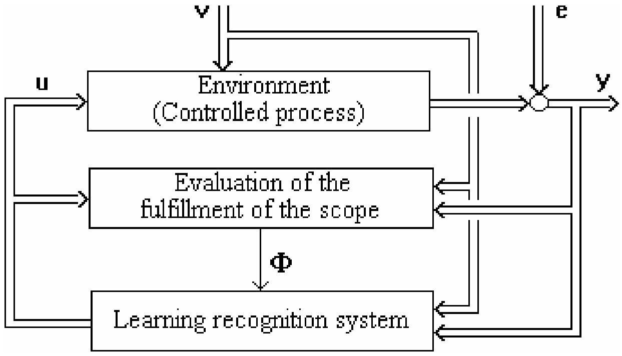


Figure 1. Control of the process, through the learning PR system

1. Output situation. Let y_i, i=1,...,p, $y \in Y_d$, be the discrete values of the output variables, out of the set Y_d of admissible values, which represent significant effects of the input variables v (measurable disturbance) and u (command). It is called ***output situation*** the set S_i of the variables observed from the "environment", w, that determine the discrete value y_i of the output, according to f, the causal input-output relationship of the controlled process:

$$S_i = \{w;\ w \xrightarrow{f} y_i;\ y_i \in Y_d\}\ . \qquad (1)$$

Let

$$w(t-1) = [y(t-1),..., y(t-n_a), u(t-k-1),..., u(t-k-n_b),$$
$$v(t-k-1),..., v(t-k-n_c)]^T \qquad (2)$$

be the vector of the observations from the environment, which is used to predict the value $y(t)$, where k is the dead time. The output situation S_i is:

$$S_i = \{w(t-1);\ \underset{j}{Min}\left\|y[w(t-1)] - y_j(t)\right\| = \left\|y[w(t-1) - y_i(t)]\right\|;\ y_j(t) \in Y_d\}\ .(3)$$

The membership of the vectors, with respect to the output situations, is given by the data recorded in the process (the discrete value of the output, at the moment t). For this reason, a supervised learning is possible, requiring no human teacher. The recognition of the output situations may be used for the predictive evaluation of the effects of the commands addressed to the controlled process.

2. Command situation. Let us assume there are p admissible values of the command input. They form the set U_d of the discrete admissible commands: u_i, i=1,...,p. Using the output, the command and the measurable disturbances, the vector of observations is defined:

$$z(t)=[y(t),...,y(t-n_a+1),u(t-k-1),...,u(t-k-n_b),v(t-k),...,v(t-k-n_c)]^T, \quad (4)$$

where n_a, n_b and n_c are finite integers and k is the dead time (expressed in sampling periods). The learning system has to make use of the "experience" accumulated up to the moment t, with the purpose of determining the function that assigns to any vector of observations, $z(t)$, the discrete command, $u_i(t)$, i=1,...,p, that maximizes the indicator Φ. It is called ***command situation*** the set S_i of the vectors z, for which the discrete command u_i is optimum, regarding the fulfillment of the objective:

$$S_i=\{z:\operatorname*{Max}_j \Phi(u_j,z)=\Phi(u_i,z),u_j\in U_d\}. \quad (5)$$

Consequently, to determine the control law means to deduce, through learning, the discriminant functions of the command situations S_i. The essential problem is to build the teacher of the PR controller. In some cases, the teacher may be a human expert or a decision system, based on pre-existing control systems. However, in the general case, the teacher can be a predictor that recognizes output situations.

Teaching the teacher and teaching the controller are performed simultaneously, in the frame of a dual control procedure. This one will be presented with the assumption that k=0, in order to simplify the expressions. At the current moment, j, the following operations are performed:

a – The generation of a first approximation of the current command by recognition of the command situation. The reference for the next moment, j+1, is already known: $y_r(j+1)$. The vector

$$z(j)=[y_r(j+1),y(j),...,y(j-n_a+1),u(j-1),...,u(j-n_b),v(j),...,v(j-n_c)]^T \quad (6)$$

will be assigned (by recognition) to one of the classes S_i^c, $i=1,...,p_c$. Let $u(j)=u_s$, $u_s\in U_d$, be the discrete command associated to the recognized command situation. The recognition controller did not yet learn, so the chosen command u_s is considered to be the answer of the "student" that has to be compared to that of the "teacher".

b – Based on the experience accumulated up to the current moment, j, the predictor recognizes output situations. It assigns the vectors

$$w_i(j)=[y(j),...,y(j-n_a+1),u_i,u(j-1),...,u(j-n_b),v(j),...,v(j-n_c)]^T \quad (7)$$

($i=1,...,p_c$) to the classes S_k^o, $k=1,...,p_o$, corresponding to the discrete values y_k of the prediction $\hat{y}(j+1)$. Making use of these predictions, the answer of the teacher is determined, according to the decision rule:

$$\{\underset{i}{Min}\| y(w_i(j))-y_r(j+1)\|=\| y(w_m(j))-y_r(j+1)\|\}\Rightarrow u(j)=u_m . \quad (8)$$

c – The responses of the PR controller ($z(j)\in S_s^c$, that is $u(j)=u_s$) and of the teacher ($u(j)=u_m$) are used for enriching the instruction set of the controller and for teaching it.

d – The command $u(j)=u_m$ is effectively applied to the process and the response $y(j+1)$ is recorded. This response is used to enrich the instruction set of the teacher, for predicting the output situations.

The PRCS can be applied to controlled processes having unknown, nonlinear (possible variant) dynamics, whose objective is to minimize a quadratic criterion ([Ceangă et al., 1984, 1985a]. It can have the expression:

$$J=E\{(y(t+k)-y_r(k))^2+\rho u^2(t)/t\}, \quad (9)$$

where ρ is a weighting factor for the command effort. As previously, the block for the evaluation of the objective fulfillment (see figure 1) contains a PR predictor. This one recognizes the output situations and teaches the learning controller, that recognizes the command situations.

3. Adaptation situation. Let us consider an environment that changes its properties slowly. The fulfillment of the control objective requires the use of a control law

$$u=\Psi(z(t)), \quad (10)$$

where $z(t)$ is defined by (4). The command (10) corresponds to particular properties of the controlled process. When the environment evolves, it is necessary to adapt the control law, by a finite number of adjustments, $A_i[\Psi(z)]$. The efficiency of the control law is determined by a set of measurable variables that form the "influence" vector, q. It is called ***adaptation situation*** ([Ceangă et al., 1991]) the set S_i of vectors q, corresponding to the particular adjustment of the control law, $A_i(.)$, which allows maintaining the control efficiency:

$$S_i=\{q:\underset{k}{Max}\,\Phi[A_k(z),q]=\Phi[A_i(z),q]\}. \quad (11)$$

Adaptation through recognition of such situations belongs to the family of multi-model techniques ([Narendra and Balakrishnan, 1997], [Dumitrache and Mărgărit, 1999]). It represents an alternative to the classical adaptive control solutions, which frequently require complex and risky recursive computations.

4. Strategic situation. Let us consider a control system that can use, depending on the properties of the environment, more control strategies:

$$u^j(t) = T_j(z(t)), \quad j = 1,...,r \ , \tag{12}$$

where the function $T_j(.)$ defines the j-th command strategy and $z(t)$ is the vector of observations (for instance, that in (4)). It is called a ***strategic situation*** (from [Ceangă, 1969a, 1985b]) the set of vectors of observations $z(t)$, corresponding to the best control strategy, out of the r possible strategies:

$$S_j = \{z(t) / \max_k \Phi[T_k(z(t))] = \Phi[T_j(z(t))]\}\ , \tag{13}$$

where $\Phi(.)$ measures the fulfillment of the objective.

The strategic situations may be defined mainly for complex systems, having various interactions and constraints, such as the biotechnological, economical or production systems. Every strategy concerns a particular tactical objective, which temporary gets the priority in order to fulfill the global objective.

5. Diagnosis situation. The behaviour of the dynamic systems may be evaluated by a set of indicators, called local criteria, which can make use of discrete information extracted by PR techniques. A global evaluation criterion may also be added, for the entire system. Let r be the vector of local criteria. It is called ***diagnosis situation*** the set of vectors in the space of local criteria, corresponding to a particular global evaluation of the controlled process. Some examples of discrete evaluations, defining the diagnosis situations, are: "admissible", "warning situation i", "emergency situation j", "damage regime", etc., where i or j have particular meanings for the supervised process.

Figure 2 presents the hierarchical structure of a control system, based on control situations recognition (adapted from [Saridis, 1989]). Every hierarchical level requires the fulfillment of a different objective, expressed in concepts with different levels of abstraction. The abstraction level increases along with the hierarchy level because, according to the IPDI principle, superior levels do not require precision. The elimination of the details (reducing the entropy) may be performed by processes similar to those implied by the formation of general concepts, starting from a set of less general ones. In the following two chapters, the reasons for using PRCS and their particular learning algorithms will be presented separately, for the execution level and for the upper levels.

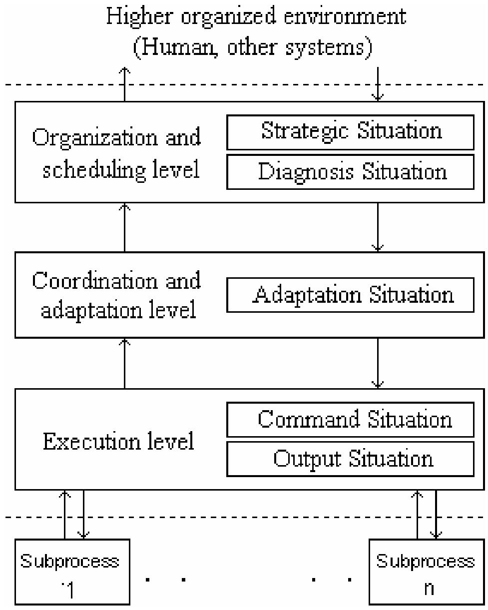


Figure 2. Hierarchical structure of a control system, recognizing ***control situations***

3. Control situations for the upper hierarchical levels

3.1. Systems based on control situation recognition

The control techniques based on adaptation, strategic or diagnosis (supervision) situations are already used in some papers. They appear explicitly or implicitly (that is, without using the terms introduced in this paper).

A. The control structures that make use of the adaptation situation recognition have the advantage of a fast adaptation of the controller. [Frangu, 2001] presents two applications where the structure of adaptation through situation recognition corresponds to that in figure 3. Let

$$x(t) = [y(t),..., y(t - n_a + 1), u(t - 1),..., u(t - n_b)]^T , \quad x \in X , \qquad (14)$$

be the vector of observations, assigned by the recognition automaton to one of the adaptation situations S_i^a, $i = 1,..., p$. The result of the classification aims at selecting the best fit controller, according to the recognized situation.

In one of the mentioned applications, the controlled object is an elastic mechanical transmission, built at Laboratoire d'Automatique de Grenoble (figure 4). It is used as benchmark for robust and adaptive control techniques, in [Landau et al., 1995]. Depending on the mechanical load, the multiresonant frequency response modifies considerably its shape (figure 5).

This behaviour limits the performances obtained by robust and adaptive control techniques. Instead, the recognition of the adaptation situation is proposed.

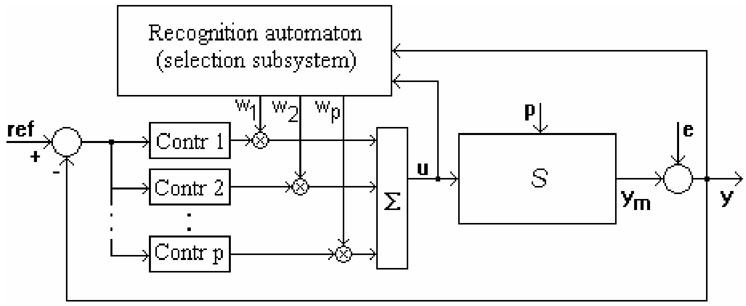


Figure 3. Pattern Recognition Adaptation Structure (analogous with gain-scheduling)

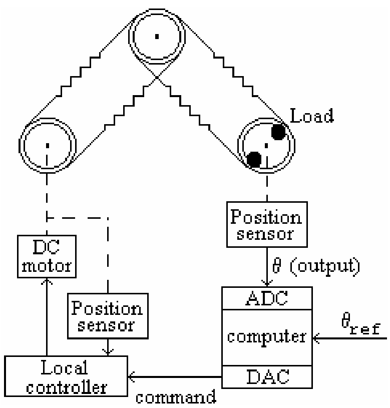


Figure 4. Structure of the position control system

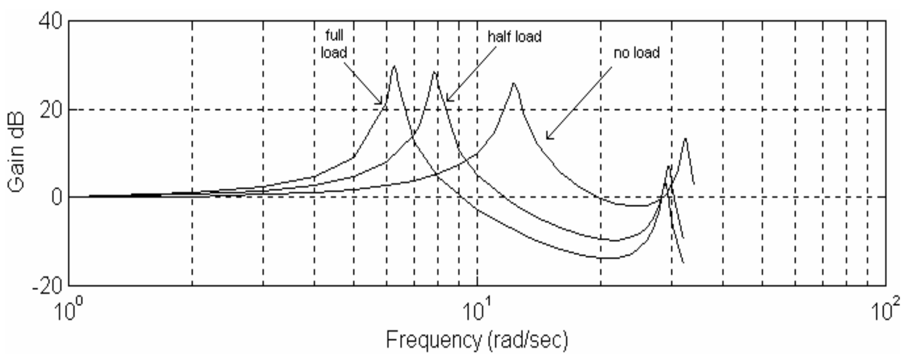


Figure 5. Frequency response of the controlled system, for three different dynamics

The vector of observations is:

$$x(t) = [y(t), y(t-1), y(t-2), y(t-3), u(t-2), u(t-3)]^T \qquad (15)$$

and the classification algorithm is based on the minimum distance to the prototypes of the classes. In addition, the selection automaton requires a supervisor, based on diagnosis situation recognition. When the supervisor detects the stationary regime, it doesn't allow the controller switching (the adaptation situations overlap in this regime, so the selection automaton runs out of information, as in any identification problem, when the identified dynamic object lacks excitation).

In [Ronco, 1997a, 1997b] a "Local Model Network" (LMN) structure is used. It contains local linear models and corresponding local controllers, one pair for each functioning regime of the process. The recognition of the local model and, implicitly, that of the controller, is based on the vector $\Phi \in X_{\Phi}$, which is part of the vector of observations (14). Within this method, the adaptation situations correspond to a pre-established partition. The recognition is performed by a set of RBF neurons, centered in a uniform net of points of X_{Φ}. A different approach, proposed in [Jordan and Jacobs, 1994], is called "Adaptive Mixture of Experts". Here, the partition of the space X_{Φ} is performed through learning, by a multilayer perceptron. Consequently, the domains of the classes (adaptation situations) are not equal, but depend on the approximation ability of each local model.

B. The idea of switching the strategies, which justifies the concept of strategic situation, assumes that **the current control objective can change**, during the control of the process. The objective can change as a result of the evolution of the subprocesses (for instance: changes in behaviour, failure of the local control loops or even failure of the superior level) or of the interaction with the higher organized environment (for instance: human). This change can require to switch the controller, at the execution level, or to switch the method for the coordination of the subprocesses, at the middle (coordination and adaptation) level. Intuitively, the strategies may be switched through instructions like: "switch to a survival strategy, because a local loop is temporarily unavailable", "switch from cooperation with other agents to competition", "switch from stimulation of the bioreactor population to the rejection of the parasitic population", "switch from emergency medical care to the convalescence recovery method". The corresponding switching decisions may belong to the upper level (organization and scheduling) or to the middle level. There already are some well grounded papers, which investigate the properties of such systems, possibly in uncertain conditions. Among these, [Kuipers, 1994] studies the validation of heterogeneous control laws and [Johansson, 1996, 1997] present an analysis method for the stability of the heterogeneous controlled systems, regardless the type of the controller. The method makes use of piecewise defined Lyapunov functions, one for each validity domain of local controllers. The result obtained in this approach is also useful for the analysis of the systems based on adaptation situation and strategic situation recognition.

Among the problems raised by the heterogeneous command, those concerning the switching decision lead to using a strategic situation recognition automaton. The advantages of including such a subsystem are:

- The automaton is a learning one, meaning it has the ability to learn from the examples, including unsupervised learning. In this variant, the recognition decision requires no information provided by the human expert, but uses the similarity of the examples, based on a measure of the objective's fulfillment.

- It can model complex discrete approximation functions, whose analytical computation may be unreachable. To materialize these functions, both classical recognition algorithms and discrete output neural networks can be used.

A simple example of using the strategic situations is presented in [Frangu, 2002], starting from the known benchmark, the “backer truck”. Obviously, there are initial positions of the truck who don’t accept solutions, such as the positions where the backside of the truck faces the wall, at low distance (figure 6). If starting from these positions, the backwards docking fails, regardless the chosen controller. In order to find out whether the initial position allows a solution or not, a recognition automaton will learn from the experience accumulated during the previous docking attempts (including the unsuccessful ones). The objective of the automaton is to predict if the current controller succeeds to dock, using the information of the initial position. If the automaton predicts the failure, the control system has to switch to a controller having a different objective than immediate docking (in this case, to drive forward, to the distance that allows secure docking).

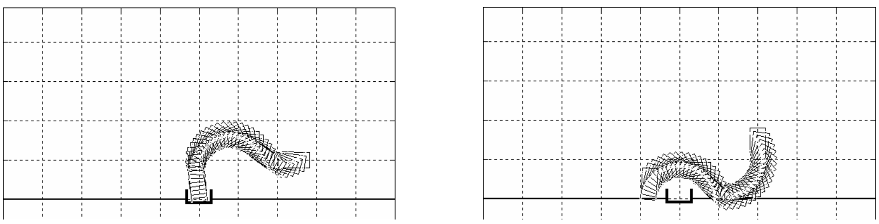


Figure 6. Docking with start from initial positions (2, 1, 0) and (2, 1, π/2)

The initial position contains the truck backside coordinates: x (along the wall), y (distance to the wall) and orientation angle with respect to the wall, θ. Figure 6 presents two docking attempts, whose initial coordinates x and y are identical, but presenting different initial angles and different docking results. The vector of observations is the initial position:

$$z = [x, y, \theta]^T \; . \tag{16}$$

The strategic situation to be recognized is the success or failure of the docking process, starting from that initial position. Using simulation experiments, the learning set was formed. The data structure analysis showed that the boundary between classes has not a simple shape. Consequently, the classification was made through the potential functions method [Aizerman, 1964].

C. The diagnosis (or supervision) situations are currently used to diagnose the systems. However, the field of using the diagnosis situations is even larger (for instance, the detection of the stationary regime, in [Frangu, 2001]). In [Bivol et al., 1976], a complex energetic boiler plant, which includes more interconnected control loops, is considered. The evaluation of the quality of the dynamic regimes of the system considers more local criteria in the individual control loops, such as: overshot, damping factor, etc. The diagnosis situation called “normal” is defined in the space of local criteria, based on a recorded set of states, previously diagnosed by human operators. An automaton learns to recognize this situation and its opposite; it will be used to real-time diagnosis of the plant.

On the other hand, each diagnosed state is associated to the known vector of parameters of the multivariable controller. Through learning, the situation “normal” in the space of local criteria is assigned to a domain defined in the space of parameters. The discriminant function of this domain may be understood as the membership function of the fuzzy set “normal”, in the space of parameters. It is used to solve the problem of optimization with constraints in the space of parameters (during the design stage).

3.2. Clusters' anatomy and learning algorithms

In the case of adaptation situations, the data structure is similar to that of the output and command situations (will be analyzed next section). In the case of strategic and diagnosis situations, the data structure can be complex, with unconnected clusters, etc. Consequently, strong and general recognition methods, able to work with poor initial information, are necessary. One of these is the potential functions method, developed in the '60s by the team of Aizerman [Aizerman, 1964], at the Control Institute of the Moscow Academy. The adaptation and use of this method to the recognition of control situations appeared in [Drăgan and Ceangă, 1968], [Ceangă, 1969b]. The method is based on memorizing the "alien prototypes" or "poles", i.e. the vectors differently classified by the PRCS and by the teacher. A potential function $K(\bar{x}, x)$ is assigned to each of the memorized poles, $\bar{x}$.

Despite its generality and efficiency qualities, the method of potential functions cannot be used when the structure of clusters changes in time, because the adaptation of the discriminant functions would indefinitely increase the number of memorized poles. This drawback was noted in [Ceangă, 1969b] and two new recognition structures were proposed. They are also based on potential functions; they preserve the general character of the initial method and provide the ability of adapting to slow changes of the data structure.

The first method, called the method of floating poles ([Ceangă, 1969b]), contains a first stage of pre-learning, using the method of potential functions for the configuration of the learning system structure. In this stage, M poles are memorized. They mainly lie near the boundaries of the classes, where the recognition automaton is usually wrong. The continuation of the learning and the adaptation of the PRCS to the changes in boundaries are performed by adjusting the position of the poles, according to the "floating poles algorithm (FPA)". Essentially, this algorithm is similar to that presently used by the Kohonen neural networks.

In the second method, presented in [Ceangă, 1969a, 1969b], [Bumbaru, 1970], the potential functions form the input layer of a recognition automaton (during the next decade, the structure was called RBF, when it was implemented by neural networks). As in the previous case (FPA), a first learning stage is used, when the poles x_j, $j = 1, M$, are memorized. The poles are assigned the potential functions $K(x_j, x)$, $j = 1, \ldots, M$, which form the input layer of the recognition automaton. During the second stage, the learning implies adjusting the weights of each potential function.

In many papers, strategic and diagnosis situations are implicitly recognized, using neural networks (such as multilayer perceptrons). The drawback of the neural networks is the lack of transparency (no explanation about the clusters is provided). The advantage of the presented methods (disregarding the classical or neural implementation) consists in the selection of relevant poles; their position suggests the structure of the clusters, without disturbing the properties of generality and efficiency of the recognition algorithm.

4. Control situations for the execution level

4.1. Arguments for using recognition systems at the inferior level

In order to prove that controllers who recognize output and command situations are suitable, the informational approach for intelligent systems is useful. The concepts introduced by Saridis in [Saridis, 1989] (machine intelligence, knowledge flow, etc.) are used in [Frangu, 2001] to demonstrate the following property: for a system with a known level of uncertainty, there is a limit of the resolution of the discrete command, beyond which the entropy of the knowledge flow cannot increase. The same property applies to the recognition of discrete values of the output. Some practical examples, involving uncertainty, are presented in the sequel.

1. The actuator is, in most cases, the lowest precision element of a control loop. Excepting some particular cases, this affects the precision of mechanical positioning in the industry processes. Because of the important uncertainty, adopting a discrete set of values for the command becomes natural. If the distance between the discrete values is comparable to the uncertainty level, this operation does not lower the positioning precision.
2. The reference of the loops is often chosen according to uncertain technological requirements. In this case, maintaining the controlled object within the boundaries determined by the uncertainties of the reference is a satisfactory objective.

The two mentioned examples are also reasons for another modern control approach: hybrid systems, with continuous/discrete interface (HSCDI, [Antsaklis, 1994]). The comparative analysis of the two approaches in [Frangu, 2001] (PRCS and HSCDI) led to the following conclusions:

1. In HSCDI the partition is applied to the state space of the controlled object; the PR approach is based on the partition of the space of observations (see section 2).
2. The HSCDI approach requires the analytical state model of the controlled process, whereas the PR approach considers this model partially or totally unknown.
3. The synthesis of HSCDI requires the partition chosen by the human designer; the PRCS do not require predefined boundaries, because these result by learning.
4. There is not much knowledge about how to choose the partition of the continuous state space, in HSCDI. The number of discrete states is not equal to that of the discrete commands. This raises some particular problems, such as: how to determine the resolution of the partition, how to determine the masked states and absorbing states, etc. ([Oltean, 1998]). In PRCS, the number of classes is equal to the number of discrete commands, but the classes may contain more clusters. The learning solves this problem, assigning the same command to the clusters belonging to a class.
5. The HSCDI synthesis requires a complex sequence of design operations. Instead, the PR approach determines by learning the function that assigns the vectors of observations to the discrete commands.

4.2. Clusters' anatomy and learning algorithms

The clusters' structure for the output and command situations have a stripe-like shape: the clusters lie in compact and adjacent domains of the space of observations, with similar boundaries. There is an order of the clusters, corresponding to the order of the discrete values of the variable that generated the classes (the output or the command input). Some of the observed properties of the clusters are mentioned in [Frangu, 2001] (some of them are even demonstrated, for Hammerstein type systems). Among these:

- if the controlled system is linear, the boundaries of the clusters are parallel hyperplanes, the clusters are adjacent and disposed in order;
- the boundaries may be hyperplanes even for a larger class of nonlinear systems; to illustrate this property, four examples are presented in figure 7a-d, in a two-dimensional space of observations; the structure 7b appeared in a control problem for a biotechnological process, developed in a bioreactor [Frangu, 2001a];
- if the system does not contain hysteresis or discontinuous functions, every cluster is connected; in the contrary case, the clusters may become unconnected, subjects of some sort of space shearing (fig. 7*c*);
- there is a unique curve in the space of observations, which corresponds to the stationary regime and crosses all the classes; the low frequency excitation of the controlled process determines observation vectors lying in the neighborhood of this curve, whereas the vectors determined by a richer dynamics of the process are situated farther;
- if the noise disturbing the output is zero averaged, the boundaries obtained by learning converge to cluster's true position, even if the noise is not white.

For the presented types of clusters, the appropriate recognition algorithms are those based on the minimum distance with respect to the skeleton of the clusters. If the boundaries are linear, the skeleton is the own regression line, obtainable through batch processing, which has guaranteed convergence.

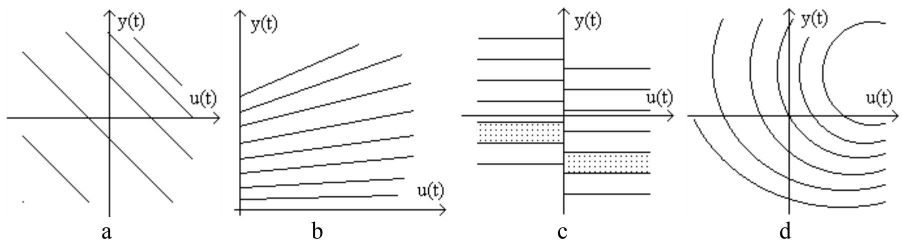


Figure 7. Possible shapes of the clusters, for the output situations

5. Conclusions and new research directions

Intelligent control implies generalization and abstraction operations, which lead to synthetic concepts, necessary for operating in uncertain conditions or for the superior levels of the hierarchical control structure. These operations are performed through learning and aim at forming patterns with different levels of abstraction. In this work, the patterns generically called "control situations" are: output, command, adaptation, strategic and diagnosis situations. They allow obtaining solutions of the control problems through situation recognition, in classical or neural implementation. Some of them are illustrated in the paper.

Some new research directions in PRCS, in the opinion of the authors:

1. The investigation of the systems' stability, when the recognition of the adaptation situation determines the switching of the control laws or of their parameters. Considering this approach as a more general multimodel method can be the starting point [Ronco, 1997a, 1997b].
2. The investigation of the structure in figure 1, where a PR system is taught to recognize command situations, by a learning teacher, who recognizes output situations (teaching the controller by an uncertain teacher). The beginning was made by Nikolic and Fu, in 1966, when they analyzed the convergence of the learning algorithms, for the controller and for the uncertain teacher, but for a particular variant of the problem.
3. Reconsidering the PRCS approach, according to [Goertzel, 1993], where a new mathematical model of the intelligence is proposed. It integrates specific operations, like perception, induction, deduction, analogy, etc., within a network (the "master network") that operates with patterns. Induction is defined as "*the construction based on the patterns recognized in the past, of a coherent model of the future*". The perception is introduced as "*the network of pattern recognition processes through which an intelligence builds a model of the world*" and "*the perceptual hierarchy is composed of a number of levels, each one recognizing patterns in the output of the level below it*". Deduction is also introduced by pattern recognition theory. Based on the previous concepts, "*intelligence is defined as the ability to optimize complex functions of unpredictable environments*". The theoretical framework of Goertzel's model of mind can generate new ideas and research directions in the field of intelligent control.

References

Aizerman, M.A., Braverman, E.M., Rozonoer, L.I. (1964) Teoreticeskie osnovy metoda potentialnyh functii i zadace ob obucenii avtomatov raspoznovaniiu situatii na classy, *Avtomatika i Telemehanika*, 1964, **25**, pp. 917-937.

Aizerman, M.A. (1965) Zadacia ob obucenii avtomatov razdeleniiu vhodnyh situatii na klassy (raspoznavaniiu obrazov) II IFAC Diskretnye i samonastraivaiushcesia sistemy, Iz. Mir, Moscow.

Aizerman, M.A., Braverman, E.M., Rozonoer, L.I. (1966) Problema obuceniia mashin raspoznavaniiu vneshnih situatii, Samoobuciaiushciesia avtomaticeskie sistemy, Moscow.

Albus, J.S. (1991). Outline for the theory of intelligence. *IEEE Transactions on Systems, Man and Cybernetics*, **21.**

Antsaklis, P.J., Stiver, J.A., Lemmon, M. (1994) A Logical DES Approach to the Design of Hybrid Control Systems. *Technical Report of the ISIS group at the University of Notre-Dame*, ISIS-94-011, Notre-Dame, Indiana, SUA, 46556.

Bivol, I., Bumbaru, S., Ceangă, E. (1976) Defining the admissible domain of large systems, in the space of parameters, through recognition methods (in Romanian), Probleme de Automatizare: "Sisteme automate şi de prelucrare a informaţiilor", Ed. Academiei, Bucharest.

Bumbaru, S., Ceangă, E., Bivol, I. (1970) Control of the Processes with Slow-Moving Characteristics by Pattern Recognizing Systems, Proceedings 6-th Int. Symposium on Information Processing, Bled, Yugoslavia, 1970, paper H1, pp.1-4.

Ceangă, E. (1969a) Methods for characterizing a multidimensional state set of the system, through learning automata (in Romanian). PhD thesis, University Politehnica of Bucharest.

Ceangă, E., Bivol, I. (1969b) Pattern-recognizing control systems in rolling technology, Proceedings Yugoslav International Symposium on Information Processing, Bled, Yugoslavia, 1969, pp. 241-243.

Ceangă, E., Bivol, I. (1971a) Experimental research in control of complex processes, using situation recognition learning automata (in Romanian), *Automatică şi Electronică*, vol. 15, **1**, 1971, pp. 8-12.

Ceangă, E., Necula, N. (1971b) Choosing the pattern space in the problem of control through learning systems (in Romanian), *Automatica şi Electronica*, Bucharest, vol. 15, 4, pp. 153-156.

Ceangă, E., Bumbaru, S., Bivol, I. (1981) About the control of processes with incomplete a priori information, using learning systems (in Romanian), Lucrările Simpozionului Naţional de Cibernetică, Ed. Academiei, Bucharest, pp. 76-81.

Ceangă, E., Bumbaru, S., Bivol, I., Vasiliu, C. (1984) Learning Controller for Automatic Control of Processes with Unknown Characteristics, The Annals of the University of Galaţi, Romania, fasc. III (Electrotechnics), II, pp. 5-10.

Ceangă, E., Vasiliu, C., Bumbaru, S. (1985a) Learning Systems with Control Situation Recognition. *Proceedings of IFAC Symposium on System Optimization and Simulation*, Berlin, pp. 384-389.

Ceangă, E., Bumbaru, S., Bivol, I. (1985b) Concepts and situation recognition techniques in the cybernetics systems (in Romanian), *Cibernetica Aplicată*, Ed. Academiei, Bucharest, pp. 48-51.

Ceangă, E., Bumbaru, S. (directors) (1991) Learning and artificial intelligent systems for diagnosis and control of the processes (in Romanian), grant 23.02.07/1991 of the Romanian Department of Education and University of Galaţi.

Drăgan, P., Ceangă, E. (1968) Research on automatic classification learning systems (in Romanian), *Studii şi cercetări matematice*, Ed. Academiei, Bucharest, nr. 3, pp. 321-333

Dumitrache, I., Mărgărit, L. (1999) Model reference adaptive control using multiple models, Proceedings of Control Systems and Computer Science 12, May 26-29, Bucharest, vol 1, pp. 150-155.

Frangu, L. (2001) Advanced control systems, based on pattern recognition, neural networks and image processing (in Romanian), PhD thesis, University "Dunărea de Jos" of Galaţi, Romania.

Frangu, L., Caraman, S., Ceangă, E. (2001a) Model Based Predictive Control using Neural Network for Bioreactor Process Control, *Control Engineering and Applied Informatics*, **3**, April, pp.29-38.

Frangu, L., Ceangă, E., Caraman, S., Boutallis, Y. (2002) A pattern recognition approach to intelligent behaviour. Switching the strategies, IEEE Symposium - Intelligent Systems, September 10-12th, 2002, Varna, Bulgaria, pp. 369-372.

Goertzel, B., (1993) The Structure of Intelligence. A new Mathematical Model of Mind. Springer.

Grigore, O., Grigore, O. (2000) Control of Nonlinear Systems Using Competitive Neural Networks, Proceedings of OPTIM (Optimization of Electrical and Electronical Equipment), May, 11, 2000, Braşov, pp. 671-674.

Jordan, M.I., Jacobs, R.A. (1994) Hierarchical mixtures of experts and the em algorithm, Neural Computation, 6, pp. 181-214.

Johansson, M., Rantzer, A. (1996) Computation of Piecewise Quadratic Lyapunov Functions for Hybrid Systems, Technical report (June 1996), Lund Institute of Technology, Sweden.

Johansson, M., Malmborg, J., Rantzer, A., Bernhardsson, B., Arzen, K. (1997) Modelling and Control of Fuzzy, Heterogenous and Hybrid Systems, 3-rd IFAC Symposium on Intelligent Measuring Instruments for Control Applications, SICICA'97, Annecy, France, June 1997, pp. 33-38.

Kuipers, B., Astrom, K. (1994) The Composition and Validation of Heterogenous Control Laws, Automatica, Pergamon Press, 30, pp. 233-249.

Landau, I.D., Rey, D., Karimi, A., Voda, A., Franco, A. (1995) A flexible Transmission System as a Benchmark for Robust Digital Control, European Journal of Control, 1995, 1, pp. 77-96.

Narendra, K., Balakrishnan, J. (1997) Adaptive control using multiple models, IEEE Trans. On Automatic Control, vol. 42, 1997, 2, pp. 171-187.

Nikolic, Z.J., Fu, S.K. (1966) An algorithm for learning without external supervision and its application to learning control systems, *IEEE Trans. on Automatic Control*, **7**

Oltean, V.E. (1998) Contributions to modelling, analysis and synthesis of hybrid systems with continuous/discrete interface (in Romanian). PhD thesis, University "Dunărea de Jos" of Galați, Romania.

Ronco, E., Gawthrop, P.J. (1997a) Incremental model reference adaptive polynomial controllers network, Technical report CSC 97003, University of Glasgow, Dept. of Mech. Eng. (available at www.mech.gla.ac.uk).

Ronco, E., Gawthrop, P.J. (1997b) Neural networks for modelling and control, Technical report CSC 97008, University of Glasgow, Dept. of Mech. Eng. (available at www.mech.gla.ac.uk).

Saridis, G. N. (1979). Towards the Realization of Intelligent Controls. *Proceedings of the IEEE*, **67**, pp. 1115-1133.

Saridis, G. N. (1988). Entropy Formulation of Optimal and Adaptive Control. *IEEE Transactions on Automatic Control*, **33**, pp. 713-721.

Saridis, G. N. (1989). Analytic Formulation of the Principle of Increasing Precision with Decreasing Intelligence for Intelligent Machines. *Automatica*, **25**, pp. 461-467.

Seem, J.E., Nesler, C.N. (1996) A new Pattern Recognition Adaptive Controller, IFAC Congress, San Francisco.

Taylor, W.K. (1963) Samonastraivaiuşciesia ustroistvo upravleniia, ispolizuiuşcee raspoznavanie obrazov, IFAC Congress, Basel (Diskretnîe i samonastraivaiuşciesia sistemî, Nauka, Moskva, 1965).

Truță, A.M. (2002) Modeling the learning process at the human cortex level (in Romanian), PhD thesis (in Romanian), University Politehnica of Bucharest, 2002.

Waltz, M.D., Fu, S.K. (1966) A heuristic approach to reinforcement learning control systems, *IEEE Transactions on Automatic Control*, **4.**

Widrow, B. (1962) Generalization and information storage in networks of adaline 'neurons', in "Self Organizing Systems 1962", Yovitz, Jacobi, Goldstein, Eds. Washington DC, Spartan Books, pp. 435-461.

Widrow, B. (1964) Pattern Recognition and Adaptive Control, *IEEE Trans. Appl. in Industry*, **9.**

THE DISTURBANCE ATTENUATION PROBLEM FOR A GENERAL CLASS OF LINEAR STOCHASTIC SYSTEMS

Vasile Dragan and Teodor Morozan

Institute of Mathematics of the Romanian Academy
P.O. Box 1-764, Bucharest, Romania, email: vdragan@fx.ro

Adrian Stoica

University "Politehnica" of Bucharest
Faculty of Aerospace Engineering
Str. Splaiul Independentei No. 313
RO-060042 Bucharest, Romania, e-mail: amstoica@fx.ro

Abstract The aim of this paper is to present a solution to the disturbance attenuation problem for stochastic systems subjected both to multiplicative white noise and to Markovian jumps. Based on a Bounded Real Lemma type result for the considered class of stochastic systems, necessary and sufficient solvability conditions are derived in terms of the solutions of a specific system of matrix inequalities.

Keywords: stochastic systems, multiplicative white noise, Markov jumps, γ-attenuation problem, Bounded Real Lemma

1. Introduction and problem formulation

The disturbance attenuation problem plays an important role in a wide variety of control applications. It is a well-known fact that the sensitivity reduction, the robust stabilization with respect to various type of uncertainty, tracking and filtering problems, to mention only a few of such applications, can be converted into disturbance attenuation problems. In the deterministic framework this problem is solved by various H^{∞} control techniques (see *e.g.* (Doyle, *et al*., 1989; Gahinet and Apkarian, 1994) and their references). In the stochastic case, corresponding state-space solutions have been also derived. Thus, for systems with multiplicative white noise such results can be found for instance in (Boyd, *et al*., 1994; Hinrichsen and Pritchard, 1998) and (Dragan and Morozan, 1997) for the time-varying case. The disturbance attenuation problem for systems with Markov jumps has been addressed, too. Corresponding theoretical developments in this case are given for example in (Dragan and Morozan, 2001; Dragan, *et al*., 1998; Stoica and Yaesh, 2002).

The aim of this paper is to present a solution to the disturbance attenuation problem for stochastic systems subjected both to multiplicative white noise and to Markovian jumps. Before stating the problem, some notations are introduced and some preliminary results are briefly recalled.

Consider the stochastic system subjected both to multiplicative white noise and to Markov jumps:

$$dx(t) = A_0\left(\eta(t)\right)x(t)dt + \sum_{k=1}^{r} A_k\left(\eta(t)\right)x(t)dw_k(t), \tag{1}$$

where $A_k(i) \in R^{n\times n}$, $0 \le k \le r$ are given, $w(t) = \left(w_1(t),\ldots,w_r(t)\right)^*$ is a standard Wiener process (see *e.g.* (Friedman, 1975)) and $\eta(t)$, $t \ge 0$ is a right continuous homogeneous Markov chain with the state space the set $\mathfrak{D} = \{1,\ldots,d\}$ and the probability transition matrix $P(t) = e^{Qt}$, where $Q = \left[q_{ij}\right]$, $\sum_{j=1}^{d} q_{ij} = 0,\ i \in \mathfrak{D}$ and $q_{ij} \ge 0$ if $i \ne j$. The σ-algebras $\mathfrak{F}_t = \sigma(w(s), s \in [0,t])$ and $\mathfrak{G}_t = \sigma(\eta(s), s \in [0,t])$ are assumed independent for all $t \ge 0$. By $\mathcal{H}_t$ it is denoted the smallest σ-algebra containing $\mathfrak{F}_t$ and $\mathfrak{G}_t$. Throughout the paper $L_{\eta,w}{}^2\left([0,\infty),\mathbf{R}^{\ell}\right)$ stands for the space of all functions $u(t)$ measurable non-anticipative with respect to the family of the σ-algebras $\mathcal{H}_t$ with values in $\mathbf{R}^{\ell}$ and with $E\left[\int_0^{\infty}\left|u(t)\right|^2 dt\right] < \infty$, where E denotes as usual the expectation. By $E[x \mid \eta(t) = i]$ it is denoted the conditional expectation on the event $\eta(t) = i$. The space of all $(n \times n)$ symmetric matrices is denoted by $\mathcal{S}_n$, $\mathcal{S}_n{}^d$ represents the space of all $H = \left(H(1),\ldots,H(d)\right)$ with $H(i) \in \mathcal{S}_n, i \in \mathfrak{D}$ and $\mathcal{M}_{n,m}{}^d$ stands for the space of $A = \left(A(1),\ldots,A(d)\right)$ where $A(i) \in \mathbf{R}^{n\times n}, i \in \mathfrak{D}$.

Definition 1. The zero solution of the stochastic system (1) is called exponentially stable in mean square (ESMS), or equivalently $(A_0,\ldots,A_r;Q)$ is stable if there exist $\alpha > 0$ and $\beta \geq 1$ such that

$$E\left[\left|x(t,t_0,x_0)\right|^2 \mid \eta(t_0)=i\right] \leq \beta e^{-\alpha(t-t_0)}\left|x_0\right|^2$$

for any $t \geq t_0 \geq 0,\ x_0 \in \mathbf{R}^n,\ i \in \mathfrak{D}$, where $x(t,t_0,x_0)$ denotes the solution of the system (1) with the initial condition $x_0 \in \mathbf{R}^n$ at t_0.□

Consider the stochastic controlled system:

$$\begin{aligned} dx(t) &= \left[A_0(\eta(t))x(t) + B_0(\eta(t)u(t))\right]dt \\ &\quad + \sum_{k=1}^{r}\left[A_k(\eta(t))x(t) + B_k(\eta(t)u(t))\right]dw_k(t) \\ y(t) &= C(\eta(t))x(t) + D(\eta(t))u(t), \end{aligned} \tag{2}$$

where $u(t) \in \mathbf{R}^m$ is the control variable and $y(t) \in \mathbf{R}^p$ denotes the output. If $(A_0,\ldots,A_r;Q)$ is stable then one can consider the input-output operator $\mathfrak{T}$ associated to (2), defined on $L_{\eta,w}^2([0,\infty),\mathbf{R}^m)$ with values on $L_{\eta,w}^2([0,\infty),\mathbf{R}^p)$, as $(\mathfrak{T}u)(t) = y_u(t)$, where

$$y_u(t) = C(\eta(t))x_u(t) + D(\eta(t))u(t),$$

$x_u(t)$ denoting the solution of the first equation (2) with the initial condition $x_u(0)=0$. The norm of this linear bounded input-output operator is denoted by $\|\mathfrak{T}\|$.

With the above notations and definitions, one can now state the disturbance attenuation problem. Consider the two-input, two-output stochastic system:

$$\begin{aligned} dx(t) &= \left[A_0(\eta(t))x(t) + G_0(\eta(t))v(t) + B_0(\eta(t))u(t)\right]dt \\ &\quad + \sum_{k=1}^{r}\left[A_k(\eta(t))x(t) + G_k(\eta(t))v(t) + B_k(\eta(t))u(t)\right]dw_k(t) \\ z(t) &= C_z(\eta(t))x(t) + D_{zv}(\eta(t))v(t) + D_{zu}(\eta(t))u(t) \\ y(t) &= C_0(\eta(t))x(t) + D_0(\eta(t))v(t), \end{aligned} \tag{3}$$

where the input variable $v(t) \in \mathbf{R}^{m_1}$ denotes the exogenous signals,

$u(t) \in \mathbf{R}^{m_2}$ denotes the control variable, $z(t) \in \mathbf{R}^{p_1}$ is the regulated output and $y(t) \in \mathbf{R}^{p_2}$ represents the measured output. The matriceal coefficients $A_k(i), G_k(i), B_k(i), 0 \le k \le r$, $C_z(i), D_{zv}(i), D_{zu}(i), C_0(i), D_0(i), i \in \mathfrak{D}$, are given matrices of appropriate dimensions. The class of admissible controllers has the state-space equations:

$$\begin{aligned} dx_c(t) &= \left[A_c(\eta(t))x_c(t) + B_c(\eta(t))y(t)\right]dt \\ u(t) &= C_c(\eta(t))x_c(t) + D_c(\eta(t))y(t), \end{aligned} \tag{4}$$

where $x_c(t) \in \mathbf{R}^{n_c}$ with $n_c > 0$ being a given integer. When coupling the controller (4) to the system (3) one obtains the resulting closed-loop system:

$$\begin{aligned} dx_{cl}(t) &= \left[A_{0cl}(\eta(t))x_{cl}(t) + G_{0cl}(\eta(t))v(t)\right]dt \\ &\quad + \sum_{k=1}^{r}\left[A_{kcl}(\eta(t))x_{cl}(t) + G_{kcl}(\eta(t))v(t)\right]dw_k(t) \\ z(t) &= C_{cl}(\eta(t))x_{cl}(t) + D_{cl}(\eta(t))v(t), \end{aligned} \tag{5}$$

where

$$\begin{aligned} A_{0cl}(i) &= \begin{bmatrix} A_0(i) + B_0(i)D_c(i)C_0(i) & B_0(i)C_c(i) \\ B_c(i)C_0(i) & A_c(i) \end{bmatrix}, \\ A_{kcl}(i) &= \begin{bmatrix} A_k(i) + B_k(i)D_c(i)C_0(i) & B_k(i)C_c(i) \\ 0 & 0 \end{bmatrix}, 1 \le k \le r, \\ G_{0cl}(i) &= \begin{bmatrix} G_0(i) + B_0(i)D_c(i)D_0(i) \\ B_c(i)D_0(i) \end{bmatrix}, \\ G_{kcl}(i) &= \begin{bmatrix} G_k(i) + B_k(i)D_c(i)D_0(i) \\ 0 \end{bmatrix}, 1 \le k \le r, \\ C_{cl}(i) &= \left[C_z(i) + D_{zu}(i)D_c(i)C_0(i) \quad D_{zu}(i)C_c(i)\right], \\ D_{cl}(i) &= D_{zv}(i) + D_{zu}(i)D_c(i)D_0(i), \ i \in \mathfrak{D}. \end{aligned} \tag{6}$$

Then the disturbance attenuation problem is formulated as follows: given $\gamma > 0$ find necessary and sufficient conditions for the existence of a controller (4) of prescribed order n_c such that $(A_{0cl}, \dots, A_{rcl}; Q)$ is stable and the input-output operator $\mathcal{T}_{cl}$ associated with (5) has the norm less than γ.

2. Stochastic Bounded Real Lemma

The following result proved in (Drăgan *et al.*, 2003) is a stochastic version of the well-known Bounded Real Lemma of the deterministic framework and it plays a key role for the main result presented in the next section.

Theorem 1 (Bounded Real Lemma) *The following are equivalent:*

(i) The system $(A_0, A_1, ..., A_r; Q)$ *is stable and* $\|\mathcal{T}\| < \gamma$ *;*

(ii) It exists $\hat{X} = (\hat{X}(1), ..., \hat{X}(d)) \in \mathcal{S}_n{}^d$, $\hat{X}(i) > 0$ *satisfying the following LMI on* $\mathcal{S}_{n+m}{}^d$ *:*

$$\mathcal{N}(\hat{X}, \gamma) < 0,$$

$\mathcal{N}(\hat{X}, \gamma)$ *denoting the generalized dissipation matrix associated with the system (2) and with the parameter* γ *, namely* $\mathcal{N}(X) = (\mathcal{N}_1(X, \gamma), ..., \mathcal{N}_d(X, \gamma))$, *where*

$$\mathcal{N}_i(X,\gamma) = \begin{bmatrix} \mathcal{N}_{11}{}^i(X,\gamma) & \mathcal{N}_{12}{}^i(X,\gamma) \\ \left(\mathcal{N}_{12}{}^i\right)^*(X,\gamma) & \mathcal{N}_{22}{}^i(X,\gamma) \end{bmatrix},$$

with:

$$\begin{aligned} \mathcal{N}_{11}{}^i(X,\gamma) &= A_0^*(i)X(i) + X(i)A_0(i) \\ &+ \sum_{k=1}^{r} A_k^*(i)X(i)A_k(i) + \sum_{j=1}^{d} q_{ij}X(j) + C^*(i)C(i) \\ \mathcal{N}_{12}{}^i(X,\gamma) &= X(i)B_0(i) + \sum_{k=1}^{r} A_k^*(i)X(i)B_k(i) + C^*(i)D(i) \\ \mathcal{N}_{22}{}^i(X,\gamma) &= -\gamma^2 I_m + D^*(i)D(i) + \sum_{k=1}^{r} B_k^*(i)X(i)B_k(i). \end{aligned}$$

(iii) There exists $Y = (Y(1), ..., Y(d)) \in \mathcal{S}_n{}^d$, $Y > 0$ *such that*

$$\begin{bmatrix} \mathcal{W}_{0,0}(Y,i) & \cdots & \mathcal{W}_{0,r}(Y,i) & \mathcal{W}_{0,r+1}(Y,i) & \mathcal{W}_{0,r+2}(Y,i) \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ \mathcal{W}_{0,r}{}^*(Y,i) & \cdots & \mathcal{W}_{r,r}(Y,i) & \mathcal{W}_{r,r+1}(Y,i) & \mathcal{W}_{r,r+2}(Y,i) \\ \mathcal{W}_{0,r+1}{}^*(Y,i) & \cdots & \mathcal{W}_{r,r+1}{}^*(Y,i) & \mathcal{W}_{r+1,r+1}(Y,i) & \mathcal{W}_{r+1,r+2}(Y,i) \\ \mathcal{W}_{0,r+2}{}^*(Y,i) & \cdots & \mathcal{W}_{r,r+2}{}^*(Y,i) & \mathcal{W}_{r+1,r+2}{}^*(Y,i) & \mathcal{W}_{r+2,r+2}(Y,i) \end{bmatrix} < 0, \quad (7)$$

$i \in \mathcal{D}$, *where*

$$\mathcal{W}_{0,0}\left(Y,i\right)=\left(A_0\left(i\right)+q_{ii}I_n/2\right)Y\left(i\right)+Y\left(i\right)\left(A_0\left(i\right)+q_{ii}I_n/2\right)^*+B_0\left(i\right)B_0^*\left(i\right)$$
$$\mathcal{W}_{00,k}\left(Y,i\right)=Y\left(i\right)A_k^*\left(i\right)+B_0\left(i\right)B_k^*\left(i\right),\ k=1,...,r$$
$$\mathcal{W}_{0,r+1}\left(Y,i\right)=Y\left(i\right)C^*\left(i\right)+B_0\left(i\right)D^*\left(i\right)$$
$$\mathcal{W}_{0,r+2}\left(Y,i\right)=\left(\sqrt{q_{i1}}Y\left(i\right),...,\sqrt{q_{i,i-1}}Y\left(i\right)\ \ \sqrt{q_{i,i+1}}Y\left(i\right),...,\sqrt{q_{i,d}}Y\left(i\right)\right),$$
$$\mathcal{W}_{l,k}\left(Y,i\right)=B_l\left(i\right)B_k\left(i\right),1\le l,k\le r,l\ne k,$$
$$\mathcal{W}_{l,l}\left(Y,i\right)=B_l\left(i\right)B_l^*\left(i\right)-Y\left(i\right),\ 1\le l\le r,$$
$$\mathcal{W}_{l,r+1}\left(Y,i\right)=B_l\left(i\right)D^*\left(i\right),\ 1\le l\le r,$$
$$\mathcal{W}_{r+1,r+1}\left(Y,i\right)=D\left(i\right)D^*\left(i\right)-\gamma^2I_p$$
$$\mathcal{W}_{l,r+2}\left(Y,i\right)=0,1\le l\le r+1$$
$$\mathcal{W}_{r+2,r+2}\left(Y,i\right)=\mathrm{diag}\left(-Y\left(1\right),...,-Y\left(i-1\right)\ \ -Y\left(i+1\right),...,-Y\left(d\right)\right).$$

3. Main result

In this section the disturbance attenuation problem with an imposed level of attenuation $\gamma>0$ is considered. The developed approach is based on an LMI technique and it extends to the stochastic framework the existing results in the deterministic context. The following known fact (see *e.g.* (Boyd *et al.*, 1994)) will be used to derive necessary and sufficient conditions guaranteeing the existence of a γ-attenuating controller.

Lemma 1 (Projection Lemma) *Let* $\mathcal{Z}\in\mathbf{R}^{\nu\times\nu},\mathcal{Z}=\mathcal{Z}^*,\mathcal{U}\in\mathbf{R}^{\nu_1\times\nu}$ *and* $\mathcal{V}\in\mathbf{R}^{\nu_2\times\nu}$ *with* ν,ν_1,ν_2 *positive integers. Consider the following basic linear inequality:*

$$\mathcal{Z}+\mathcal{U}^*\Theta\mathcal{V}+\mathcal{V}^*\Theta^*\mathcal{U}<0 \tag{8}$$

with the unknown variable $\Theta\in\mathbf{R}^{\nu_1\times\nu_2}$. *Then the following are equivalent:*
(i) There exists $\Theta\in\mathbf{R}^{\nu_1\times\nu_2}$ *solving (8);*

(ii)
$$\mathcal{W}_{\mathcal{U}}^*\mathcal{Z}\mathcal{W}_{\mathcal{U}}<0 \tag{9}$$
and
$$\mathcal{W}_{\mathcal{V}}^*\mathcal{Z}\mathcal{W}_{\mathcal{V}}<0, \tag{10}$$

where $\mathcal{W}_{\mathcal{U}}$ *and* $\mathcal{W}_{\mathcal{V}}$ *denote any bases of the null spaces* $\mathrm{Ker}\,\mathcal{U}$ *and* $\mathrm{Ker}\,\mathcal{V}$*, respectively.*

The next result provides necessary and sufficient conditions for the existence of a controller of type (4) solving the disturbance attenuation problem for the system (3).

Theorem 2 *For a* $\gamma > 0$*, the following are equivalent:*

(i) There exists a controller of order $n_c > 0$ *solving the disturbance attenuation problem with the level of attenuation* $\gamma > 0$ *for the system (5);*

(ii) There exist:

$$X = (X(1),...,X(d)) \in \mathcal{S}_n^{\ d}, X(i) > 0, i \in \mathfrak{D},$$

$$Y = (Y(1),...,Y(d)) \in \mathcal{S}_n^{\ d}, Y(i) > 0, i \in \mathfrak{D},$$

$$S = (S(1),...,S(d)) \in \mathcal{S}_n^{\ d}, S(i) > 0, i \in \mathfrak{D},$$

$$N = \left(N(1),...,N(d)\right),\ N \in \mathcal{M}_{n,n_c}^{\ d}$$

such that:

$$\begin{bmatrix} V_0^*(i) & V_1^*(i) \end{bmatrix} \mathcal{N}_i(X) \begin{bmatrix} V_0(i) \\ V_1(i) \end{bmatrix} < 0, \tag{11}$$

$$\begin{bmatrix} \Pi_{0,0}(i) & \Pi_{0,1}(i) & -U_1^*(i)N(i) & \cdots & -U_r^*(i)N(i) & \Pi_{0,r+1}(i) \\ \Pi_{0,1}^*(i) & -\gamma^2 I_{m_1} & 0 & \cdots & 0 & 0 \\ -N^*(i)U_1(i) & 0 & -S(i) & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ -N^*(i)U_r(i) & 0 & 0 & \cdots & -S(i) & 0 \\ \Pi_{0,r+1}^*(i) & 0 & 0 & \cdots & 0 & \Pi_{r+1,r+1}(i) \end{bmatrix} < 0 \tag{12}$$

$$\operatorname{rank} \begin{bmatrix} X(i) & I_n & 0 \\ I_n & Y(i) & N(i) \\ 0 & N^*(i) & S(i) \end{bmatrix} = n + n_c, \tag{13}$$

where

$$\begin{bmatrix} V_0(i) \\ V_1(i) \end{bmatrix}$$

is a basis of $\operatorname{Ker}\begin{bmatrix} C_0(i) & D_0(i) \end{bmatrix}$,

$$\begin{bmatrix} U_0(i) \\ \vdots \\ U_{r+1}(i) \end{bmatrix}$$

denotes a basis of $\operatorname{Ker}\begin{bmatrix} B_0^*(i) & \cdots & B_r^*(i) & D_{zu}^*(i) \end{bmatrix}$,

$$\mathcal{N}_i(X,i)=\begin{bmatrix}\mathcal{N}_{11}(X,i) & \mathcal{N}_{12}(X,i)\\ \mathcal{N}_{12}^*(X,i) & \mathcal{N}_{22}(X,i)\end{bmatrix},$$

$$\mathcal{N}_{11}(X,i)=A_0^*(i)X(i)+X(i)A_0(i)+\sum_{k=1}^{r}A_k^*(i)X(i)A_k(i)$$
$$+\sum_{j=1}^{d}q_{ij}X(j)+C_z^*(i)C_z(i)$$

$$\mathcal{N}_{12}(X,i)=X(i)G_0(i)+\sum_{k=1}^{r}A_k^*(i)X(i)G_k(i)+C_z^*(i)D_{zv}(i)$$

$$\mathcal{N}_{22}(X,i)=-\gamma^2 I_{m_1}+D_{zv}^*(i)D_{zv}(i)+\sum_{k=1}^{r}G_k^*(i)X(i)G_k(i)$$

$$\Pi_{0,0}(i)=U_0^*(i)\left[A_0(i)Y(i)+Y(i)A_0^*(i)+q_{ii}Y(i)\right]U_0(i)$$
$$+\sum_{k=1}^{r}U_0^*(i)Y(i)A_k^*(i)U_k(i)+U_0^*(i)Y(i)C_z^*(i)U_{r+1}(i)$$
$$+U_{r+1}^*(i)C_z(i)Y(i)U_0(i)+\sum_{k=1}^{r}U_k^*(i)A_k(i)Y(i)U_0(i)$$
$$-\sum_{k=1}^{r}U_k^*(i)Y(i)U_k(i)-U_{r+1}^*(i)U_{r+1}(i),$$

$$\Pi_{0,1}(i)=\sum_{k=0}^{r}U_k^*(i)G_k(i)+U_{r+1}^*(i)D_{zv}(i),$$

$$\Pi_{0,r+1}(i)=U_0^*(i)\left[I_n\ \ 0\right]\left[\sqrt{q_{i1}}\tilde{Y}(i)\ \cdots\ \sqrt{q_{i,i-1}}\tilde{Y}(i)\ \sqrt{q_{i,i+1}}\tilde{Y}(i)\ \cdots\ \sqrt{q_{i,d}}\tilde{Y}(i)\right]$$

$$\Pi_{r+1,r+1}(i)=-\mathrm{diag}\left(\tilde{Y}(1)\cdots\tilde{Y}(i-1),\tilde{Y}(i+1)\cdots\tilde{Y}(d)\right),$$

$$\tilde{Y}(i)=\begin{bmatrix}Y(i) & N(i)\\ N^*(i) & S(i)\end{bmatrix},\ i\in\mathcal{D}.$$

Proof. The outline of the proof is similar with the one in the deterministic framework. The stochastic feature of the considered system does not appear explicitly in the following developments of the proof. This feature appears only in the specific formulae of the Bounded Real Lemma.

(i)$\Rightarrow$(ii) Assume that it exists a controller of form (4) stabilizing the system (3) such that $\|\mathcal{T}_{cl}\|<\gamma$. Using the implication (i)$\Rightarrow$(ii) of Theorem 1 (Bounded Real Lemma) for the closed loop system it results that there exist:

$$X_{cl}=(X_{cl}(1),...,X_{cl}(d))\in\mathcal{S}_{n+n_c}{}^d,\ X_{cl}(i)>0$$

such that

$$\mathcal{N}_i(X_{cl},\gamma)<0\,, \tag{14}$$

where

$$\mathcal{N}_i\left(X_{cl},\gamma\right)=\begin{bmatrix}\left(\mathcal{L}_{cl}^{*}X_{cl}\right)(i)+C_{cl}^{*}(i)C_{cl}(i) & \mathcal{P}_i^{*}\left(X_{cl}\right)\\ \mathcal{P}_i\left(X_{cl}\right) & \mathcal{R}_i\left(X_{cl}\right)\end{bmatrix},$$

$$\left(\mathcal{L}_{cl}^{*}X_{cl}\right)(i)=A_{0cl}^{*}(i)X_{cl}(i)+X_{cl}(i)A_{cl}(i)$$
$$+\sum_{k=1}^{r}A_{kcl}^{*}(i)X_{cl}(i)A_{kcl}(i)+\sum_{j=1}^{d}q_{ij}X_{cl}(j),$$

$$\mathcal{P}_i\left(X_{cl}\right)=G_{0cl}^{*}(i)X_{cl}(i)+\sum_{k=1}^{r}G_{kcl}^{*}(i)X_{cl}A_{kcl}(i)+D_{cl}^{*}(i)C_{cl}(i),$$

$$\mathcal{R}_i\left(X_{cl}\right)=-\gamma^2 I_{m_1}+\sum_{k=1}^{r}G_{kcl}^{*}(i)X_{cl}(i)G_{kcl}(i).$$

Based on Schur complement arguments it is easy to see that (14) is equivalent with:

$$\begin{bmatrix}\left(\mathcal{L}_0^{*}X_{cl}\right)(i) & X_{cl}(i)G_{ocl}(i) & A_{1cl}^{*}(i)X_{cl}(i) & \cdots & A_{rcl}^{*}(i)X_{cl}(i) & C_{cl}^{*}(i)\\ G_{0cl}^{*}(i)X_{cl}(i) & -\gamma^2 I_{m_1} & G_{1cl}^{*}(i)X_{cl}(i) & \cdots & G_{rcl}^{*}(i)X_{cl}(i) & D_{cl}^{*}(i)\\ X_{cl}(i)A_{1cl}(i) & X_{cl}(i)G_{1cl}(i) & -X_{cl}(i) & \cdots & 0 & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\ X_{cl}(i)A_{rcl}(i) & X_{cl}(i)G_{rcl}(i) & 0 & \cdots & -X_{cl}(i) & 0\\ C_{cl}(i) & D_{cl}(i) & 0 & \cdots & 0 & -I_{p_1}\end{bmatrix}<0, \tag{15}$$

where $\left(\mathcal{L}_0^{*}X_{cl}\right)(i)=A_{0cl}^{*}(i)X_{cl}(i)+X_{cl}(i)A_{0cl}(i)+\sum_{j=1}^{d}q_{ij}X_{cl}(j).$

Let us introduce the following notations:

$$\tilde{A}_k(i)=\begin{bmatrix}A_k(i) & 0\\ 0 & 0\end{bmatrix},\tilde{G}_k(i)=\begin{bmatrix}G_k(i)\\ 0\end{bmatrix},\ 0\le k\le r,$$

$$\tilde{B}_0(i)=\begin{bmatrix}0 & B_0(i)\\ I_{n_c} & 0\end{bmatrix},\tilde{B}_k(i)=\begin{bmatrix}0 & B_k(i)\\ 0 & 0\end{bmatrix},\ 1\le k\le r,$$

$$\tilde{C}_0=\begin{bmatrix}0 & I_{n_c}\\ C_0(i) & 0\end{bmatrix},\tilde{C}_z(i)=\begin{bmatrix}C_z(i) & 0\end{bmatrix},$$

$$\tilde{D}_{zu}(i)=\begin{bmatrix}0 & D_{zu}(i)\end{bmatrix},\ \tilde{D}_0(i)=\begin{bmatrix}0\\ D_0(i)\end{bmatrix},$$

$$\Theta_c(i)=\begin{bmatrix}A_c(i) & B_c(i)\\ C_c(i) & D_c(i)\end{bmatrix},\ i\in\mathfrak{D}.$$

Using (6) one obtains:

$$\begin{aligned}
A_{kcl}(i)&=\tilde{A}_k(i)+\tilde{B}_k(i)\Theta_c(i)\tilde{C}_0(i)\\
G_{kcl}(i)&=\tilde{G}_k(i)+\tilde{B}_k(i)\Theta_c(i)\tilde{D}_0(i),\ 0\le k\le r,\\
C_{cl}(i)&=\tilde{C}_z(i)+\tilde{D}_{zu}(i)\Theta_c(i)\tilde{C}_0(i)\\
D_{cl}&=D_{zv}(i)+\tilde{D}_{zu}(i)\Theta_c(i)\tilde{D}_0(i), i\in\mathfrak{D}.
\end{aligned}$$

With the above notations one can easily see that (15) can be rewritten in the basic linear matrix inequality form:

$$\mathcal{Z}(i)+\mathcal{U}^*(i)\Theta_c(i)\mathcal{V}(i)+\mathcal{V}^*(i)\Theta_c^*(i)\mathcal{U}(i)<0,\ i\in\mathfrak{D}, \tag{16}$$

where

$$\mathcal{Z}(i)=\begin{bmatrix}
\left(\tilde{\mathfrak{L}}_0^*X_{cl}\right)(i) & X_{cl}(i)\tilde{G}_o(i) & \tilde{A}_1^*(i)X_{cl}(i) & \cdots & \tilde{A}_r^*(i)X_{cl}(i) & \tilde{C}_z^*(i)\\
\tilde{G}_0^*(i)X_{cl}(i) & -\gamma^2 I_{m_1} & \tilde{G}_1^*(i)X_{cl}(i) & \cdots & \tilde{G}_r^*(i)X_{cl}(i) & D_{zv}^*(i)\\
X_{cl}(i)\tilde{A}_1(i) & X_{cl}(i)G_1(i) & -X_{cl}(i) & \cdots & 0 & 0\\
\vdots & \vdots & \vdots & \ddots & \vdots & \vdots\\
X_{cl}(i)\tilde{A}_r(i) & X_{cl}(i)\tilde{G}_r(i) & 0 & \cdots & -X_{cl}(i) & 0\\
\tilde{C}_z(i) & D_{zv}(i) & 0 & \cdots & 0 & -I_{p_1}
\end{bmatrix} \tag{17}$$

$$\mathcal{U}(i)=\begin{bmatrix}\tilde{B}_0^*(i)X_c(i) & 0_{(m_2+n_c)\times m_1} & \tilde{B}_1^*(i)X_{cl}(i) & \cdots & \tilde{B}_r^*(i)X_{cl}(i) & \tilde{D}_{zu}^*(i)\end{bmatrix}$$

$$\mathcal{V}(i)=\begin{bmatrix}\tilde{C}_0(i) & \tilde{D}_0(i) & 0_{(p_2+n_c)\times\left[p_1+r(n+n_c)\right]}\end{bmatrix},\ i\in\mathfrak{D},$$

and

$$\left(\tilde{\mathfrak{L}}_0^*X_{cl}\right)(i)=\tilde{A}_{0cl}^*(i)X_{cl}(i)+X_{cl}(i)\tilde{A}_{0cl}(i)+\sum_{j=1}^{d}q_{ij}X_{cl}(j).$$

Therefore the existence of a stabilizing γ-attenuation controller for (3) is equivalent with the solvability of (16). Based on Lemma 1, (16) is feasible if and only if there exist:

$$\mathcal{W}_{\mathcal{U}}^{*}(i)\mathcal{Z}(i)\mathcal{W}_{\mathcal{U}}(i) < 0 \quad (18)$$

$$\mathcal{W}_{\mathcal{V}}^{*}(i)\mathcal{Z}(i)\mathcal{W}_{\mathcal{V}}(i) < 0,\ i \in \mathcal{D}, \quad (19)$$

where $\mathcal{W}_{\mathcal{U}}(i), \mathcal{W}_{\mathcal{V}}(i)$ denote bases of the null spaces of $\mathcal{U}(i)$ and of $\mathcal{V}(i)$, respectively. It is easy to see that a basis of the null space of $\mathcal{U}(i)$ is:

$$\mathcal{W}_{\mathcal{U}}(i) = \mathcal{X}^{-1}(i)\mathcal{W}_{\tilde{\mathcal{U}}}(i), \quad (20)$$

where

$$\mathcal{X}(i) = \text{diag}(X_{cl}(i), I_{m_1}, X_d(i), \ldots, X_{cl}(i), I_{p_1})$$

and $\mathcal{W}_{\tilde{\mathcal{U}}}(i)$ is a basis of the null subspace of the matrix:

$$\tilde{\mathcal{U}}(i) = \left[\tilde{B}_0^{*}(i)\ \ 0_{(m_2+n_c)\times m_1}\ \ \tilde{B}_1^{*}(i) \cdots \tilde{B}_r^{*}(i)\ \tilde{D}_{zu}^{*}(i)\right].$$

A basis of the null subspace of $\tilde{\mathcal{U}}(i)$ is

$$\mathcal{W}_{\tilde{\mathcal{U}}}(i) = \begin{bmatrix} T_0(i) & 0 & 0 & \cdots & 0 \\ 0 & I_{m_1} & 0 & \cdots & 0 \\ T_1(i) & 0 & L & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ T_r(i) & 0 & 0 & \cdots & L \\ U_{r+1}(i) & 0 & 0 & \cdots & 0 \end{bmatrix}, \quad (21)$$

where

$$T_k(i) := \begin{bmatrix} U_k(i) \\ 0 \end{bmatrix},\ 0 \le k \le r,\ L = \begin{bmatrix} 0 \\ I_{n_c} \end{bmatrix}$$

and

$$\begin{bmatrix} U_0(i) \\ \vdots \\ U_{r+1}(i) \end{bmatrix}$$

is a basis of the null subspace of the matrix:

$$\left[B_0^{*}(i)\ B_1^{*}(i) \cdots B_r^{*}(i)\ D_{zu}^{*}(i)\right].$$

A suitable choice for $\mathcal{W}_V(i)$ is the following:

$$\mathcal{W}_V(i)=\begin{bmatrix} V_0(i) & 0 \\ 0 & 0 \\ V_1(i) & 0 \\ 0 & I_{p_1+r(n+n_c)} \end{bmatrix}, \tag{22}$$

where $\begin{bmatrix} V_0(i) \\ V_1(i) \end{bmatrix}$ is a basis of the null subspace of the matrix $\begin{bmatrix} C_0(i) & D_0(i) \end{bmatrix}$.

Consider the partition of $X_{cl}(i)$:

$$X_{cl}(i)=\begin{bmatrix} X(i) & M(i) \\ M^*(i) & \tilde{X}(i) \end{bmatrix}$$

with $X(i)\in\mathbf{R}^{n\times n}$. Then by direct computations one obtains:

$$\mathcal{W}_V^*(i)\mathcal{Z}(i)\mathcal{W}_V(i)=\begin{bmatrix} \Psi_{0,0}(i) & \Psi_{0,1}(i) & \cdots & \Psi_{0,r}(i) & \Psi_{0,r+1}(i) \\ \Psi_{0,1}^*(i) & -X_{cl}(i) & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \Psi_{0,r}^*(i) & 0 & \cdots & -X_{cl}(i) & 0 \\ \Psi_{0,r+1}^*(i) & 0 & \cdots & 0 & -I_{p_1} \end{bmatrix}, \tag{23}$$

where

$$\Psi_{0,0}(i)=V_0^*(i)\left[A_0^*(i)X(i)+X(i)A_0(i)+\sum_{j=1}^{d}q_{ij}X(j)\right]V_0(i)$$
$$+V_0^*(i)X(i)G_0(i)V_1(i)+V_1^*(i)G_0^*(i)X(i)V_0(i)-\gamma^2V_1^*(i)V_1(i),$$
$$\Psi_{0,k}(i)=\left(\begin{bmatrix} V_0^*(i) & 0\end{bmatrix}\tilde{A}_k^*(i)+V_1^*(i)\tilde{G}_k^*(i)\right)X_{cl},\ 1\le k\le r,$$
$$\Psi_{0,r+1}(i)=V_0^*(i)C_z^*(i)+V_1^*(i)D_{zv}^*(i).$$

Using again Schur complement arguments it follows that condition (19) together with (23) is equivalent with:

$$\Psi_{0,0}(i)+\sum_{k=1}^{r}\Psi_{0,k}(i)X_{cl}^{-1}(i)\Psi_{0,k}^*(i)+\Psi_{0,r+1}(i)\Psi_{0,r+1}^*(i)<0.$$

Detailing the coefficients in the above inequality, (11) directly follows.

In order to explicit the condition (18) one first computes:

$$\mathfrak{X}^{-1}(i)\mathfrak{Z}(i)\mathfrak{X}^{-1}(i)=$$

$$=\begin{bmatrix} \left(\mathfrak{L}_0^*\tilde{Y}\right)(i) & \tilde{G}_0(i) & \tilde{Y}(i)\tilde{A}_1^*(i) & \cdots & \tilde{Y}(i)\tilde{A}_r^*(i) & \tilde{Y}(i)\tilde{C}_z^*(i) \\ \tilde{G}_0^*(i) & -\gamma^2 I_{m_1} & \tilde{G}_1^*(i) & \cdots & \tilde{G}_r^*(i) & D_{zv}^*(i) \\ \tilde{A}_1(i)\tilde{Y}(i) & \tilde{G}_1(i) & -\tilde{Y}(i) & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \tilde{A}_r(i)\tilde{Y}(i) & \tilde{G}_r(i) & 0 & \cdots & -\tilde{Y}(i) & 0 \\ \tilde{C}_z(i)\tilde{Y}(i) & D_{zv}(i) & 0 & \cdots & 0 & -I_{p_1} \end{bmatrix}, \quad (24)$$

where

$$\left(\mathfrak{L}_0^*\tilde{Y}\right)(i)=\tilde{A}_0(i)\tilde{Y}(i)+\tilde{Y}(i)\tilde{A}_0^*(i)+\sum_{j=1}^{d}q_{ij}\tilde{Y}(i)\tilde{Y}^{-1}(j)\tilde{Y}(i) \quad (25)$$

$$\tilde{Y}(i)=X_{cl}^{-1}(i). \quad (26)$$

Introduce the following notation:

$$\tilde{Y}(i)=\begin{bmatrix} Y(i) & N(i) \\ N^*(i) & S(i) \end{bmatrix}, Y(i)\in R^{n\times n}.$$

Using (21), (24), (25) and (20), (18) becomes:

$$\begin{bmatrix} \tilde{\Pi}_{0,0}(i) & \Pi_{0,1}(i) & -U_1^*(i)N(i) & \cdots & -U_r^*(i)N(i) \\ \Pi_{0,1}^*(i) & -\gamma^2 I_{m_1} & 0 & \cdots & 0 \\ -N^*(i)U_1(i) & 0 & -S(i) & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -N^*(i)U_r(i) & 0 & 0 & \cdots & -S(i) \end{bmatrix}<0, \quad (27)$$

where

$$\tilde{\Pi}_{0,0}(i)=U_0^*(i)\Big\{A_0(i)Y(i)+Y(i)A_0^*(i)+\sum_{j=1}^{r}q_{ij}\begin{bmatrix} Y(i) & N(i)\end{bmatrix}\tilde{Y}^{-1}(j)\begin{bmatrix} Y(i) \\ N^*(i)\end{bmatrix}\Big\}U_0(i)$$

$$+\sum_{k=1}^{r}U_0^*(i)Y(i)A_k^*(i)U_k(i)+U_0^*(i)Y(i)C_z^*(i)U_{r+1}(i)+U_{r+1}^*(i)C_z(i)Y(i)U_0(i)$$

$$+\sum_{k=1}^{r}U_k^*(i)A_k(i)Y(i)U_0(i)-\sum_{k=1}^{r}U_k^*(i)Y(i)U_k(i)-U_{r+1}^*(i)U_{r+1}(i).$$

By Schur complement arguments it follows that (27) is equivalent with (12) Further, taking into account that:

$$\operatorname{rank}\begin{bmatrix} X(i) & I & 0 \\ I & Y(i) & N(I) \\ 0 & N^*(i) & S(i) \end{bmatrix} =$$

$$\operatorname{rank}\begin{bmatrix} X(i)-\left(Y(i)-N(i)S^{-1}(i)N^*(i)\right)^{-1} & 0 & 0 \\ 0 & Y(i)-N(i)S^{-1}(i)N^*(i) & 0 \\ 0 & 0 & S(i) \end{bmatrix},$$

$S(i) > 0$ and $Y(i)-N(i)S^{-1}(i)N^*(i) > 0$, it follows that (26) gives

$$X(i) = \left(Y(i)-N(i)S^{-1}(i)N^*(i)\right)^{-1},$$

from which (13) directly follows.

(ii)$\Rightarrow$(i) Assume that there exist $X(i), Y(i), N(i)$ and $S(i)$ verifying (11)-(13). By (12) it follows that $\Pi_{r+1,r+1}(i) < 0$ namely

$$\tilde{Y}(i) = \begin{bmatrix} Y(i) & N(i) \\ N^*(i) & S(i) \end{bmatrix} > 0$$

and therefore $\tilde{Y}(i)$ is invertible. Moreover $\tilde{Y}^{-1}(i)$ has the structure

$$\begin{bmatrix} X(i) & * \\ * & * \end{bmatrix},$$

where by $*$ the irrelevant entries have been denoted. From the developments performed to prove the implication (i)$\Rightarrow$(ii) it follows that (18) and (19) are verified by

$$X_{cl}(i) = \tilde{Y}^{-1}(i)$$

and hence (16) has a solution which fact guarantees the existence of a stabilizing and γ -attenuating controller. Thus the proof ends.■

When the existence conditions stated in part ii) are accomplished, the construction of the γ-attenuating controller of imposed order n_c is made according with the proof of part ii)$\Rightarrow$i), by solving (16) with respect to $\Theta_c(i), i \in \mathfrak{D}$.

4. Conclusions

In this paper the disturbance attenuation problem for stochastic systems subjected both to multiplicative white noise and to Markovian jumps has been considered. The solution of this problem is determined in the set of deterministic controllers of fixed order. Necessary and sufficient solvability conditions are derived in terms of some specific matrix inequalities which solution allows computing the γ -attenuating controller.

References

Boyd, S., L. El-Ghaoui, E. Feron and V. Balakrishnan (1994). Linear matrix inequalities in systems and control theory. *Studies in Applied Mathematics, SIAM*, Philadelphia, PA, **Vol. 15**.

Doyle, J.C., K. Glover, P. Khargonekar and P. Francis (1989). State-space solutions to standard H_2 and H_∞ control problem. *IEEE-Trans. Autom. Control*, **Vol. 34**, 831-848.

Drăgan, V. and T. Morozan (1997). Global solutions to a game-theoretic Riccati equation of stochastic control. *Journal of differential equations*, Vol. 138, 328-350.

Drăgan, V., A. Halanay and A. Stoica (1998). A small gain and robustness for linear systems with jump Markov perturbations. *Proceedings of Mathematical Theory of Networks and Systems (MTNS), Padova, Italy, 6-10 July.*

Drăgan, V. and T. Morozan (2001). Game-theoretic coupled Riccati equations associated to controlled linear differential systems with jump Markov perturbations. *Stochastic Analysis and Applications*, **Vol. 19**, 715-751.

Drăgan, V., T. Morozan and A. Stoica (2003). A stochastic version of the Bounded Real Lemma. To appear in *Mathematical Repports.*

Friedman, A. (1975). Stochastic differential equations and applications. *Academic Press.*

Gahinet, P. and P. Apkarian (1994). A linear matrix inequality approach to H_∞ control. *International Journal Robust and Nonlinear Control*, **No. 4**, 421-448.

Hinrichsen, D. and A.J. Pritchard. Stochastic H^∞ (1998). *SIAM J. Control and Optim.*, **36**, 1504-1538.

Petersen, I.R., V.A. Ugrinovskii and A.V. Savkin (2000). Robust Control Design Using H^∞ Methods. *Springer Verlag.*

Stoica, A. and I. Yaesh (2002). Robust H^∞ control of wings deployment loops for an uninhabited air vehicle-the jump Markov model approach. *Journal of Guidance Control and Dynamics*, **Vol. 35**, 407-411.

CONCEPTUAL STRUCTURAL ELEMENTS REGARDING A SPEED GOVERNOR FOR HYDROGENERATORS

Toma Leonida Dragomir and Sorin Nanu

"Politehnica" University of Timisoara
Dept. of Automation and Industrial Informatics
Bd. V. Parvan 2, RO-1900 Timisoara, Romania
E-mai : dragomir@aut.utt.ro

Abstract: This paper was conceived in the circumstances that literature references regarding speed governors for hydro-generators are either qualitative when come from producers, or uppermost theoretical when come from academic environment. The paper applies to governors produced by U.C.M. Reşiţa that are in operation since many years, designed jointly with Dep. of Automation from "Politehnica" University in Timişoara. Are presented base elements that allowed designing of governor that stands on actual international norms, and elements regarding a previous analysis using simplified models for external blocks.

Keywords: hydro-generator, speed governor, structure, design

1. Introduction

In principle, the speed control relies on two kinds of mechanic-electric interactions that concern the turbine-generator group: interactions that originate in changing of mechanical power (opening of blades and impact process with water) and interactions that originate in exchanging the electrical power with exterior. Hereby:

- The turbine blades opening-closing operation, performed by wicked gates has, from a systemic point of view, the meaning of control action. It determines the changing of power both carried to group p_m as a result of interaction with exterior, and transmitted inside the group p_G. The control variable is the stroke y of wicked gates.

- The changing of electrical power p_G exchanged by the generator with exterior, as a result of the external processes of the generator, can be considered a disturbing, measurable action of load type.

In order to maintain the speed variations of turbine-generator group, determined by the electrical stress of generators, i.e. the variation of power p_G, between normal limits, the automatic change of the position of stator blades is performed. The operation is accomplished by a control system that automatically strikes in the position of stator blades, with the aid of wicked gates. A simplified block diagram of such system is depicted in fig.1.

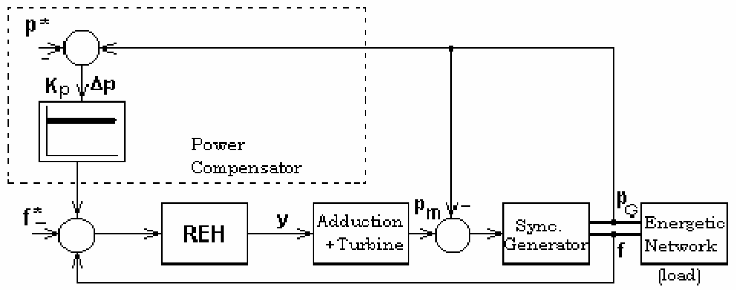


Figure 1. Fundamental block diagram for speed control by power compensator

The automatic intervention of electro-hydraulic controller (REH) is performed, on one side, depending on the rotational speed, (sensed as changing in frequency f from prescribed value f^*), and on other side, depending on the changing of power p_G (measured) from prescribed value p^*. The intervention depending on power has two characteristics: the first is an anticipative compensator characteristic, used for attenuation and damping of the effects of rotational speed variations, the second is the locking of power variations in a vicinity of prescribed value for it.

In the steady state regime, the two deviations that determine the control of wicked gates, i.e. frequency deviation $\Delta f = f\text{-}f^*$ and the issued power deviation $\Delta p = p_G - p^*$, are in equilibrium. The ratio of deviations is characterized by a proportionality coefficient b_{pp}.

$$\Delta p = -b_{pp}^{-1}\,\Delta f \ . \qquad (1)$$

The relation (1) reflects the compensation effect mentioned before: Δp represents the deviation of power issued from prescribed value, due to the deviation Δf of frequency in energetic network from prescribed value. For speed control, the possibility of adjusting b_{pp} and maintaining the proportionality (1) are of fundamental importance. Within this context, power speed droop is considered.

In accordance to fig.1, the automatic interventions depending on frequency / rotational speed deviations and power deviations can be achieved with the aid of feedbacks brought from the group shaft, respectively from generator bars. At the start-up, up to the moment of connecting the generator to network, because there is no power issued, the feedback after the blades position is used instead of power feedback (fig.2). In this case, the problem of droop is considered relatively to

$$\Delta y = -b_p \, \Delta f \tag{2}$$

between frequency deviation and the variation of blades position $\Delta y = y - y^*$. The opening speed droop is considered in this case.

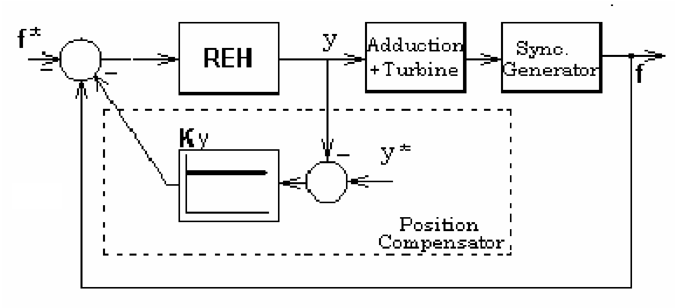


Figure 2. A structure of speed control system with compensating block after opening

The usual structures are variable; they switch from a start-up structure, with an opening speed droop, to a permanent structure, with power speed droop. In principle, is possible to maintain both feedbacks even after start-up, the contribution of opening feedback being much reduced that power feedback.

In order to increase transient performances of the controller with variable structure, i.e. the accelerating of feedback, the signals Δf and y are additionally processed. The result is a so-called tachometric with transient droop structure. They present, besides the permanent speed droop, an extra dynamic feedback, proportional to the stroke of wicked gates of main servomotor. Its effect exponentially dumps, reaching a zero value in steady state. The output of this block (temporary droop signal) is used for dynamic correction of the position of electro-hydraulic servo-system. In fig.3 the corresponding extension of fig.2 is presented. The simplest implementation block diagram for opening temporary speed droop is presented in fig.4.

In accordance to (IEC, 1970), (IEC, 1997), the opening temporary speed droop coefficient b_t is the slope of static characteristic of speed depending on opening of the controller, in a point of operation, when permanent droop is zero. In steady state regime, the diagram in fig.5 accomplishes:

$$\Delta y = (b_t + b_p)^{-1} \, \Delta f \tag{3}$$

and for b_p=0 it becomes

$$\Delta y = b_t^{-1} \Delta f . \quad (4)$$

In diagram, T_d represents the temporary time constant.

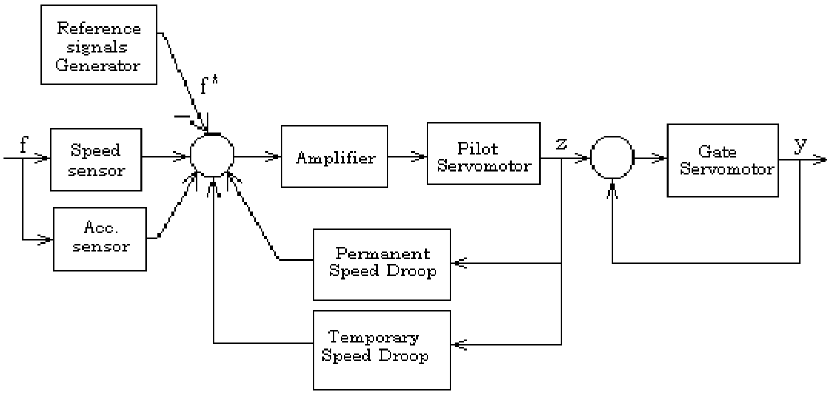


Figure 3. Tachometric structure with opening temporary speed droop

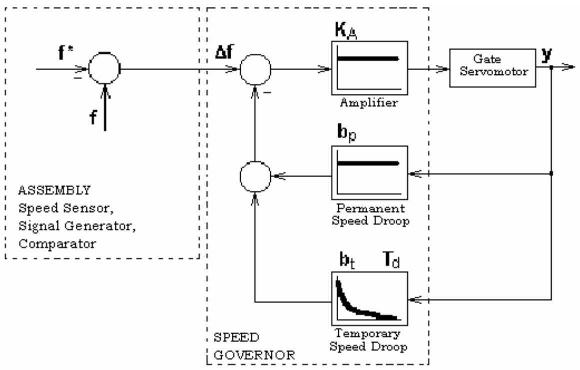


Figure 4. Speed control structure with opening permanent and temporary speed droop

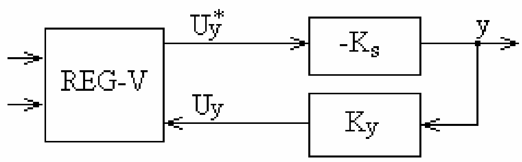


Figure 5. Block diagram of the ensemble speed governor-stabilized servomotor

Due to practical reasons for correct adaptation of group to particular conditions in working node, is compulsory for the coefficients b_p, b_t, and T_d to be adjustable. This represents a complex problem in analogical implementations, but yet simple in digital implementations.

In (Nanu, 2003) is represented a broad analysis of speed controller's structure, expanded more or less detailed in bibliographic sources as: international standards and norms, company ads and offers, articles. The approaches vary from principles, general structures to detailed, theoretical or particular structures, but irrelevant due to their simplicity or linearization considered.

Within this context, a presentation of some conceptual aspects of a speed governor structure, developed in "Politehnica" University of Timişoara together with UCM Resița is considered useful. The structure was used for hydro-generators in 3 MVA – 7 MVA and 35 MVA – 55 MVA domains of power. Initially it was implemented in analogical variant that is already in use since 4 years. Consequently, it was developed in digital variant. The purpose of this paper is to present the ensemble of the controller, together with some details, containing original elements.

2. The main structural conceptual elements

2.1. Design hypothesis

The basic hypothesis the speed governor structure was designed on (the structure contains the blocks that realizes the opening permanent droop, opening temporary droop and power permanent droop), are:

- Electro-hydraulic servo-system is stabilized with a schematic like in fig.3. The details of block "Servo-system stabilizer" do not represent the purpose of this paper. In steady state, the stabilized servo-system is corresponding to a schematic like in fig.5, constituted by two proportional elements with gain K_S and K_y. The input signal U_y^* is the reference from speed governor REG-V, and the feedback U_y is the output from measurement element for y.
- The REG-V structure has to allow the independent adjustment of b_{pp}, b_p, b_t, and T_d parameters.
- The dependency between the opening of wicked gates and power at constant speed has the aspect in fig.6. In reality, there is a family of complex non-linear characteristics. The idealization like in fig.6 permits the understanding of phenomena.

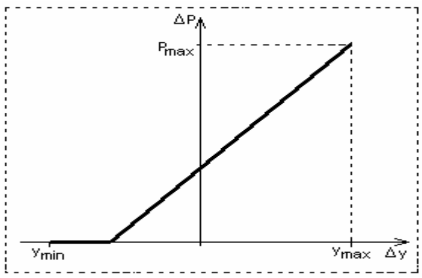


Figure 6. Dependency between the power issued by generator and stroke

- No constructive elements that protect (technical and partially conceptual) the governor and control system are detailed: the limiting elements on different channels of governor, anti-wind-up elements, the protective element for bumps when switching from a structure to another (e.g. from the manual to automatic regime, from opening permanent droop to power permanent droop structure).

2.2. Integrator type-loop and droop providing

The governor structure developed is presented in fig.7. (Vancea, 1998). Following will be explained how this structure ensures, for the loop containing block c_2, an integrator behavior that means in steady state automatically $U_a = 0$. Practically, the value of U_a will be very small and hence the (1), (2) and (3) conditions will be carried out.

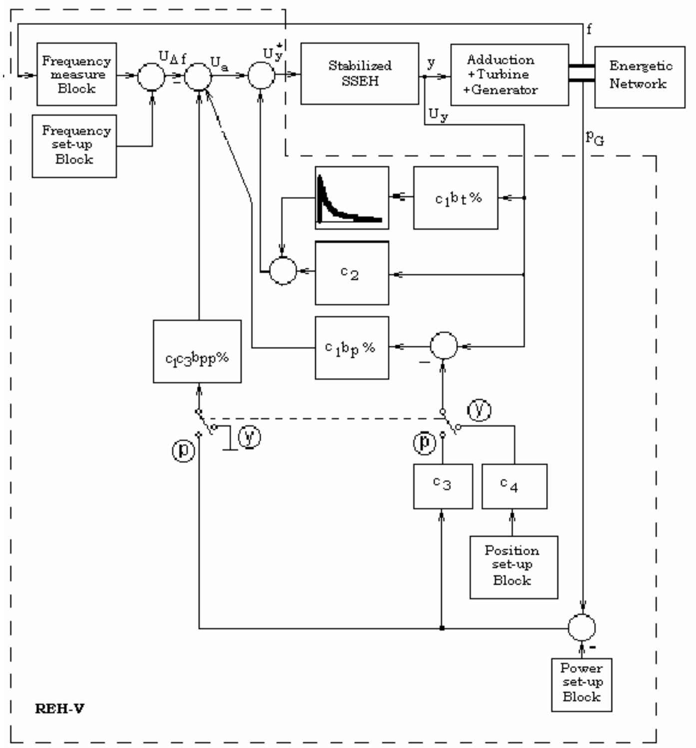


Figure 7. Implemented governor structure

An integrator type system behaves in principle like an integrator block, and the idea of using them in this application is due to the necessity of fulfilling the IEEE steady-state conditions. (IEC, 1997). The integrator character can be accomplished in different manners. One of this starts from structure in fig.8a, where the blocks S_1 and S_2 are each stable and have the transfer coefficients K_1 and K_2 so that

$$K = K_1 \ K_2 = 1 . \tag{4}$$

K_1 and K_2 are constants if S_1 and S_2 are linear. If the systems are non linear, then K_1 and K_2 depend on the point of operation. Hence, if the S_1 and S_2 have the characteristics $y_1=f_1(u_1)$, respectively $y_2=f_2(u_2)$, then, in order to fulfill condition (4), f_2 will be the inverse of f_1 ($f_2=f_1^{-1}$). Must be observed that the feedback is positive (a more detailed elaboration is in (Dragomir, *et al*., 2001), (Dragomir, 2002)).

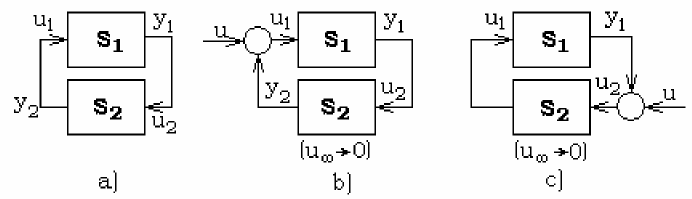


Figure 8. Integrator type connection between two blocks

Qualitatively, the integrator character is associated to the fact that gain K of the loop in fig.8a is 1, and, consequently, when an external signal u is applied like in fig.8b or 8c, the loop can reach an equilibrium state, or can be in a steady state only if the input u is zero. Otherwise, theoretically, the signals in system can vary with infinite amplitude, but practically- in case of physical systems- they will determine the saturation of the output of at least one system S_1 or S_2. Quantitatively, the problem can be treated more general, more or less harshly, as the systems are linear or non-linear.

The loops with integrator character can be used, in principle, to control any v signal that contributes to serial transfer of information within the loop. The principle is depicted in fig.9a where S_1 and S_2 are the systems in fig.8b or 8c, $v = y_1$ is the signal that must be controlled and w is the prescribed signal. In this case, the difference w-v plays the role of u.

$$u = w - v \tag{5}$$

If, with the feedback from y_1, the scheme obtained is stable, then for K = 1 it automatically fulfils, in steady state, the condition:

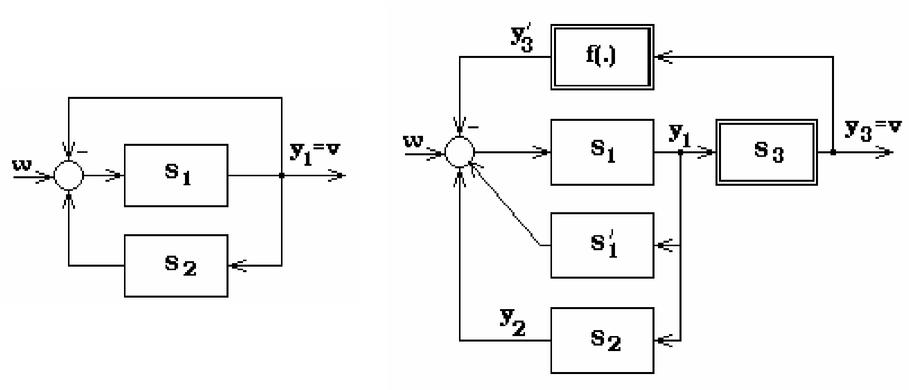
$$v = y_1 = w, \tag{6}$$

and v takes exactly the prescribed value. Practically, $K \neq 1$ so that the schematic accomplishes in steady state

$$v = y_1 = \left(1 - \frac{1-K}{K_1} \right)^{-1} w , \tag{7}$$

that means a control error in steady state

$$\varepsilon_v = \frac{w - v}{w} = \frac{K - 1}{K_1}. \tag{8}$$



a b

Figure 9. Integrator character loops for stationary control of a signal in serial chain

In fig.9b there is an extension of idea from fig.8 used for building the governor in fig.7. S_3 is a stable, inertial subsystem, with the steady state transfer characteristic

$$y_3 = f_3(y_1). \tag{9}$$

The structure serves to control either $v = y_3$ to value

$$v = y_3 = f^{-1}(w) \tag{10}$$

using a command

$$y_1 = f_3^{-1}(y_3) \tag{11}$$

generated by integrator type S_1-S_2 loop, or y_1 to value

$$y_1 = f_3^{-1}(f^{-1}(w)). \tag{12}$$

In the same time, the extension in fig.9b appends to structure in fig.9a the derivative system S_1', intended to accomplish the opening temporary droop. With this structure the dynamic of integrator type loop is modified while the dependencies (10) and (11) or (12) are maintained. A qualitative analysis of these structures is provided by (Nanu, 2003)

In accordance to IEE standard and equation (2), the opening permanent droop is calculated with the formula:

$$b_p = -\left[\Delta(\frac{f}{f_n})\right]^{-1}\left[\Delta(\frac{y}{y_n})\right]. \tag{2'}$$

The controller in fig.7 allows the controlled adjustment by operator of the b_p droop with the formula (13) where $b_{p\%} = 100\ b_p$, and $\alpha \cdot b_{p\%} = 1.6\ f_n / \Delta y_n$.

$$\Delta y = -\frac{1}{\alpha \cdot b_{p\%}}\, u_{\Delta f}\,. \tag{13}$$

This formula depicts the dependency between the voltage variation $u_{\Delta f}$ at the output of frequency measurement block, which calculates the network frequency variation, and the stroke Δy of the wicked gates servomotor. Within this dependency, the $b_{p\%}$ droop is adjustable and it influences in inverse proportion the variations of y. The values of $b_{p\%}$ were {1,2,10} in the application.

Alike formula (13) and in accordance to definition (4) of $b_{t\%}$ the diagram in fig.7 accomplishes, for $b_{p\%} = 0$ and $T_d \to \infty$, for the steady state dependency $b_{t\%} = 100\ b_t$:

$$\Delta y = -\frac{1}{\alpha \cdot b_{t\%}} \cdot \Delta u_{\Delta f}\ . \tag{14}$$

Here, $b_{t\%} = 100\ b_t$. In accordance to main operation regimes (loaded and unloaded) that require different dynamic for speed governor, in the applications

- $b_{tg,s\ \%} \in$ {5%, 10%, ...,50%} and $T_d \in$ {0.1, 0.2, ..., 4.9s} for loaded operation mode,
- $b_{tg,s\ \%} \in$ { 10%, 20%, ..., 100%} and $T_d \in$ {0.1, 0.2, ..., 9.9s} for unloaded operation mode.

The block that achieves the power droop was designed after the following principle: the power issued by generator is a function of opening y, in accordance to a functional dependency p(y), which is specific for every dam (fig.6). Considering that the characteristic can be practically identified with a proper precision in order to be inversed over the active domain, the inverse dependency $y=f^{-1}(p)$ can be used to carry out an extra power feedback, that compensate the opening permanent droop, no matter about its value. Using this feedback together with power droop feedback, a diagram is obtained that ensures either the power droop as per standards, and allows the shockless switching of governor from opening droop to power droop. The switching is performed at the moment the loading of generator is started. The system has the structure in fig.10. The behavior of scheme as a structure with only power droop is obtained by proper design of K_{pp} (parting into two branches, one for power droop and the other for opening droop).

K_{mp} is the gain of the power measuring element and K_{pp} is the gain inside governor for power. Starting point in building K_{pp} is the use of schematic in fig.11 instead of fig.10 due to the use of p = f(y) dependency. With $K_{(\bullet)}$

different gains were marked. In order to emphasize the fact that in real system there is no physical connection Δy→p a dotted line was used.

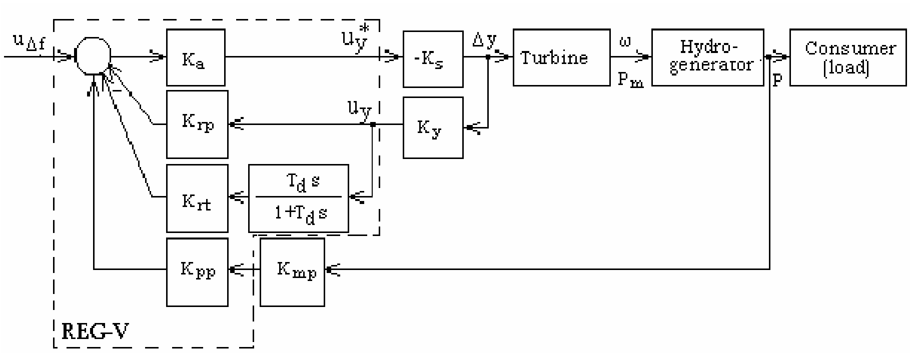


Figure 10. Structure of REG-V with opening and power feedback

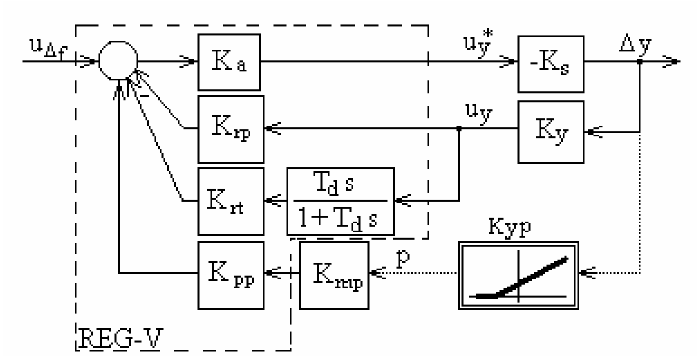


Figure 11. The structure of REG-V with opening and power feedback using p=f(y)

In (Nanu, 2003) is shown that with

$$K_{pp} = \frac{\alpha}{K_{mp} \cdot K_{yp}} \cdot b_{pp}\% - \frac{K_y}{K_{mp} \cdot K_{yp}} \cdot \frac{\alpha}{K_y} \cdot b_p\%\,, \tag{15}$$

where

$$b_{pp\%} = 100 \cdot b_{pp}\,, \tag{16}$$

and considering the interactions inside the speed control system, the structure in fig.12 is obtained.

In this structure appear only the channels that carry out the power droop according to (17) (and the opening transient droop). It is adjusted through coefficient $b_{pp\%}$.

$$\Delta y = -\frac{1}{\alpha \cdot b_{pp\%}}\,\Delta u_{\Delta f}\,. \tag{17}$$

The values of $b_{pp\%}$ are in the set {1%, 2%, ..., 12%}.

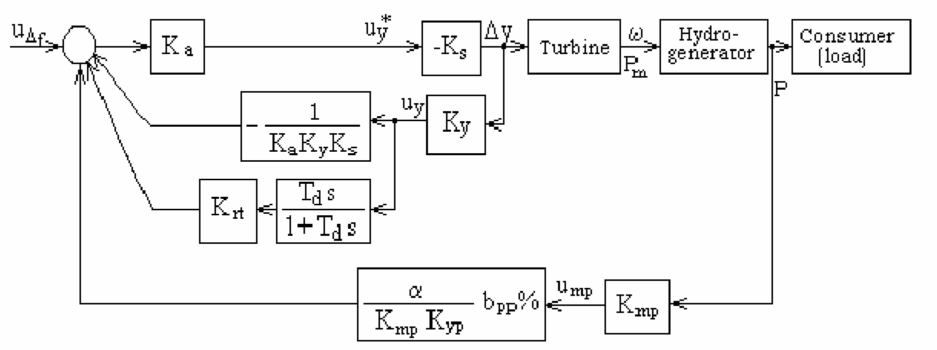


Figure 12. Structure of K_{pp} block

3. The analysis of speed (frequency) control system

In order to simulate the operating of speed governor, its model was included into the control system presented in fig.13. In this figure, the following abbreviations were used: SSEH- stabilized servo-system, SAT-adduction-turbine system, GS-synchronous generator. Other notations: C_m – active torque, Cr-component of resistant torque corresponding to interaction with electro-energetic system, Cr1- component of resistant torque corresponding to local consumers, U- the voltage at generator bars, ω- the angular frequency of U, u_{wy} and u_{wp} - references for y and p, issued by “opening assignment” and “power assignment” in fig.7.

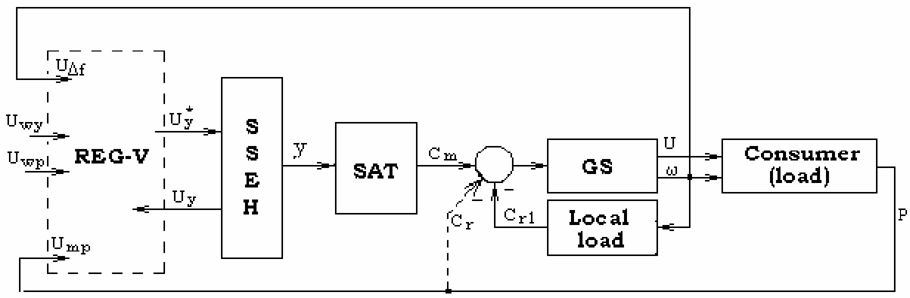


Figure 13. Structure used for studying of REG-V behavior

In fig.14 is presented a Simulink model used for simulation of system in fig.13. Obviously, it is a simplified model. The signals accompanied by symbol „ ^ ” are represented in norm values. The signal u_{wx} corresponds to references u_{wy} and u_{wp}. The block REG-V and SSEH includes both the structure REG-V presented before and stabilized electro-hydraulic servo-system.

The Simulink model of SSEH is presented in fig.15. The block SAT is dived into two parallel channels SAT-1 and SAT-2. It is a non-minimum phase system and the division offered simulations facilities. The consumer and local load are included in blocks Ext.Net., Inertia and Ext.Net. and Elas (elastically). Signal df^v represents the frequency shocks that can appear in system and wdf signal is the frequency reference. "Norm" block is meant to norm the signal y, while "G.S.-Inertia" block characterizes the rotor inertia. The electro mechanic time constant considered is 6.9 seconds.

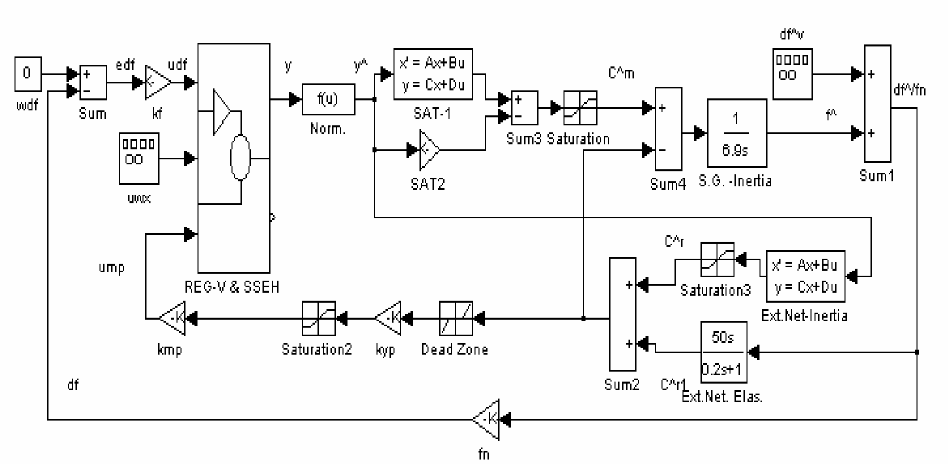


Figure 14. Simulink model used for simulation of structure in fig.13

The structure in fig.15 depicts the stabilization manner of electro-hydraulic servo-system. The model, as the one in fig.14, is strong non-linear. The significance of parameters is like in (Dragomir, *et al*., 1996). In comparison to this paper, some changes concerning integrator blocks not connected to wind-up phenomena were performed (Nanu, 1997).

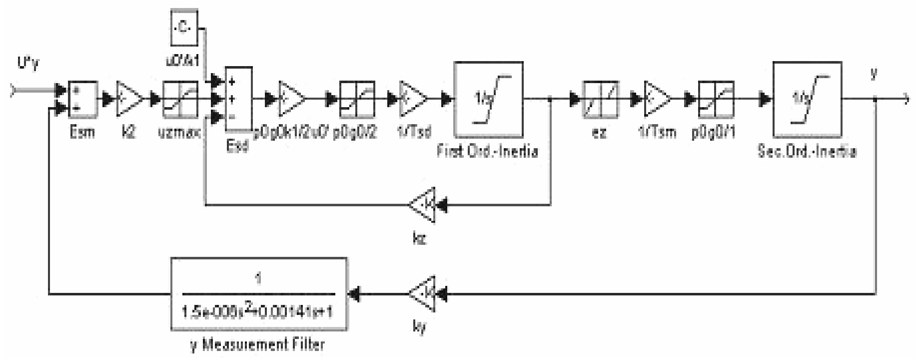
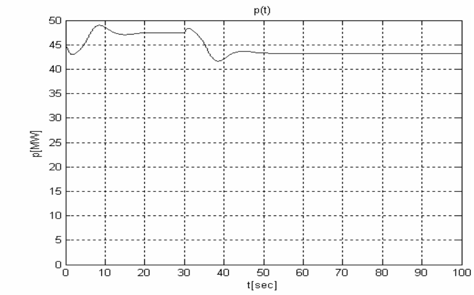


Figure 15. Simulink model of SSEH

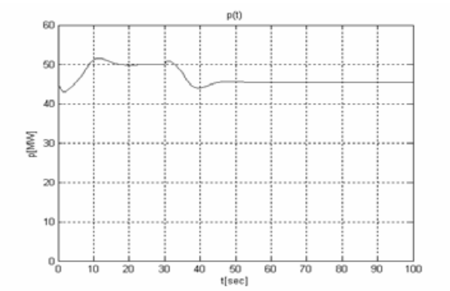
In next figures some results obtained by simulation are presented. They correspond fully qualitative and most quantitative to real behavior of control systems.

Fig.16 and fig.17 depicted the behavior of the system when $b_{pp\%} = 11\%$, respectively $b_{pp\%} = 10\%$ and the power reference changes from 44MW to

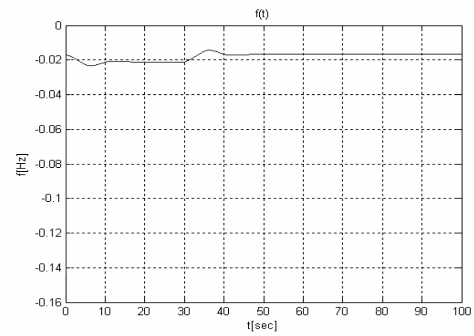
47MW at t=0 sec. and back to 44 MW at t=30 sec. Are recorded the issued power (p), norm of frequency (f^), and the opening of wicked gates (y). Some highlights: non-minimum phase character, hysteretic effect at SSEH level (fig. 16a and 17a), appearance of steady state errors (fig. 16b and 17b) due to considering an external power system comparable to the generator, capability of SSEH to bear the stress without touching the saturation (fig.16c and fig.17c) [-0.18 m, 0.18 m].



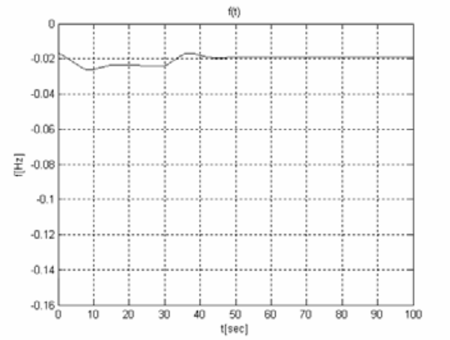
- a -



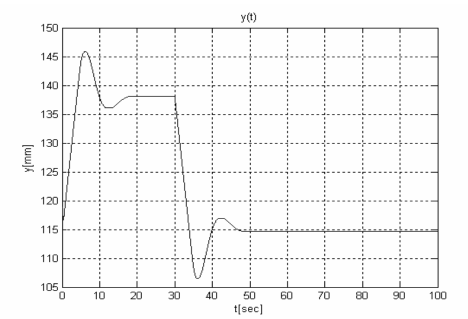
- a -



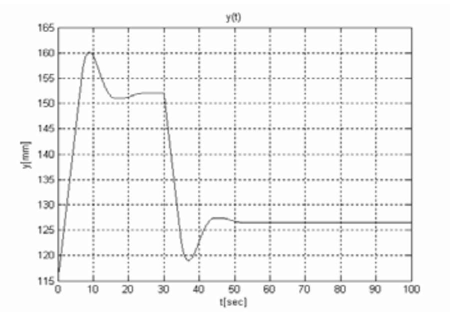
- b -



- b -



- c -

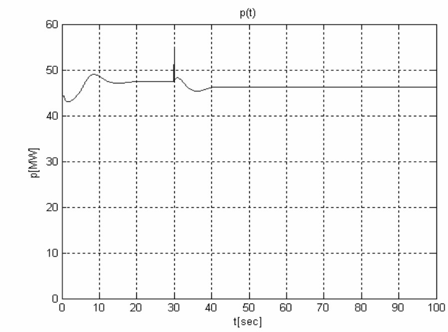


- c -

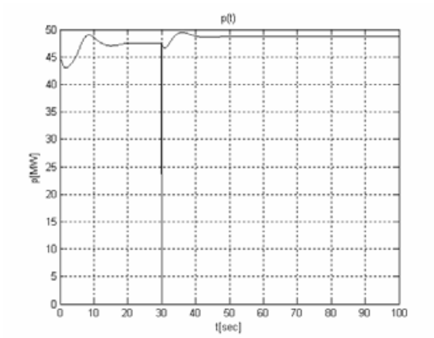
Figure 16. System behavior for $b_{pp} = 11\%$ *Figure 17.* System behavior for $b_{pp} = 10\%$

Fig.18 and fig.19 show the behavior of system in following conditions: initially, system operates at equilibrium, at nominal power of 44MW and frequency 49.984 Hz. At t=0 moment, the prescribed power changes to 47

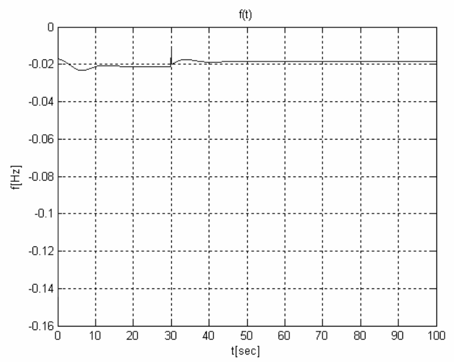
MW and at t=30 sec a rapid change is produced in network frequency with 1Hz (fig.18) or –0.5Hz (fig.19). Up to 30 seconds, the behavior is already known. Following-up, can be noticed, aside from non minimum phase system behavior, the capability of group to contribute to frequency correction (the group has no the capability to correct by itself the frequency deviation) and also the changes of y away from saturation, that means its ability to maintain speed to variations even greater of network frequency.



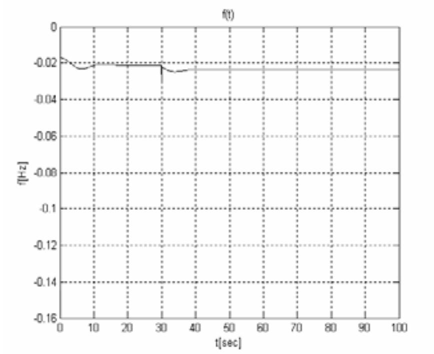
- a -



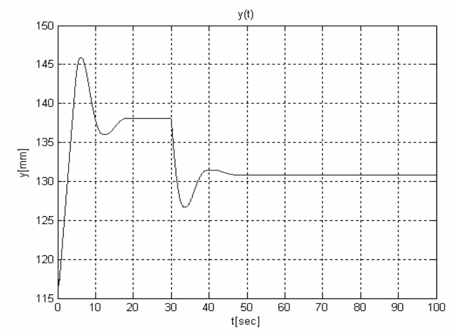
- a -



- b -

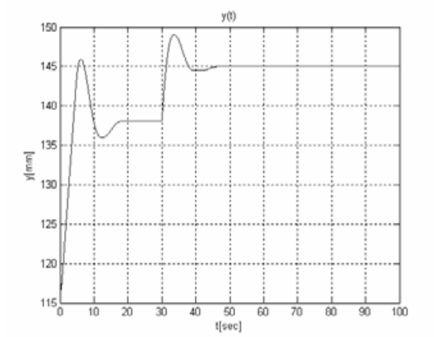


- b -



- c -

Figure 18. System behavior for +1 Hz change



- c -

Figure 19. System behavior for -0.5 Hz change

4. Conclusions

In order to obtain a speed controller for small medium power hydro generators, different structures of governors are used. Generally, they are qualitatively reasoned and companies protect information.

The paper presents a structure of governor used in application by U.C.M. Reşiţa DCP in two hydroelectric power plants in Romania and Turkey. Structure is based on stabilizing the electro-hydraulic servo-system and on the principle of integrator type main loop. This principle allows the control of every parameter imposed by norms: power speed droop, opening permanent speed droop, opening transient speed droop.

The design results were validated in practice. For theoretical validation, governor was simulated through a model very near to practical one. The adduction-turbine, synchronous generator and external (load) system have an approximate model. The simulation scenarios show an allowed behavior.

Acknowledgements

The authors want to acknowledge the team from U.C.M. Reşiţa –DCP, especially Mr. Ioan Grando and Stefan Lozici, for the possibility offered to cooperate in achieving (design, production, testing, calibration) the speed governors, respectively for the special quality of communication.

References

Dragomir, T.L., Dranga, O., and Nanu, S. (1996). About an Electro hydraulic speed governor structure*,* In *Proceedings of National Conference on System Theory*, SINTES 8, 87-95, Craiova.

Dragomir, T.L.and Nanu, S. (2001). About some integrator feature structures, In *Buletinul Stiintific al UPT*, *Seria Automatica si Calculatoare*, 46 (60),, 11-16.

Dragomir, T.L. (2002). Proportional Elements And Proportional Feedbacks, *Periodica Politechnica, Transactions on AC & CS*, 47 (61), 89-92.

International Electrotechnical Commission IEC, (1970). *International code for testing of speed governing systems for hydraulic turbines*, Bureau Central de la Commision Electrotechnique Internationale, Geneve, Suisse.

International Electrotechnical Commission, IEC61362 (1997). *Guide to specification of hydroturbine control system, Final Draft International Standard, International code for testing of speed governing systems for hydraulic turbines*, Bureau Central de la Commision Electrotechnique Internationale, Geneve, Suisse.

Nanu, S. (1997). Double integrator element with limitation protected against reset wind-up, *Buletinul Stiintific UPT*, 42 (56), 91-97.

Nanu, S. (2003). Contribuţii la dezvoltarea unor structuri de regulatoare de viteză pentru hidrogeneratoare, Phd Thesis, UPT.

Vancea, F. (1998). Digital speed governor for hydro-generators, Diploma Thesis, UPT.

TOWARDS INTELLIGENT REAL-TIME DECISION SUPPORT SYSTEMS FOR INDUSTRIAL MILIEU

F. G. Filip, D. A. Donciulescu
The National Institute for R&D in Informatics-ICI
Bucharest, Romania

Cr. I. Filip
Academy of Economic Studies, School of Management, Bucharest
Bucharest, Romania

Abstract Decision support systems (DSS) are human-centered information systems meant to help managers placed on different authority levels to make more efficient and effective decisions for problems evincing an imperfect structure. These systems are very suitable information tools to apply to various management and control problems that are complex and complicated at the same time. Several issues concerning the modern trends to build anthropocentric systems are reviewed. Then the paper surveys several widely accepted concepts in the field of decision support systems and some specific aspects concerning real-time applications. Several artificial intelligence methods and their applicability to decision-making processes are reviewed next. The possible combination of artificial intelligence technologies with traditional numerical models within advanced decision support systems is discussed and an example is given.

Keywords: artificial intelligence, decision, human factors, manufacturing, models

1. Introduction

The role and place of the human operator in industrial automation systems started to be seriously considered by engineers and equally by psychologists towards the middle of the 7th decade. Since then, this aspect has been constantly and growingly taken into consideration in view of famous accidents of highly automated systems and of incomplete fulfillment of hopes put in CIM systems [Martenson, 1996; Johanson, 1994].

The evaluation of the place of man in the system has known a realistic evolvement, triggered not only by practical engineer experience but also by the debates from academia circles. A long cherished dream of automatic engineers, that of developing "completely automated systems where man would be only a consumer" or "unmanned factories", tends to fade away – not only due to *ethical* or *social motivations*, but more important because the technical realization of this dream proved to be *impossible.*

A possible solution seems to be the use of artificial intelligence methods (such as knowledge based systems) in the control of industrial systems, since these methods minimize the thinking effort in the left hemisphere of the human brain. Artificial neural networks, functioning similar with the right hemisphere of the human brain, became since 1990 also increasingly attractive, especially for problems that cannot be efficiently formalized with present human knowledge. Even so, "on field", due to strange combinations of external influences and circumstances, rare or new situations may appear that were not taken into consideration at design time. Already in 1990 Martin et al showed that "although AI and expert systems were successful in solving problems that resisted to classical numerical methods, their role remains confined to support functions, whereas the belief that evaluation by man of the computerized solutions may become superfluous is a very dangerous fallacy". Based on this observation, Martin et al (1991) recommend "appropriate automation", integrating technical, human, organizational, economical and cultural factors.

This paper aims at surveying from an anthropocentric perspective several concepts and technologies for decision support systems with particular emphasis on real time applications in manufacturing systems.

2. Anthropocentric systems

2.1. Anthropocentric manufacturing systems

Anthropocentric manufacturing systems (AMS) emerged from convergent ideas with roots in the social sciences of the '50s. Kovacs and Munoz (1995) present a comparison between the anthropocentric approach (A) and the technology-centered approach (T) along several directions: *a) role of new technologies*: complement of human ability, regarding the increase of production flexibility, of product quality and of professional life quality (A), versus decrease of worker number and role (T); *b) activity content at operative level*: autonomy and creativity in accomplishing complex tasks at individual or group level (A), versus passive execution of simple tasks (T); *c) integration content and methods*: integration of enterprise components through training, development of social life, of communication and co-operation, increased accesses to information and participation in decision taking (A), versus integration of enterprise units by means of computer-aided centralization of information, decision and control (T); *d) work practice*: flexible, based on decentralization principles, work multivalence, horizontal and vertical task integration and on participation and co-operation (A), versus rigid, based on centralization, strict task separation at horizontal and vertical level associated with competence specialization (T).

2.2. Human-centered information systems

Johanson (1994) shows that "failure and delay encountered in the implementation of CIM concepts" must be sought in organizational and personnel qualification problems. It seems that not only CIM must be considered but also HIM (human integrated manufacturing)". In a man-centered approach integration of man at all control levels must be considered starting with the early stages of a project.

In Filip (1995), 3 *key questions* are put from the perspective of the "man in system" and regarding the man – information tool interaction: a) does the information system help man to better perform his tasks? b) what is the impact of man- machine system on the performance of the controlled object? c) how is the quality of professional life affected by the information system?

Most of the older information systems were not used at the extent of promises and allocated budget because they were *unreliable*, *intolerant* (necessitating a thread of absolutely correct instructions in order to fulfill their functions), *impersonal* (the dialogue and offered functions were little personalized on the individual user) and *insufficient* (often an IT specialist was needed to solve situations). It is true that most of this problems have been solved by IT progress and by intense training, but nevertheless the problem of personalized systems according to the individual features of each user (such as temperament, training level, experience, emotional state) remains an open problem especially in industrial applications.

The second question requires an analysis of *effectiveness* (supply of necessary information) and *efficiency* (supply of information within a clear definition of user *classes* – roles – and real performance evaluation for *individuals* – actors – who interact with the information tool along the dynamic evolution of the controlled object). In the case of industrial information systems, the safety of the controlled object may be more important than productivity, effectiveness or efficiency. As Johanson (1994) pointed out, "in a technology-oriented approach the trend to let the information system take over some of the operator tasks may lead to disqualification and even to boredom under normal conditions and to catastrophic decisions in crisis situations".

This last observation is also a part of the answer to the last question, which answer holds an ethical and social aspect besides the technical one. Many years ago, Briefs (1981) stated, rather dramatically, that the computerization of intellectual work seem to imply "a major threat to human creativity and to the conscious development". This remark was motivated by "the trend to polarize people into two categories. The first one groups IT specialists, who capitalize and develop their knowledge and creativity by making more and more sophisticated tools. The second one represents the broad mass of users, who can accomplish their current tasks quickly and easy, without feeling tempted to develop an own in-depth perception of the new and comfortable means of production".

As Filip (1995) noticed, "it is necessary to elaborate information systems that are not only precise, easy to use and attractive, all at a reasonable cost, but also stimulating to achieve new skills and knowledge and eventually to adopt new work techniques that allow a full capitalization of individual creativity and intellectual skills". The aim to develop anthropocentric information systems applies today as well, but the designer finds little use in generally formulated objectives with no methods to rely on. It is possible to formulate derived objectives representing values for various attributes of information systems: a) *broad service range* (not "Procustian") – for the attribute "use "; b) *transparency* of system structure in regard to its capability to supply explanations – for the attribute "structure ", and c) growing adaptability and learning capabilities – for the attribute "construction"

3. DSS - basic concepts

The DSS appeared as a term in the early '70ies, together with managerial decision support systems. The same as with any new term, the significance of DSS was in the beginning a rather vague and controversial notion. While some people viewed it as a new redundant term used to describe a subset of MISs, some other argued it was a new label abusively used by some vendors to take advantage of a term in fashion. Since then many research and development activities and applications have witnessed that the DSS concept definitely meets a real need and there is a market for it (Holsapple and Whinston, 1996; Power, 2002)

3.1. Decision- making process

Decision- making (DM) process is a specific form of information processing that aims at setting- up an action plan under specific circumstances. There are some examples: setting-up an investment plan, sequencing the operations in a shop floor, managing a technical emergency a.s.o. Several models of a DM are reviewed in the sequel.

Nobel Prize winner H. Simon identifies three steps of the DM process, namely: a) “intelligence”, consisting of activities such as data collection and analysis in order to recognize a decision problem, b) “design”, including activities such as problem statement and production and evaluation of various potential solutions to the problem, and c) “choice”, or selection of a feasible alternative to the implementation.

If a decision problem cannot be entirely clarified and all possible decision alternatives cannot be fully explored and evaluated before a choice is made then the problem is said to be “unstructured “ or “semi-structured”. If the problem were completely structured, an automatic device could have solved the problem without any human intervention. On the other hand, if the problem has no structure at all, nothing but hazard can help. If the problem is semi-structured a computer-aided decision can be envisaged.

The ‘econological’ model of the DM assumes that the decision-maker is fully informed and aims at extremizing one or several performance indicators in a rational manner. In this case the DM process consists in a series of steps such as: problem statement, definition of the criterion (criteria) for the evaluation of decision alternatives, listing and evaluation of all available alternatives, selection of the “best” alternative and its execution.

It is likely that other DM models are also applicable such as: a) the *“bounded rationality”* model, that assumes that decision-making considers more alternatives in a sequential rather than in a synoptic_way, use heuristic rules to identify promising alternatives and make then a choice based on a “satisfying” criterion instead of an optimization one; b) the *“implicit favorite”* model, that assumes that the decision-maker chooses an action plan by using in his/her judgment and expects the system to confirm his choice (Bahl, Hunt, 1984).

While the DSS based on the “econological” model are strongly normative, those systems that consider the other two models are said to be “passive”.

In many problems, decisions are made by a group of persons instead of an individual. Because the *group decision* is either a combination of individual decisions or a result of the selection of one individual decision, this may not be “rational” in H. Simon's acceptance. The group decision is not necessarily the best choice or combination of individual decisions, even though those might be optimal, because various individuals might have various perspectives, goals, information bases and criteria of choice. Therefore, group decisions show a high “social” nature including possible conflicts of interest, different visions, influences and relations (De Michelis, 1996). Consequently, a group DSS needs an important communication facility.

4. DSS technology

4.1. General issues

A distinction should be made between a specific (application-oriented) DSSs (SDSS) and DSS tools. The former is used by particular decision-makers (“final users”) to perform their specific tasks. Consequently, the systems must possess *application-specific* knowledge. The latter are used by “system builders” to construct the application systems. There are two categories of tools: integrated tools and basic tools. The *integrated* tools, called DSS “generators” (DSSG), are prefabricated systems oriented towards various application domains and functions and can be personalized for particular applications within the domain provided they are properly customized for the application characteristics and for the user's specific needs. The DSS *basic construction tools* can be general-purpose or specialized information technology tools. The first category covers hardware facilities such as PCs, workstations, or software components such as operating systems, compilers, editors, database management systems, spreadsheets, optimization libraries, browsers, expert system shells, a.s.o. *Specialized technologies* are hardware and software tools such as sensors, specialized simulators, report generators, etc, that have been created for building new application DSSs or for improving the performances of the existing systems. An application DSS can be developed from either a system generator, to save time, or directly from the basic construction tools to optimize its performances.

The generic framework of a DSS, first proposed by Bonczek, Holsapple, and Whinston (1980) and refined later (Holsapple and Whinston, 1996), is quite general and can accommodate the most recent technologies and architectural solutions. It is based on three essential components: *Language [and Communications] Subsystem* (LS), b) *Knowledge Subsystem* (KS) and c) *Problem Processing Subsystem* (PPS

Recently, Power (2002) expanded Alter's DSS taxonomy and proposed a more complete and up-to-date framework to categorize various DSS in accordance with one main dimension (the dominant component) and three secondary dimensions (the target user, the degree of generality, and the enabling technology)

4.2. Real time DSS for manufacturing

Most of the developments in the DSS domain have addressed business applications not involving any real time control. In the sequel, the real time decisions in industrial milieu will be considered. Bosman (1987) stated that control problems could be looked upon as a "natural extension" and as a "distinct element" of planning *decision making processes* (DMP) and

Sprague (1987) stated that a DSS should support communication, supervisory, monitoring and alarming functions beside the traditional phases of the problem solving process.

Real time (RT) DMPs for control applications in manufacturing are characterized by several particular aspects such as: a) they involve continuous monitoring of the dynamic environment; b) they are short time horizon oriented and are carried out on a repetitive basis; c) they normally occur under time pressure; d) long-term effect are difficult to predict (Charturverdi et al, 1993). It is quite unlikely that an "econological" approach, involving optimization, be technically possible for genuine RT DMPs. Satisfying approaches, that reduce the search space at the expense of the decision quality, or fully automated DM systems (corresponding to the 10th degree of automation in Sheridan's (1992) classification), if taken separately, cannot be accepted either, but for some exceptions.

At the same time, one can notice that genuine RT DMPs can come across in "crisis" situations only. For example, if a process unit must be shut down, due to an unexpected event, the production schedule of the entire plant might turn obsolete. The right decision will be top take the most appropriate compensation measures to "manage the crisis" over the time period needed to recomputed a new schedule or update the current one. In this case, a satisfying decision may be appropriate. If the crisis situation has been previously met with and successfully surpassed, an almost automated solution based on past decisions stored in the *information system* (IS) can be accepted and validated by the human operator. On the other hand, the minimization of the probability of occurrences of crisis situations should be considered as one of the inputs (expressed as a set of constraints or/and objectives) in the scheduling problem. For example in a pulp and paper mill, a *unit plant* (UP) stop may cause drain the downstream *tank* (T) and overflow the upstream tank and so, shut/slow down the unit plants that are fed or feed those tanks respectively. Subsequent UP starting up normally implies dynamic regimes that determine variations of product quality. To prevent such situations, the schedule (the sequence of UP production rates) should be set so that stock levels in Ts compensate to as large extent as possible for UP stops or significant slowing down (Filip, 1995).

To sum up those ideas, one can add other specific desirable features to the particular subclass of information systems used in manufacturing control. An effective *real time DSS for manufacturing* (RT DSSfM) should support decisions on the preparation of "good" and "cautious" schedules as well as "ad hoc", pure RT decisions to solve crisis situations (Filip, 1995).

5. AI based decision- making

As discussed in the previous section, practical experience has shown that, in many cases, the problems are either too complex for a rigorous mathematical formulation, or too costly to be solved by using but optimization and simulation techniques. Moreover, an optimization-based approach assumes an "econological" model of the DM process, but in real life, other models of DM, such as "bounded rationality" or "implicit favorite" are frequently met. To overcome these difficulties several alternatives based on artificial intelligence are used (Dhar and Stein, 1997, Filip, 2002). The term Artificial Intelligence (AI) currently indicates a branch of computer science aiming at making a computer reason in a manner similar to human reasoning.

5.1. Expert systems

The Expert System (ES) is defined by E. Feigenbaum (the man who introduced the concept of "knowledge engineering") as "intelligent computer programs that use knowledge and inference procedures to solve problems that are difficult enough to require significant human expertise for their solution". As in the case of the DSS, one can identify several categories of software products in connection with ES: application ES or "Knowledge Based Systems" (KBS), that are systems containing adequate domain knowledge which the end user resorts to for solving a specific type of problem; system "shells", that are prefabricated systems, valid for one or more problem types to support a straightforward knowledge acquisition and storage; basic tools such as the specialized programming languages LISP, PROLOG or object-oriented programming languages.

One can easily notice the similarity of the ES and DSS as presented in Section 4. Also several problem types such as prediction, simulation, planning and control are reported to be solved by using both ESs and traditional DSSs. At the same time, one can notice that while there are some voices from the DSS side uttering that ESs are only tools to incorporate into DSSs, the ES fans claim that DSS is only a term denoting applications of ESs. Even though those claims can be easily explained by the different backgrounds of tool constructors and system builders, there is indeed a fuzzy border between the two concepts. However a deeper analysis (Filip and Barbat, 1999) can identify some real differences between typical ESs and typical DSSs such as: a) the application domain is well-focused in the case of ES and it is rather vague, variable, and, sometimes, unpredictable in the case of the DSS; b) the information technology used is mainly based on symbolic computation in the ESs case and is heavily dependent on numerical models and database, in traditional cases; c) the user's initiative and attitude towards the system are more creative and free in the DSs case in contrast with ESs case, when the solution may be simply accepted or rejected.

5.2. Case-based reasoning

The basic idea of *Case-Based Reasoning* (CBR) consists in using solutions already found for previous similar problems to solve current decision problems. CBR assumes the existence of a stored collection of previously solved problems together with their solutions that have been proved feasible and acceptable. In contrast with the standard expert systems, which are based on deduction, CBR is based on induction.

The operation of CBR systems basically includes the first or all the three phases: a) selection from a knowledge base of one or several cases (decision situations) similar to the current one by using an adequate similarity measure criterion; b) adaptation of the selected cases to accommodate specific details of the problem to solve. This operation is performed by an expert system which is specialized in adaptation applications; "differential" rules are used by the CBR system to perform the reasoning on differences between the problems; c) storing and automatically indexing of the just processed case for further learning and later use.

5.3. Artificial neural networks

Artificial Neural Networks (ANN), also named connectionist systems, are apparently a last solution to resort to when all other methods fail because of a pronounced lack of the structure of a decision problem. The operation of ANN is based on two fundamental concepts: the parallel operation of several independent information processing units, and the learning law enabling processors adaptation to current information environment

Expert systems and ANNs agree on the idea of using the knowledge, but differ mainly on how to store the knowledge. This is a rather explicit (mainly rules or frames), understandable manner in the case of expert systems and implicit (weights, thresholds) manner, incomprehensible by the human in case of connectionist systems. Therefore while knowledge acquisition is more complex in case of ES and is simpler in case of ANN, the knowledge modification is relatively straightforward in case of ES but might require training from the very beginning in case a new element is added to ANN. If normal operation performance is aimed at, ANNs are faster, more robust and less sensitive to noise but lack “explanation facilities”.

6. Knowledge based DSS

6.1. Combined technologies

It has been noticed that some DSS are “oriented” towards the left hemisphere of the human brain and some others are oriented towards the right hemisphere. While in the first case, the quantitative and computational aspects are important in the second, pattern recognition and the reasoning based on analogy prevail. In this context, there is a significant trend towards combining the numerical models and the models that emulate the human reasoning to build advanced DSS.

Over the last three decades, traditional numerical models have, along with databases, been the essential ingredients of DSS. From an information technology perspective, their main advantages (Dutta, 1996) are: compactness, computational efficiency (if the model is correctly formulated) and the market availability of software products. On the other hand, they present several disadvantages. Because they are the result of intellectual processes of abstraction and idealization, they can be applied to problems that possess a certain structure, which is hardly the case in many real-life problems. In addition, the use of numerical models requires that the user possesses certain skills to formulate and experiment the model. As it was shown in the previous section, the AI-based methods supporting decision-making are already promising alternatives and possible complements to numerical models. New terms such as "tandem systems", or "expert DSS-XDSS" were proposed to name the systems that combine numerical models with AI based techniques. A possible task assignment is given in Table 1 (inspired from Dutta, 1996). Even though the DSS generic framework (mentioned in Section 4.2) allows for a conceptual integration of AI based methods, for the time being, the results reported mainly refer specific applications and not general ones, due to technical difficulties arising from the different ways of storing data or of communicating parameters problems, and from system control issues (Dutta, 1996).

Table 1. A possible task assignment in DSS

	H	NM	ES	ANN	CBR
Intelligence					
Perception of DM situation	I/E		P		
Problem recognition	I/P				I
Design					
• Model selection	M/I		I		I
• Model building	M		I	P	
• Model validation	M				
• Model solving		E		P	
• Model experimentation	I/M		M/I		
Choice					
Solution adoption and release	E		P		

Legend for Table 1. NM - numerical model, ES - rule based expert system, ANN - artificial neural network, CBR - case based reasoning, H - human decision-maker, P - possible, M - moderate, I - intensive, E - essential

6.2. Example

DISPATCHER is a series of DSSs, developed over a twenty-year time period, to solve various decision-making problems in the milieu of continuous 'pure material' process industries. The system initially addressed the short-term production-scheduling problem. Then it evolved in both function set supported and new technologies used in order to satisfy users' various requirements (see Figure 1). New supported functions such as tank sizing, maintenance planning and even acceptance and planning of raw materials or/and utility purchasing allow a certain degree of integration of functions within the [extended] enterprise (Filip, Bărbat, 1999).

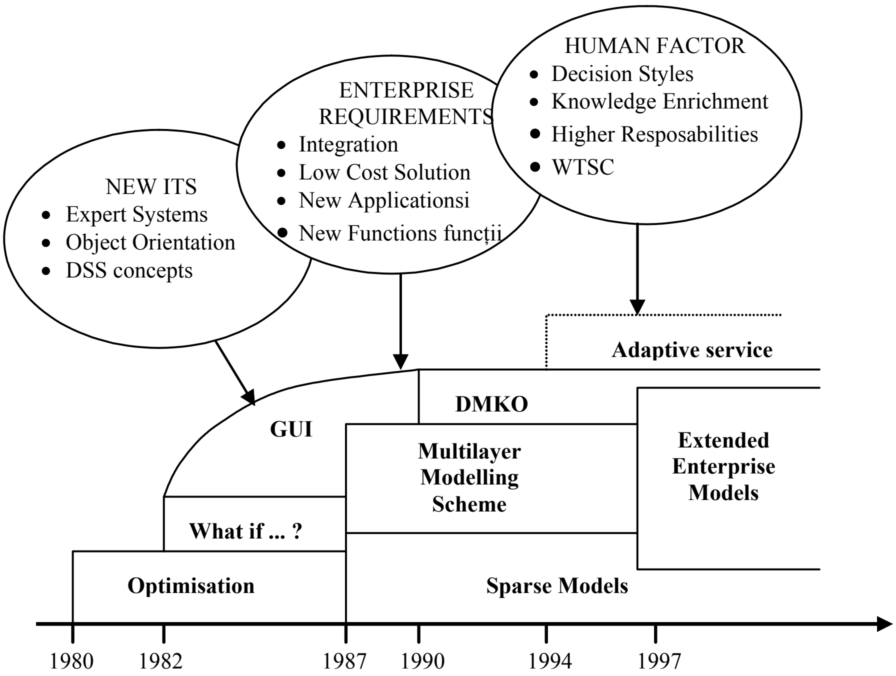


Figure 1. The evolution of the DSS DISPATCHER line [Filip, Barbat,1999)]

Numerous practical implementations of the standard version of DISPATCHER helped draw interesting conclusions. First, the system has been considered by most users as being flexible enough to support a wide range of applications and, in some cases, its utilization migrated from the originally intended one. It has been used in crisis situations (mainly due to significant deviation from the schedule, to equipment failures or other emergencies) as well as in normal operation or in training applications. However, though the system is somehow transparent, and the users have sound *domain* ("what"- type) *knowledge* (DK), they have behaved in a "wise" or even "lazy" (Rasmunsen, 1983) manner, mainly trying to keep their mental load under an average *willing to spend capacity* (WTSC). This can be explained by the initial lack of *tool* ("how"-type) *knowledge* (TK) as well as by insufficient work motivation.

To fight the lack of TK and to stimulate users' creativity and quest for new skills, a *declarative model* of an "ambitious" and *knowledgeable operator* (DMKO) was proposed (Filip, 1993). DMKO is one component of a multilayer, modeling scheme that also includes: the external model (formulated in user's terms, b) the conceptual model (addressing the system builder's needs), and c) internal (performance model (meant for the use of the "toolsmith" programmer. It supports a) model building for various decision contexts, b) problem feasibility testing to propose corrective measures (for example limit relaxation or transformation of fixed/known perturbations into free variables etc.), c) automatically building the internal model from the external description, choosing the appropriate solving algorithm, d) experimenting the problem model, for example by producing a series of alternatives through modifying various parameters in answer to qualitative assessments (made by the user) of the quality of simulated solutions, followed by due explanations. To handle the complexity and diversity of the technologies used, object orientation has been adopted.

Efforts have been made to introduce new intelligence into the system, especially for evaluating user's behavior so that DMKO (originally meant for supporting a certain "role") could dynamically adapt to specific needs of particular "actors", in an attempt at rendering the system less impersonal.

Of course, there are other reported results combining traditional numeric methods with KBS to build "hybrid" or "tandem" DSSfM. Apparently such systems are primarily meant for making numerical computation easier, including heuristics so that the space search for optimization/simulation algorithms is adapted / reduced. It should be noted that the approach presented here is mainly human factor- centered and aims at increasing system acceptance rather than improving its computational performance.

7. Conclusions

Several important issues on the design of anthropocentric modern information systems were reviewed. DSS, as a particular kind of human-centered information system, was described with particular emphasis on real-time applications in the industrial milieu. The possible integration of the AI-based methods within DSS with the view to evolve DSS from simple job aids to sophisticated computerized decision assistants was discussed

Several further developments have been foresighted such as:

- Incorporation and combination of newly developed numeric models and symbolic/sub-symbolic (connectionist) techniques in advanced, user-friendly DSS will continue; also the use of "fuzzy logic" methods are expected to be intensively used in an effort to reach the "unification" of man, numerical models, expert systems and artificial neural networks;

- Largely distributed group decision support systems that intensively use new, high-performance computer networks will be created so that an ever larger number of people from various sectors and geographical locations are able to communicate and make "co-decisions" in real-time in the context of new enterprise paradigms;
- Mobile communications and web technology will be ever more considered in DSS, thereby people will make co-decisions in "virtual teams", no matter where they are temporarily located;
- Other advanced information technologies such as virtual reality techniques (for simulating the work in highly hostile environments) or "speech computers" are likely to be utilized.

References

Bahl, H.C., R.G. Hunt(1984). Decision making theory and DSS design. *DATABASE, 15*(4), 10-14.

Bonczek, R.H., C.W. Holsapple, A.B. Whinston (1980). *Foundations of Decision Support Systems.* Academic Press, New York.

Briefs, V. (1981). Re-thinking industrial work: computer effects on white-collar workers. *Computers in Industry*, 2*,* 76-89.

Bosman, A.(1987). Relations between specific Decision Support Systems. *Decision Support Systems,3,* 213-224.

Chartuverdi, A. R., G.K. Hutchinson, D.L. Nazareth (1993). Supporting complex real-time decision-making through machine learning. *Decision Support Systems, 10,* 213-233.

De Michelis, G. (1996). Co-ordination with co-operative processes. In: *Implementing Systems for Support Management Decisions.* (P. Humphrey, L. Bannon, A. Mc. Cosh, P. Migliarese, J. Ch. Pomerol Eds). 108-123, Chapman & Hall, London.

Dhar, V., R. Stein (1997). *Intelligent Decision Support Methods: The Science of Knowledge Work.* Prentice Hall, Upper Saddle River, New Jersey.

Dutta, A (1996). Integrated AI and optimisation for decision support: a survey. *Decision Support Systems,* 18*,* 213-226.

Filip, F. G. (1995). Towards more humanised real-time decision support systems. In: *Balanced Automation Systems; Architectures and Design Methods.* (L. M. Camarinha – Matos and H. Afsarmanesh, Eds.), 230-240, Chapman & Hall, London.

Filip, F.G. (2002). *Computer Aided Decision-Making: Basic Methods and Tools.* Expert Publishers and Technical Publishers (In Romanian).

Filip, F. G., B. Barbat (1999). *Industrial Informatics.* Technical Publishers, Bucharest (In Romanian).

Holsapple, C. W., A.B. Whinston (1996). *Decision Support Systems: A Knowledge - Based Approach.* West Publishing Company, Minneapolis/St. Paul.

Johannsen,G. (1994). Integrated systems engineering the challenging cross discipline. In: *Preprints IFAC Conf. On Integrated Syst. Engng.,* 1-10, Pergamon Press, Oxford.

Kovacs, J., A. B. Munoz (1995). Issues in anthropocentric production systems. In: *Balanced*

Automation Systems; Architectures and Methods (L. Camarinha – Matos, H. Afsarmanesh, Eds.), 131-140, Chapman & Hall, London.

Martenson, L. (1996). Are operators in control of complex systems? In: *Preprints, IFAC 13th World Congress* (J. Gertler, J.B. Cruz; M.Peshkin, Eds.), vol. B, 259 – 270.

Martin, T., J. Kivinen, J.E. Rinjdorp, M.G. Rodd, W.B. Rouse (1980). Appropriate automation integrating human, organisation and culture factors. In: *Pre-prints IFAC 11th World Congress vol.* 1 *, 47 – 65.*

Power, D. J. (2002) *Decision Support Systems: Concepts and Resources for Managers*. Quorum Press, Westport, Connecticut.

Sheridan,T.B.(1992*). Telerobotics, Automation, and Human Supervisory Control.* MIT Press.

Sprague Jr., R.H. (1987). DSS in context. *Decision Support Systems, 3*, 197-202.

Rasmunsen, J. (1983). Skills, roles and knowledge, signal signs and symbols and other distinctions in human performance model. *IEEE Trans. On Syst., Man and Cybern.-SMS, 13*, 257-266.

NON-ANALYTICAL APPROACHES TO MODEL-BASED FAULT DETECTION AND ISOLATION

Paul M. Frank

Universität Duisburg – Essen

Bismarckstr. 81, 47057 Duisburg, Germany

e-mail: p.m.frank@uni-duisburg.de

Abstract The paper deals with the treatment of modeling uncertainties in model-based fault detection and isolation (FDI) systems using different kinds of non-analytical models which allow *accurate* FDI under even imprecise observations and at reduced complexity.

Keywords: fault detection and isolation, modeling uncertainty, robustness, qualitative models

1. Introduction

All real systems in nature – physical, biological and engineering systems – can malfunction and fail due to faults in their components. The chances for failures are increasing with the systems' complexity. The complexity of engineering systems is permanently growing due to the growing size of the systems and the degree of automation, and accordingly increasing are the chances for faults and aggravating their consequences for man and environment.

Therefore, increased attention has to be paid to reliability, safety and fault tolerance in the design and operation of engineering systems. But obviously, compared to the high standard of perfection that nature has achieved with its self-healing and self-repairing mechanisms in complex biological organisms, the fault management in engineering systems is far behind the standards of their technological capabilities and is still in its infancy, and much work is left to be done.

In automatic control systems, defects may happen in sensors, actuators, components of the controlled object, or within the hardware or software of the control framework. A fault in a component may develop into a failure of the whole system. This effect can easily be amplified by the closed loop, but the closed loop may also hide an incipient fault from being observed until a situation is reached in which failing of the whole system is unavoidable. Even making the closed loop *robust* or *reliable* (by using *robust* or *reliable* control algorithms) can not solve the problem in full. It may help to make the closed loop continue its mission with the desired or a tolerable degraded performance, despite the presence of faults, but when the faulty device continues to malfunction, it may cause damage to man and environment due to the persistent impact of the faults (i.e., leakage in gas tanks or in oil pipes etc.). So, both robust control and reliable control exploiting the available hardware or software redundancy of the system may be efficient ways to maintain the functionality of the control system, but it can not guarantee safety or environmental compatibility.

A realistic fault management has to guarantee *dependability* which includes both reliability *and* safety. Dependability is a fundamental requirement in industrial automation, and a cost-effective way to provide dependability is ***f**ault-**t**olerant **c**ontrol* (*FTC*). The key issue of FTC is to prevent local faults from developing into system failures that can end the mission of the system and cause safety hazards for man and environment. Because of its increasing importance in industrial automation, FTC has become an emerging topic of control theory.

Fault management in engineering systems has many facets. Safety-critical systems, where no failure can be tolerated, need redundant hardware to accomplish fault recovery. *Fail-operational* systems are insensitive to any single component fault. *Fail-safe* systems perform a controlled shut-down to a safe state with graceful degradation when a critical fault is detected. *Robust* and *reliable* control ensures stability or pre-assigned performance of the control system in the presence of *continuous* or *discrete* faults, respectively. *Fault-tolerant* control (FTC) provides online supervision of the system and appropriate remedial actions to prevent faults from developing into a failure of the whole system. In advanced FTC systems, this is attained with the aid of *fault detection and isolation* (*FDI)* in order to detect the faulty components, followed by appropriate system reconfiguration.

Not only that FDI has become a key issue in FTC, it is also the core of *fault-tolerant measurement* (*FTM*). The goal of FTM is to ensure reliability of the measurements in a sensor platform by replacing erroneous sensor readings by reconstructed signals due to the existing analytical redundancy. FDI has also become a basic tool for offline tasks such as *condition-based maintenance and repair* carried out according to the information from early fault monitoring.

The backbone of *modern FDI systems* is the *model-based* approach, where the model contains what is known under the term analytical redundancy. Making use of *dynamic* models of the system under consideration allows us to detect small faults and perform high-quality fault diagnosis by determining time, size and cause of a fault during all phases of dynamic system operations. The classical approach to model-based FDI makes use of functional models in terms of an *analytical* (“*parametric”*) representation.

A fundamental difficulty with analytical models is that there are always modeling uncertainties due to unmodeled disturbances, simplifications, idealizations and parameter mismatches which are basically unavoidable in the mathematical modeling of a real system. They may be subsumed under the term *unknown inputs* and are not mission-critical. But they can obscure small faults, and if they are misinterpreted as faults they cause false alarms which can make an FDI system totally useless. Hence, the most essential requirement for an analytical model-based FDI algorithm is to provide *robustness* w. r. t. the different kinds of uncertainties. This problem is well recognized in the control community, and analytical approaches to robust FDI schemes that enable the detection and isolation of faults in the presence of modeling uncertainties have attracted increasing research attention in the past two decades, and there is both a great number of different solutions of this problem with a good theoretical foundation [11, 12, 14, 32, 33, 38]. More relevant literature on analytical approaches to robust FDI can be found in the books of Patton, Frank and Clark [30, 31], Gertler [16] and Chen and Patton [7].

Surprisingly, much less attention has been paid to the use of *qualitative* models in FDI systems, also known as knowledge-based redundancy methods, in which case the parameter uncertainty problem does inherently not appear. The appeal of the qualitative approaches lies in the fact that qualitative models permit accurate FDI decision making even under imperfect system modeling and imprecise measurements. Moreover, qualitative model-based approaches may end up in less complex FDI systems than comparably powerful analytical model-based approaches. At present, increased research is going on in this field of FDI using non-analytical modeling including computational intelligence, and there is a good deal of publications with most encouraging results, see, for example, [1, 10, 11, 13, 15, 21, 23, 24, 25, 26, 33, 37, 39, 46].

In this paper, we focus our attention on how to cope with modeling uncertainties and imprecise measurements by using non-analytical, i.e., qualitative, structural, data-based and computationally intelligent models. Our intention is to stress the fact that modeling abstraction enables us to make accurate decisions for FDI with less complexity even in the face of large modeling uncertainty, measurement imprecision and lack of system knowledge.

2. The model-based approach to FDI

2.1. Diagnostic strategy

The basic idea of the model-based approach to FDI is to compare the behavior of the actual system with that of its functional model. The diagnostic strategy can follow either of the two policies:

1) *If the measurements of outputs are inconsistent with those of a fault-free model with the same input, this indicates that a fault has occurred.*
2) *If the measurements are consistent with the model behavior corresponding to a certain fault scenario, f_i, then the fault scenario, f_i is declared.*

The diagnostic strategy depends on the kind of model used. In the first case the *nominal*, *fault-free* behavior of the system is modeled, and the inconsistency of the actual system behavior with that of the model indicates a fault. Alternatively one can model a *faulty* behavior for a particular pre-assigned fault scenario; if such a fault model is used the consistency of the actual system with the model indicates that the assumed fault scenario has occurred. In this paper we will only discuss the more common approach of using a fault-free ("nominal") reference model.

In general, the FDI task is accomplished by the following two-step procedure (Fig.1):

1) *Residual*/s*ymptom generation*. This means to generate residuals/ symptoms that reflect the faults of interest from the measurements or observations of the actual system. If the individual faults in a set of faults are to be *isolated*, one has to generate properly *structured* residuals or *directed* residual vectors.
2) *Residual*/s*ymptom evaluation.* This is a logical decision making process to determine the time of occurrence of faults (fault *detection*) and to localize them (fault *isolation*). If, in addition, faults are to be identified, this requires the determination of the type, size and cause of a fault (fault *analysis*).

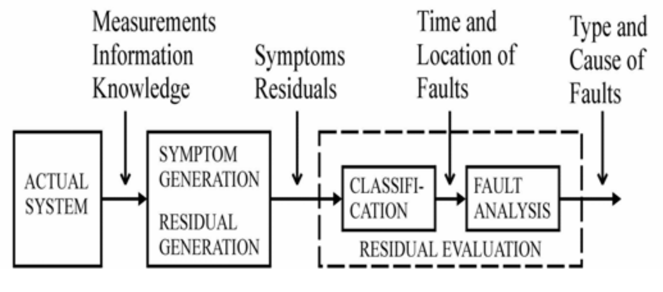


Figure 1. The two-step process of residual generation and evaluation

2.2. Types of models for residual generation

It has been mentioned earlier that any kind of model that reflects the faults can be used for residual generation. The most appropriate model is the one which allows a accurate fault decision at a minimum false alarm rate and low complexity. There is a variety of different kinds of non-analytical models that can be used for this task. The types of models can roughly be classified into four categories, namely analytical (quantitative), qualitative, knowledge-based (statistical, fuzzy, computationally intelligent), data-based (fuzzy, neural), structural. The classification of the corresponding *residual generation* methods is shown in Figure 2.

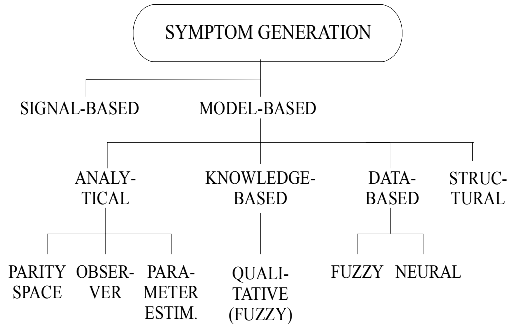


Figure 2. Classification of different model-based approaches to residual generation

Analytical models with usual uncertainties are problematic for FDI unless one can do without those parts of the model which carry substantial uncertainty. To get rid of the uncertain part is the main problem of all robust FDI strategies. It means that finally one has to concentrate on the certain part of the model which reflects the faults of interest, and neglect the uncertain part.

3. FDI with non-analytical models

3.1. The power of abstraction

The best way to overcome model uncertainties is to avoid them from the very beginning. That is to say, to use such kinds of models that are not precisely (analytically) defined in terms of parameters. The use of non-analytical models, such as qualitative or structural models, and dealing with symptoms rather than signals means an increase of the degree of *abstraction*, which plays a fundamental role, however, in reaching accurate results for FDI. Logically, achieving accurateness in FDI implies that the check of the reference model must be accurate, i.e., it must be in agreement with the observations of the fault-free system even if these observations are imprecise. This is possible with an according degree of *abstraction* of the model. In addition, abstraction may reduce the complexity of the model and consequently of the resulting FDI system.

Figure 3 shows the typical relationship between model complexity, measurement imprecision and modeling uncertainty of an *accurate* model for FDI depending on different kinds of modeling. It can be seen that, to reach accuracy, the required complexity is maximum for precise, i.e. quantitative analytical models, and it decreases considerably with the degree of abstraction obtained by the use of non-analytical models. This means that accurate decisions are possible even in case of imprecise observations if abstract (non-analytic) modeling is applied, or, in other words:

A reduction of complexity of robust FDI algorithms can be obtained by increasing the degree of abstraction of the model

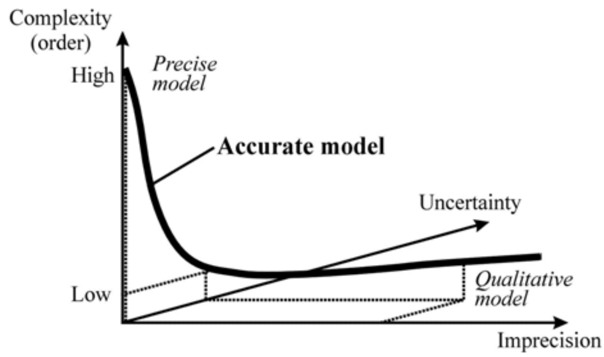


Figure 3. Complexity of an accurate model for FDI versus uncertainty and imprecision

3.2. FDI based on qualitative models

3.2.1. Qualitative approaches to FDI

Qualitative models reduce the resolution of the representations by introducing tolerances in order to emphasize relevant distinctions and ignore unimportant or unknown details. Under imprecise observations this description represents the systems accurately if a *set* of values rather than single values become the primitives of representation.

In the last decade, the study of applying qualitative models to system monitoring and FDI received much attention, see, e.g., [10, 21, 22, 23, 34], and the concept of *qualitative (knowledge based) observer* was introduced [13]. Typical qualitative descriptions of variables are signs [9], intervals [20], [23] or fuzzy sets [35]. As a fuzzy set can be divided into a series of intervals, the use of the α-cut identity principle proposed by Nguyen [29] allows to reduce fuzzy mappings into interval computations. Therefore, intervals are the fundamental representations in qualitative modeling. The rough representation of variables leads to the imprecision of the qualitative model which relates the variables to each other.

According to the available information about a system, there are different possibilities for a qualitatively representation of the information of the dynamic process. Basically, a qualitative simulation method should be responsible for retaining the accuracy of the represented system behavior (so called soundness property following the definition of Kuipers [20]), so that the FDI approaches based on them can avoid false alarms. The most important types of representation known are:

- Qualitative differential equations (QDE) [20, 35]
- Envelope behavior (e.g.), [5, 18]
- Stochastic qualitative behavior [23, 46].

Other relevant methods to qualitative models for fault diagnosis are, e.g., signed directed graphs [22], logical based diagnosis [24] and structural analysis [36]. Dynamic behaviors are not emphasized in these methods, their main concern is the causality or correlativity among various parts of the systems, which are useful for performing fault isolation and fault analysis.

3.2.2. FDI using qualitative observers based on QDE

Conceptually, a qualitative differential equation can be considered as the extension of an ordinary differential equation

$$\dot{x} = g(x,\ u,\ \theta)\,, \tag{1}$$

where x, u and θ denote the vectors of state variables, known inputs and parameters with the dimension of n, r and s, respectively. However, in a QDE, the variables take intervals as their values and the variant of the non-linear function g(.) is allowed to include various imprecise representations: e.g., interval parameters, non-analytical functions empirically represented by IF-THEN rules and even, in the algorithm QSIM of Kuipers [18], unknown monotonic functions. If the non-linear function g(.) is rational, its corresponding QDE can be readily derived from it by using the natural interval extension of the real function [28]. Qualitative simulation procedures that are composed of the two main steps “generation” and “test/exclusion” are basically different from the numerical ones. The behavior of continuous variables is discretely represented by a branching tree of qualitative states.

The resulting *qualitative* observer (QOB) based on QDE is an extension of a qualitative simulator, and it functions in further reducing the number of irrelevant behaviors (including spurious solutions) to the system under consideration [39] as illustrated in Fig. 4**.** The principle of observation filtering is that the simulated qualitative behavior of a variable must cover its counterpart of the measurements obtained from the system itself; otherwise the simulated behavioral path is inconsistent and can be eliminated. Since these procedures do not lead to the violation of the accuracy of the qualitative behavior under fault free condition, the output of QOB is the refined prediction behavior in this case.

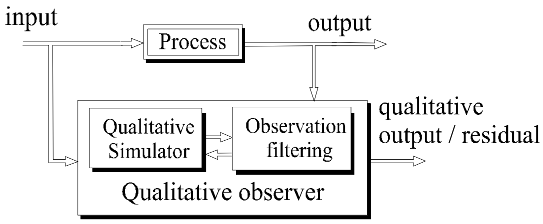


Figure 4. Qualitative observer

However, when a fault occurs and causes a significant deviation of the system output such that no consistent predicted counterpart of the output could be generated, the output of the QOB becomes an empty set, which indicates the fault occurrence. Following this principle, fault detection and sensor fault isolation can be implemented [39]. It is important to note that, in exchange with the advantage of requiring weaker process knowledge in this method, one has to put up with an increase in computational complexity and less sensitivity to small faults.

3.2.3. Fault detection based on envelope behaviors

A key issue of improving the small fault detectability when applying qualitative methods is that the qualitative system behavior should be predicted as precisely as possible. Different from the qualitative model and the simulation method presented above, the model considered in this and the next sections is of less ambiguity. In other words, imprecision in equation (1) is caused only by interval parameters and interval initial states, the structure of g(.) is considered to be fully known. While qualitative behaviors here are interval values of system variables against time, qualitative simulation aiming at producing all possible dynamic behaviors means the generation of their envelope. Once the envelope is generated, the fault detection task is a direct comparison between the envelope and the measurements. In fault-free case, the measurements are contained in the envelope; otherwise, it indicates a fault.

Recently, many efforts have been made to increase the efficiency of classical qualitative simulation, i.e., to avoid unnecessary conservativeness. More quantitative information is brought into the model representation [3], and simulation methods tend to be more constructive. Kay and Kuipers [18] and Verscovi et al. [38] propose approaches based on standard numerical methods to obtain the bounding behavior. In [5, 19] Bonarini et al. and Keller et al. treat the interval parameters and the state variables as a super-cube, whose evolution at any time is specified by its external surface. Armengo et al. [1] present the computation of envelopes making use of modal interval analysis.

3.2.4. Residual generation via stochastic qualitative behaviors

Another qualitative representation of system behaviors is the stochastic distribution under partitioned state and output spaces. Beginning with the similar model assumptions as in section 3.3.1, the parameter vector is in θ and the initial state is uniformly distributed within a prescribed area, say cell *0*. $X_i(t)$ and Y_i (t) denote the probabilities that the trajectories of the respective state and output variables, which start from all initial states in cell *0*, fall into the *i*-th cell at any time *t*. The behavior can be approximately represented by a Markov chain [46]. It turns out that the new state and output variables *X* and *Y* can be described by the following discrete hidden Markov model (HMM):

$$X(k+1) = A(u,\theta)X(k) + V(k) \tag{2}$$

$$Y(k+1) = C(\theta)X(k+1)\ , \tag{3}$$

where *V* represents the influence of spurious solutions.

A fault detection scheme based on the HMM is shown in Fig. 5 [46]. A qualitative observer (QOB) aiming at attenuating the effect of *V* and watching over the possible abnormal behavior of measurements is applied. The residual *r* and its credibility ν can be calculated, the latter reflects the degree of spurious solutions.

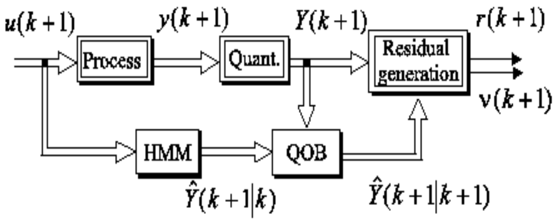


Figure 5. Observer-based residual generation using HMM

3.3. Residual generation employing computational intelligence

In the case of fault diagnosis in complex systems, one is faced with the problem that no or insufficiently accurate mathematical models are available. The use of data-model-based (neural) diagnosis expert systems or in combination with a human expert, is then a much more appropriate way to proceed. The approaches presented in the following section employ computational intelligence techniques such as neural networks, fuzzy logic, genetic algorithms and combinations of them in order to cope with the problem of uncertainty, lacking analytical knowledge and non-linearity [15].

3.3.1. Neural observer-based residual generation

Neural networks can be used as non-linear multiple-input single-output (MISO) models of ARMA type to set up different kinds of observer schemes [15, 27]. The neural networks replace the analytical models of observer-based FDI. If instead of a single multiple-input multiple-output structure a separate neural network is taken for each output, a set of smaller neural networks can be used for each class of system behavior.

The type of neural network employed for this task is of a mixed structure called *dynamic multi-layer perceptron (DMLP-MIX)* integrating three generalized structures of a DMLP [25]. These three are: the DMLP with synaptic generalized filters, which have each synapse represented by an ARMA filter with different orders for denominator and numerator, the DMLP with internal generalized filters [2] integrating an ARMA filter within the neurons before the activation function, and the DMLP with a connectionist hidden layer, which has a partially recurrent structure interconnecting only the hidden units. The mixed structure is implemented selecting either a basic architecture or a combination of them. The training of the DMLP-MIX neural network is performed by applying dynamic back propagation, the problem of structural optimization is solved with the help of a genetic algorithm [26]. Two types of observer schemes for actuator, component and instrument fault detection have been proposed by Marcu et al. [27]: the neural single observer scheme (NSOS) and the neural dedicated observer scheme (NDOS).

3.3.2. Fuzzy observer-based residual generation

There are many ways of using fuzzy logic to cope with uncertainty in observer-based residual generation [15]. The resulting type of fuzzy observer depends upon the type of the fuzzy model used. Fuzzy modeling can roughly be classified into four categories: fuzzy rule-based, fuzzy qualitative, fuzzy relational and fuzzy functional (Tagaki-Sugeno type).

3.3.3. Residual generation with hierarchical fuzzy neural networks

Here the fault diagnosis system is designed by a knowledge-based approach and organized as a hierarchical structure of fuzzy neural networks (FNN) [6]. FNNs combine the advantage of fuzzy reasoning, i.e. being capable of handling uncertain and imprecise information, with the advantage of neural networks, i.e. being capable of learning from examples. The neural nets consist of a fuzzification layer, a hidden layer and an output layer. Fault detection is performed through the knowledge-based system, where the detection rules are generated from knowledge obtained from the structural decomposition of the overall system into subsystems and operational experience. After detecting a fault the diagnostic module is triggered, which consists of a hierarchical structure (usually three layers) of FNNs. The number of FNNs is determined by the number of faults considered. The lower level only contains one FNN, which processes all measured variables. The FNNs on the medium level are fed by all measurements but also by the outputs of the previous level. The upper level consists of an OR operation on the outputs of the medium level. This hierarchical structure can cope with multiple simultaneous faults under highly uncertain conditions.

3.3.4. Fuzzy residual evaluation

Fuzzy logic is especially useful for decision making under considerable uncertainty. The three main categories of current residual evaluation methods are: classification (clustering) or pattern recognition, inference or reasoning, and threshold adaptation. Although all approaches employ fuzzy logic, the first one is actually data-based while the other two are knowledge-based.

3.3.5. Fuzzy clustering

The approach of fuzzy clustering actually consists of a combination of statistical tests to evaluate the time of occurrence of the fault and the fuzzy clustering to provide isolation of the fault [8]. The statistical tests are based on the analysis of the mean and the variance of the residuals, e.g., the CUSUM test [17]. The subsequent fault isolation by means of fuzzy clustering consists of the two following steps: In an online phase the characteristics of the different classes are determined. A learning set which contains residuals for all known faults is necessary for this online phase. In the online phase the membership degree of the current residuals to each of the known classes is calculated. A commonly used algorithm is the fuzzy C-means algorithm [4].

3.3.6. Fuzzy reasoning

The basic idea behind the application of fuzzy reasoning for residual evaluation is that each residual is declared as *normal*, *high* or *low* with respect to the nominal residual value [8, 37]. These linguistic attributes are defined in terms of fuzzy sets, and the rules among the fuzzy sets are derived from the dynamics of the system. For fault detection, the only relevant information is whether or not the residual has deviated from the fault free value, and hence it is only necessary to differentiate between normal and abnormal behavior. However, if isolation of faults is desired, it may be necessary to consider both the direction and magnitude of the deviation.

3.3.7. Fuzzy threshold adaptation

Fuzzy reasoning has been applied with great success to threshold adaptation [13, 33]. In the case of poorly defined systems it is difficult or even impossible to determine adaptive thresholds. In such situations the fuzzy logic approach is much more efficient. The relation for the adaptive threshold can be defined as a function of input u and output y by

$$T(u,y) = T_0 + \Delta T(u,y) \tag{4}$$

Here $T_o = T_0(u_0, y_0)$ denotes a constant threshold for nominal operation at the operational point (u_o, y_0) where only the effects of the stationary disturbances including measurement noise are taken into account. The increment $\Delta T(u,y)$ represents the effects of $u(t)$ and $y(t)$ caused by the modeling errors. These effects are described in terms of IF-THEN rules and the variables by fuzzy sets (e.g. SMALL, MIDDLE, LARGE, etc.) that are characterized by proper membership functions.

As a typical example of an industrial application we consider the residual evaluation via fuzzy adaptive threshold of a six-axis industrial robot (Manutec R3) [13, 33]. Let the goal be to detect a collision of the robot by checking the moments of the drives. A model of the robot is available, but without knowledge of the friction of the bearings, which is highly uncertain. It is known, however, that the residual of the moment is heavily distorted by the friction which strongly depends on the arm acceleration. This knowledge can be formulated by rules. For example for the third axis the following rules apply:

IF{speed small}, THEN{threshold middle}

IF{acceleration high}, THEN{threshold large}

IF{acceleration very high}, THEN {threshold very large}

IF {acceleration of any other axis very high}, THEN {threshold middle}.

The linguistic variables *small, middle, high, very high, large, very large* are defined by proper membership functions [33], they are assigned intuitively based on the experience of the operators or the manufacturers of the robot.

Figure 6 shows the time shape of the threshold together with the shape of the residual of axis 3 for a particular maneuver of the robot. Note that at t = 4,5 sec the heavy robot which can handle 15 kg objects in its gripper, hits an obstacle which causes a momentum of about 5 Nm. As can be seen, this small fault can be detected at high robustness to the uncertainty caused by the neglected unknown friction.

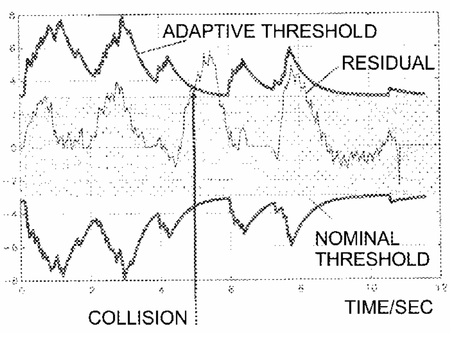


Figure 6. Obstacle detection of a robot with fuzzy adaptive threshold

3.4. FDI based on structural models

The use of structural system models together with structural analysis is another way of abstraction of the modeling of the system behavior in order to increase the robustness of the FDI algorithm to model uncertainties. Here we only consider the *structure* of the constraints, i.e., the existence of links between variables and parameters rather than the constraints themselves [36]. The links are usually represented by a *bi-partite graph*, which is independent of the nature of the constraints and variables (quantitative, qualitative, equations, rules, etc.) and of the values of the parameters. Structural properties are true almost everywhere in the system parameter space.

This represents indeed a very low-level easy-to-obtain model of the system behavior, which is logically extremely insensitive to changes in the system parameters but, of course, also to parametric faults. The important tasks of structural analysis are solved with the aid of the analysis of the system structural graph and its canonical decomposition. An important factor in the canonical decomposition is the property of causality which complements the bi-partite graph with an orientation. FDI is performed with the aid of analytical redundancy relations based on a structural analysis and the generation of structured residuals.

Note that the use of structural models together with the strong decoupling approach solves automatically the robustness problem in structurally observable systems.

4. Conclusion

The paper reviews the methods of handling modeling uncertainties, incomplete system knowledge and measurement imprecision in model-based fault detection and isolation by using non-analytical models. It is pointed out that abstract non-analytical models may be superior over analytical models with respect to uncertainty, imprecision and complexity. The paper outlines the state of the art and relevant on-going research in the field approaching the modeling uncertainty and measurement imprecision problem in FDI by various types of non-analytical models.

References

[1] Armengo, J., L. Travé-Massuyès, J. Vehi and M. Sáinz (1999). Semi-qualitative simulation using modal interval analysis. In: *Proc. of the 14th IFAC World Congress 99,* Beijing, China.

[2] Ayoubi, M. (1994). Fault diagnosis with dynamical neural structure and application to a turbo-charger. In: *IFAC Symp. SAFE-PROCESS'94*, Espoo, Finland, Vol. 2.

[3] Berleant, D. and B. Kuipers (1992). Qualitative-numeric simulation with q3. In: *Recent Advances in Qualitative Physics* (B. Faltings and P. Struss, Eds.). MIT Press Cambridge, Mass.

[4] Bezdek, J. C. (1991). Pattern Recognition with Fuzzy Objective Functions Algorithms. *Plenum Press*, New York.

[5] Bonarini, A. and G. Bontempi (1994). A qualitative simulation approach for fuzzy dynamical models. *ACM transactions on Modeling and Computer Simulation* 4(4), pp. 285-313.

[6] Calado, J. M. F. and J. M. Sa da Costa (1999). Online fault detection and diagnosis based on a coupled system. In: *Proc. ECC'99*, Karlsruhe.

[7] Chen, J and R. Patton (1999). Robust model-based fault diagnosis for dynamic systems. *Kluwer Academic Publishers.*

[8] Dalton, T., N. Kremer and P. M. Frank (1999). Application of fuzzy logic based methods of residual evaluation to the three tank benchmark. In: *ECC'99,* Karlsruhe.

[9] de Kleer, J. and J.S. Brown (1984). Qualitative physics based on confluences. *Artificial Intelligence* 24, pp. 7-83.

[10] Dvorak, D. and B. Kuipers (1989). Model-based monitoring of dynamic systems. In: *Proc. 11th Intern. Joint Conf. on Art. Intell.*, Detroit, MI. pp. 1238-1243.

[11] Frank, P. M. (1990). Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy - A survey and some new results. *Automatica* 26, pp. 459-474.

[12] Frank, P.M. (1994). Enhancement of robustness in observer-based fault detection. *Int. J. Control*, 59, pp. 955-981.

[13] Frank, P.M. (1996). Analytical and qualitative model-based fault diagnosis – A survey and some new results. *Europ. J. Control* 2, pp. 6-28.

[14] Frank, P.M., G. Schreier and E. Alcorta Garcia (1999). Nonlinear observers for fault detection and isolation. In: *New Directions in Nonlinear Observer Design* (Ed.: H. Nejmeijer et al.). Springer, pp. 399-422.

[15] Frank, P.M. and T. Marcu (2000). Diagnosis strategies and systems: Principles, Fuzzy and neural approaches. In: *Teodorescu, H.-N. et al. (Eds.). Intelligent Systems and Interfaces*. Kluwer Akademic Publishers.

[16] Gertler, J.J. (1998). Fault Detection and Diagnosis in Engineering Systems. *Marcel Dekker.*

[17] Gomez, R. E. (1995). Modellbasierte Fehlererkennung und -diagnose in Mehrgroessensystemen mit Hilfe statistischer Methoden. PhD Thesis Univ. Bochum Germany.

[18] Kay, H. and B. Kuipers (1993). Numerical behavior envelopes for qualitative models. In: *Proc. of the 7th National Conf. on Artificial Intelligence*, pp. 606-613.

[19] Keller, U., T. K. Wyatt and R. R. Leitch (1999). Frensi - a fuzzy qualitative simulator. In: *Proc. of Workshop on Applications of Interval Analysis to Systems and Control*, Girona, Spain. pp. 305-313.

[20] Kuipers, B. (1986). Qualitative simulation. *Artificial Intelligence* 66(29), pp. 289-338.

[21] Leitch, R., Q. Shen, G. Conghil, M. Chantler and A. Slater (1994). Qualitative model-based diagnosis of dynamic systems. In: *Colloquium of the Institution of Measurement and Control*, London.

[22] Leyval, L., J. Montmain and S. Gentil (1994). Qualitative analysis for decision making in supervision of industrial continuous processes. *Mathematics and computers in simulation* 36, pp. 149-163.

[23] Lunze, J. (1994). Qualitative modeling of linear dynamical systems with quantized state measurements. *Automatica* 30(3), pp. 417-431.

[24] Lunze, J. (1995). Künstliche Intelligenz für Ingenieure, Band 2: Technische Anwendungen. *Oldenbourg Verlag*, München.

[25] Marcu, T., L. Mirea and P. M.Frank (1998). Neural observer schemes for robust detection and isolation of process faults. In: *UKACC Int. Conf. CONTROL'98*, Swansea, UK. Vol. 2.

[26] Marcu, T., L. Ferariu and P. M. Frank (1999a). Genetic evolving of dynamical neural networks with application to process fault diagnosis. In: *ECC'99,* Karlsruhe.

[27] Marcu, T., M. H. Matcovschi and P. M. Frank (1999b). Neural observer-based approach to fault detection and isolation of a three-tank system. In: *ECC'99*, Karlsruhe.

[28] Moore, R. (1979). Methods and applications of Interval analysis. *SIAM*. Philadelphia.

[29] Nguyen, H.T. (1978). A note on the extension principle for fuzzy sets. *Journal of Mathematical Analysis and Applications* 64, pp.369-380.

[30] Patton, R., P. M. Frank and R. N. Clark (1989). Fault diagnosis in dynamic systems. *Prentice Hall.*

[31] Patton, R., P. M. Frank and R. N. Clark (2000). Issues of fault diagnosis for dynamic systems. *Springer Verlag.*

[32] Rambeaux, F. F. Hammelin and D. Sauter (1999). Robust residual generation via LMI. In: *Proc. of the 14th IFAC World Congress 99,* Beijing, China.

[33] Schneider, H. Implementation of fuzzy concepts for supervision and fault detection of robots. In: *Proc .of EUFIT'93,* Aachen, pp. 775-780.

[34] Shen, L. and P. Hsu (1998). Robust design of fault isolation observers. *Automatica, 34, pp. 1421-1429.*

[35] Shen, Q and R. Leitch (1993). Fuzzy qualitative simulation. *IEEE Trans. SMC* 23(4).

[36] Staroswiecki, M., J. P. Cassar and P. Declerck (2000). A structural framework for the design of FDI system in large scale industrial plants. In: *Issues of Fault Diagnosis for Dynamic Systems* (Patton, Frank and Clark, Eds.). Springer, pp. 453-456.

[37] Ulieru, M. (1994). Fuzzy reasoning for fault diagnosis. *2nd Int. Conf. on Intelligent Systems Engineering*.

[38] Verscovi, M., A. Farquhar and Y. Iwasaki (1995). Numerical interval simulation. In: *International Joint Conference on Artificial Intelligence IJCAI* 95. pp. 1806-1813.

[39] Zhuang, Z. and P.M. Frank (1997). Qualitative observer and its application to fault detection and isolation systems. *J. of Systems and Contr. Eng., Proc. of Inst. of Mech. Eng.*, Pt. I 211(4), pp. 253-262.

[40] Zhuang, Z. and P.M Frank (1999). A fault detection scheme based on stochastic qualitative modeling. In: *Proc. of the 14th IFAC World Congress 99*, Beijing, China.

CONTROL OF DVD PLAYERS FOCUS & TRACKING CONTROL LOOP

Bohumil Hnilička and Alina Besançon -Voda
ST author Laboratoire d'Automatique de Grenoble,
ENSIEG, BP. 46, 38402 Saint Martin d'Hères, France
Fax: +33.4.76.82.63.88, Tel: +33.4.76.82.62.32
Email: Bohumil.Hnilicka@lag.ensieg.inpg.fr,
Alina.Besancon@lag.ensieg.inpg.fr

Geampaolo Filardi
STMicroelectronics, Optical Disc Business Unit,
BP.217, 38019 Grenoble, France
Fax: 33.4.76.58.61.27, Tel:33.4.76.58.46.34
Email: Geampaolo.Filardi@st.com

Abstract This paper is devoted to the design of a control system applied to an industrial DVD-video drive. For high-speed players the disturbance rejection is more difficult problem than for CD players because of higher performance requirements. A combined pole placement/sensitivity function shaping methodology is used for control design purposes to reduce the effect of repetitive disturbances. Controller order reduction is performed to allow its practical implementation. Experimental results, obtained on a real system in STMicroelectronics laboratories, illustrate the performance of the proposed algorithms for both the focus and tracking control loop.

Keywords: DVD player, focus loop, tracking loop, pole placement, controller order reduction

1. Introduction

Optical disk drives are widely used today to hold music, store data or to record digital movies. Even though improvements are observed in obtaining shorter data time access, higher storage capacity and information density on the disk, one of the major obstacles for reliability of readout data is given by the internal and external disturbances affecting the disk. The most important disturbances are optical disk imperfections, mechanical vibration and shocks, and position sensing noise which show up at the optical detector used to measure the error signals for controllers.

Mainstream of DVD/CD players have no passive mechanics control of the vertical distance between the read/write element and the recording media in contrast to the hard disks. Because of this, a vertical axis (focus) control must be applied in DVD/CD systems and therefore the size of optical pick-up unit (OPU) is large. Nevertheless, the important advantage of DVD/CD players is fact that the optical disk is a removable medium with random access to the recorded data. The main control loops here are three:

- **a focus loop** to ensure the distance between the objective lens and the media;
- **a coarse (low frequency) tracking loop** to roughly position the optical pick-up unit, pulling so-called sledge, in the vicinity of the desired tracks;
- **a fine (high frequency) tracking loop** to lock the focused laser beam onto the track position.

A laser beam is used to read the recorded digital data from the optical disk. In order to retrieve the data correctly, the laser beam must be focused on the data layer surface of the disk, and must follow the track, both with high precision. This is difficult in the presence of disturbances. Therefore, the most critical control loops are especially the focus and the fine tracking loop (in next called shortly the **tracking or radial loop**).

Concerning DVD players, to our knowledge, only recently few papers have been dedicated to a narrow-band disturbance suppression in a high-speed DVD players system. In [3], the notch and funnel filters are used for the estimation of the rotating speed of a DVD player, in order to implement a control scheme which selectively cancels narrow-band disturbances. The work [2] proposes a control architecture for track following using the notch filtering and multirate control. Paper [14] presents a control system using sliding mode control to handle shock and vibration disturbances. Repetitive control is in [10] are devoted to disturbance observer design. A combined pole placement/sensitivity function shaping methodology have not been used to control design till now although this method offers the possibilities on the periodic disturbance rejection.

This paper treats mainly the control design of the focus/tracking loop. The controller is based on experimental results obtained on a real system in STMicroelectronics laboratories.

Shortly, the aims of this paper are the following:

- to provide a methodology for control design the robust fixed low-order controllers used in the focus/tracking control loop;

- to present the results obtained from testing these controllers on a real system.

The present paper is organized as follows: in section 2, an overview of the general principles for DVD/CD players is presented. Section 3 describes the disturbances sources incoming to the system.

Specification requirements on the radial/focus control loops are given in section 4. A control methodology is described in section 5. Section 6 shows control design and performance analysis is explained in section 7.

2. Physical system description

2.1. Optical pick-up organization

In fig. 1 a schematic view of the DVD mechanism, using so-called DVD-5 (single layer single side optical disk), is shown. The system is composed of an optical pick-up unit that retrieves data from the disk. The optical disk is turned by a DC turntable motor with the spindle rotation frequency f_{rot}.

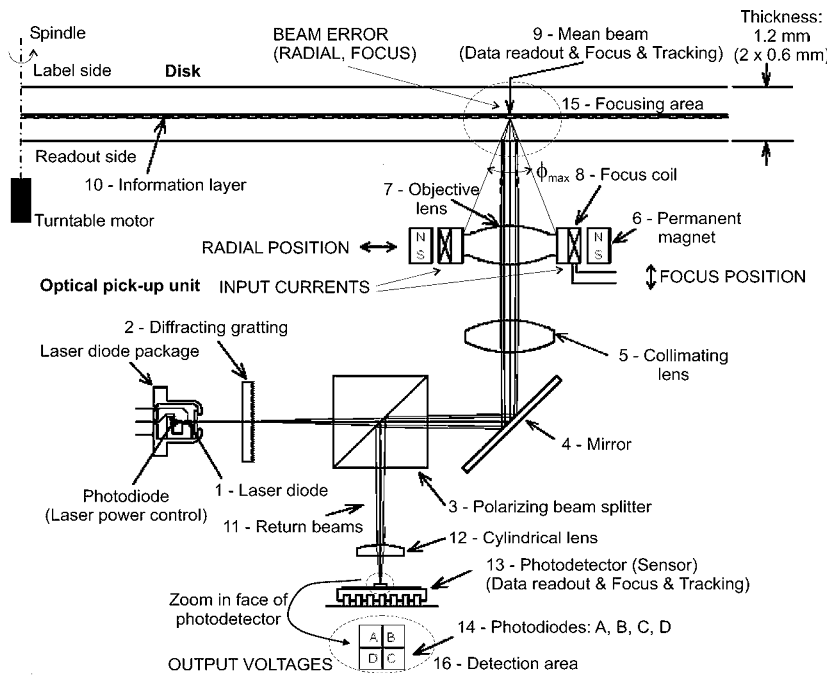


Figure 1. Optical pick-up unit organization of the DVD mechanism using single layer single side optical disk

A laser diode (1) located in the pick-up unit emits a laser beam which is guided through the optics elements (2, 3, 4, 5, 7) to the disk information layer (10). An objective lens (7) is the last optical element for laser beam focusing on the disk information layer (10). The objective lens (7) can be moved in vertical direction, to give focusing action, and in radial direction, to perform track following. It is suspended by leaf springs and its position is controlled by the electromagnets (6) in the vertical and the radial directions. To focus the incident beams on the disk, a focus coil (8) is placed in the electromagnetic field of a permanent magnet (6).

The mean beam (9) of incident rays is reflected from the information layer (10) at a focusing area (15). The return beams (11) pass through the objective lens (7) and the optics elements (5, 4). The return rays are splitted by a polarizing beam splitter (3) in perpendicular direction to the rays emitted by the laser diode (1). Therefore the return beams (11) pass through a cylindrical lens (12), shaping the laser spot that is falling on a photodetector (13). The photodetector (13) that is composed of four photodiodes (14) (A, B, C, D) is a detection area (16) for the return beams (11). The photodiodes (14) measure the light intensity of the laser spot.

Since only light touches the disk information layer (10), four photodiodes (14) generate the voltages which are used to generate the focus, tracking and redout signals.

2.2. Optical pick-up unit control loops

In fig. 2 a block diagram of the focus loop is shown. A peculiar feature of this servo loop is that the absolute position of the optical pick-up can not be measured. The only measure available for control purposes is the focus error e_{F}, in the neighborhood of the information layer. Moreover, notice that input/output measures of the plant u_{F}, e_{F} can be collected only in closed-loop working condition.

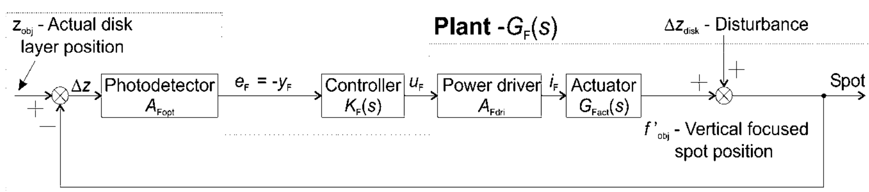


Figure 2. Configuration of the focus control loop (basic schema)

The more complex equivalent block diagram for the optical pick-up unit control of the DVD player is illustrated in fig. 3 as a multi-input multi-output (MIMO) system. For a clear explanation, figs. 4 and 5 show the single-input single output (SISO) focus and tracking control loops, respectively.

To move the objective lens (7) (see fig. 1) in vertical/radial direction, the focus / radial coils, which are used to generate the electromagnetic

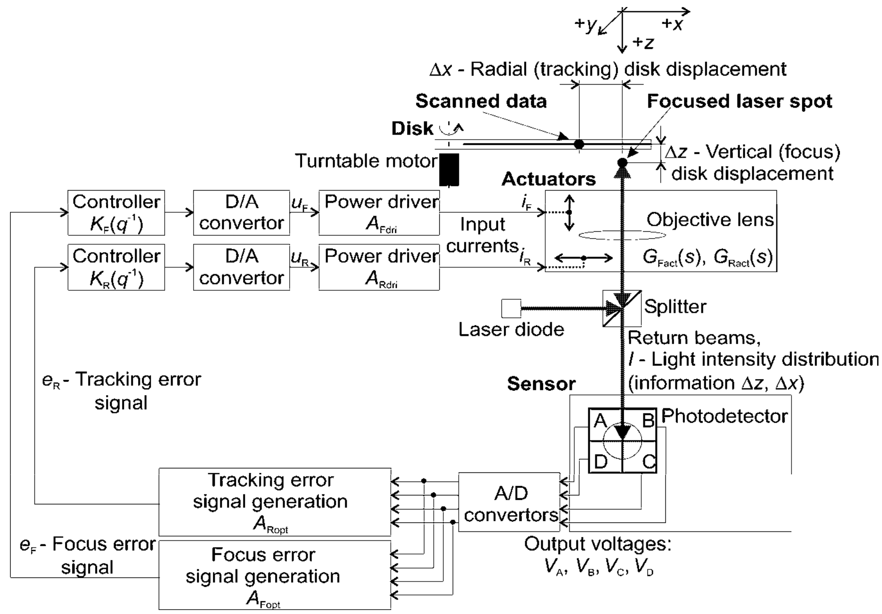


Figure 3. The equivalent block diagram for control of the DVD player of fig. 1

field, are supplied by the input currents i_F, i_R. These currents i_F, i_R are therefore the *inputs to the mechanical actuators* described by continuous-time transfer functions $G_F(s)$, $G_R(s)$. The mean beam (9) of incident rays, generated by the laser diode (1), is reflected by the disk information layer (10). Therefore the light intensity of return beams (11), which is the *input to the sensor*, contains the information about the actual disk vertical/radial displacement Δz, Δx, respectively.

The vertical/radial disk displacement Δz, Δx is defined in the orthogonal coordinate system x, y, z, (see fig. 3), where its origin is placed in the middle of the objective lens (7) in fig. 1. $\Delta z = z_{\text{obj}} - f'_{\text{obj}}$ is the vertical disk displacement between the disk information layer position z_{obj} and the position of the objective lens focus $f'_{\text{obj}} = -f_{\text{obj}}$. f_{obj} is the objective lens focal length and $f_{\text{obj}} < 0$ in case of converging spherical lens that is used in DVD/CD players. $\Delta x = x_{\text{track}} - x_{\text{foc}}$ is the horizontal disk displacement between actual scanned track position x_{track} and actual focused laser spot position x_{foc} in radial direction x as is shown in fig. 3.

The equivalent path between the input currents i_F, i_R and output voltages from the photodetector (13) V_A, V_B, V_C, V_D in fig. 1 is realized by two blocks (nominally *actuators* and *sensor*) in fig. 3.

Fig. 1, fig. 4 and fig. 5 show the whole closed loops, where the outputs from the plants y_F, y_R, used for the controllers, are obtained from the measured voltages V_A, V_B, V_C, V_D. Firstly, the voltages V_A, V_B, V_C,

V_D are converted by the analog to digital convertors. Secondly, some calculations are done to generate the focus/tracking error signals e_F, e_R from the measured voltages V_A, V_B, V_C, V_D.

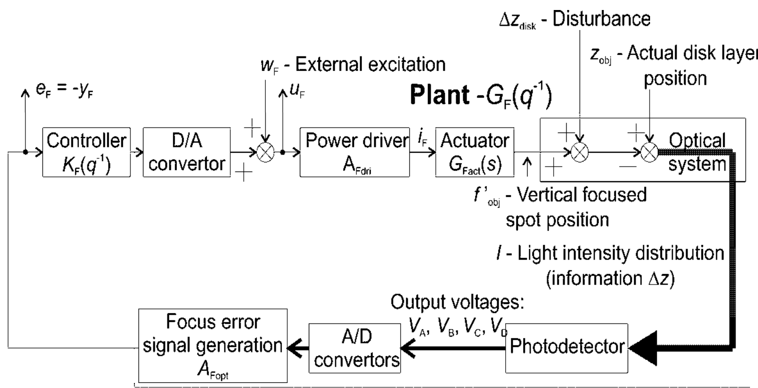


Figure 4. The equivalent block diagram of the focus control loop of fig. 3

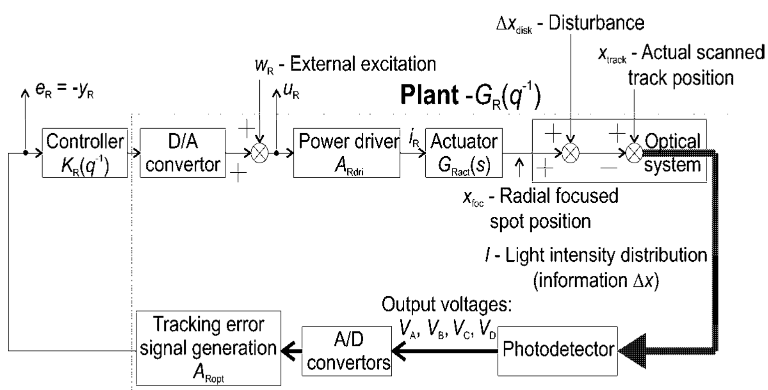


Figure 5. The equivalent block diagram of the tracking control loop of fig. 3

Since the outputs from the plants y_F, y_R are directly error signals e_F, e_R and since the aim of control design is to minimize the vertical/radial displacements Δz, Δx despite the presence of internal and external disturbances, the reference position errors signals r_F, r_R are always considered zero.

$K_F(q^{-1})$, $K_R(q^{-1})$ denote the digital controller's transfer functions which digital outputs are converted to analog signals by the analog to digital converters. Finally, the power drivers, which gains are $K_F(q^{-1})$, $K_R(q^{-1})$, amplify the controllers outputs voltages u_F, u_R in order to supply the focus/radial coils by sufficient currents i_F, i_R.

3. Disturbances sources

The disturbance sources affecting the DVD systems are different and consist mostly of optical imperfections, see fig. 6, which show up at the photodetector (13), see fig. 1. This means that any imperfection, influencing the incident laser beam on the photodetector (13), and any photodetector imperfections (e.g. the displacement, non-uniform sensitivity, noise) cause the disturbance of the focus and radial error signals e_F, e_R.

Laser noise (fig. 6a) gives a background high frequency noise which may alias into the servo frequencies if an anti-aliasing filter is not used. Optical misalignment and optical skew (fig. 6b&c) cause asymmetry and cross coupling between the focus and tracking error signal. The detectors are subject dead-zones between segments (fig. 6d). Lens and groove imperfections show up as spurious signals on the detectors (fig. 6e). Disk warping (fig. 6g), so-called a vertical deviation, puts a large repetitive error into the focus loop. Disk misalignment on the spindle (fig. 6f) causes a large repetitive error at the spindle frequency f_{rot} into the tracking loop. Disk thickness variations and disk tilt also affect both the focus e_F and tracking e_R error signals (fig. 6h).

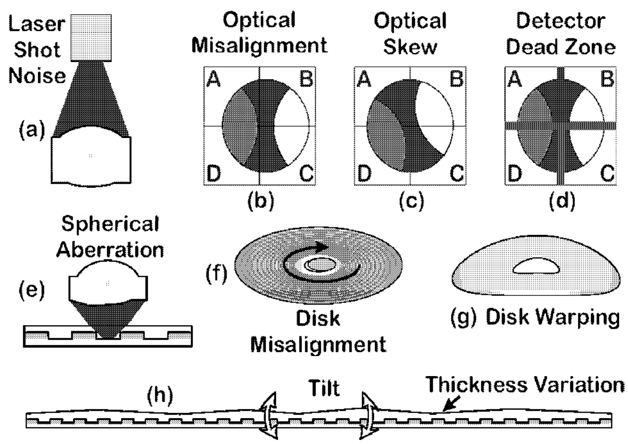


Figure 6. Sources of optical disk noises

The numerical aperture of the lens is defined as $NA = \sin(\phi_{\mathrm{max}}/2)$, where ϕ_{max} is the full angle of the cone of the light rays that can pass through the objective lens (7), (see fig. 1): $NA = 0.45$ is for CD players while $NA = 0.6$ is for DVD systems.

The higher numerical aperture NA, presented in DVD players, induces the following phenomena:

1 Increase the focused spot size, a diameter of the Airy disk, [4], is given by $d_{\mathrm{Airy}} \approx 1.22\lambda/NA$. λ is wavelength of the laser beam in

air. $\lambda = 780\,\mathrm{nm}$ is for CD players while $\lambda = 650\,\mathrm{nm}$ is for DVD systems.

2 Decreases the focus depth, defined as $\Delta z_{\max} = \lambda/2NA^2$.

3 Increases the coma nonlinearly where the coma is the primal aberation of objective lens causing the non-uniform intensity distribution of the focused laser spot, [4].

4 Causes problems with tilt and disk thickness variations, pushing higher density systems towards thinner protective layers and less removability.

3.1. External disturbances

The external disturbances affecting drives are typically environmental shocks and vibrations, whether from a moving vehicle, a factory floor environment, a computer under a desk being kicked, or simply the motion of a laptop computer. For streaming media such as DVD/CD, they are usually overcome by data buffering in the portable players or cars. Therefore, the use of accelerometers in DVD/CD players has been limited. Nevertheless, disturbance cancellation is used in random access applications, as hard disks.

3.2. Internal disturbances

All DVD and last CD drivers rely on mechanical concepts similar to those depicted in fig. 7. The focus and radial actuators are dedicated to keep the laser spot in focus and on track. They can perform fine displacements along the focus and respectively radial direction relative to the disk while being positioned by a sledge at a raw radial location. The sledge forms a rigid body together with the turntable motor and turntable itself, being further consolidated on what is called the baseplate. In this structure, the baseplate and the housing itself are considered as one body. Therefore, the internal disturbances are largely stimulated by the spindle rotation of the disk and the actuators reaction forces on the drive baseplate and housing.

A large percentage of the internal disturbances are synchronous with the spindle frequency f_{rot} and its harmonics including oscillations in the disk media. The power spectrum of the radial error signal e_{R}, at given rotational frequency $f_{\mathrm{rot}} = 15\,\mathrm{Hz}$, has been acquired on the real DVD system in STMicroelectronics laboratories. It is shown in fig. 22. One can see that main disturbances are given by the first and the third harmonic components of disk rotational frequency f_{rot}.

The optical cross coupling between the focus and tracking loops, see fig. 6c, have been minimized for the DVD players. The CD mechanism,

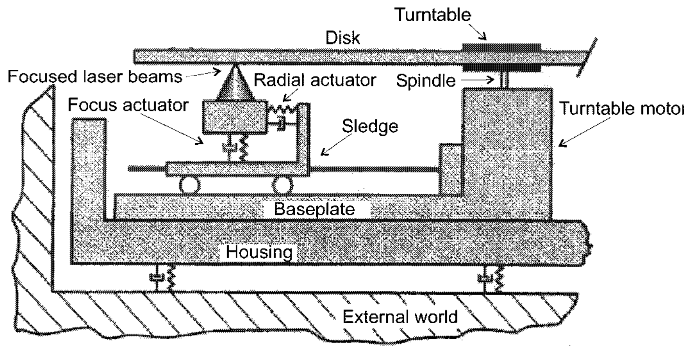


Figure 7. Schematic representation of the mechanical construction of the DVD/CD players

shown in fig. 8, usually consists of a radial arm in order to follow the spiral track of the disk (a course tracking loop). This induces the optical cross coupling between the focus and tracking loops, caused by a changing of skew angle. Instead of the radial arm, all DVD drives use a linear actuator to roughly position the optical head assembly in the vicinity of the desired tracks, see fig. 9. This improvement prevents the changing of skew angle. Therefore, the optical cross coupling between the focus and tracking loops is suppressed.

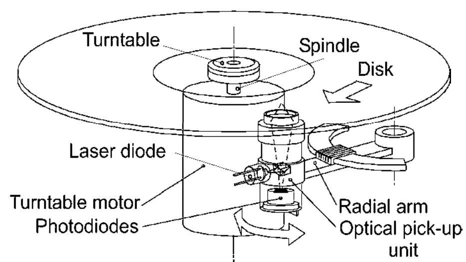


Figure 8. Schematic view of the CD mechanism

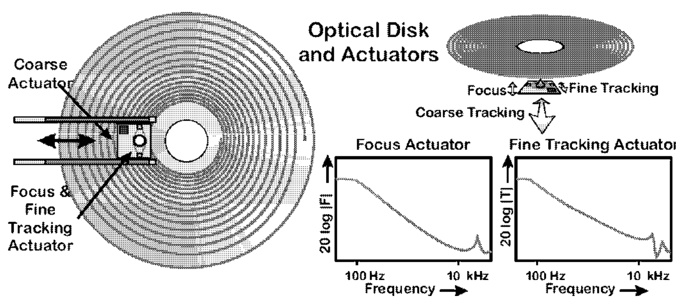


Figure 9. Optical disks and actuators of the DVD player

More detail information on the presented disturbances sources of both the DVD/CD and hard disks systems are in [1] and its references.

4. Specification of control loops

4.1. Focus control loop

The laser spot position must follow the disk track (in the focus/radial directions) despite the presence of disturbances, originating inside and outside the DVD drive as explained in section 1.

To specify the servo system for focus, a function H_s, (1), is used. It specifies the nominal values of the open-loop transfer function for so-called *reference servo* between the focus error signal e_F and output signal y_F of the plant, (see fig. 4). Specification of the open loop H_s is given in the frequency range 23,1 Hz to 10 kHz.

$$H_s(i\omega) = \frac{1}{3} \times \left(\frac{\omega_c}{i\omega}\right)^2 \times \frac{1 + \frac{3i\omega}{\omega_c}}{1 + \frac{3i\omega_c}{\omega}}, \qquad (1)$$

where $\omega = 2\pi f$, $\omega_c = 2\pi f_c$ and $i = \sqrt{-1}$. f is frequency in general. f_c is the 0 dB crossover frequency of the open loop transfer function H_s.

The specification contained in [7], mainly for the over-speed factor $N = 1$, prescribes two maximal deviations Δz_{low}, Δz_{high} from the nominal position of the disk and the maximal acceleration a of the scanning point at given frequencies, as briefly presented in table 1.

Table 1. Focus servo specification for the DVD, $N = 1$, with normalized servo.

	Parameter	Range (Hz)	Value
Δz_{low}	Max. deviation from nominal position	$f_{\text{rot}} \leq 23.1$	± 0.3 mm
a	Max. vertical acceleration	$f_{\text{rot}} > 23.1$	$8\ \text{m/s}^2$
Δz_{high}	Max. deviation from nominal position	$f_{\text{rot}} > 2000$	$\pm 0.23\ \mu\text{m}$
f_c	Crossover frequency of the loop		2 kHz
f^{BW}	Desired closed loop bandwidth		$\approx f_c$
x_{ini}	Min. radius	$f_{\text{rot}} = 23.1$	24 mm
x_{fin}	Max. radius	$f_{\text{rot}} = 9.6$	58 mm
v_a	Scanning velocity		3.49 m/s
$\Delta z_{\max}$	Focus depth		$0.903\ \mu\text{m}$

f_c is specified by (2), where the coefficient α is equal to 1.5 , (1), in order to increase the maximal axial acceleration a.

The desired closed loop bandwidth f^{BW} is not specified in [7] but it is approximately equal to the cross-over frequency of the open loop f_c.

$$f_c = \frac{N}{2\pi}\sqrt{\frac{3\,\alpha\,a}{\Delta z_{\mathrm{high}}}} = \frac{1}{2\pi}\sqrt{\frac{3\times 1.5\times 8}{0.23\times 10^{-6}}} = 2\,\mathrm{kHz}. \tag{2}$$

An additional information, concerning the closed loop performances, can be given from the rise time t_r of the closed loop step response, defined as the time it takes for the output y_F to first reach 90 % of its final value. t_r usually verifies the following equation

$$t_r \simeq \frac{2.3}{2\pi f_c} = \frac{2.3}{2\pi\, 2\times 10^3} = 0.183\,\mathrm{ms}. \tag{3}$$

Another important parameter for focusing is so-called focus depth Δz_{max} that establish maximal disk displacement between actual disk information layer (10) in fig. 1 and the position of the focused laser spot of the objective lens (7) in fig. 1. Δz_{max} is defined by

$$\Delta z_{\mathrm{max}} = \frac{\lambda}{2NA^2}. \tag{4}$$

The numerical value $\Delta z_{\mathrm{max}} \doteq 0.903\,\mu\mathrm{m}$ is given by DVD parameters: laser wavelength $\lambda = 650\,\mathrm{nm}$ and numerical aperture of the objective lens $NA = 0.6$. The focus servo should therefore control the objective lens within $\pm\Delta z_{\mathrm{max}}$ to avoid loosing the data read-out signal at every time during playing.

The actual (spindle) rotational frequency f_{rot} of the disk is given by

$$f_{\mathrm{rot}} = \frac{N v_a}{2\pi x}, \tag{5}$$

where v_a is the scanning velocity and x is the distance between the disk rotational axis and the falling laser beam, $x \in \langle x_{\mathrm{ini}},\, x_{\mathrm{fin}}\rangle$, (see figs. 11 and 12).

Finally, fig. 10 shows the upper and lower limits of the disturbance rejection requirements on the output harmonic disturbances in frequency domain, taken from specification [7], for the focus control loop, $N = 1$. Here, the curve L_1 illustrates the requirements for manufacturing of the disk only. The curve L_2 is an upper and curve L_3 is a lower limit on the output sensitivity function S_{yp} (defined in section 2) for controller design. L_2 and L_3 are defined by percentage variation from the open-loop transfer function H_s at the given frequency ranges, see [7].

For an open loop transfer function of the real system H_{real}, $|1 + H_{\mathrm{real}}|$ is limited by the shaded surface in fig. 10.

This specification is equivalent to give a template for the output sensitivity function S_{yp}. More detailed explanation on how to use these requirements in control procedure directly will be described in section 13.

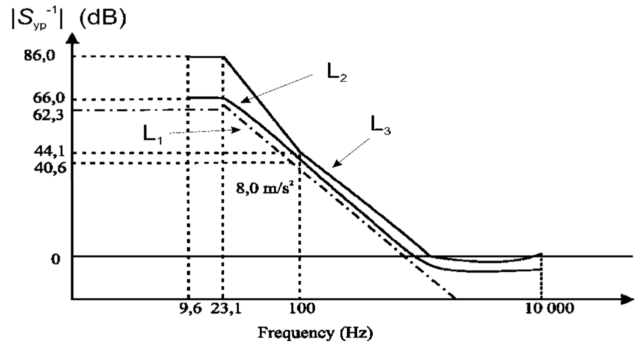


Figure 10. Specification of the focus control loop, $N = 1$

4.2. Radial control loop

Radial servo specifications are given in a similar way in [7]. For an over-speed factor $N = 1$, they prescribe two maximal values Δx_{low}, Δx_{high} for the maximal eccentricity of the track radius and maximal axial acceleration a of the disk at given frequencies, as briefly presented in table 2.

Table 2. Radial servo specification for the DVD, $N = 1$, with normalized servo.

	Parameter	Range (Hz)	Value
Δx_{low}	Max. eccentricity of the track radius	$f_{\text{rot}} \leq 23.1$	$\pm 50\,\mu\text{m}$
a	Max. axial acceleration	$f_{\text{rot}} > 23.1$	$1.1\,\text{m/s}^2$
Δx_{high}	Max. eccentricity of the track radius	$f_{\text{rot}} > 2400$	$\pm 0.022\,\mu\text{m}$
f_c	Crossover frequency of the open loop		$2.4\,\text{kHz}$
f^{BW}	Desired closed loop bandwidth		$\approx f_c$
$\Delta x_{\max}$	Max. radial displacement		$0.074\,\mu\text{m}$

5. Control design methodology

5.1. The aims of the controllers design

5.1.1. Focus loop

The goal of the focus controller design is to minimize the amplitude of the focus displacement Δz, measured by the focus error signal e_{F}, during playback.

The periodical disturbances are mainly given by the rotational frequency f_{rot}, that is not constant during playback, (see expression (5)), because a constant linear velocity (CLV) is used to read the data recorded on the disk. These disturbances are mainly caused by the disk eccentricity Δx_{disk}, disk warping, disk thickness variation. Fig. 11, that is a zoom on the *optical block* in fig. 4, shows the geometry of the vertical deviation sources.

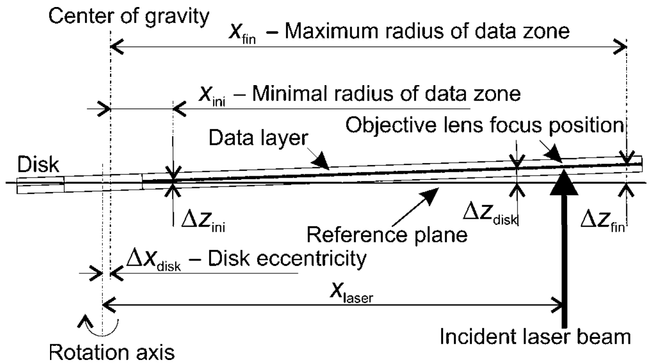


Figure 11. Sources of vertical deviation on the DVD disk, mainly caused by disk warping and disk tracking eccentricity

In real DVD/CD players, the disk is not balanced ideally because there is a non-zero distance, so-called disk eccentricity $\Delta x_{\text{disk}} \neq 0$, between the center of gravity of the disk and the rotation axis of the spindle, see fig. 11.

It is clear that the vertical deviation of the turning disk Δz_{disk} is not constant along the whole scanned track on the disk, (Δz_{disk} depends on the laser beam position x_{laser} from the rotation axis of the spindle).

Therefore at the starting radius of the data zone x_{ini}, (where $f_{\text{rot}} = 23.1\,\text{Hz}$), the vertical deviation Δz_{ini} is *smaller* than the vertical deviation Δz_{fin} at the maximum radius of the data zone x_{fin} (where $f_{\text{rot}} = 9.6\,\text{Hz}$).

Hence, the different requirements on the disturbance rejection can be given in frequency domain. The sensitivity function shaping method is a useful tool to design controllers satisfying these system requirements.

5.1.2. Radial loop

The goal of the radial controller design is to minimize the amplitude of the radial displacement Δx, measured by the radial error signal e_{R}, during playback.

The periodical disturbances are also mainly given by the rotational frequency f_{rot}. The disturbances are mainly caused by the disk eccentricity Δx_{disk} and disk tilt.

The geometry of the radial deviation sources is illustrated in fig. 12. It is clear that the radial eccentricity Δx_{disk} does not depends on the laser beam position from the spindle rotation axis x_{laser}.

Therefore, during disk rotation, the disk eccentricity Δx_{disk} is still constant from the starting radius of the data zone x_{ini}, (where $f_{\text{rot}} =$

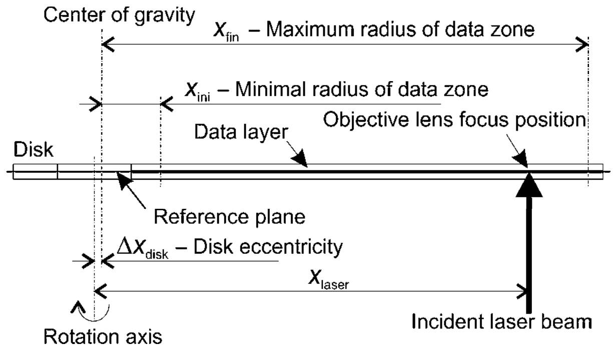


Figure 12. The main source of radial deviation on the DVD disk—track eccentricity

23.1 Hz), to the maximum radius of the data zone x_{fin} (where $f_{\text{rot}} = 9.6\,\text{Hz}$).

Nevertheless, the controller design should take also into account information that on the DVD/CD disk the data area increases proportionally to the square of the disk radius x_{fin}. Therefore the disturbance rejection should be larger near the maximal disk radius x_{fin} then near the minimal disk radius x_{fin}.

The sensitivity function method to control design is also useful in the radial control loop because the requirements on the disturbance rejection are given in frequency domain too.

5.2. Combined pole placement/sensitivity function shaping

The standard digital control configuration obtained with polynomial RST controller, [9], is presented in fig. 13. Part T of RST structure has been omitted because the control design in the focus/tracking loop of DVD/CD players only contains the disturbance rejection.

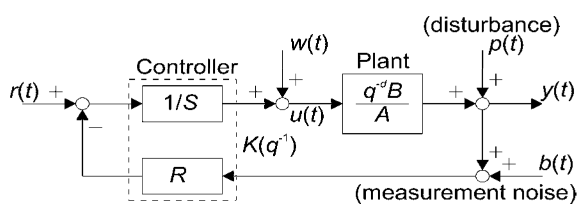


Figure 13. The closed loop system with RS controller

Pole placement method have been used to design the controller's parts RS.

Fig. 13 is a simpler version (general structure) of the detailed block diagrams of the focus/tracking control loops illustrated in figs. 4 and 5, respectively. y, y_{F}, y_{R} are the plant outputs. u, u_{F}, u_{R} are the plant inputs. r, z_{obj}, x_{track} are the reference signals. $K(q^{-1})$, $K_{\text{F}}(q^{-1})$, $K_{\text{R}}(q^{-1})$ are the controllers.

The signal $w(t)$ in fig. 13 presents both the input disturbance of the plant (i.e. the D/A convertor noise in our case) and the excitation signal which is used for the identification purposes, like shown in fig. 4. However, here $w(t)$ is a sampled signal and in fig. 4 w_{F} is an continuous time signal. So, here it can be considered as an equivalent signal of the physical signal w_{F} from fig. 4.

The disturbance signal p in fig. 13 presents all disturbances incoming into the plant, i.e. the power driver noise, actuator non-linearity, cross coupling, disk eccentricity, disk warping, disk thickness, disk tilt, spherical abberation, groove distortion, vibrations, shocks, photodetector displacement and dead zone.

The measured noise b in fig. 13 presents the sensing noise, A/D convertors noise and non-linearity in function of focus/tracking error signal generation.

The linear-time-invariant model of the plant is described by the transfer function

$$G(q^{-1}) = \frac{q^{-d}\,B(q^{-1})}{A(q^{-1})} = \frac{q^{-d}\left(b_1\,q^{-1} + \cdots + b_{n_B}\,q^{-n_B}\right)}{1 + a_1\,q^{-1} + \cdots + a_{n_A}\,q^{-n_A}}, \tag{6}$$

where q^{-1} is the backward time shift operator, d is the pure time delay, T_{s} = the sampling period and sampling frequency $f_{\text{s}} = 1/T_{\text{s}}$.

Experience shows that simple linear model of DVD/CD system leads to sufficient high-performance controllers, [2] given by specifications presented in section 9 and in [7]. A more complex controllers, based on the high order model, have higher performance but they are practically useless. It is given from the constrainers during controllers implementation where the low-order controllers are still preferred.

The RS controller has the following transfer function

$$\begin{aligned} K(q^{-1}) &= \frac{R(q^{-1})}{S(q^{-1})} = \frac{R'(q^{-1})\,H_{\text{R}}(q^{-1})}{S'(q^{-1})\,H_{\text{S}}(q^{-1})} \\ &= \frac{r_0 + r_1\,q^{-1} + \cdots + r_{n_{\text{R}}}\,q^{-n_{\text{R}}}}{1 + s_1\,q^{-1} + \cdots + s_{n_{\text{S}}}\,q^{-n_{\text{S}}}}, \end{aligned} \tag{7}$$

where $H_{\text{R}}(q^{-1})$ and $H_{\text{S}}(q^{-1})$ denote the fixed parts of the controller (either imposed by the design or introduced in order to shape the sensitivity functions). $R'(q^{-1})$ and $S'(q^{-1})$ are solutions of the Bezout equation

$$AS'H_{\text{S}} + BR'H_{\text{R}} = P, \tag{8}$$

where P represents the characteristic polynomial (closed loop poles).

The four sensitivity functions $S_{ij}(K, G)$ are defined as follows. The output sensitivity function is $S_{\mathrm{yp}}(q^{-1}) = \dfrac{y(t)}{p(t)} = \dfrac{AS'H_S}{P}$, the input sensitivity function is $S_{\mathrm{up}}(q^{-1}) = \dfrac{u(t)}{p(t)} = -\dfrac{AR'H_R}{P}$, the output sensitivity function with respect to an input disturbance $S_{\mathrm{yw}}(q^{-1}) = y(t)/w(t) = = q^{-d}BS'H_S/P$ and the complementary sensitivity function is $S_{\mathrm{yr}}(q^{-1}) = y(t)/r(t) = q^{-d}BR'H_R/P = -S_{\mathrm{yb}}(q^{-1})$. Sensitivity functions play an important role in the robustness analysis of the closed loop system with respect to modelling errors. One can see that the equation $S_{\mathrm{yp}} - S_{\mathrm{yb}} = 1$ is valid between the sensitivity functions.

In our case, not only robustness (the modulus margin ΔM, delay margin $\Delta\tau$ and phase margin $\Delta\phi$, [9]) but also the performances specifications, [7], have to be check. Fortunately, the robustness requirements are related to sensitivity functions. Therefore the sensitivity function shaping is a useful tool to the controller design in case of DVD/CD players.

5.3. Controller order reduction

Controller order reduction is a very important issue in many control application, either because the size of the controller is limited by hardware and computation time or because simpler controllers are easier to implement. What is most important is that controller reduction should aim to preserve the required closed loop properties as far as possible. One of useful methodology, which is used here, is the balanced reduction [13] method in state space domain that using a Gramian of the balanced state-space realization.

Unfortunately, the given structure of controller in DSP is not general and therefore only some designed the 4rd/3rd order controllers are implementable.

5.4. Generalized stability margin

The resulting reduced order controller should stabilize the nominal model and should give sensitivity functions which are close to the nominal ones in the critical frequency regions, to ensure performance and robustness. One way to verify the stability margin of the whole system is the generalized stability margin $b(K, G)$, [12], defined from all sensitivity functions

$$b(K,G) = \begin{cases} \left\|\mathbf{T}(K,G)\right\|_{\infty}^{-1} & \text{if } (K,H) \text{ is stable,} \\ 0 & \text{otherwise,} \end{cases} \tag{9}$$

where

$$\mathbf{T}(K,G) = \begin{bmatrix} S_{\mathrm{yr}} & S_{\mathrm{yw}} \\ -S_{\mathrm{up}} & S_{\mathrm{yp}} \end{bmatrix},$$

in which S_{yr}, S_{yw}, S_{up} and S_{yp} have been defined is section 12. The generalized stability margin gets higher the large value of $b(K,G)$ is achieved.

6. Control design

An application of introduced methodology will be presented in more detail for the tracking loop. Nevertheless, the final results are given for both the focus and tracking control loop.

6.1. Plant model in tracking loop

A simplified, linear transfer function of the tracking control loop $G_{\mathrm{R}}(s)$ is derived from the physical equations of radial system as follows. In table 3 its parameters are presented.

Table 3. Parameters of the plant in radial control loop

Symbol	Parameter	Nominal value
R_{R}	Coil resistance	$6.5\,\Omega$
L_{R}	Coil inductance	$18\,\mu$H
K_{Re}	Back efm constant	0.061 Vs/m
M_{R}	Actuator moving mass	0.33×10^{-3} kg
D_{R}	Damping constant	0.014 Ns/m
K_{Rs}	Elastic constant	35.2 N/m
K_{Rf}	Force constant	0.061 N/A
A_{Rdri1}	First power driver gain	3.13 V/V
A_{Rdri2}	Second power driver gain	4 V/V
A_{Ropt}	Optical gain & remanent gains	1.447×10^{6} V/m

$$G_{\mathrm{R}}(s) = \frac{\frac{K_{\mathrm{Rf}}}{M_{\mathrm{R}}L_{\mathrm{R}}} A_{\mathrm{Rdri1}} A_{\mathrm{Rdri2}} A_{\mathrm{Ropt}}}{s^3 + \left(\frac{R_{\mathrm{R}}}{L_{\mathrm{R}}} + \frac{D_{\mathrm{R}}}{M_{\mathrm{R}}}\right) s^2 + \left(\frac{D_{\mathrm{R}}R_{\mathrm{R}}}{M_{\mathrm{R}}L_{\mathrm{R}}} + \frac{K_{\mathrm{Rs}}}{M_{\mathrm{R}}} + \frac{K_{\mathrm{Rf}}K_{\mathrm{Re}}}{M_{\mathrm{R}}L_{\mathrm{R}}}\right) s + \frac{K_{\mathrm{Rs}}R_{\mathrm{R}}}{M_{\mathrm{R}}L_{\mathrm{R}}}}. \tag{10}$$

The discrete transfer function of the radial system $G_{\mathrm{R}}(q^{-1})$ is given by conversion from the continuous-time to the discrete-time using a zero-order hold and the sampling period T_{s}, (a clock period of the Digital Signal Processor (DSP)), by the following expression:

$$G_{\mathrm{R}}(q^{-1}) = \frac{b_1 q^{-1} + b_2 q^{-2} + b_3 q^{-3}}{1 + a_1 q^{-1} + a_2 q^{-2} + a_3 q^{-3}}. \tag{11}$$

Here $d = 0$, see (6). The other numerical values for $B(q^{-1})$ and $A(q^{-1})$ polynomial coefficients are not presented for reason of confidentiality. The radial closed loop system is shown in fig. 14. In this figure, disturbances have been added, namely the radial eccentricity $\Delta x_{\mathrm{disk}}(t)$ of the disk.

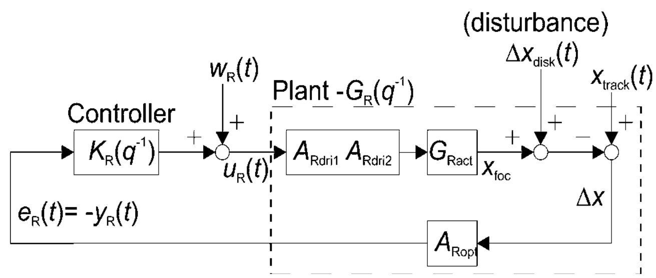


Figure 14. Radial tracking closed loop of the DVD

6.2. Nominal and uncertainty model

To verify the robustness of the proposed control systems, different models of radial actuator, based on the parameters specification and their variation, have been created. The nominal model of the plant is obtained considering the values of physical parameters of a DVD pick-up, which is used in the industrial application. Then, the model set is created by taking into account the variation of each physical parameter in an interval of values, as indicated in the pick-up data-sheet, [8]. The nominal values of the actuator physical parameters, together with their maximum percentage variation, are shown in table 4. f_{Rn} and K_{RDC} are the values of the actuator resonance frequency and DC sensitivity, that are used to compute the values of K_{Re}, D_{R}, K_{Rs} and K_{Rf} in the model transfer function (10).

Table 4. Values of the radial actuator physical parameters together with their maximum percentage variation

Symbol	Nominal value	Variation
R_{R}	6.5 Ω	±15 %
L_{R}	18 μH	±33 %
f_{Rn}	52 Hz	±5 %
M_{R}	0.33 g	±10 %
K_{RDC}	$0.27 \cdot 10^{-3}$	±20 %

6.3. Standard controller

The second order lead-lag controller $K_{\text{act}}(q^{-1})$, given by (12), is used in many actual DVD/CD applications as a standard controller structure.

$$K_{\text{act}}(q^{-1}) = g_0 \frac{\left(1 - c_1\, q^{-1}\right)\left(1 - c_2\, q^{-1}\right)}{\left(1 - d_1\, q^{-1}\right)\left(1 - d_2\, q^{-1}\right)}. \tag{12}$$

This standard second order lead-lag controller ($K_{\text{act}}(q^{-1}) : n_{\text{R}} = 2, n_{\text{S}} = 2$) is actually replaced by a third order ($K_{\text{RS3}}(q^{-1})$: $n_{\text{R}} = 3, n_{\text{S}} = 3$) or a fourth order ($K_{\text{RS4}}(q^{-1})$: $n_{\text{R}} = 4, n_{\text{S}} = 4$) controller in order to meet higher performance on the disturbance rejection.

The aim of this work is to provide a methodology to design the third (or fourth) order controllers that improve the actual performance on the disturbance rejection and fulfil the realization constrains in actual DVD platform, see section 13.

6.4. New controller design

The third order RS controller $K_{\text{RS3}}(q^{-1})$, ($n_{\text{R}} = 3, n_{\text{S}} = 3$), and the fourth order RS controller $K_{\text{RS4}}(q^{-1})$, ($n_{\text{R}} = 4, n_{\text{S}} = 4$), are designed for the over-speed factor $N = 1.5$.

The spindle rotational frequency is $f_{\text{rot}} = 34.7\,\text{Hz}$ at the data zone starting radius x_{ini} while $f_{\text{rot}} = 14.4\,\text{Hz}$ at the data zone maximum radius x_{fin}. The controller design has taken also into account that on the DVD/CD disk the data area increases proportionally to the square of the disk radius x_{laser} in fig. 12.

When the over-speed factor N is bigger than 1, the break rotational frequency in performance specification $f_{\text{cor}} = 23.1\,\text{Hz}$, must be linearly shifted by the same factor N. Therefore the minimum sensitivity

$$S_{\text{low}} = 20 \log \left(\frac{|\Delta x_{\text{high}}|}{|\Delta x_{\text{low}}|} \right) = -67.13\,\text{dB} \tag{13}$$

is required at frequency $f_{\text{rot}} = 23.1 \cdot 1.5 = 34.7\,\text{Hz}$.

To suppress periodical disturbances, mainly caused by disk eccentricity Δx_{disk} at rotational frequency f_{rot}, a slight modification on the $|S_{\text{yp}}|$ template has been done:

1 $f_{\text{rot}} = 34.7\,\text{Hz}$:
$|S_{\text{yp}}| = S_{\text{low}} - 27\,\text{dB} = -94.13\,\text{dB}$;

2 $f_{\text{rot}} = 14.4\,\text{Hz}$:
$|S_{\text{yp}}| = 20 \log \left(\frac{|\Delta x_{\text{high}}|}{|\Delta x_{\text{low}}|} \cdot \frac{x_{\text{ini}}}{x_{\text{fin}}} \right) - 27\,\text{dB} = -101.80\,\text{dB}$;

3 $f_{\text{rot}} \in \langle 14.4\,\text{Hz}, 34.7\,\text{Hz} \rangle$:
$|S_{\text{yp}}|$ = linear interpolation between $|S_{\text{yp}}|$ at two given frequencies: $f_{\text{rot}} = 14.4\,\text{Hz}$ and $f_{\text{rot}} = 34.7\,\text{Hz}$;

4 $f_{\mathrm{rot}} > 34.7\,\mathrm{Hz}$: $|S_{\mathrm{yp}}|$ is given by specification in [7] where the upper and lower limits on the $|1/(1 + H_s)|$ are defined.

The low limit at the rotational frequency $f_{\mathrm{rot}} = 34.7\,\mathrm{Hz}$ have been toughened up rather 27 dB. The low limit at the rotational frequency $f_{\mathrm{rot}} = 14.4\,\mathrm{Hz}$ have been linearly increased with respect on the ratio of the data zone radiuses $(x_{\mathrm{ini}}/x_{\mathrm{fin}})$, to take into account data distribution on the disk, and shifted rather 27 dB too.

The specification requirements (defined by the DVD disk and by normalized servo) with our modification are illustrated as the output sensitivity function modulus $|S_{\mathrm{yp}}|$ templates in fig. 15.

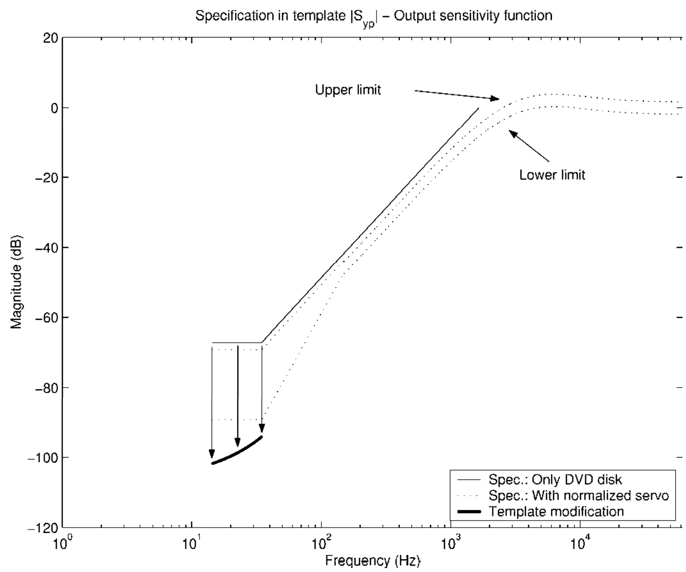


Figure 15. Desired template for the modulus of the output sensitivity function $|S_{\mathrm{yp}}|$ for radial tracking (radial loop) of the DVD, $N = 1.5$

The controller design of the fourth order controller K_{RS4} has been realized using the following specifications:

- P (closed loop poles):
  - a pair of complex poles near the model's slowest vibration frequency $f_{\mathrm{D}} = 52\,\mathrm{Hz} \rightarrow 170\,\mathrm{Hz}$ but well damped $\xi_{\mathrm{D}} = 0.068 \rightarrow 0.936$;
  - two multiple real poles $\gamma_{\mathrm{F}} = 0.9$ for keeping in the $|S_{\mathrm{yp}}|$, $|S_{\mathrm{up}}|$ templates;
  - one complex pole $f_{\mathrm{F}} = 13800\,\mathrm{Hz}$, $\xi_{\mathrm{F}} = 0.927$ to restrain the controller action in higher frequencies where the gain of the system is low;
- H_{S}: a pair of complex poles $f_{\mathrm{S}} = 19\,\mathrm{Hz}$, $\xi_{\mathrm{S}} = 0.4$ to ensure disturbance rejection in the frequency range $f_{\mathrm{rot}} \in \langle 14.4\,\mathrm{Hz}, 34.7\,\mathrm{Hz}\rangle$.

- H_R: a real zero $\gamma_R = 0.1$ to lower the magnitude of the input sensitivity function $|S_{up}|$ at high frequencies where the gain of the system is low.
- The resulting controller has the orders $n_R = 5$ and $n_S = 5$. Therefore the balanced reduction method [13] has been used to obtain a controller structure $n_R = 4$ and $n_S = 4$.
- Check the sensitivity functions $|S_{yp}|$, $|S_{up}|$, $|S_{yr}|$ again. End the design procedure if the requirements on the sensitivity functions were satisfied.

The same procedure has been used to the third order RS controller K_{RS3} design except balanced reduction. Nevertheless, the restrictions on the resulting controller gives smaller performance in the frequency range $f_{rot} \in \langle 14.4\,\text{Hz}, 34.7\,\text{Hz}\rangle$. The controllers are designed using the recently developed software tool "ppmaster" [11], developed in MATLAB®environment.

7. Performance analysis

7.1. Focus loop: Simulation experiments

The results are presented for the final RS controllers of the 3rd/4th order (K_{RS3}, K_{RS4}) in table 6 and 5. More detailed results are given in [5].

Table 5. Comparison of the controller order reduction and generalized stability margin, $N = 1.5$, focus loop.

$K(q^{-1})$	n_R (−)	n_S (−)	$\widehat{n_R}$ (−)	$\widehat{n_S}$ (−)	$b(K,G)$ (−)	$\widehat{b(K,G)}$ (−)
K_{act}	3	3	3	3	0.15563	
K_{RS3}	3	3	4	4	0.16476	0.16476
K_{RS4}	4	4	5	5	0.23351	0.23294

Table 6. Comparison of the various reduced controllers, $N = 1.5$, focus loop.

$K(q^{-1})$	n_R (−)	n_S (−)	$\lvert S_{yp}\rvert_{max}$ (dB)	$\lvert S_{yp}\rvert_{14.4}$ (dB)	$\lvert S_{yp}\rvert_{34.7}$ (dB)	$\lvert S_{up}\rvert_{max}$ (dB)	t_r (μs)
K_{act}	3	3	3.02	−73.5	−74.9	16.1	72.9
K_{RS3}	3	3	3.28	−80.2	−82.5	15.5	72.9
K_{RS4}	4	4	3.45	−88.0	−80.3	12.3	81.0
Spec.			3.85	−66.0	−66.0		122.5

7.2. Focus loop: Real-time measurements

The disturbance elements of the focus error signal e_F, in frequency domain at two different rotational frequencies f_{rot}, are illustrated in figs. 16 and 17, where the power spectrum of the focus error signal e_{F} has been acquired on the real DVD system with the 3rd-order controller K_{RS3}. These results point out that the obtained improvements are still influenced by disk rotational frequency f_{rot}.

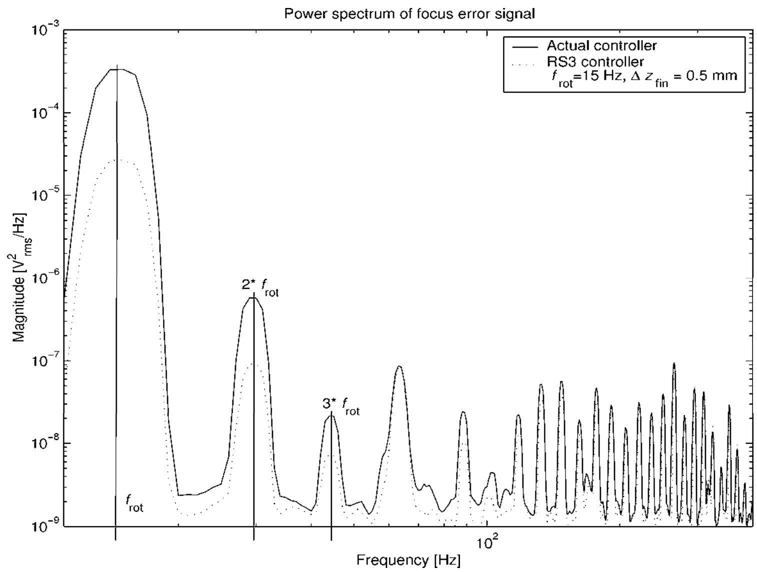


Figure 16. Closed-loop with K_{act}, K_{RS3}. The measured power spectrum density of the focus error signal e_{F} for tested disk with very small disk eccentricity Δx_{disk}, but with high disk vertical deviation at the disk outer edge $\Delta z_{\mathrm{fin}} = 0.5\,\mathrm{mm}$, $f_{\mathrm{rot}} = 15\,\mathrm{Hz}$

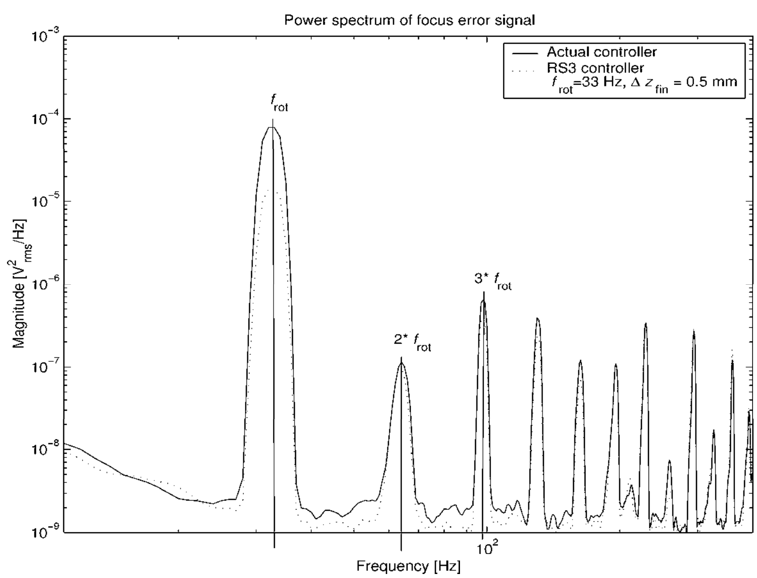


Figure 17. The same as in fig. 16 but for $f_{\mathrm{rot}} = 33\,\mathrm{Hz}$

7.3. Tracking loop: Simulation experiments

Results are shown for the actual controller $K_{\rm act}$ and the designed fourth order RS controller $K_{\rm RS4}$. A final comparison is also done with the third order RS controller $K_{\rm RS3}$. In fact, actual controller $K_{\rm act}$ is the second order controller because some constants are set to zero in the implemented structure.

The disturbance rejection is illustrated by the output sensitivity function modulus $|S_{\rm yp}|$ in fig. 18. Notice that the perturbations suppression at $f_{\rm rot} = 19\,{\rm Hz}$ have been achieved by the $H_{\rm S}$ polynomial choice.

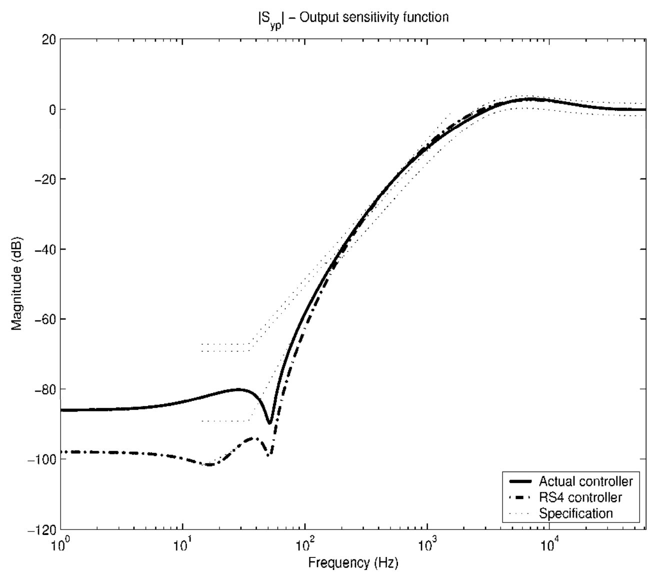


Figure 18. Output sensitivity function, radial loop

Fig. 19 presents the input sensitivity function $|S_{\rm up}|$. The lower peak in $|S_{\rm yp}|$ and lower values of $|S_{\rm up}|$ in high frequencies for the controller $K_{\rm RS4}$ than ones for actual controller $K_{\rm act}$ are seen in table 8. However, in case of the fourth order controller $K_{\rm RS4}$, higher values of $|S_{\rm up}|$ for low frequencies is a problem. It is a trade-off between the disturbance rejection and the robustness requirements.

A good control order reduction methodology and a good generalized stability margin of the third/fourth order designed controllers $K_{\rm RS3}/K_{\rm RS4}$ are shown in table 7, where the parameters of reduced controllers are $n_{\rm R}$, $n_{\rm S}$, $b(K,G)$ and the ones of non-reduced controllers are $\widehat{n_{\rm R}}$, $\widehat{n_{\rm S}}$, $\widehat{b(K,G)}$.

Figs. 20, 21 illustrate the envelopes of the output sensitivity functions modulus that have been calculated for the plant model set and the actual/third order designed controller. One can see that the stability templates and the desired performances are fulfilled in both cases.

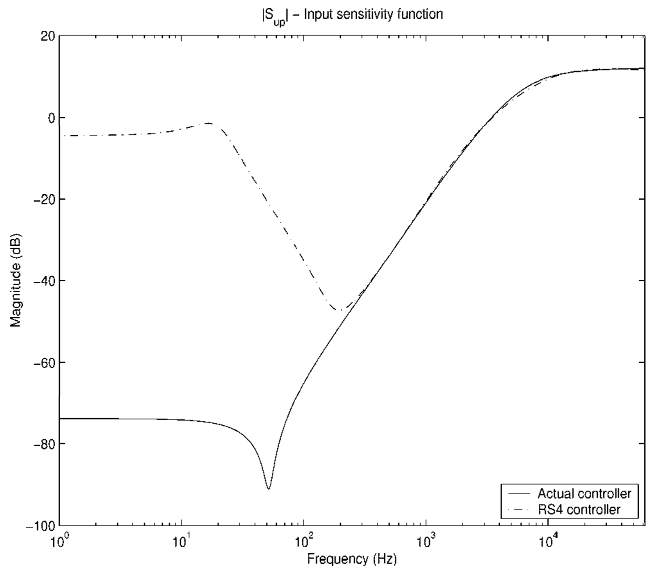


Figure 19. Input sensitivity function, radial loop

Table 7. Comparison of the controller order reduction and generalized stability margin, $N = 1.5$, radial loop.

$K(q^{-1})$	n_R (−)	n_S (−)	$\widehat{n_R}$ (−)	$\widehat{n_S}$ (−)	$b(K,G)$ (−)	$\widehat{b(K,G)}$ (−)
K_{act}	2	2	2	2	0.2127	
K_{RS3}	3	3	3	3	0.2156	
K_{RS4}	4	4	5	5	0.2253	0.2151

Table 8. Comparison of the various reduced controllers, $N = 1.5$, radial loop.

$K(q^{-1})$	n_R (−)	n_S (−)	$\|S_{yp}\|_{max}$ (dB)	$\|S_{yp}\|_{14.4}$ (dB)	$\|S_{yp}\|_{34.7}$ (dB)	$\|S_{up}\|_{max}$ (dB)	t_r (μs)
K_{act}	2	2	2.95	−82.3	−80.6	12.1	64.8
K_{RS3}	3	3	2.88	−94.1	−90.8	12.0	64.8
K_{RS4}	4	4	2.67	−101.3	−94.5	11.8	64.8
Spec.			3.86	−69.2	−69.2		102

7.4. Tracking loop: Real-time measurements

The disturbance elements of the radial error signal e_R, in frequency domain at two different rotational frequencies f_{rot}, are illustrated in figs. 22 and 23, where the power spectrum of the radial error signal e_R has been acquired on the real DVD system with the 3rd-order controller K_{RS3}. These results point out that the obtained improvements are influenced by disk rotational frequency f_{rot}. Some other results, obtained with the third order controller K_{RS3}, are also presented in [6].

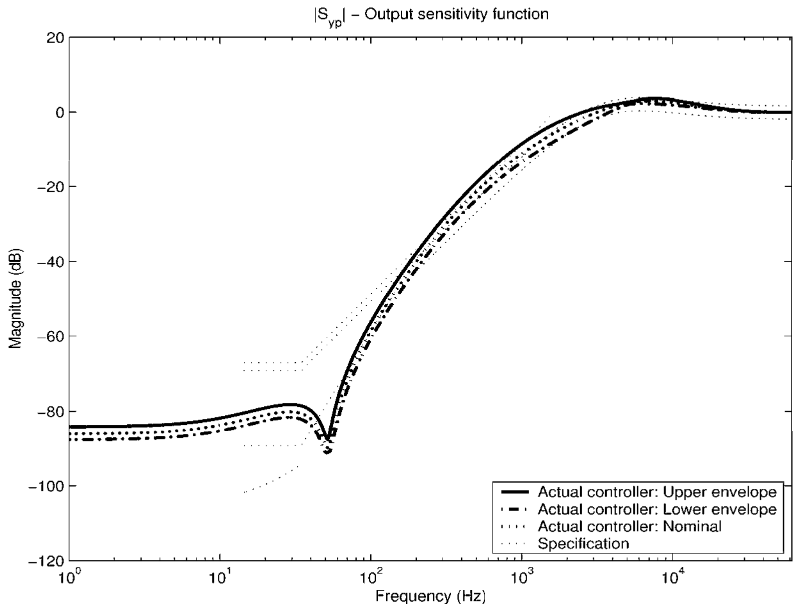


Figure 20. Envelopes of the output sensitivity functions, actual controller, radial loop

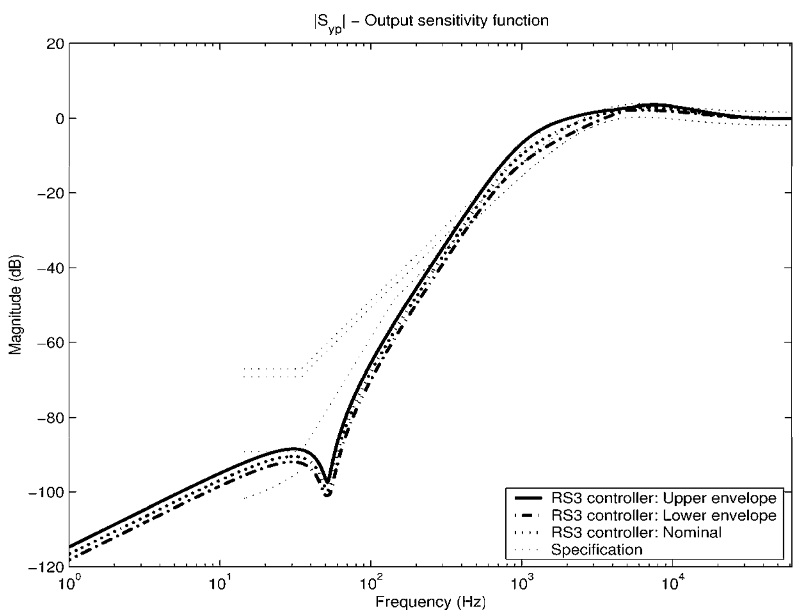


Figure 21. Envelopes of the output sensitivity functions, RS3 controller, radial loop

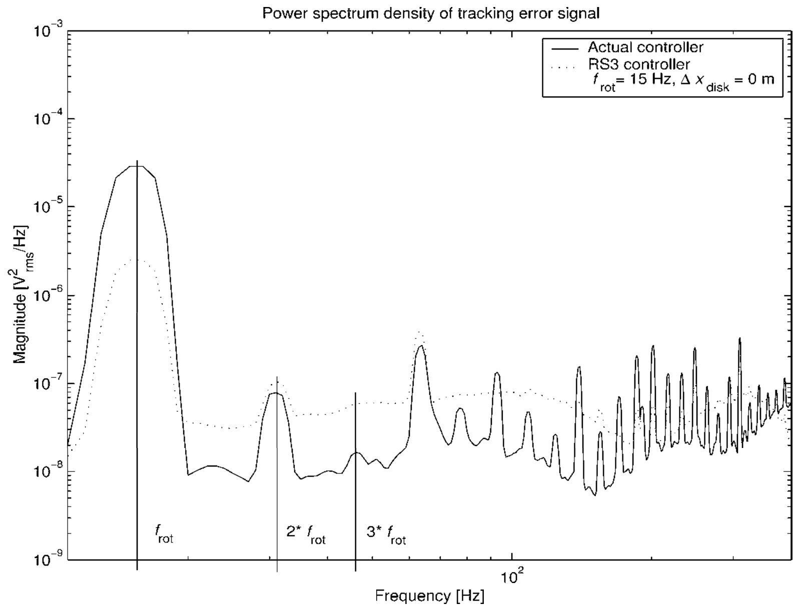


Figure 22. Closed-loop with K_{act}, K_{RS3}. The measured power spectrum density of the radial error signal e_{R} for tested disk with very small disk vertical deviation at the disk outer edge x_{fin}, disk eccentricity $\Delta x_{\mathrm{disk}} = 0\,\mu\mathrm{m}$, $f_{\mathrm{rot}} = 15\,\mathrm{Hz}$

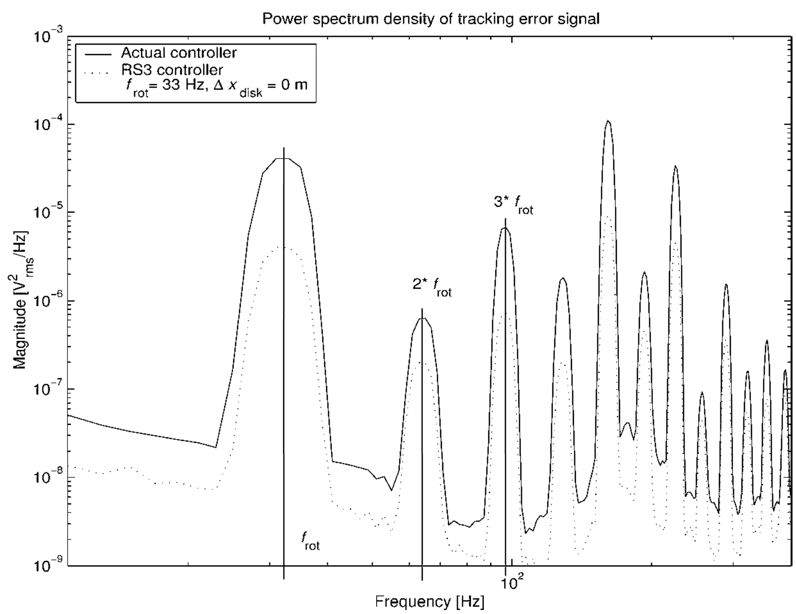


Figure 23. The same as in fig. 22 but for $f_{\mathrm{rot}} = 33\,\mathrm{Hz}$

8. Conclusion

This paper has dealt with combined pole placement/sensitivity function shaping control design connected with controller order reduction for focus/tracking control loop of DVD player, under industrial performance specifications. Modifications on the output sensitivity template have been explained and directly fulfill in controllers design to obtain improvement in term of periodic disturbances.

Robustness analysis point out that the achieved the third/fourth order controllers remain stable for large uncertainty of the actuator physical parameters, and that performance specifications are met. Final comparison of the actual and designed controllers illustrate that new controllers provide better system performances and robustness than the actual controllers. Nevertheless, the presented disturbance rejection is impossible to obtain with more controller order reduction.

Acknowledgements

The work of B. Hnilička was supported by the European Commission (Marie Curie Fellowship) and by the Czech Grant Agency GAČR under the grant number GAČR 102/01/1485 Environment for Development, Modelling and Application of Heterogenous Systems and by the MŠMT research plan CEZ MSM: 260000013 Automation of Technological and Manufacturing Processes.

References

[1] D. Abramovitch. Magnetic and optical disk control: Parallels and contrast. *Proceedings of the 2001 American Controls Conference (ACC)*, pages 421–428, Arlington, Virginia, USA, 2001.

[2] S. Bittanti, F. Dell'Orto, et al. Radial tracking in high-speed DVD players. *Proceedings of the 40th IEEE Conference on Decision and Control (CDC)*, 4705–4710, Orlando, Florida, USA, 2001.

[3] S. Bittanti and S. M. Savaresi. Safe estimate of sinusoidal signals for control applications. *Proceedings of the 38th IEEE Conference on Decision and Control (CDC)*, 2827–2832, Phoenix, Arizona, USA, 1998.

[4] M. Born and E. Wolf. *Principles of Optics*. Pergamon, New York, 6th edition, 1987.

[5] B. Hnilička, A. Besançon-Voda, et al. Pole placement/sensitivity function shaping and controller order reduction in DVD players (Focus control loop). *European Control Conference (ECC)*, Cambridge, U.K., 2003.

[6] B. Hnilička, A. Besançon-Voda, et al. Pole placement/sensitivity function shaping and controller order reduction in DVD players (Tracking control loop). *Computational Engineering in Systems Applications multiconference (CESA)*, Lille, France, 2003.

[7] ECMA. Standard ECMA-267: 120 mm DVD-read-only disk. Technical report, ECMA, 2001.

[8] G. Filardi, O. Sename, et al. Robust H_∞ control of a DVD drive under parametric uncertainties. *European Control Conference (ECC)*, Cambridge, U.K., 2003.

[9] I. D. Landau. *Commande des systèmes – conception, identification et mise en œuvre.* Hèrmes Science Publications, Paris, 3rd edition, 2002.

[10] J.-H. Moon, M.-N. Lee, and M. J. Chung. Repetitive control for the track-following servo system of an optical disk drive. *IEEE Transaction on Control Systems Technology*, 6(5):663–670, 1998.

[11] H. Procházka and I.D. Landau. Logiciel pour l'enseignement et le calcul du placement de pôles robuste. *Conférence Internationale Francophone d'Automatique (CIFA)*, pages 694–698, Nantes, France, 2002.

[12] G. Vinnicombe. Frequency domain uncertainty and the graph topology. *IEEE Trans. on Automatic Control*, 38:1371–1383, 1993.

[13] K. Zhou and J.C.Doyle. *Essentials of robust control.* Prentice Hall, New Jersey, 1998.

[14] Y. Zhou, M. Steinbuch, et al. Estimator-based sliding mode control of an optical disc drive under shock and vibration. *Proceedings of the 2002 IEEE International Conference on Control Applications (CCA)*, pages 631–636, Glasgow, Scotland, U.K., 2002.

ON THE STRUCTURAL SYSTEM ANALYSIS IN DISTRIBUTED CONTROL

Corneliu Huțanu and Mihai Postolache
"Gh. Asachi" Technical University
Department of Automatic Control and Industrial Informatics
Blvd. Mangeron 53A, 700050 Iasi, Romania
Fax, Phone: +40 - 32 - 23 07 51
E-mail: chutanu@delta.ac.tuiasi.ro, mpostol@delta.ac.tuiasi.ro

Abstract Large scale systems often consist of several relatively autonomous subsystems sharing common resources, material or energy flows, and informational networks. Distributed control of such systems requires the use of some decomposition, modeling and analysis techniques in order to stabilize the global system and to fulfill further design requirements. In the paper some techniques based on structural investigations trying to infer the properties of the interconnected system (eigenvalues and fixed modes) from the properties of its constituent subsystems are discussed.

Keywords: structural analysis, system decomposition, interconnected systems, decentralized control

1. Centralized VS distributed control systems

Consider the centralized control problem of a given plant P

$$P : U \to Y, \quad y = f(u), \; u \in U, \; y \in Y, \tag{1}$$

where *u, y* are their input and output, respectively. Finding a (single) control unit is required, whose main tasks are *i)* to ensure asymptotic stability of the closed loop system for a given class of command signals and disturbances, and *ii)* to meet dynamical I/O behavior of the system as specified. This approach turns out to be not suitable for many of the modern plants at present. Today's plants are highly integrated systems based on the cooperation between several machines and industrial robots. All these constituents are more or less interconnected by information networks, material and energy flows, or share common resources.

Such large scale systems are characterized by at least one of the following properties: *i)* large size and high dimension of the plant model; *ii)* presence of uncertainties regarding the model and the plant structure; *iii)* restrictions concerning the information access due to the geographical distribution of several component subsystems; *iv)* an interconnection structure that links together the constituent subsystems, either directly (at the subsystems level) or indirectly, by means of a common subsystem or by common resources (at the system level). Each of these properties causes a centralized control unit of the large scale system to be hard to find at reasonable costs.

A distributed control approach of such a large scale system is more appropriate due to the intrinsic isomorphism that can be established between the structure of the controlled system and the structure of the distributed control unit. This way the global control problem splits into several smaller but still interrelated control problems that are easier to solve by smaller control units providing communication capabilities.

This is the so called "off-line" phase of decentralized control. It can be the subject of some decomposition techniques applied to the global system, as well as to the global control problem. Also, redefining of the control problem has often to be made at this level, to take into account new properties such as flexibility, reliability, or robustness of the control unit. Such aims may become more important than a limited, short term optimum behavior, and they have not been addressed primarily by the classical multivariable control theory.

A certain methodology is needed in order to deal with large scale systems: *i)* specification of the system objectives; *ii)* system decomposition in N interconnected subsystems; *iii)* analysis of the isolated subsystems to reveal their qualitative and quantitative properties; *iv)* inferring the overall system properties from those of the isolated subsystems taking into account the interconnection structure.

1.1. Decomposition of large scale systems

Let S be a large scale system defined by

$$\begin{gathered} S = \{U,\ X,\ Y,\ f,\ g,\ t\} \\ U \times X \xrightarrow{f} X, \quad X \xrightarrow{g} Y, \end{gathered} \tag{2}$$

where U, X, and Y are the input, state, and output sets, respectively.

The system S is to be decomposed in N sub-systems S_i, i=$\overline{1,N}$,

$$\begin{gathered} S = \{U_i,\ X_i,\ Y_i,\ f_i,\ g_i,\ t\} \\ U_i \times X_i \xrightarrow{f_i} X_i, \quad X_i \xrightarrow{g_i} Y_i \end{gathered} \tag{3}$$

with $U_i \subset U$, $X_i \subset X$, $Y_i \subset Y$.

There are two ways to have the system S split in subsystems: *i)* horizontally, and *ii)* vertically, respectively. Horizontal decomposition can be made based on structural criteria as well as based on mathematical conversions followed by partitioning and parametric decomposition (Lunze, 1992). When direct coupling links between subsystems are used, the result of the horizontal decomposition phase is an interconnected system as shown in Figure 1. Its interconnected structure can be expressed as

$$s = H(z) = Lz, \tag{4}$$

where L is the interconnection matrix, $L=[l_{ij}]$, $l_{ij}\in\{0,1\}$, $\forall i,j=\overline{1,N}$.

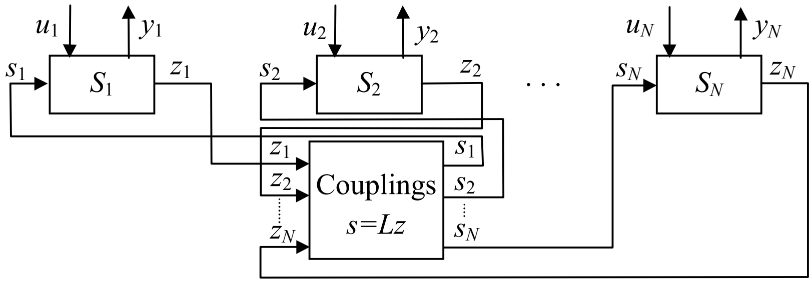


Figure 1. System decomposition into several directly coupled subsystems

Further vertical decomposition refers mainly to the control unit. It leads to a hierarchical structure, usually consisting of up to three layers (Ionescu, 1982): *i)* operational layer; *ii)* planning layer, and *iii)* strategic layer. The operational layer is the one resulting from the horizontal decomposition process.

1.2. Linear models for decentralized systems

As stated before, a global model such as the well known centralized linear state model:

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx + Du, \; x(0) = x_0, \end{cases} \tag{5}$$

with $x\in\mathbf{R}^n$, $u\in\mathbf{R}^m$, $y\in\mathbf{R}^r$, A, B, C, and D constant matrices of suitable sizes, is useless for the large-scale system control even if such a model could be found. Instead of it, smaller size order models of the subsystems together with the interconnection structure model should be used when the analysis and design of the decentralized control unit is in view.

The I/O Oriented Model. It allows only for a decentralized I/O characterization of the large scale system while state information continues to remain centralized. It can be derived from (5) where input and output variables are partitioned into N smaller size components

$$u = [u_1 \quad u_2 \quad \dots \quad u_N]^T \;;\; y = [y_1 \quad y_2 \quad \dots \quad y_N]^T \tag{6}$$

$$\dim u_i = m_i,\ \dim y_i = r_i,\ i = \overline{1, N},$$

as well as the input matrix B and the output matrices C and D

$$\begin{cases} \dot{x} = Ax + \sum_{i=1}^{N} B_{Si} \cdot u_i, \quad x(0) = x_0, \\ y_i = C_{Si} \cdot x + \sum_{j=1}^{N} D_{ij} \cdot u_j, \quad i = \overline{1, N}. \end{cases} \tag{7}$$

The Interaction Oriented Model. Each subsystem is considered a standalone system, having its own state, input, and output variables. It is described by a centralized smaller order linear model

$$\begin{cases} \dot{x}_i = A_i x_i + B_i u_i + E_i s_i \\ y_i = C_i x_i + D_i u_i + F_i s_i \\ z_i = C_{zi} x_i + D_{zi} u_i + F_{zi} s_i, \ i = \overline{1, N}, \ x_i(0) = x_{i0}, \end{cases} \tag{8}$$

where $\dim x_i = n_i$, $\dim u_i = m_i$, $\dim y_i = r_i$, $\dim s_i = m_{si}$, $\dim z_i = r_{zi}$.

Consider the subsystem couplings being fully described by the algebraic equation

$$s = Lz, \tag{9}$$

where $s = [s_1 \ s_2 \ \dots \ s_N]^T$, $z = [z_1 \ z_2 \ \dots \ z_N]^T$ with $\dim s = m_s = \sum_{i=1}^{N} m_{si}$, $\dim z = r_z = \sum_{i=1}^{N} r_{zi}$.

Due to the absence of direct couplings from the local inputs u_i to the local outputs y_i as well as between the interaction inputs s_i and outputs z_i., the decentralized I/O model of the overall system can be written

$$\begin{cases} \dot{x}_i = A_{ii} x_i + \sum_{\substack{j=1 \\ j \neq i}}^{N} A_{ij} x_j + B_i u_i, \quad x_i(0) = x_{i0}, \\ y_i = C_i x_i, \quad i = \overline{1, N}, \end{cases} \tag{10}$$

where

$$A = [A_{ij}]_{N \times N}, \text{ where } A_{ii} = A_i + E_i L_{ii} C_{zi},\ A_{ij} = E_i L_{ij} C_{zj},\ i \neq j$$

$$B = \operatorname{diag} B_i,\ C = \operatorname{diag} C_i,\ D = 0. \tag{11}$$

2. Analysis of decentralized systems

The following three theorems are well known from the multivariable feedback control of centralized systems described by the state space linear model

$$\begin{cases} \dot{x} = Ax + Bu \\ y = Cx, \;\; x(0) = x_0 \, . \end{cases} \tag{12}$$

The solution of the differential equation (12) may be expressed as

$$y(t) = C\sum_{i=1}^{n} v_i w_i e^{\lambda_i t} x_0 + \int_0^t \sum_{i=0}^{n} Cv_i \cdot w_i B_i \cdot e^{\lambda_i (t-\theta)} \cdot u(\theta) d\theta \, , \tag{13}$$

where λ_i are the n different eigenvalues of the system (12), and v_i, w_i are the corresponding right and left eigenvectors, respectively.

Theorem 1 *The input mode* $m_{Bi} = w_i e^{\lambda_i t}$ *and the eigenvalue* λ_i *is said to be controllable if one of the following equivalent conditions is satisfied*

$$w_1 \left[A - \lambda_i I \;\vdots\; B \right] \neq 0, \forall i = \overline{1,n} \tag{14}$$

$$\text{rank} \left[A - \lambda_i I \;\vdots\; B \right] = n, \forall i = \overline{1,n}.$$

Theorem 2 *The output mode* $m_{Ci} = v_i e^{\lambda_i t}$ *and the eigenvalue* λ_i *is said to be observable if one of the following equivalent conditions is satisfied*

$$\begin{bmatrix} A - \lambda_i I \\ \cdots\cdots \\ C \end{bmatrix} v_i \neq 0, \;\; \text{rank} \begin{bmatrix} A - \lambda_i I \\ \cdots\cdots \\ C \end{bmatrix} = n, \forall i = \overline{1,n} \, . \tag{15}$$

Theorem 3 *The set of centralized fixed eigenvalues is identical to the set of the uncontrollable and unobservable eigenvalues of the system (12).*

The centralized fixed modes cannot be changed even by any dynamic output feedback. Negative values of the real part of such fixed eigenvalues make difficult to fulfill some design requirements, while positive values is the worst case: the system is unstable and cannot be stabilized. In the following, the I/O oriented model (7) of the system (12) is considered together with the decentralized static control law

$$u = -K_y y \, , \tag{16}$$

where K_y is the decentralized feedback matrix

$$K_y = \text{diag}(K_{y1}, K_{y2}, \ldots , K_{yN}) \, . \tag{17}$$

2.1. Decentralized fixed modes and eigenvalues

Decentralized fixed modes are defined in a similar manner as the centralized ones, corresponding to those eigenvalues of the closed loop system that remain unchanged for any decentralized feedback applied to the system (Corfmat et al. 1973). First, according to (17) $K_y \in \mathbf{R}^{m \times r}$, so the centralized fixed modes are all decentralized fixed modes too. However, further decentralized fixed modes may occur due to the decentralized control structure, even if the closed loop system is completely controllable and observable.

Theorem 4 (Anderson *et al.*, 1981) *The eigenvalue $\lambda[A]$ is a decentralized fixed one if and only if a disjoint partition of the index set $\mathcal{J}=\{1, 2,..., N\}$ exists, consisting of the sets $\mathcal{D} = \{i_1, i_2,..., i_k\}$ and $\mathcal{H} = \{i_{k+1}, i_{k+2},..., i_N\}$, $\mathcal{D} \cap \mathcal{H} = \emptyset$, $\mathcal{D} \cup \mathcal{H} = \mathcal{J}$, so that the following condition is satisfied*

$$\operatorname{rank}\left[\begin{array}{c|c} A-\lambda I & B_D \\ \hline C_H & 0 \end{array}\right] < n\,, \tag{18}$$

where

$$B_D = \left[B_{Si_1} \; B_{Si_2} \; \dots \; B_{Si_k} \right], \quad B_H = \left[B_{Si_{k+1}} \; B_{Si_{k+2}} \; \dots \; B_{Si_N} \right],$$

$$C_D = \left[C^T_{Si_1} \; C^T_{Si_2} \; \dots \; C^T_{Si_k} \right]^T, \quad C_H = \left[C^T_{Si_{k+1}} \; C^T_{Si_{k+2}} \; \dots \; C^T_{Si_N} \right]^T.$$

It is obvious that the eigenvalue $\lambda[A]$ is not a decentralized fixed one if there exists at least one channel (u_i, y_i), $i \in \mathcal{J}$ so that $\lambda[A]$ is both controllable and observable through it. Decentralized fixed eigenvalues are either uncontrollable or unobservable or both of them at the same time through any of the decentralized I/O control channels.

Theorem 4 provides a necessary and sufficient condition for an eigenvalue to be a decentralized fixed eigenvalue. It assures that if the eigenvalue $\lambda[A]$ is uncontrollable through the I/O channels i, $\forall i \in \mathcal{D}$, then it cannot be made controllable by any decentralized feedback control at the I/O control channels i, $\forall i \in \mathcal{H}$ by which it is observable. Conversely, it assures that if the eigenvalue $\lambda[A]$ is unobservable through the I/O channels i, $\forall i \in \mathcal{H}$, then it cannot be made observable by any feedback control at the I/O control channels i, $\forall i \in \mathcal{D}$ by which it is controllable.

The condition (18) is independent by the decentralized control matrix K_y. It reveals a property the system has to be provided with for a decentralized fixed mode to exist. However, exploiting the condition provided by the Theorem 4 to find the decentralized fixed modes leads to complex mathematical operations with the centralized model of the large-scale system. In the following, equivalent conditions will be given that establish a connection between the decentralized fixed modes and the properties of the subsystems described by the interaction oriented model (10) in a simplified form

$$\begin{cases} \dot{x}_i = A_i x_i + B_i u_i + E_i s_i \\ y_i = C_i x_i, \ z_i = C_{zi} x_i, \quad i = \overline{1,N}, x_i(0) = x_{i0}, \\ s = Lz. \end{cases} \tag{19}$$

If all the subsystems $i \in \mathcal{J}$ are completely controllable and observable through their local I/O channel (u_i, y_i), then the eigenvalues of the isolated subsystems can all be changed by an appropriate decentralized I/O feedback. However, even if all subsystems and the overall system are completely controllable and observable, the overall system may still have decentralized fixed modes, as exemplified in (Lunze, 1992).

Thus, sufficient conditions for a subsystem eigenvalue to be a decentralized fixed eigenvalue of the overall system can be tailored to the interaction-oriented model. To begin with, it is clear that a subsystem eigenvalue $\lambda[A_i]$ is a decentralized fixed eigenvalue of the overall system (19) if it is either uncontrollable through both the inputs u_i and s_i or unobservable through both the outputs y_i and z_i.

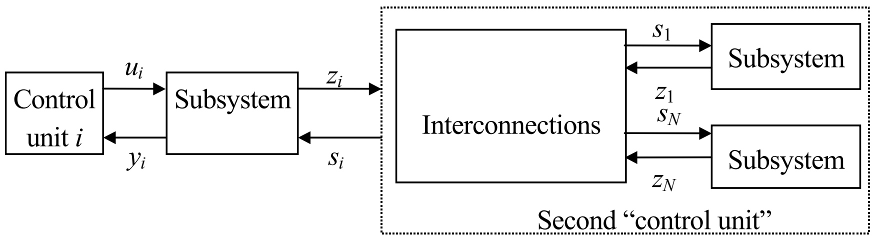


Figure 2. The i-th subsystem and its I/O channels: (u_i, y_i), (s_i, z_i)

For the subsystem i assimilated to a plant with two control units, as shown in Fig.2 the following theorem is given.

Theorem 5 *A subsystem eigenvalue $\lambda[A_i]$ is a decentralized fixed eigenvalue of the overall system (19) if at least one of the following conditions holds*

$$\operatorname{rank}\begin{bmatrix} A_i - \lambda I & E_i \\ C_i & 0 \end{bmatrix} < n_i, \ \operatorname{rank}\begin{bmatrix} A_i - \lambda I & B_i \\ C_{zi} & 0 \end{bmatrix} < n_i. \tag{20}$$

These conditions can be derived directly from Theorem 4 applied to the subsystem i while its control structure is considered as in Fig.2. In other words, $\lambda[A_i]$ is a decentralized fixed eigenvalue of the overall system (19) if it is not and cannot be made either simultaneously controllable through s_i

and observable through y_i, nor simultaneously controllable through u_i and observable through z_i.

Finally, Theorem 4 can also be applied to the subsystem *i* with embedded interconnections, having *N* I/O channels, while the matrices *A*, B_D şi C_H are aggregated from the subsystem matrices A_i, B_i, C_i, E_i, C_{zi}, and L_{ij}, using the relations (11), as described in (Lunze, 1992).

A subsystem eigenvalue that turns out to be a decentralized fixed value does not depend on the interconnections between subsystems. These values remain unmodified even if some subsystems are decoupled from the overall system or other subsystems are later connected.

2.2. Decentralized structural fixed modes and eigenvalues

Large scale systems control often cannot make use of a precise linear model. However, some properties of the system can be determined in spite of the uncertainties regarding the parameters and the structure of the plant. Such structural properties are valid for a relatively large area of numerical values instead of being strongly dependent on some singular parameter values (Sezer, 1981).

The structure of a system *S* (*A, B, C*) may be described through the mean of a structure matrix or graph. All the matrices of the system (1) *A*, *B*, and *C* may be converted as structure matrices S_a= [*A*], S_b= [*B*], and S_c= [*C*], respectively, by using the following notation

$$\left[a_{ij}\right] = \begin{cases} 0, & a_{ij} = 0, \\ *, & a_{ij} \neq 0. \end{cases} \tag{21}$$

The class *S* consists of all systems having the same structure matrices S_a, S_b, and S_c

$$S\,(S_a, S_b, S_c) = \{(A, B, C) \,/\, [A] = S_a, [B] = S_b, [C] = S_c\}\ . \tag{22}$$

The graph representation of class *S* can be derived from the structure matrix *Q* attached to the system

$$\begin{matrix} & X & U & Y & \\ Q = & \begin{bmatrix} S_a & S_b & 0 \\ 0 & 0 & 0 \\ S_c & 0 & 0 \end{bmatrix} & & & \begin{matrix} X \\ U \\ Y \end{matrix} \end{matrix}\ . \tag{23}$$

There are n=dim$x$ state vertices x_i, m=dim u input vertices u_i, and r=dim$y$ output vertices y_i within the structure graph *G* (*Q*). An edge from one vertex v_i to another vertex v_j exists if and only if the corresponding entry q_{ij} = * in matrix *Q* in (23).

Definition 1 For a structure matrix S_a, a set of independent entries is defined as a set of indeterminate '*' entries, no two of which lie on the same row or column.□

Definition 2 The structural rank (*s-rank*) of S_a is defined as being the maximum number of elements contained in a set of independent entries.

A typical relation between numerical and structural investigations can be defined. The structural rank of a structure matrix S_a is equal to the maximum rank of all admissible matrices having the same structure

$$s\text{-rank } S_a = \max_{A \in S_a} \text{rank } A\,.\square \tag{24}$$

Equation (24) specifies that almost all matrices $A \in S_a$ have a numerical rank equal to the structural rank. In other words, rank $A < s$-rank S_a holds only for some exceptional matrices A, whose entries lie on a hyper surface, and, thus, they are relatively infrequent.

Definition 3 A class S of systems is said to be structurally controllable (s-controllable) or structurally observable (s-observable) if there exists at least one admissible realisation $(A, B, C) \in S$ which is completely controllable or completely observable, respectively. As a consequence, the s-controllability and s-observability of a class S are necessary conditions for the numerical controllability and observability, respectively, for almost all systems from class S.□

Definition 4 A class of systems S is said to be input connectable if for every state vertex v in the graph G (Q) there is a path from at least one of the input vertices to v. It is said to be output-connectable if for every state vertex v a path to at least one output vertex exists.□

Input-connectivity and output-connectivity guarantees that $s\text{-rank}\left[A - \lambda I \;\vdots\; B\right] = n$ and $s\text{-rank}\begin{bmatrix} A - \lambda I \\ \cdots \\ C \end{bmatrix} = n$, for $\lambda \neq 0$ and almost all admissible systems. Taking these into account, the structural counterpart of Theorems 1, 2 is

Theorem 6 *A class S of systems is s-controllable if and only if it is input-connectable, and* s-rank $[S_a\ S_b] = n$. *The class S is s-observable if and only if it is output-connectable, and* s-rank $[S_a^T\ S_c^T]^T = n$.

Definition 5 A class S of systems has structurally fixed modes if all the admissible systems of that class have fixed modes.□

As a result, a given system (A, B, C) has fixed modes if its container class $S([A],[B],[C])$ has structural fixed modes. On the other hand, a class S has no structural fixed modes if at least one system completely controllable and observable $(A, B, C) \in S$ exists, i.e. if the class S is s-controllable and

s-observable. Subsequently, the presence of structural fixed modes can be checked using the s-controllability and s-observability conditions established by Theorem 6.

Theorem 7 *A class S of systems has structurally fixed modes if and only if at least one of the conditions of Theorem 6 is not satisfied. Structural fixed modes of type I occur due to the missing input or output connectivity, otherwise structural fixed modes, if exists, are said to be of type II.*

Although the conditions of Theorem 6 refer to the open loop system, a good graphical representation for them can be made based on the closed loop system which takes into account the static output feedback $u = -K_y y$. The structure matrix Q_0 of the closed loop system is

$$Q_0 = \begin{bmatrix} S_a & S_b & 0 \\ 0 & 0 & E \\ S_c & 0 & 0 \end{bmatrix}, \quad E = \begin{bmatrix} * & \cdots & * \\ \vdots & \ddots & \vdots \\ * & \cdots & * \end{bmatrix}. \tag{25}$$

Compared to $G\,(Q)$, the closed loop graph $G\,(Q_0)$ has supplemental edges from all the output nodes to all the input ones, expressing the information flow across the feedback links. As any system mode can be changed only if it is placed within a closed loop in the graph $G\,(Q_0)$, structural fixed modes exist only if there are vertices in $G\,(Q_0)$ that cannot be embedded in such closed loops. As any input can be linked to any output using a proper static feedback, state vertices are part of no closed loop only if they are disconnected from all the inputs or all the outputs.

The second condition may be examined starting from the following considerations. The condition s-rank $[S_a\ S_b] = n$ may be equally expressed as the requirement that every state vertex must have at least an edge from different state or input vertices (or any line of $[S_a\ S_b]$ must have at least one indeterminate entry placed on different columns). Similar meanings can be given to the condition s-rank $[S_a^T\ S_c^T]^T = n$: every state vertex must have at least an edge to different state or output vertices (or any column of $[S_a\ S_b]$ must have at least one indeterminate entry placed on different lines). These two conditions are simultaneously satisfied if there are disjoint closed loops or cycles in $G\,(Q_0)$ with all the state vertices embedded within at least one of them. All these closed loops yields to a cycle family, whose dimension is given by the number of state vertices included.

The necessary and sufficient condition given by Theorem 7 can be expressed in the equivalent form of

Theorem 8 *A class S of nth-order systems has structurally fixed modes if and only if at least one of the conditions is satisfied for graph G (Q_0):*

I. S is neither input-connectable nor output-connectable.

II. There does not exist a cycle family of width n.

Structural fixed modes as they were defined for the centralized model (5) have been generalized for interconnected systems described by the decentralized I/O model (10). Thus, the graph for the decentralized closed loop system may be derived from the following structure matrix

$$
Q_d = \left[\begin{array}{ccc|ccc|ccc}
[[A_{11}]] & \cdots & [[A_{1N}]] & [[[B_1]]] & \cdots & 0 & & & \\
\vdots & \ddots & \vdots & \vdots & \ddots & \vdots & & 0 & \\
[[A_{N1}]] & \cdots & [[A_{NN}]] & 0 & \cdots & [[B_N]] & & & \\
\hline
 & & & & & & * & \cdots & 0 \\
 & 0 & & & 0 & & \vdots & \ddots & \vdots \\
 & & & & & & 0 & \cdots & * \\
\hline
[[C_1]] & \cdots & 0 & & & & & & \\
\vdots & \ddots & \vdots & & 0 & & & 0 & \\
0 & \cdots & [[C_N]] & & & & & &
\end{array}\right]. \tag{26}
$$

Applying Theorem 8 for the decentralized feedback system matrix Q_d one can claim that for the class S_d of N interconnected subsystems there are structurally fixed modes if at least one of the following conditions is satisfied for graph $G\,(Q_d)$:

I. There exists a subsystem vertex, which is not connectable to any channel (u_i, y_i).
II. There does not exist a cycle family of width N.

Another way to establish if an interconnected system has structural fixed modes can be derived by investigating each subsystem together with its couplings to and from the other subsystems. Thus, while the subsystem i is considered as shown in Fig.2, the attached graph $G\,(Q_{di})$

$$
Q_{di} = \left[\begin{array}{c|cc|cc}
[A_i] & [B_i] & [E_i] & 0 & 0 \\
\hline
0 & 0 & 0 & * & 0 \\
0 & 0 & 0 & 0 & * \\
\hline
[C_i] & 0 & 0 & 0 & 0 \\
[C_{zi}] & 0 & 0 & 0 & 0
\end{array}\right] \tag{27}
$$

may be used in order to check if an eigenvalue of the subsystem is a structurally fixed eigenvalue for the overall system.

Given the class S_d of systems whose ith subsystem structure is described by Q_{di}, the class S_d has structurally fixed modes if there exists an index i so that at least one of the following conditions are satisfied for the graph $G\,(Q_{di})$:

I. There exists a state vertex, which is not connectable to any channel (u_i, y_i).
II. There does not exist a cycle family of width n_i.

3. Conclusions

Structural investigations concerning the existence of decentralized fixed modes are important due to their impact on the stabilizability of the plant. A decentralized stabilizability theorem may be given which is similar to Theorem 3. The eingenvalues of the plant and their associated modes can be all modified using a decentralized static feedback $u= -K_y y$ if and only if there are no decentralized fixed modes.

Making all the decentralized eigenvalues to be controllable and observable trough the same channel leads to a centralized design procedure of the decentralized control problem. Other solutions presented in the literature (Lunze, 1992) are decentralized dynamical compensation, and the replacement of a centralized state feedback by a decentralized output feedback without changing the eigenvalues established at the previous stage.

Having the control units interconnected via an industrial control network is a different approach to solve the overall control problem in a distributed way. This approach may overcome many of the constraints imposed by the decentralized structure of the controller (Wang *et al.*, 1978). Structural analysis of the large scale system still offers useful information as long as it is focused on the centralized fixed modes of the system.

References

Anderson, B.D.O. and D.J. Clements (1981). Algebraic characterization of fixed modes in decentralized control, Automatica, 17, 703-71.

Corfmat, J.P. and A.S. Morse (1973). Stabilization with decentralized feedback control, IEEE Trans. Autom. Control , AC-18, 679-681.

Lunze, I. (1992). Feedback Control of Large Scale Systems. Prentice Hall, Dresden.

Ionescu, T. (1982). Sisteme şi echipamente pentru conducerea proceselor. Editura Didactică şi Pedagogică, Bucureşti.

Sezer, M.E. and D.D. Šiljac (1981). Structurally fixed modes, Syst. Control Letters, 1, 60-64.

Wang, S.H. and E.J. Davison (1978). Minimization of transmission cost in decentralized control systems, Int. J. Control, 28, 889-896.

ON THE DYNAMICAL CONTROL OF HYPER REDUNDANT MANIPULATORS

Mircea Ivanescu
Automatic and Computer Department
University of Craiova
A. I. Cuza Str. No. 13, Craiova 1100, Romania
E-mail:ivanescu@robotics.ucv.ro

Abstract The control problem of the spatial tentacle manipulator is presented. The difficulties determined by the complexity of the nonlinear integral - differential equations are avoided by using a very basic energy relationship of this system. Energy-based control laws are introduced by using only the relations that determine the energy stored in the system. A PD controller and a fuzzy controller are discussed. Numerical simulations for spatial and planar tentacle models are presented in order to prove the efficiency of the method.

Keywords: tentacle, distributed parameter system, fuzzy controller

1. Introduction

An ideal tentacle manipulator is a non-conventional robotic arm with an infinite mobility. It has the capability to take sophisticated shapes and to achieve any position and orientation in a 3D space. These systems are also known as Hyper-Redundant Manipulators or Hyper-Degree-Of-Freedom (HDOF) Manipulators and, over the past several years, there has been a rapidly expanding interest in the study and construction of them.

The control of these systems is very complicated and a great number of researchers tried to offer solutions for this difficult problem. Hemami (1984) analyzed the control by cables or tendons meant to transmit forces to the elements of the arm in order to closely approximate the arm as a truly continuous backbone. Also, Mochiyama *et al.* have investigated the problem of controlling the shape of an HDOF rigid-link robot with two-degree-of-freedom joints using spatial curves (Mochiyama and Kobayashi, 1999;

Mochiyama *et al.*, 1998). Important results were obtained by Chirikjian and Burdick (1990, 1992, 1995) who laid the foundations for the kinematical theory of hyper-redundant robots. Their results are based on a “backbone curve” that captures the robot’s macroscopic geometric features. The inverse kinematical problem is reduced to determining the time varying backbone curve behavior. New methods for determining “optimal” hyper-redundant manipulator configurations based on a continuous formulation of kinematics are developed. Gravagne (2000) analyzed the kinematical model of “hyper-redundant” robots, known as “continuum” robots. Robinson and Davies (1999) present the “state of art” of continuum robots, outline their areas of application and introduce some control issues.

In other papers (Suzumori *et al.*, 1991; Cieslak and Moreki, 1994; Shigoma, 1996) several technological solutions for actuators used in hyper-redundant structures are presented and conventional control systems are introduced.

All these papers treat the control problem from the kinematical point of view and few researchers focus their efforts on the dynamic problem of these systems. The dynamic models of these manipulators are very complicated. Chirikjian (1993b) proposed a dynamic model for hyper-redundant structures as an infinite degree-of-freedom continuum model and some computed torque control systems are introduced. Ivanescu (1984) presented a dynamic model for an ideal planar tentacle system and discussed optimal control solutions. Ivanescu and Stoian (1995) proposed a sequential distributed control for a tentacle manipulator actuated by electrorheological fluids.

The difficulty of the dynamic control is determined by integral-partial-differential models with high nonlinearities that characterize the dynamic of these systems. In Appendix 1 of this paper the dynamic model of an ideal spatial tentacle manipulator is presented and the difficulties to obtain a control law are very clear.

In this paper we treat the control problem by using a very basic energy relationship of these models and avoid the difficulties determined by the complexity of the dynamic model. The energy-based controller (Ge, *et al.*, 1996; Wang, *et al.*, 2001) determines the control law by using only the relations that determine the energy stored in the system. By this method, a class of controllers that can assure the motion of the manipulator to a desired position with good performances is proposed. The method is verified for an ideal spatial tentacle manipulator and the control laws are numerically simulated. The paper is organized as follows: section 2 reviews the basic principles of a tentacle manipulator; section 3 presents the general relationship of the energy for these systems; section 4 introduces the control law; section 5 verifies by computer simulations the control laws for a 2D and 3D tentacle manipulator. In Appendix 1 the dynamic model of a 3D manipulator is inferred and in Appendix 2 and 3 the control laws are demonstrated.

2. Background

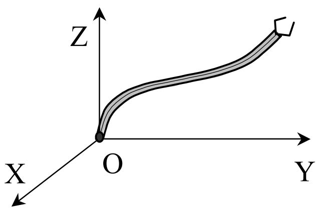
We will consider an ideal tentacle arm, with a uniformly distributed mass and torque, with ideal flexibility that can take any arbitrary shape (Figure 1). Technologically, we will analyze a backbone structure with peripheral cells that can determine the shape of the arm by an appropriate control. We will neglect friction and structural damping.

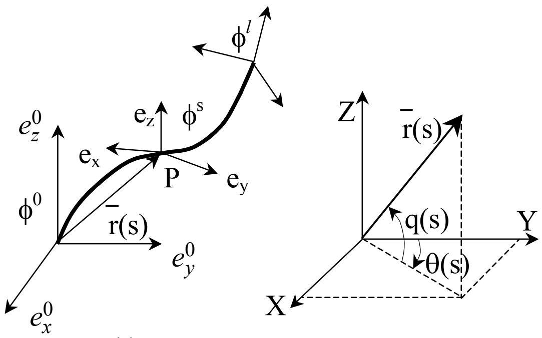
The essence of the tentacle model is a 3-dimensional backbone curve C that is parametrically described by a vector r(s) $\in \mathbf{R}^3$ and an associated frame Φ(s) $\in \mathbf{R}^{3\times 3}$ whose columns create the frame bases (Figure 2a). The independent parameter s is related to the arc-length from the origin of the curve. We denote by *l* the total length of the arm on curve C.

The position of a point s on curve C is defined by the position vector,

$$\bar{r} = \bar{r}(s), \tag{1}$$

where s $\in$ [0, *l*]. For a dynamic motion, the time variable will be introduced, $\bar{r} = \bar{r}(s,t)$.





(a) (b)

Figure 1. Tentacle model *Figure 2.* a) Backbone structure; b) Backbone parameters

We used a parameterization of the curve C based upon two "continuous angles" θ(s) and q(s) (Chirikjian and Burdick, 1990, 1992, 1995), (Figure 2b). At each point $\bar{r}(s,t)$, the robot's orientation is given by a right-handed orthonormal basis vector $\{\bar{e}_x, \bar{e}_y, \bar{e}_z\}$ and its origin coincides with point $\bar{r}(s,t)$. The set of backbone frames can be parameterized as

$$\Phi^s(t) = \left(\bar{e}_x(s,t) \quad \bar{e}_y(s,t) \quad \bar{e}_z(s,t)\right) \tag{2}$$

$$\Phi^s(t) = \begin{bmatrix} c_\theta(s,t) & s_\theta(s,t)c_q(s,t) & -s_\theta(s,t)s_q(s,t) \\ -s_\theta(s,t) & c_\theta(s,t)c_q(s,t) & -c_\theta(s,t)s_q(s,t) \\ 0 & s_q(s,t) & c_q(s,t) \end{bmatrix} \tag{3}$$

with $c_\theta = \cos\theta,\ s_\theta = \sin\theta,$ etc.

For a small variation ds along curve C,

$$\bar{r}(s,t)+d\bar{r}(s,t)=\bar{r}(s+ds,t) \tag{4}$$

the new frame is given by

$$\Phi^{s+ds}(t)=\left(\bar{e}_x(s+ds,t)\ \ \bar{e}_y(s+ds,t)\ \ \bar{e}_z(s+ds,t)\right) \tag{5}$$

The position vector on curve C is given by

$$\bar{r}(s,t)=[x(s,t)\ \ y(s,t)\ \ z(s,t)]^T, \tag{6}$$

where

$$x(s,t) = \int_0^s \sin\theta(s',t)\cos q(s',t)ds' \tag{7}$$

$$y(s,t) = \int_0^s \cos\theta(s',t)\cos q(s',t)ds' \tag{8}$$

$$z(s,t) = \int_0^s \sin q(s',t)ds' \tag{9}$$

with $s' \in [0,s]$. We can adopt the following interpretation (Chirikjian and Burdick, 1990, 1992, 1995; Gravagne and Walker, 2000): at any point s the relations (6)-(9) determine the current position and the matrix Φ^s contains the robot's orientation, and the robot's shape is defined by the behavior of functions $\theta(s)$ and $q(s)$. The robot "grows" from the origin by integrating to get $\bar{r}(s,t)$.

3. Energy - work relationship

The method developed in this paper is based on the energy-work relationship of the tentacle manipulator. To simplify, we will consider the (OYZ) planar model of an ideal tentacle arm without friction and structural damping. For this model, the main parameter is the angle between the tangent to the curve and axis Y, at time t (Figure 3),

$$q = q(s,t).$$

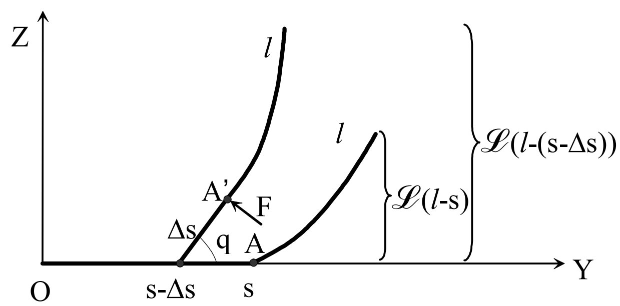


Figure 3. Energy-work relationship

We consider as initial position the horizontal position when total energy, kinetic and potential, is zero.

We will assume that the mechanical work required to move the $(l - s)$ – length arm from the horizontal position (initial position) to the motion position is $\mathscr{L}(l - s)$. If an element Δs is moved by a torque M to a new position defined by the angle q, at a time t, the mechanical work will be

$$\mathscr{L}(l-(s-\Delta s)) = \mathscr{L}(l-s)+\int_0^t M(s,\tau)\dot{q}(s,\tau)d\tau, \tag{10}$$

where

$$\dot{q}(s,t) = \frac{\partial q}{\partial t}(s,t)$$

but

$$M(s,t) = F(s,t)\Delta s \tag{11}$$

and (10) becomes

$$\mathscr{L}(l-(s-\Delta s)) = \mathscr{L}(l-s)+\int_0^t F(s,\tau)\dot{q}(s,\tau)d\tau\Delta s. \tag{12}$$

We can define the derivative of the mechanical work as

$$\frac{d\mathscr{L}(s)}{ds} = \lim_{\Delta s\to 0}\frac{\mathscr{L}((l-s)+\Delta s)-\mathscr{L}(l-s)}{\Delta s} \tag{13}$$

or

$$\frac{d\mathscr{L}(s)}{ds} = \int_0^t F(s,\tau)\dot{q}(s,\tau)d\tau \quad . \tag{14}$$

By integration, the mechanical work will be

$$\mathscr{L}(s) = \int_s^l\int_0^t F(s',\tau)\dot{q}(s',\tau)d\tau ds',. \tag{15}$$

where $s' \in [s,l]$.

For all the arm, l – length, it results

$$\mathscr{L} = \int_0^l\int_0^t F(s',\tau)\dot{q}(s',\tau)d\tau ds', \quad s' \in [0,l] \tag{16}$$

We can extend this result for the 3-dimensional model which means the motion controlled by two angles θ and q,

$$\mathscr{L} = \int_0^l\int_0^t \left(F_\theta(s,\tau)\dot{\theta}(s,\tau)+F_q(s,\tau)\dot{q}(s,\tau)\right)d\tau ds, \tag{17}$$

where $F_\theta(s,t)$, $F_q(s,t)$ represent the distributed forces on the length of the arm that determine motion and orientation in the 3-dimensional space.

Thus, from the energy-work relationship, we have the following equation

$$\left[W_K(t)+W_P(t)\right]-\left[W_K(0)+W_P(0)\right]=\int_0^1\int_0^t\left(F_\theta(s,\tau)\dot{\theta}(s,\tau)+F_q(s,\tau)\dot{q}(s,\tau)\right)d\tau ds\,,\quad(18)$$

where $W_K(t)$ and $W_K(0)$, $W_P(t)$ and $W_P(0)$ are the total kinetic energy and total potential energy of the system at time t and 0, respectively.

From (18), we have

$$\dot{W}_K(t)+\dot{W}_P(t)=\int_0^1\left(F_\theta(s,\tau)\dot{\theta}(s,\tau)+F_q(s,\tau)\dot{q}(s,\tau)\right)ds\,. \quad (19)$$

4. Control laws

The classical methods are often impossible to apply to this manipulator with hyper-redundant configurations. The great number of parameters, theoretically an infinite number of parameters, the complexity of the dynamical model make the application of the classical algorithms to obtain the control law very difficult. For example, in Appendix 1 we determined the dynamical model of a 3D spatial tentacle manipulator,

$$\begin{aligned}&\int_0^s\int_0^s G_q\left(\ddot{q},\ddot{\theta},\dot{q},\dot{\theta},q,\theta\right)ds'ds''=F_q\\&\int_0^s\int_0^s G_\theta\left(\ddot{q},\ddot{\theta},\dot{q},\dot{\theta},q,\theta\right)ds'ds''=F_\theta\,,\end{aligned}\quad(20)$$

where G_q, G_θ are nonlinear functions of the motion parameters (the exact forms and notations are presented in Appendix 1) and F_q, F_θ are distributed forces along the arm in the q-plane and θ-plane, respectively.

The dynamical model of this system is determined as a nonlinear integral differential equation and the difficulty of finding a control law is well-known. Ivanescu (1984) determined an optimal control for minimum energy criterion, Chirikjian and Burdick (1990, 1992, 1995) use the approximation methods and Hemami (1984), Gravagne and Walker (2000), Mochiyama and Kobayashi (1999), Mochiyama *et al.* (1998), Robinson and Davies (1999), Suzumori *et al.* (1991), Cieslak and Moreki (1994) analyze the kinematical position control. In all these papers the simplified procedures are treated or the difficult components are neglected in order to generate a particular law for position or motion control.

In contrast to these traditional methods, we will develop the dynamic control law from the basic energy-work relationship and that can generate the closed-loop stability of the system (Ge *et al.,* 1996; Wang *et al.*, 2001). This method avoids the complex problems derived by a nonlinear derivative integral model and offers an easy solution to implement an adequate controller.

The position control of the tentacle manipulator means the motion control to a desired steady position of the arm defined by a curve,

$$C_d : (\theta_d(s), q_d(s)), \quad s \in [0, l].$$

We define the motion errors as

$$e_\theta(t,s) = \theta(t,s) - \theta_d(s), \quad s \in [0,l],$$

$$e_q(t,s) = q(t,s) - q_d(s), \quad s \in [0,l].$$

Theorem 1 (PD uniform distributed control) *The closed-loop tentacle manipulator arm system is stable if the control law is given by*

$$F_\theta(s,t) = -k_\theta^1(s) e_\theta(s,t) - k_\theta^2(s) \dot{e}_\theta(s,t) \tag{21}$$

$$F_q(s,t) = -k_q^1(s) e_q(s,t) - k_q^2(s) \dot{e}_q(s,t), \quad s \in [0,l], \tag{22}$$

where $k_\theta^1(s), k_\theta^2(s), k_q^1(s), k_q^2(s)$ *are positive coefficients of the control law.*

Proof.. See Appendix 2.■

Theorem 2 (spatial weighted distributed control) *The closed-loop tentacle manipulator arm system is stable if the control law is*

$$F_\theta(s,t) = -k_\theta^1(s) e_\theta(s,t) - k_\theta^2(s) \dot{e}_\theta(s,t) - k_\theta^3(s) f_\theta(s,t) \int_0^t f_\theta(s,\tau) \dot{e}_\theta(s,\tau) d\tau \tag{23}$$

$$F_q(s,t) = -k_q^1(s) e_q(s,t) - k_q^2(s) \dot{e}_q(s,t) - k_q^3(s) f_q(s,t) \int_0^t f_q(s,\tau) \dot{e}_q(s,\tau) d\tau \tag{24}$$

where $k_\theta^1(s), k_\theta^2(s), k_\theta^3(s), k_q^1(s), k_q^2(s), k_q^3(s),$ *are positive coefficients distributed along the arm and* $f_\theta(s,t)$ *and* $f_q(s,t)$ *represent the spatial weighted functions.*

Proof. See Appendix 3.■

The control system proposed by Theorems 1 and 2 is presented in Figure 4.

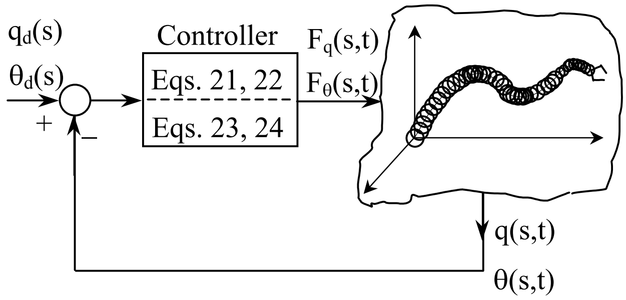


Figure 4. Control system

Equations (21), (22) and (23), (24) define a generalized PD controller with good performances of the position control for an ideal tentacle model without friction and internal damping and for a control criterion defined by a steady desired position.

5. Discussion

a. The stability proofs are independent of the system dynamics and thus the problems associated with model-based controllers (Ivanescu, 1984; Ivanescu and Stoian 1995; Ge *et al.,* 1996) are avoided. Also, the controllers (21), (22) and (23), (24), respectively, are independent of the system parameters and thus possess stability robustness to system parameter uncertainties.

b. The infinite dimensionality of the system determines difficulties in the selection of the control parameters $k_\theta^1, k_\theta^2, k_q^1, k_q^2$. Of course, the closed-loop system stability requires only as $k_\theta^1 > 0,\ k_\theta^2 > 0,\ k_q^1 > 0,\ k_q^2 > 0$ but a practical experiment or simulation imposes a procedure in order to reach an adequate performance. Certainly, this parameter selection had to be associated with dynamic model of the system. Difficulty of the problem determines methods, rather heuristic, to evaluate the control coefficients.

c. We will suggest an approximate method for evaluating the control parameters. We assume that:

A1. The arm motion is a "small" motion that verifies the condition:

$$\left|q(s',t) - q(s'',t)\right| < \delta, \quad s', s'' \in [0,s], \ t \in [0, t_f], \tag{25}$$

where δ is a positive constant, sufficiently small.

A2. A sequential spatial control is assumed, the elements of the arm achieve the desired position step by step: the first element achieves the desired position, then the second, and so on.

The control system (21), (22), by the conditions A1, A2, can be approximated in the error space by the equation

$$\ddot{e}_i + \frac{k_i^2}{\rho\Delta^2}\dot{e}_i + \frac{k_i^1}{\rho\Delta^2}e_i + \frac{g}{\Delta}h_i e_i = 0, \quad i = 1,2,\ldots,N, \tag{26}$$

where $h_i = h_i\left(q_{d_i}\right)$ represents the nonlinear term determined in the error space by the gravitational component and k_i^1, k_i^2 corresponds to $k_{q_i}^1, k_{q_i}^2$ or $k_{\theta_i}^1, k_{\theta_i}^2$, in the q-plane or θ-plane, respectively. (The procedure is presented in the section 6).

Equation (26) can be rewritten in the classic terms of the damping ratio ζ_i and the natural frequency ω_{n_i},

$$\ddot{e}_i + 2\zeta_i \omega_{n_i} \dot{e}_i + \omega_{n_i}^2 e_i + h_i^* e_i = 0, \tag{27}$$

where

$$\omega_{n_i} = \left(\frac{k_i^1}{\rho \Delta^2} \right)^{\frac{1}{2}} \tag{28}$$

$$\zeta_i = \frac{k_i^2}{2\left(\rho \Delta^2 k_i^1\right)^{\frac{1}{2}}} \tag{29}$$

$$h_i^* = \frac{g}{\Delta} h_i . \tag{30}$$

For this model, we suggest a method based by the sliding mode control in which the trajectory is forced along the switching line, directly to the origin, by the control of damping ratio ζ_i (Figure 5).

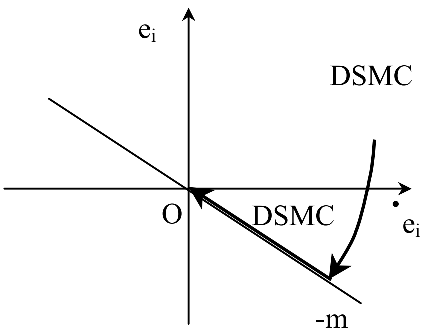


Figure 5. Direct Sliding Mode Control

This special control is named DSMC (Direct Sliding Mode Control) and was introduced for linear systems. Ivanescu and Stoian (1995) presented an extension for nonlinear systems.

The DSMC control can be obtained if the damping ratio ζ_i verifies the conditions (see the following Section):

$$\zeta_i^2 > 1 + \frac{h_i^*}{\omega_{n_i}^2} \tag{31}$$

$$h_i^* > -\omega_{n_i}^2 . \tag{32}$$

The increasing of ζ_i determines an over damped motion but we appreciate that this control ensures a good robustness of the global system.

d. This parameter selection is based by the approximate system but it can be used in order to establish the main domains of the control coefficients.

Simulation examples presented in the following Section will confirm this procedure.

e. The functions $f_\theta(s,t), f_q(s,t)$ can be introduced in order to achieve good performances when the desired trajectory has a "non-smooth form", with corners and group forms. We can use as f_θ, f_q the functions of the distance between the real position and the corner points (or terminal points) of the desired trajectory.

6. Simulations

In this section, some numerical simulations are carried out as 3D and 2D tentacle manipulators.

Test 1. We consider a spatial tentacle manipulator that operates in OXYZ space. The mechanical parameters are: linear density ρ=2.2 kg/m and the length of the arm $l = 0.6$ m.

The initial position of the arm is assumed to be horizontal (OY-axis),

$$\theta(s,0) = 0; \quad q(s,0) = 0; \quad s \in [0,\ 0.6], \tag{33}$$

and the desired position is represented by a curve C1 in OXYZ frame that is defined in terms of motion parameters as

$$\text{C1:}\ \theta_d(s) = \frac{\pi}{12}s; \quad q_d(s) = \frac{\pi}{10}s\,. \tag{34}$$

The control law is chosen as (21), (22) where the proportional and derivative coefficients are selected as

$$\begin{aligned} k_\theta^1(s) = k_q^1(s) = 12.5 \\ k_\theta^2(s) = k_q^2(s) = 1.58\,. \end{aligned} \tag{35}$$

(The selection of coefficients will be explained in the following Test).

To solve the integral-differential system (A.1.9), (A.1.10) with the control law (21), (22), (27) we used a discretization of the s-space, with an increment Δ = 0.1 m,

$$s_i = i \cdot \Delta\,,\ i = 1,2,\ \ldots,\ 6,$$

and a MATLAB system is used for simulation.

The error for the global system is defined as

$$e(t) = \int_0^l \left(\left(q(s,t) - q_d(s)\right)^2 + \left(\theta(s,t) - \theta_d(s)\right)^2 \right) ds \tag{36}$$

$$\dot{e}(t) = \int_0^l \left(\frac{\partial q}{\partial t}(s,t) + \frac{\partial \theta}{\partial t}(s,t) \right) ds\,. \tag{37}$$

The result of simulation is presented in Figure 6. We selected only most significant five intermediary positions of the motion and the phase portrait has the form presented in Figure 7. We see the stability of motion and error convergence to zero.

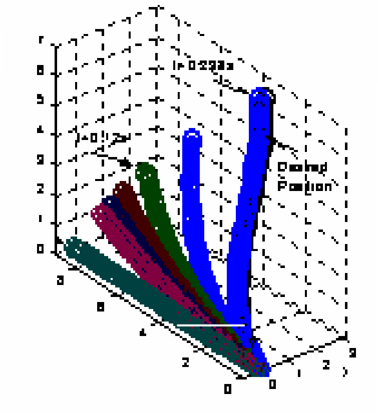


Figure 6. 3D motion simulation

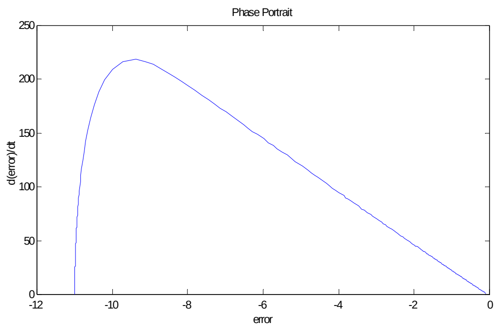


Figure 7. 3D phase portrait

Test 2. A better understanding of the control can be obtained for 2D tentacle arm. We analyze now the case of a planar tentacle model in OXZ plane.

The dynamic model is obtained from the equation (A1.9), (A1.10) for $\theta = 0$,

$$\rho\int_0^s\int_0^s\left[\ddot{q}'\cos(q'-q'')+\dot{q}'^2\sin(q''-q')-\dot{q}'\dot{q}''\sin(q''-q')\right]ds'ds''+ \\ +\rho g\int_0^s \cos q'ds' = F_q\,. \tag{38}$$

The control law is reduced to the form

$$F_q(s,t) = -k_q^1\left(q(s,t)-q_d(s)\right)-k_q^2\dot{q}(s,t). \tag{39}$$

A spatial discretization $s = i\Delta$, i = 1, 2, …, 6, is introduced.

The system (38), (39) can be rewritten in the error space by using the constraint A1 (Section 5) as

$$\rho\Delta^2\sum_{i=1}^m \ddot{e}_i + k_{q_i}^2\dot{e}_i + k_{q_i}^1\dot{e}_i + \rho g\sum_{i=1}^m h_i(q_d)e_i = 0\,, \tag{40}$$

where

$$e_i(t) = e(i\Delta,t) = q(i\Delta,t)-q_d(i\Delta) \tag{41}$$

$$h_i(q_d) = \left(\frac{\partial H}{\partial q}\right)_{\substack{q=q_d \\ s=i_\Delta}} \tag{42}$$

and $H = H(q)$ is determined by the gravitational component

$$H(q) = \cos q \,. \tag{43}$$

A sequential spatial control strategy (constraint A2 – Section 5) determines a decomposition of the global motion. For the first element, m = 1,

$$\rho\Delta^2 \ddot{e}_1 + k_{q_1}^2 \dot{e}_1 + k_{q_1}^1 e_1 + \rho g h_1 e_1 = 0 \,, \tag{44}$$

the control law determines

$$\lim_{t\to\infty} e_1(t) = 0 \,. \tag{45}$$

Then, for the second element, m = 2,

$$\rho\Delta^2 \ddot{e}_2 + k_{q_2}^2 \dot{e}_2 + k_{q_2}^1 e_2 + \rho g h_2 e_2 = 0 \,, \tag{46}$$

we obtain

$$\lim_{t\to\infty} e_2(t) = 0 \,. \tag{47}$$

We repeat the procedure for each element, m = 3, 4, … .
In this case, we can use for each element the dynamic model

$$\ddot{e}_i + 2\zeta_i \omega_{n_i} \dot{e}_i + \omega_{n_i}^2 e_i - \frac{g}{\Delta} \sin q_{d_i} \cdot e_i = 0, \;\; i = 1, 2, \ldots, 6. \tag{48}$$

The DSMC control (Figure 5) imposes the condition

$$\frac{\dot{e}_i}{e_i} = -m \,. \tag{49}$$

The condition for convergence of the motion to zero, on the switching line (49), can readily be found as

$$\zeta_i^2 > 1 - \frac{g \cdot \sin q_{d_i}}{\Delta \omega_{n_i}^2} \tag{50}$$

$$\frac{g}{\Delta} \sin q_{d_i} < \omega_{n_i}^2 \,. \tag{51}$$

For the simulation test we choose the initial position of the arm as the vertical line (OZ-axis)

$$D2: q(s,0) = \frac{\pi}{2}, \quad s \in [0, 0.6], \tag{52}$$

and the desired position is a semicircle, that is approximated by

$$C2: q_d(s_i) = \frac{\pi}{3} - s_i \frac{\pi}{6}, \quad i = 1,2,...,6. \tag{53}$$

We choose an uniform natural frequency $\omega_{n_i} = 24$, i=1, 2, …, 6, and for verifying the conditions (50), (51), (53) we select

$$\zeta_i = 1.5, \quad i=1, 2, …, 6. \tag{54}$$

From (28), (29) we obtain the control coefficients

$$k_i^1 = 12.5, \quad k_i^2 = 1.58, \quad i=1,2,…,6. \tag{55}$$

The results of the simulations are presented in Figure 8 and the phase portrait is plotted in Figure 9.

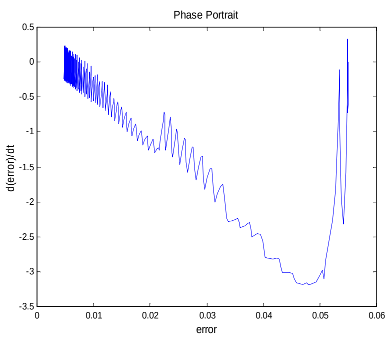


Figure 8. 2D phase portrait

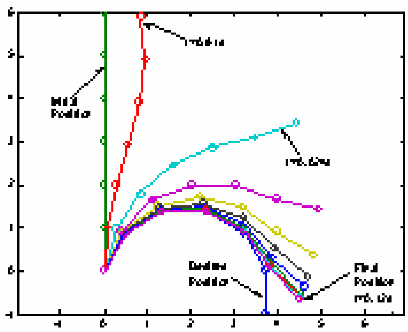


Figure 9. 2D motion simulation

7. Conclusions

The paper treats the control problem of a tentacle manipulator. In order to avoid the difficulties generated by the complexity of the nonlinear integral-differential equations that define the dynamic model of this system, the control problem is avoided by using a very basic energy relationship of this dynamic model.

The energy relationships of a tentacle manipulator are inferred. An energy-based control law is introduced by using only the relations that determine the energy stored in the system.

Two controllers are proposed that generate the PD control laws and a procedure of control coefficient selection is discussed.

In Appendix 1 is obtained the dynamic model of a spatial tentacle arm that allows for the checking of the control laws.

Numerical simulations for spatial and planar tentacle models illustrate the efficiency of the method.

Appendix 1

We will consider a spatial tentacle model, an ideal system, neglecting friction and structural damping. We assume a uniformly distributed mass with a linear density ρ [kg/m]. We will consider a non-extensible arm with a constant length *l* (Figure 2a, 2b).

We will use the notations:

$$q = q(s,t), \quad s \in [0,l], \quad t \in [0,t_f],$$

$$\theta = \theta(s,t), \quad s \in [0,l], \quad t \in [0,t_f],$$

$$q' = q(s',t), \quad s' \in [0,s], \quad t \in [0,t_f],$$

$$\dot{q} = \frac{\partial q(s,t)}{\partial t}, \quad s \in [0,l], \quad t \in [0,t_f],$$

$$\dot{q}' = \frac{\partial q(s',t)}{\partial t}, \quad s' \in [0,s], \quad t \in [0,t_f],$$

$$\ddot{q}' = \frac{\partial^2 q(s',t)}{\partial t^2}, \quad s' \in [0,s], \quad t \in [0,t_f],$$

$$\ddot{q}'' = \frac{\partial^2 q(s'',t)}{\partial t^2}, \quad s'' \in [0,s], \quad t \in [0,t_f],$$

. . . . ,

$$F_q = F_q(s,t), \quad s \in [0,l], \quad t \in [0,t_f],$$

. . . .

The position of a point P is

$$\begin{aligned} x &= \int_0^s \cos q' \sin\theta' ds' \\ y &= \int_0^s \cos q' \cos\theta' ds' \\ z &= \int_0^s \sin q' ds' \end{aligned} \qquad (A.1.1)$$

and the velocity components are

$$\begin{aligned} v_x &= \int_0^s \left(-\dot{q}' \sin q' \sin\theta' + \dot{\theta}' \cos q' \cos\theta'\right) ds' \\ v_y &= \int_0^s \left(-\dot{q}' \sin q' \cos\theta' - \dot{\theta}' \cos q' \sin\theta'\right) ds' \\ v_z &= \int_0^s \dot{q} \cos q' ds' . \end{aligned} \qquad (A.1.2)$$

For an element dm, kinetic and potential energy will be

$$dW_k = \frac{1}{2} dm \cdot v^2 \tag{A.1.3}$$

$$dW_p = dm \cdot g \cdot z, \tag{A.1.4}$$

where

$$dm = \rho ds. \tag{A.1.5}$$

From (A.1.2)-(A.1.5) we obtain,

$$W_k = \frac{1}{2}\rho\int_0^l \left(\left(\int_0^s \left(-\dot{q}\sin q' \sin\theta' + \dot{\theta}' \cos q' \cos\theta' \right) ds' \right)^2 + \right.$$
$$\left. + \left(\int_0^s \left(-\dot{q}' \sin q' \cos\theta' - \dot{\theta}' \cos q' \sin\theta' \right) ds' \right)^2 + \left(\int_0^s \dot{q}' \cos q' ds' \right)^2 \right) ds \tag{A.1.6}$$

$$W_p = \rho g \int_0^l \int_0^s \sin q' ds' ds . \tag{A.1.7}$$

The dynamic model is obtained by using Lagrange equations of motion

$$\frac{d}{dt}\left(\frac{\partial W_k}{\partial \dot{q}} \right) - \frac{\partial W_k}{\partial q} + \frac{\partial W_p}{\partial q} = F , \tag{A.1.8}$$

where $\frac{\partial W}{\partial q}$ denotes a functional partial (variational) Gateaux derivative (Wang, 1965) that is defined as the variation of the functional W with respect to the function q at a point $s \in [0, l]$. From (A.1.6), (A.1.7) it results,

$$\rho \int_0^s \int_0^s \left(\ddot{q}' \left(\sin q' \sin q'' \cos(q' - q'') + \cos q' \cos q'' \right) - \ddot{\theta}' \cos q' \sin q'' \sin(\theta'' - \theta') + \right.$$
$$+ \dot{q}'^2 \left(\cos q' \sin q'' \cos(\theta' - \theta'') - \sin q' \cos q'' \right) + \dot{\theta}'^2 \cos q' \sin q'' \cos(\theta' - \theta'') -$$
$$\left. - \dot{q}' \dot{q}'' \sin(q'' - q') \right) ds' ds'' + \rho g \int_0^s \cos q' ds' = F_q \tag{A.1.9}$$

$$\rho \int_0^s \int_0^s \left(\ddot{q}' \sin q' \cos q'' \sin(\theta'' - \theta') + \ddot{\theta}' \cos q' \cos q'' \cos(\theta'' - \theta') - \right.$$
$$- \dot{q}'^2 \cos q' \cos q'' \sin(\theta'' - \theta') + \dot{\theta}' \cos q' \cos q'' \sin(\theta'' - \theta') -$$
$$\left. - \dot{\theta}' \dot{q}' \sin q' \cos q'' \cos(\theta'' - \theta') \right) ds' ds'' = F_\theta . \tag{A.1.10}$$

Appendix 2

We consider the following Lyapunov function (Ivanescu, 1984)

$$V(t) = W_k(t) + W_p(t) + \frac{1}{2}\int_0^l \left(k_\theta^1(s)e_\theta^2(s,t) + k_q^1(s)e_q^2(s,t)\right)ds\,. \quad (A.2.1)$$

V(t) is positive definite because the terms that represent the energy W_K and W_P are always

$$W_K(t) \geq 0,\ W_P(t) \geq 0\,.$$

From (17) we obtain (for a steady desired position),

$$\dot{V}(t) = \int_0^l \left(F_\theta(s,t)\dot{e}_\theta(s,t) + F_q(s,t)\dot{e}_q(s,t) + k_\theta^1(s)e_\theta(s,t)\dot{e}_\theta(s,t) + k_q^1(s)e_q(s,t)\dot{e}_q(s,t)\right.. \quad (A.2.2)$$

If we use the control law defined by the relations (21) and (22), we will have

$$\dot{V}(t) = -\int_0^l \left(k_\theta^1(s)(\dot{e}_\theta(s,t))^2 + k_q^1(s)(\dot{e}_q(s,t))^2\right)ds \quad (A.2.3)$$

$$\dot{V}(t) \leq 0\,. \qquad \text{(Q.E.D)}$$

Appendix 3

We extend the Lyapunov function (A.2.1) as

$$V(t) = W_k(t) + W_p(t) + \frac{1}{2}\int_0^l \left(k_\theta^1(s)e_\theta^2(s,t) + k_q^1(s,t)e_q^2(s,t) + \right.$$
$$\left. + k_\theta^3(s)\left(\int_0^t \dot{e}_\theta(s,\tau)\cdot f_\theta(s,\tau)d\tau\right)^2 + k_q^3(s)\left(\int_0^t \dot{e}_q(s,\tau)\cdot f_q(s,\tau)d\tau\right)^2\right) ds\,. \quad (A.3.1)$$

In this case,

$$\dot{V}(t) = \int_0^1 \left(F_\theta(s,t)\dot{e}_\theta(s,t) + F_q(s,t)\dot{e}_q(s,t) + \right.$$
$$+ k_\theta^1(s)e_\theta(s,t)\dot{e}_\theta(s,t) + k_q^1(s)e_q(s,t)\dot{e}_q(s,t) +$$
$$+ k_\theta^3(s)\dot{e}_\theta(s,t)f_\theta(s,t)\int_0^t \dot{e}_\theta(s,\tau)f_\theta(s,\tau)d\tau +$$

$$+k_q^3(s)\dot{e}_q(s,t)f_q(s,t)\int_0^t \dot{e}_q(s,\tau)f_q(s,\tau)d\tau\Bigg)ds \qquad \text{(A.3.2)}$$

and by using the control laws (23) and (24) we obtain

$$\dot{V}(t) = -\int_0^l \left(k_\theta^1(s)\left(\dot{\theta}(s,t)\right)^2 + k_q^1(s)\left(\dot{q}(s,t)\right)^2\right)ds \qquad \text{(A.3.3)}$$

$$\dot{V}(t) \le 0 \ . \quad \text{(Q.E.D)}$$

References

Chirikjian, G. S. and J. W. Burdick (1990) An Obstacle Avoidance Algorithm for Hyper-Redundant Manipulators, *Proc. IEEE Int. Conf. on Robotics and Automation*, Cincinatti, Ohio, May, pp. 625 – 631.

Chirikjian, G. S. and J. W. Burdick (1992). Kinematically Optimal Hyper-Redundant Manipulator Configurations, *Proc. IEEE Int. Conf. on Robotics and Automation*, Nice, May, pp. 415 – 420.

Chirikjian, G. S. (1993a). A General Numerical Method for Hyper-Redundant Manipulator Inverse Kinematics, *Proc. IEEE Int. Conf. on Robotics and Automation*, Atlanta, May, pp. 107 – 112.

Chirikjian, G. S. (1993b). A Continuum Approach to Hyper-Redundant Manipulator Dynamics, *Proc. 1993 Int. Conf. on Intelligent Robots and Systems*, Yokohama, Japan, July, pp. 1059-1066.

Chirikjian, G. S. and J. W. Burdick (1995). Kinematically Optimal Hyper-Redundant Manipulator Configurations, *IEEE Trans. Robotics and Automation*, **vol. 11**, no. 6, Dec., pp. 794 – 798.

Cieslak, R. and A. Moreki (1994). Technical Aspects of Design and Control of Elastic's Manipulator of the Elephants Trunk Type, *5th World Conference on Robotics Research*, Sept. 27-29, Cambridge, Massachusetts, pp. 1351-1364.

De Laurentis, K.I. (2002). Optimal Design of Shape Memory Alloy Wire Bundle Actuators, *Proc. of the 2002 IEEE Int. Conf. on Robotics and Automation*, Washington, May, pp. 2363-2368.

Ge, S.S., T.H. Lee and G. Zhu (1996). Energy-Based Robust Controller Design for Multi-Link Flexible Robots, *Mechatronics*, **vol. 6**, no. 7, pp. 779-798.

Gravagne, I.D. and I. D. Walker (2000). On the kinematics of Remotely - Actuated Continuum Robots, *Proc. 2000 IEEE Int. Conf. on Robotics and Automation*, San Francisco, April, pp. 2544-2550.

Gravagne, I.D. and I. D. Walker (2001). Manipulability, Force, Compliance Analysis for Planar Continuum Manipulators, *Proc. IEEE / RSI International Conference o Intelligent Robots and Systems*, pp. 1846-1867.

Hale, J.K. and J.P. Lasalle (1967). *Differential Equations and Dynamical Systems*, Academic Press.

Hemami, A. (1984). Design of Light Weight Flexible Robot Arm, *Robots 8 Conference Proceedings,* Detroit, USA, June 4-7, pp. 1623-1640.

Ivanescu, M. (1984). Dynamic Control for a Tentacle Manipulator, *Proc. Int. Conf. on Robotics and Factories of the Future*, Charlotte, pp. 315-327.

Ivanescu, M. and V. Stoian (1995). A Variable Structure Controller for a Tentacle Manipulator, *Proc. IEEE Int. Conf. on Robotics and Automation*, Nagoya, May 21-27, pp. 3155-3160.

Mishkin, E. and L. Braun Jr. (1961). *Adaptive Control Systems*, New York, McGraw Hill.

Mochiyama, H., E. Shimeura and H. Kobayashi (1998). Direct Kinematics of Manipulators with Hyper Degrees of Freedom and Serret-Frenet Formula, *Proc. 1998 IEEE Int. Conf. on Robotics and Automation*, Leuven, Belgium, May, pp. 1653-1658.

Mochiyama, H. and H. Kobayashi (1999). The Shape Jacobian of a Manipulator with HyperDegrees of Freedom, *Proc. 1999 IEEE Int. Conf. on Robotics and Automation,* Detroit, May, pp. 2837-2842.

Robinson, G. and J.B.C. Davies (1999). Continuum Robots – A State of the Art, *Proc. 1999 IEEE Int. Conf. on Robotics and Automation,* Detroit, Michigan, May, pp. 2849-2854.

Shigoma, K. B. (1996). Robot Grasp Synthesis Algorithms: A Survey, *The International Journal of Robotics Research*, **vol. 15**, no. 3, pp. 230-266.

Suzumori, K., S. Iikurav and H. Tanaka (1991). Development of Flexible Micro-Actuator and its Application to Robot Mechanisms, *IEEE Conference on Robotics and Automation,* Sacramento CA, April, pp 1564-1573.

Wang, P.K.C. (1965). Control of Distributed Parameter Systems, in *Advance in Control Systems*, by C.T. Leondes, Academic Press.

Wang, L.-X. (1999). Analysis and Design of Hierarchical Fuzzy Systems, *IEEE Trans. on Fuzzy Systems*, **vol. 7**, No. 5, Oct., pp. 617-624.

Wang, Z.P., S.S. Ge and T.H. Lee (2001). Non-Model-Based Robust Control of Multi-Link Smart Materials Robots, *Asian Conference on Robotics and its Applications*, 6-8 June, Singapore, pp. 268-273.

ROBOTS FOR HUMANITARIAN DEMINING

P. Kopacek
Institute of Handling Devices and Robotics, Vienna University of Technology
Favoritenstr. 9–11, A–1040 Vienna, Austria
kopacek@ihrt.tuwien.ac.at

Abstract "Humanitarian demining", is growing up dramatically in the last decade. Here a new idea – application of a very well known tool from production automation "advanced robots" - will be presented. These robots of the new generation offer possibilities to solve this task in a very efficient way. Finally "Humanitarian Demining Multi Agent Systems – HDMAS" an autonomous, intelligent robot swarm for cleaning minefields in the future is discussed.

Keywords: landmines, demining, robots, MAS, robot swarms

1. Introduction

According to current estimates, more than 100.000.000 anti-personnel and other landmines have been laid in different parts of the world. A similar number exists in stockpiles and it is estimated that about two million new ones are being laid each year. According to recent estimates, mines and other unexploded ordnance are killing between 500 and 800 people, and maiming 2.000 others per month (Red Cross, 1995), mainly innocent civilians who had little or no part in the conflicts for which the mines were laid. Anti-personnel mines are usually designed not to kill, but to inflict horrible injuries instead (McGrath, 1994). However, many victims eventually die of their injuries, and suffer a long and agonizing death, often with little medical attention.

Some countries have banned the use of landmines and others are supportive of a complete ban. However, their low cost ($1- $30) and the large numbers in existing stockpiles make them an attractive weapon for insurgency groups which operate in may countries with internal conflicts – the most common cause of wars today. They are used for self-defense by villages and groups of people traveling in many districts where civil law and order provide little effective protection. Many countries retain massive landmine barriers on their borders or near military installations. Some of the most severe landmine problems exist in Egypt, Angola, Afghanistan, Rwanda, Bosnia, Cambodia, Laos, Kuwait, Iraq, Chechnya, Kashmir, Somalia, Sudan, Ethiopia, Mozambique and the Falkland Islands.

2. Landmines

Landmines are usually very simple devices which are readily manufactured anywhere. There are two basic types of mines: anti-vehicle or anti-tank (AT) mines and anti-personnel (AP) mines.

AT mines are comparatively large (0.8 – 4 kg explosive), usually laid in unsealed roads or potholes, and detonate which a vehicle drives over one. They are typically activated by force (>100 kg), magnetic influence or remote control.

AP mines are much smaller (80-250g explosive, 7-15cm diameter) and are usually activated by force (3-20kg) or tripwires. There are currently over 700 known types with different designs and actuation mechanisms. We have two main categories of AP mines. A *blast mine* is usually small and detonates which a person steps on it: the shoe and foot is destroyed and fragments of bone blast upwards destroying the leg. When a *fragmentation mine* explodes, metal fragments are propelled out at high velocity causing death or serious injuries to a radius of 30 or even 100 meters, and penetrating up to several millimeters of steel if close enough. Simple fragmentation mines are installed on knee high wooden posts and activated by tripwires (stake mines). Another common type of fragmentation mine (a bounding mine) is buried in the ground. When activated, it jumps up before exploding. Mines of one type have often been laid in combination with another type to make clearance more difficult: stake mines with tripwires may have buried blast mines placed around them.

3. Demining; state of the art

First you have to find the mines and then you must destroy. Used methods for identifying mines today are:

- Manually: by humans – deminers – equipped with e.g. metal detectors.
- Dogs: using dogs that sniff the explosive contents of the mines, has significant limitations and cannot be regarded to as general-purpose solution.
- High-tech methods for mine detection: radar, infrared, magnetic tools,

touching sensors (piezo resistive sensor) …. Also GPS is used for finding the place again where a mine lies, and for the navigation of the robots.

Today used methods for destroying and removal are: brutal force methods include ploughs, rakes, heavy rolls, flails mounted on tanks.

The problems with these methods are that.

- Ploughs only can by used to clear roads for military use. Mines are only pushed to the side of the road. Some ploughs also attempt to sieve the mines from the displaced soil.
- Flails are mechanical devices, which repeatedly beat the ground, typically with lengths of chain. These chains are attached to a rotating drum and their impact on the ground causes the mines to explode, but this can cause severe damage to cultivable land.
- Rollers generally consist of a number of heavy circular discs, which are rolled along the ground in order to cause the explosion of any mines.

Before demining can start, surveys are needed to produce detailed maps of minefields to be cleared. The survey team may use specially trained dogs to narrow down the limits of a mined area, and normally verifies a one or two meter wide “safe lane” around each minefield to define the minefield which may be surrounded with unknown land or other minefields. Typical minefields are 100-200m across and 0.1-10ha in area.

Hand-prodding is today the most reliable method of mine clearing, but it is a very slow, and extremely dangerous. People performing this type of clearing can normally only perform this task for twenty minutes before requiring a rest. This method clears one square meter of land in approximately 4 minutes.

The tools of a deminer are:

1. A whisker wire which is gently swung or lifted to check for tripwires.
2. A metal detector which is swung from side to side to check for metal objects.
3. A prodder (typically a bayonet, screw driver or knife) which is used to probe the ground at an angle of about 30 degrees to the horizontal and to excavate earth from around a suspect object. Usually a prodder is used to investigate a suspect metal object. However, which dealing with minimum metal mines or large numbers of metal fragments, the entire area has to be prodded by hand.

The UN estimates the cost of removing a single mine at 300 to 1000 $. The primary factor is the local cost of labor. So in low labor-cost countries such as (Cambodia, Afghanistan, or Africa) US$ 100 per month is a high rate pay for manual work, even with the obvious risks. In contrast, the labor cost for de-mining in the former Yugoslavia may be twenty times higher.

Thinking about the number of mines is rather pointless which estimates range from a few million word-wide (including national borders) to 150 million. It is much more sensible to think in terms of the areas of land which are:

a) known to be affected by mines, and are important to local or displaced population:
   homes, food producing land, roads, infrastructure (roads, canals, power lines, water supplies etc.);
b) believed to be affected by mines;
c) known or believed to be affected by mines, but land is of no immediate importance.

The standard which is required for humanitarian demining is a guaranteed 99.6% clearance. Therefore the remaining risk to be injured or killed by a mine is 0.4%.

Mechanical mine clearance means either actuating the mine, or removing it for later destruction.

For actuating ploughs are pushed by a tank or an armored bulldozer. There is a bulldozer with a rotating cylinder in front, digging up to 50cm into the ground. The vehicle has been tested in Mozambique. Although it did not reach the 99.6% UN requirements, it removed 25.000 mines in a six-month campaign. Another demining vehicle uses the same principle, with closer teeth. It is based on a Leopard 1 main battle tank chassis to which a rotating roller is added. The tank can be remote controlled from 500m away. In normal terrain this vehicle should clear up to 20.000 square meters per hour with total safety for the mine-clearing team.

The disadvantage is that mines includes a lot of chemicals which when they detonate are put into the ground which is later used for food producing.

4. Robots for demining

4.1. State of the art

Several projects have proposed the use of autonomous robots to search for antipersonnel mines. The sensor problem is nearly solved now and it will take only little time until a combination of sensors will be available and sufficiently tested in order to give full confidence that all the mines have been discovered. There may be false alarms, but no mine must be left. Once the location of a mine is known, several manual techniques are easily applied to neutralize it. A robot can also be developed to do this easy job, which is simply to go to a precise spot, avoiding obstacles and other mine locations. Then it should deposit a shaped charge or some chemical to destroy the mine.

The necessary features of a demining robot are:

- Ability to distinguish mines from false alarms like soil clumps, rocks, bottles and tree roots.

- Operation in a variety of soil types, moisture contents and compaction states.
- Ability to detect both types or in fact variety of different mine types and sizes.
- Operation in vegetated ground cover.
- Costs may be lower then 10.000 US$ including the sensors.

Today there are some appropriate, reasonable cheap sensors available or in development based on optical technologies, acoustic and seismic detection, radio frequency resonance absorption spectroscopy, trace explosion detection. Worldwide approximately 100 companies or research institutes offers intelligent, mobile platforms but the price is too high according to the small lot sizes in production. It's only a question of time until this problem is solved.

Random navigation for covering the field and searching for mines has been proposed. Even with improved algorithms applied to a group of robots, it is difficult to accept ignoring a small proportion of uncovered areas. Systematic navigation is theoretically easy with a global positioning system (GPS), but the resolution must be better than the size of the detector, in order to be sure to cover all the area.

A robot has been designed as a light-weight autonomous robot to search for antipersonnel mines. The pressure force on the ground, 5kg, is not intended to trigger the mines. The sensor head oscillates under the alternating movement of the wheels, in order to scan a width of about 1.2 m. the project is suspended until an adequate sensor, weighing less than 4kg, can be installed inside the head.

Research groups experienced with walking robots try to suggest the use of their devices for this application. Snake robots are more attractive, since they should be able to crawl close to the ground inside dense vegetation. Their design is, however, quite challenging.

The advantages of robots for demining are

- Minefields are dangerous to humans; a robotic solution allows human operators to be physically removed from the hazardous area.
- Robots can be designed not to detonate mines.
- The use of multiple, inexpensive robotized search elements minimizes damage due to unexpected exploding mines, and allows the rest of the mission to be carried on by the remaining elements.
- Teams of robots can be connected so that one team is for searching and one for destroying or displacement.

This means that many robots are searching and a few or only one robot is destroying or displacing the mines.

But there are also disadvantages of using robots:

- it is very difficult for robots to operate in different frequently rough terrain;

- they are still expensive;
- you need special qualified operators.

4.2. Multi Agent Systems - MAS

A MAS consists of a number of intelligent, co-operative and communicative hardware agents – mobile robots – getting a common task. Because of the intelligence they are able to divide the task in subtasks as long as at least one agent is able to fulfill one subtask. Repeating this procedure yields the solution of the common task. Newest research goes in the direction of MMAS - Multiple Multi Agent Systems – different MAS are involved for the solution of a complex task.

A MAS get a whole task. The host computer divides the whole task in a number of different subtasks as long as a distinct subtask can be carried out by at least one agent. The agents will fulfill their subtasks in a cooperative way until the whole task is solved. Such a global task could be: assemble a car. The agents – mobile, intelligent assembly robots – have to create subtasks (e.g. assembling of wheels, windows, brakes,......) in an optimal way (equal distribution of the workload of the agents) and distribute to the agents.

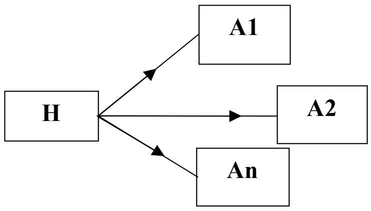
The main hurdles for MAS-research are the complexity of the whole system. This complexity is dramatically increasing by adding new agents. Therefore the interaction, communication, coordination of the tasks between agents, and control are the topics for the development of a Multi Agent System (MAS).

For heterogeneous robots it is difficult to implement the communication, because each robot has its own kinematical structure, programming language etc.. Furthermore the range of frequencies used for communication and the capability of RF modules is limited. It is necessary to develop standardized communication protocols and methods, which should be one of the works for the next years.

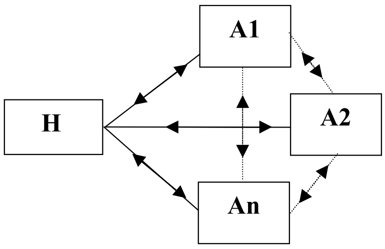
Fig.1a. shows the present situation of the communication between the host and the agents. For the future the agents should also communicate with the host and also with the other agents as shown in Fig. 1b and Fig.1c.
The characteristics of MAS are:

- each agent has incomplete information or capabilities for solving the problem and, thus, has a limited viewpoint;
- there is no system global control;
- data are decentralized;
- computation is asynchronous.

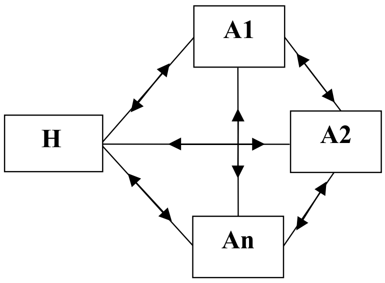
In scientific papers one uses various approaches and denotations to subdivide control strategies for autonomous mobile robots in different types. There are two fundamental describing ideas: the functional and the behavior based approach.



a. Communication only between host and agents



b. Partial communication between host and agents as well as between the agents



c. Full communication between host and agents as well as between the agents

Figure 1. Different types of MAS H : Host computer; A1, A2,.....,An : Agents;

————: full communication; ------------ : partial communication

4.3. Robot Swarms – MAS - for demining

Robot swarms improve the capacity of robotic applications in different areas where robots are already used today. Robot swarms are similar to – or a synonym for - ‘Multi Agent Systems – MAS’. These systems are very well known in software engineering – “software agents” - since more than twenty years. In the last years there are more and more works related to “hardware agents” like robots forming “robot swarms”.

Applying robots for demining there are two possibilities:

a. using mobile, intelligent multipurpose robots equipped with devices for mine detection, mine removing as well as mine transportation;
b. using three different swarms of single-purpose robots equipped either with detection devices or removing devices or transportation facilities.

Our approach is the second one – three different swarms in the minefield. The detection robots scan the field for possible mines. If a metallic part –

probably a landmine is detected one of the removal robots close to this site removes the mine and takes it over to a transportation robot with free capabilities. This robot transports the mine to a collection area out of the minefield. The whole process is fully autonomous. Operators are only needed for monitoring and of course maintenance. To achieve this goal the robots must have a high level of intelligence and must be able to communicate among themselves. Since the power supply of mobile robots is very limited there is also need for docking stations. The host computer in Fig. 2 is necessary to solve the path planning problem in a dynamic environment. Each robot represents for all other robots a dynamic obstacle which has to be avoided. The host computer controls the movements of all robots by means of wireless communication. But soon such a host computer will be obsolete (Fig.1c). Software implemented in the onboard computer of each robot will take over this task.

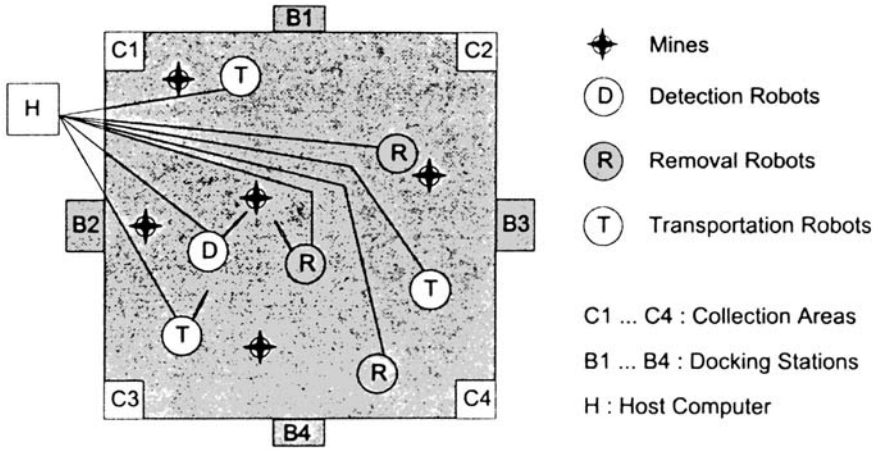


Figure 2. Humanitarian Demining Robot Swarms [Kopacek, 2002b]

Using different single purpose robots for the different tasks reduces the weight of the robots. Therefore it is much easier to design robots which are lightweight enough not to cause an explosion while crossing over a mine field.

As mentioned before the use of modular robots is perfect for the design of task-specific demining robots because of the similarities between the tasks. [Shivarov, 2001]

Since the complexity of a system raises the susceptibility to trouble exponential it is always better to keep devices as simple as possible and therefore to use simpler robots. Using smaller robots extends the operational time before re-fuelling or re-charging is necessary or at least prevents the use of bulky and heavy batteries or tanks.

On the other side the second possibility requires an increased effort in communication between the robots in the swarm. If every robot is able to perform the whole demining process by itself the communication is reduced more or less to get each other out of the way and ensure to cover all the area. However the task-specific robots have to exchange a lot more of data. The detection robots must work together since they are equipped with different detection technologies. When they have found a mine they must signal it to the removal robots and they have to inform after done work the transportation robots.

4.3.1. Detection robots

The robots for the detection of landmines are probably the most simplest of the three types. The basic composition of modules common for all three types has to be upgraded only with the detection system. There are several detection technologies in use respectively under development, none of them able to detect a mine alone by itself. The solution is to use two or more of these different sensors simultaneously. The first logical step would be to attach different sensors on one robot. Since there are weight limits, and limits in the amount of available energy, this is probably not the best solution. Some of these technologies need strong power sources and some of them are relatively heavy constructed. These facts will not help to keep the weight of the robot low, so using for each type of sensor a single robot seems to be the better solution.

The detection robots should communicate with each other, change data and coordinate their work. If one robot with one distinct detection technology has found a possible target, the area should be verified by all other detection technologies before any further action is started. Therefore the different technologies must be compatible to allow coordination. At least the data from the sensors should be assessed by the same software. Combining results from different mine detection technologies is not easy and demands special strategies. These so-called sensor fusion technologies are not only of concern for mine detection.

Another important point is the power supply of the detection swarm. One possibility is to equip detection robots with an autonomous power source. But this could complicate the recharging of the system. There would be need for extra docking stations and at the worst for each detection technology a different docking station.

This cooperation during the development and design process of the modular robot system and landmine detection sensors is of greater concern than only for an appropriate modular interface. Some of these sensor systems are extremely sensible and may drop in performance in presence of distinct materials. Using these materials for parts of the robot system which has to carry the sensor technology has to be avoided. And many of the sensor techniques work by using radiation in some range of the electromagnetic spectrum. It has to be guaranteed that systems of the robot do not jam the sensor technology or the other way round.

4.3.2. Removal robots

The removal of landmines is probably the heaviest work during the whole de-mining process. This is clearly a matter of the type of soil in which the mines are buried. But generally this task needs the highest forces and therefore the system has to be more stiff and heavy constructed.

The removal robots have also the most complex part to fulfill. While the detection robots only transport the detection technology and the transportation robots have to accomplish an advanced pick and place task, the removal robots have to, in case of buried mines, dig out a highly sensitive device, which must be handled extremely carefully, but at the same time applying relatively high forces to penetrate the soil. In addition the excavation of a mine is every time a different procedure. The main parameters which differ for each buried mine are the type and shape of the mine, the position relative to the surface and the type of soil in which the mine is buried.

Since the excavation is a complex task a dexterous robot arm with a high number of degrees of freedom is likely to be used. For the mine removal various end-effectors may be necessary. The robot arm can be equipped with a variety of standard tools which are similar to tools used for manual excavation. All forms of shovels are doubtless of interest to remove foremost close grained material. Grippers may used to sweep stones or other bigger obstacles. These tools are commercially available and well proven.

Up to the present the most removal work performed at hazardous materials was executed teleported. For that the aid of sensors is mainly limited to force and torque sensors which ensure not to apply too high forces to the sensible object. But the whole process is controlled by an operator using video cameras to lead the tools. Using a robot for autonomous removal of landmines presupposes the usage of sensors to compensate the teleoperator. Two broad classes of sensing technologies support earthmoving automation. One class allows determining the state of the robot itself, the other class concerns perception of the environment around the earthmover.

Local state is achieved by measuring displacements at the robots various joints. If the actuators are hydraulic cylinders the use of position transducers would be a good choice. An alternative is to use joint resolvers, like potentiometers, directly at rotary joints. Another form of state estimation is to locate the robot arm with respect to some fixed coordinate frame. Many sensing modalities have been used including, GPS, inertial sensors and reflecting beacons. Successful estimation schemes combine several of these techniques.

4.3.3. Transportation robots

The transportation seems to be quite simpler than the removal of a landmine. Basically the robot has to pick up the landmine, store it somewhere during the transportation and deliver it at the collection point.

An important decision in respect of the transportation robots is the number of mines the robots should be able to carry. Carrying only one mine would it make possible to use a rather simple robot. At the best it may possible to retrench the storing place for the landmine. The robot could pick up the mine with a gripper, lift it up somewhat above the ground and transport it to the collection area while holding it tight with the gripper. The use of a dexterous robot arm, like that one for the removal task, would be disproportionate. A simple 2 DOF lift onboard the mobile robot platform could be sufficient.

On the other side the application of a transportation robot with the ability to carry more than one mine is in a manner useful too. Since transportation robots are likely to be rather slow this approach is much more timesaving. The volume of saved time depends on the amount and distribution of collection areas in proportion to the field of activity as well. But establishing lesser collection areas simplifies the further strategy for the disposal of the collected landmines. To give the robot the ability to transport more than one mine it must be equipped with some sort of storage device.

One principle would be of use to make the storage device of protective material to mitigate accidentally explosions. One possibility is to use a lockable storage device. But therefore the device must be designed with regard to a maximal allowed load of explosives. An explosion inside a locked container exceeding the maximal allowed load may be worse than without any protective measures. Fragments of the blasting container could damage the robot in addition. For this reason it would be better to use a container which is opened upwards. This guarantees a way out for the pressure wave in case of an accidental explosion.

An important factor for the decision of using single or multi transport robots is the density of the minefield. If there are only few landmines per surface unit the application of single-mine transportation robots is more likely. In this case the work quota of the detection robots is much higher compared to that of the removal and transportation robots. Therefore raising the working capacity of the transportation robots would not increase the overall efficiency perceptible.

4.4. Realization

The features of the robot for these three tasks have to be quite different. For detection a light-weight robot only able to carry little load has to be developed. For removal the robot has to be more stiff and heavy constructed because removal requires force. The size of transportation robots depends on the number and kind of the mines to be transported.

Another point of view which has to be taken into account is the time necessary for these operations. Detection is usually relatively fast and is not so time consuming than removal. According to some experiences the removing time is 3 to 5 times more than the detection time. Transportation time is also relatively small.

Therefore it could be advantageous to use three different types of robots (Fig. 2): robots for detection (D), robots for removal (R) and robots for transportation (T) of the mines. One main disadvantage of this philosophy is if a robot of the swarm D (detection) has found or detected a mine it has to send a command to the host computer or to the other robots. The host computer or the other robots have to decide which of the robots of the swarm R (removal) is in the neighborhood of this mine and not busy at that time with removal operations on another mine. If a robot of the swarm R is selected this robot gets usually wireless the position data and some other information about the place of the mine. The R robot is now moving to displace and start with the removal work. After the removal of the mine it has to place the mine on the ground in a distinct position. One of the transportation robots (T) have to pick up the mines and have to carry it to a collecting place.

4.4.1. Mobile robots

Today we are in the position to develop robots of all three types mainly using commercially available mobile platforms. As pointed out earlier it is not economically feasible to develop so-called single purpose robots for each of these three types. A good approach could be a kind of a tool kit [Shivarov, 2001] of mobile robots consisting of a mobile platform and different equipments and tools compatible in hard- and software. A good approach could be to have two platforms, one with wheels or chains and one walking platform. According to the types of mines as well as the surface of the minefield these platforms could be equipped with necessary tools in a very short time.

Usually the mobile robots of both types available today are moving relatively slow. Speed for wheeled and chained robots is between 0.5 and 0.7 m/s, walking robots are usually much slower. This could be a disadvantage concerning the demining time but from the viewpoint of control and path planning it is much easier to work with such slow robots. We have in that case the usual problem of path planning of robots in a changing environment. Usually in a minefield we have fixed obstacles like trees, rocks, buildings as well as moving obstacles usually the robots of the own or other swarms.

4.4.2. Humanoid robots [Kopacek, 2003]

The main feature of a real human is the two legged movement and the two legged way of walking. In principle the stability during the walking decreases with the number of the legs. At the begin of this development there were consequently 8, 6 and 4 legged robots copied from the nature (insects, swarms,

…). In the future two legged robots should be responsible for human tasks like service applications, dangerous tasks, tasks on the production level, support of humans in everyday life …

The main advantage of legged robots is the ability to move in a rough terrain without restrictions like wheeled and chained robots. Two legged robots could work in environments which were until now reserved only for humans. In addition to walking such robot could realize other movements like climbing, jumping, swimming, …. Walking robots are much more flexible than robots with other movement possibilities. Especially fixed and moved obstacles can be surmounted by legged robots.

Two legged robots require new technologies in the field of robotics. In some cases a combination from well known methods of mechanical engineering, electrical engineering, electronics, control engineering, computer sciences, applied physics are necessary.

Currently there are worldwide two categories of two legged humanoid robots available:

- "Professional" humanoid robots.
- "Research" humanoid robots.

The humanoid robots of the first category are mostly developed but very expensive and currently not available on the market. The robots of the second category a usually prototypes

Therefore for several tasks e.g. humanitarian demining a two legged, humanoid robot should be developed. These robots could be applied for all three tasks of demining – detection, removing and transportation in the future.

5. Summary

As pointed out demining is today a very time consuming, dangerous and expensive task. Automatic demining e.g. as presented in this paper by robots, is today and will be in the future a powerful tool to solve these problems. All the existing and planed robots for humanitarian demining are only able to detect the mines. Brutal force methods destroy mines without detection – but with a not reasonable probability. In a next step our robots have to be able to remove the mines from the ground.

"Multi Agent Systems – MAS" [Kopacek, 2002a] are very well known in software engineering since more than 20 years. In the last years there are some works related to the application in production automation. A MAS consists of a number of intelligent, co-operative and communicative hardware agents e.g. robots getting a common task. Because of the intelligence they are able to divide the whole task in subtasks as long as at least one of the agents is able to fulfill one subtask.

Repeating this procedure yields the solution of the common task. Newest research goes in the direction of MMAS – Multiple Multi Agent Systems – different MAS are involved for the solution of a complex task. In a mid or long term perspective it might be possible to develop "Humanitarian Demining Multi Agent Systems – HDMAS " consisting of a number of such robots or agents [3]. Robot swarms or HDMAS for demining especially with two legged (humanoid) robots are currently only a vision but will be reality in the nearest future.

References

McGrath, R. (1994): Landmines, Legacy of Conflict: A Manual for Development Workers, Oxfam, Oxford.

Red Cross (1995): Landmines must be stopped. International Committee of Red Cross, Geneva, Switzerland.

Trevelyan, J. (1997): Robots and landmines. Industrial Robot, 24 (1997), Nr.2, p.114-125

Baudoin, Y. et.al. (2000): "Humanitarian Demining: Sensory and Robotics". Proceedings of the IMEKO World Congress 2000, Vienna, Vol. XI, p. 241 – 251.

Kopacek, P. (2000): "SWIIS – An Important Expression of IFAC's Commitment to Social responsibility", Preprints of the IFAC Workshop on Supplemental Ways for Improving International Stability – SWIIS 2000, May 2000, Ohrid, Macedonia.

Shivarov, N. (2001): "A tool kit for modular, intelligent, mobile robots". PhD. Thesis, Vienna University of Technology, 2001.

Kopacek, P. (2002a): "Demining Robots – a tool for International Stability". Proceedings of the 16th IFAC World Congress, Barcelona, July 2002.

Kopacek, P. (2002b): Robot Swarms for Demining. Proceedings of the IARP Workshop "Robots for Humanitarian Demining – HUDEM`02", Vienna, Austria, 2002; p. 39-44.

Kopacek, P. (2003): Humanoid Robots for Demining – Vision or Reality. Proceedings of the IARP Workshop "Robots for Humanitarian Demining – HUDEM`03", Pristina, Kosovo, 2003 (In publishing).

PARAMETRIZATION OF STABILIZING CONTROLLERS WITH APPLICATIONS [1]

Vladimír Kučera

Czech Technical University in Prague, Faculty of Electrical Engineering
Center for Applied Cybernetics, Prague, Czech Republic
http://www.fel.cvut.cz/dean/, kucera@fel.cvut.cz

Abstract This contribution addresses, in a tutorial and retrospective manner, the parametrization of all controllers that stabilize a given plant with rational transfer function. An account of the classical results that paved the way for the parametrization is given. The parametrization result is then derived for several definitions of stability. The parameter, which is a qualified rational function, is shown to appear in the feedback system transfer functions in an affine manner. A two-step procedure for control system synthesis is then formulated, namely to determine all stabilizing controllers first, then meet additional performance specifications by selecting the parameter. Various applications of this procedure are given and illustrated by numerous examples. Advantages as well as limitations of this approach are discussed.

Keywords: linear systems, feedback systems, stabilization, parametrization, control system synthesis

1. Introduction

The majority of control problems can be formulated using the diagram shown in Figure 1. Given a plant *S*, determine a controller *R* such that the feedback control system is (asymptotically) stable and satisfies some additional performance specifications such as reference tracking, disturbance attenuation, optimality, robustness, or system integrity.

It is natural to separate this task into two consecutive steps: (1) stabilization and (2) achievement of additional performance specifications. To do this, all solutions of the first step, i.e. *all controllers that stabilize the given plant*, must be found.

---

[1] Supported by the Ministry of Education of the Czech Republic under Project LN00B096

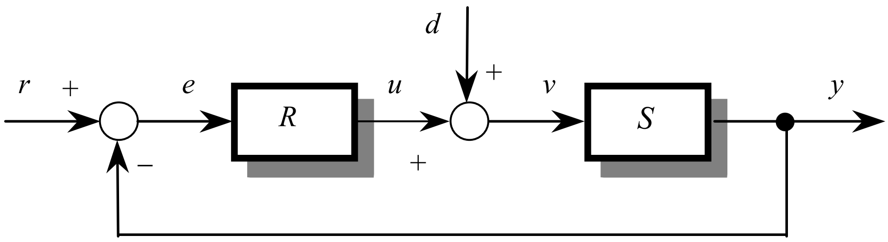


Figure 1. Feedback control system

How can one characterize such controllers? In case the plant is *stable* and one stabilizing controller R is known, then

$$V = \frac{R}{1 + SR}$$

is a stable rational function. On the other hand, if V is any given stable rational function, then the corresponding controller

$$R = \frac{V}{1 - SV}$$

must necessarily stabilize the plant S. Therefore the stabilizing controllers can be parametrized by the set of stable rational functions (Newton, *et al*., 1957).

As argued by Kučera (2002), if H_{sens} denotes the reference-to-error transfer function (sometimes called the sensitivity function) and H_{comp} the disturbance-to-control transfer function (the so called complementary sensitivity function) in the closed loop control system, namely

$$H_{\text{sens}} = \frac{1}{1 + SR}, \quad H_{\text{comp}} = \frac{SR}{1 + SR},$$

then the preceding result can be phrased as follows: the control system is stable if and only if $H_{\text{comp}} = SV$, since V can be interpreted as the reference-to-control transfer function from r to v. This means that H_{com} must absorb all the unstable zeros of the plant S. In case the plant is *unstable*, however, V is no longer arbitrary: the zeros of V must absorb all the unstable poles of S.

To derive stability conditions, one needs to know the (unstable) poles and zeros of the plant. Expressing S as the ratio of two coprime polynomials, $S = b/a$, and assuming the controller in a like form, $R = n/m$, the two closed loop transfer functions can be written as

$$H_{\text{sens}} = a\frac{m}{am + bn} := aX, \quad H_{\text{comp}} = b\frac{n}{am + bn} := bY.$$

Consequently, a stable control system calls for stable rational functions X and Y. These functions cannot be arbitrary, however, since $H_{\text{sens}} + H_{\text{comp}} = 1$. A stability equation follows (Strejc, 1967)

$$aX + bY = 1\,.$$

Any stabilizing controller can be expressed as $R = Y/X$, where X and Y is a stable rational solution pair of the stability equation (Kučera, 1974). This solution can be expressed in parametric form, furnishing in turn an explicit parametrization (Youla, *et al*., 1976a) of all stabilizing controllers

$$R = \frac{y + aT}{x - bT}\,.$$

Here x and y are any polynomials satisfying the equation $ax + by = 1$ while T is a parameter ranging over the set of stable rational functions (and bound to satisfy $x - bT \neq 0$).

The set of stabilizing controllers admits transfer functions R that are not *proper*. Example: given $S(s) = 1/s$, one calculates $x = 0$, $y = 1$ so that

$$R(s) = \frac{1 - sT}{T}\,.$$

Taking $T = 1$ leads to the stabilizing controller $R(s) = 1 - s$. The resulting feedback system is asymptotically stable but, alas, it has poles at $s = \infty$.

If impulse modes are to be eliminated, stability has to be defined in a different way. The asymptotic stability of the control system in Fig. 1 will be replaced by the requirement that any external input d, r of bounded amplitude result in the internal signals e, v (hence also u, y) of bounded amplitude. One can say that such a control system is *internally* stable. While the control system is asymptotically stable if and only if its characteristic polynomial is Hurwitz, it is internally stable if and only if the four transfer functions from d, r to e, v (or u, y) are proper (analytic at the point $s = \infty$) and stable (analytic at the closed right half plane Re $s \geq 0$). Naturally, this notion of stability does not capture hidden modes in the plant and in the controller. These modes, however, cannot be stabilized by output feedback anyway. That is why the internal stability is a natural option.

In order to study internal stability, it is convenient to express the transfer functions of unstable systems as ratios of two coprime transfer functions, each representing a stable system. Internal stability can than be told by inspection: the four transfer functions have a trivial denominator. This is a key observation in an attempt to obtain a simple condition for the internal stability of closed loop systems (Desoer, *et al*., 1980). Accordingly, the polynomial fractional representations used in the study of asymptotic stability will be replaced by fractional representations over *proper and stable rational functions*. For example, the integrator transfer function $S(s) = 1/s$ will be written in the form

$$S(s) = \left(\frac{1}{s + \lambda}\right)\left(\frac{s}{s + \lambda}\right)^{-1},$$

where λ is a positive real; the particular value of λ is irrelevant.

When studying discrete-time control systems, the typical proper stable fractional representation has the form

$$S(z)=\frac{1}{z-1}=\left(\frac{1}{z-\lambda}\right)\left(\frac{z-1}{z-\lambda}\right)^{-1},$$

where $|\lambda| < 1$. A legitimate choice is $\lambda = 0$. Proper and stable rational functions in z whose poles are all located at the point $z = 0$ can be viewed as *polynomials* in z^{-1}:

$$S(z)=\left(\frac{1}{z}\right)\left(\frac{z-1}{z}\right)^{-1}=\frac{z^{-1}}{1-z^{-1}}.$$

This representation has been in use for a long time, see Kučera (1979). The methodology explained above provides an elegant justification for the use of z^{-1} in lieu of z in the synthesis of discrete-time control systems.

2. Parametrization

We shall now derive a parametrization of all controllers that internally stabilize a plant with a given rational transfer function, which is not necessarily proper, nor stable. The derivation is a variation of the one given by Vidyasagar (1985).

Theorem 1 *Let $S = B/A$, where A and B are coprime, proper and stable rational functions. Let X and Y be two proper and stable rational functions satisfying the Bézout equation*

$$AX + BY = 1.$$

Then the set of all controllers that internally stabilize the control system shown in Fig. 1 is given by

$$R=\frac{Y+AW}{X-BW}$$

where W is a parameter ranging over the set of proper and stable rational functions such that $X-BW \neq 0$.

Proof. It consists of three steps.

1) Fist we shall show that if $S = B/A$ and $R = N/M$ are two coprime fractions of proper and stable rational functions, and if C is defined by $C := AM + BN$, then the control system is internally stable if and only if $1/C$ is proper and stable.

Indeed, the control system is internally stable if and only if the four transfer functions

$$\begin{bmatrix} v \\ y \end{bmatrix}=\frac{1}{1+SR}\begin{bmatrix} 1 & R \\ S & SR \end{bmatrix}\begin{bmatrix} d \\ r \end{bmatrix}=\frac{1}{C}\begin{bmatrix} AM & AN \\ BM & BN \end{bmatrix}\begin{bmatrix} d \\ r \end{bmatrix}$$

are proper and stable. The sufficiency part of the claim is evident: the transfer functions are all seen to be proper and stable. The necessity part is not evident: the denominator C can have zeros in Re $s \geq 0$ or at the point s = ∞ which, conceivably, might cancel in all four numerators AM, AN, BM, and BN. However, this is impossible as the pairs A, B and M, N are both coprime.

2) Further we shall show that a controller R internally stabilizes the plant $S = B/A$ if and only if it can be expressed in the form $R = M/M$ for some proper and stable rational solution pair M, N of the Bézout equation $AM + BN = 1$.

Indeed, if the equation is satisfied, then $C = 1$ and the control system is internally stable. Conversely, if some controller $R = \overline{N}/\overline{M}$ internally stabilizes S, then $C = A\overline{M} + B\overline{N}$ and the inverse $1/C$ is proper and stable. Therefore, $M = \overline{M}/C$ and $N = \overline{N}/C$ is a proper and stable rational solution pair of the Bézout equation and it defines the same controller $R = \overline{N}/\overline{M} = N/M$. The proper and stable factor C is seen to cancel from both sides of the Bézout equation.

3) Finally we shall prove that all proper and stable rational solution pairs of the equation $AM + BN = 1$ are given by

$$M = X - BW, \quad N = Y + AW,$$

where X, Y is a particular solution pair of this equation and W is a parameter that ranges over the set of proper and stable rational functions.

Indeed, M and N satisfy the Bézout equation:

$$A(X - BW) + B(Y + AW) = 1.$$

It remains to show that every solution pair of the equation has the form shown above for some proper and stable rational function W. We have

$$A(X - M) = B(N - Y).$$

Since A and B are coprime, A is a factor of $N - Y$ while B is a factor of $X - M$. Put $W := (N - Y)/A$. Then $X - M = BW$, and the claim has been proved. ■

Let us illustrate the above theorem by determining all controllers that internally stabilize the plant

$$S(s) = \frac{1}{s-1}.$$

A fractional representation of the plant transfer function is obtained as follows

$$\frac{1}{s-1} = \frac{B(s)}{A(s)}, \quad A(s) = \frac{s-1}{s+1}, \quad B(s) = \frac{1}{s+1}.$$

The Bézout equation

$$\frac{s-1}{s+1}X(s)+\frac{1}{s+1}Y(s)=1$$

has a particular solution $X(s) = 1$, $Y(s) = 2$ so that the formula for all stabilizing controllers reads

$$R(s) = \frac{1 + \frac{s-1}{s+1}W(s)}{1-\frac{1}{s+1}W(s)},$$

where W is a parameter that ranges over the set of proper stable rational functions.

The set of stabilizing controllers clearly contains controllers of any finite order. If only PI controllers are of interest, one puts

$$R(s) = \frac{k_P s + k_I}{s}.$$

These controllers correspond to the parameter

$$W(s) = \frac{(k_P - 1)s^2 + (k_P - k_I - 1)s + k_I}{s^2 + k_P s + k_I}.$$

Consequently, $k_P > 0$ and $k_I > 0$ in order for W to be proper and stable.

Theorem 1 can be applied to both continuous-time and discrete-time controllers. Accordingly, a rational function is defined to be stable if it is analytic either in $\operatorname{Re} s \geq 0$ or in $|z| \geq 1$.

In the case of discrete-time systems, additional constraints have to be imposed: the transfer functions S and R are proper (so that the plant and the controller are causal systems) and one of them is strictly proper (so that the closed loop system is causal). The chronology of samples in the control system is usually taken in such a way that R is to be strictly proper. Selecting a particular solution pair X, Y of the Bézout equation such that Y is strictly proper, and constraining the parameter W to be strictly proper and stable will achieve this requirement. Incidentally, no distinction need be made between asymptotic and internal stability in discrete-time systems – impulsive modes do not exist.

3. Control system design

The most important property of the parametrization is that all transfer functions in an internally stable control system are *affine* functions of the parameter W. In contrast, the controller R appears in a nonlinear fashion:

$$\begin{bmatrix} v \\ y \end{bmatrix} = \frac{1}{1+SR}\begin{bmatrix} 1 & R \\ S & SR \end{bmatrix}\begin{bmatrix} d \\ r \end{bmatrix} = \begin{bmatrix} A(X-BW) & A(Y+AW) \\ B(X-BW) & B(Y+AW) \end{bmatrix}\begin{bmatrix} d \\ r \end{bmatrix}$$

As R and W are in a one-to-one correspondence, it is convenient to use W in the design process and return to R subsequently. Thus the parametrization of all stabilizing controllers makes it possible to separate the design process into two steps: the determination of all stabilizing controllers and the selection of the parameter that achieves the remaining design specifications. The extra benefit is that both tasks are *linear*.

3.1. Asymptotic properties

Asymptotic properties of control systems can easily be accommodated in the sequential design procedure. These include the elimination of an offset due to step references, the ability of system output to follow a class of reference signals, or the asymptotic elimination of specific disturbances.

The design procedure is best illustrated by an example. Given a plant with transfer function

$$S(s) = \frac{1}{s-1}$$

find an internally stabilizing controller that asymptotically eliminates harmonic disturbances

$$d(s) = \frac{\alpha s + \beta}{s^2 + 100}$$

as well as the offset due to step references

$$r(s) = \frac{\gamma}{s}$$

at the plant output y. Here α and β are arbitrary constants that parametrize the amplitude and phase of the family of all harmonic signals that have frequency 10. Similarly γ serves to describe the class of step references with arbitrary magnitude.

The first step is to determine the set of all internally stabilizing controllers. Referring to the previous example,

$$R(s) = \frac{1 + \dfrac{s-1}{s+1}\,W(s)}{1 - \dfrac{1}{s+1}\,W(s)},$$

where W is a proper stable rational parameter function. The next step is to accommodate the specifications by constraining the parameter. When $r = 0$, the output y equals

$$y(s)=\frac{1}{s+1}\left(1-\frac{1}{s+1}W(s)\right)\left(\frac{s^2+100}{s^2+2s+1}\right)^{-1}\left(\frac{\alpha s+\beta}{s^2+2s+1}\right).$$

When $d = 0$, the tracking error $e = r - y$ equals

$$e(s)=\frac{s-1}{s+1}\left(1-\frac{1}{s+1}W(s)\right)\left(\frac{s}{s+1}\right)^{-1}\left(\frac{\gamma}{s+1}\right).$$

Both functions are to be proper and stable, thus the inverses must be absorbed in

$$1-\frac{1}{s+1}W(s).$$

This condition means that

$$1-\frac{1}{s+1}W(s) = \frac{s^2+100}{s^2+2s+1}W_1(s) = \frac{s}{s+1}W_2(s)$$

for some proper and stable rational functions W_1 and W_2. Taking the least common multiple, one obtains

$$1-\frac{1}{s+1}W(s) = \frac{s^3+100s}{(s+1)^3}$$

so that the simplest parameter equals

$$W(s) = \frac{3s^2+97s+1}{s^2+2s+1}.$$

The resulting controller is

$$R(s) = \frac{4s^2-97s+101}{s^2+100}.$$

3.2. Optimal control

The sequential design procedure will be further illustrated on the design of *linear-quadratic optimal* controllers. Given a plant with transfer function $S(s)$ in the form of a coprime fraction of two proper and stable rational functions, $S = B/A$. The task is to find a continuous-time controller that internally stabilizes the control system of Fig. 1 while minimizing the effect of the disturbance d on the output y in the sense of minimizing the H_2 norm of the transfer function

$$H = \frac{S}{1 + SR} \ ,$$

which is defined by

$$\|H\|^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} |H(j\omega)|^2 \, d\omega .$$

The set of all internally stabilizing controllers is described by the formula

$$R = \frac{Y + AW}{X - BW} ,$$

where $AX + BY = 1$ and W is a proper and stable rational parameter. In the internally stable control system, one has

$$H = B(X - BW) := P - QW ,$$

where P and Q are some proper stable rational functions. Consider the inner-outer factorization (Doyle, *et al.*, 1992) of Q defined as $Q = Q_i Q_0$, where Q_i has unit magnitude on the imaginary axis and Q_0 has no zeros in Re $s \geq 0$. With this factorization,

$$\|P - QW\|_2 = \|P - Q_j Q_0 W\|_2 = \left\| Q_j \left(\frac{P}{Q_j} - Q_0 W \right) \right\|_2 = \left\| \frac{P}{Q_j} - Q_0 W \right\|_2 .$$

Next P/Q_i is decomposed as

$$\frac{P}{Q_i} = \left\{ \frac{P}{Q_i} \right\}_+ + \left\{ \frac{P}{Q_i} \right\}_- ,$$

where $\{\ \cdot\ \}_+$ is analytic in Re $s \geq 0$ and $\{\ \cdot\ \}_-$ is strictly proper and analytic in Re $s \leq 0$. With this decomposition,

$$\left\| \frac{P}{Q_i} - Q_0 W \right\|_2^2 = \left\| \left\{ \frac{P}{Q_i} \right\}_- + \left\{ \frac{P}{Q_i} \right\}_+ - Q_0 W \right\|_2^2$$

$$= \left\| \left\{ \frac{P}{Q_i} \right\}_- \right\|_2^2 + \left\| \left\{ \frac{P}{Q_i} \right\}_+ - Q_0 W \right\|_2^2$$

as the cross-terms contribute nothing to the norm. The last expression is a complete square whose first part is independent of W. Hence the minimizing parameter is

$$W = \frac{\left\{ \dfrac{P}{Q_i} \right\}_+}{Q_0} W$$

and if it is indeed proper and stable, it defines the unique optimal controller. The consequent minimum norm equals

$$\min_W \|H\|_2 = \left\| \left\{ \frac{P}{Q_i} \right\}_- \right\|_2 .$$

If the minimizing W happens to be improper or unstable, then no optimal controller exists.

To illustrate, consider the following example:

$$S(s) = \frac{s-1}{s} = \left(\frac{s-1}{s+1} \right) \left(\frac{s}{s+1} \right)^{-1}$$

The Bézout equation has a particular solution $X = 2$, $Y = -1$. The class of all internally stabilizing controllers is

$$R(s) = \left(-1 - \frac{s}{s+1} W \right) \left(2 + \frac{s-1}{s+1} W \right)^{-1}$$

for an arbitrary proper stable rational W. The disturbance-to-output transfer function is

$$H(s) = 2 \frac{s-1}{s+1} - \left(\frac{s-1}{s+1} \right)^2 W .$$

The inner-outer factorization yields

$$Q_i = \left(\frac{s-1}{s+1} \right)^2 , \qquad Q_0 = 1 ,$$

and the stable-antistable decomposition is

$$\frac{P}{Q_i} = 2 \frac{s+1}{s-1} = 2 + \frac{4}{s-1} .$$

The H_2 norm of H attains minimum for $W = 2$. The corresponding optimal controller is

$$R(s) = \frac{s-1}{4} .$$

Note that R is not proper. Nevertheless the control system is internally stable: the impulsive mode of the controller cannot be excited in the closed loop. From the practical point of view, however, the control system will not perform satisfactorily at high frequencies and a suboptimal strictly proper controller may be preferable.

3.3. Robust stabilization

Generally speaking, the notion of *robustness* means that some characteristic of the feedback system holds for every plant in a set. There are three ingredients in this definition. Firstly, robustness refers to some particular characteristic of the control system, like stability, asymptotic tracking, suboptimal level of performance, or some other performance condition. Secondly, the characteristic is to hold for every plant in the set. The ultimate goal is that it holds for the actual plant. The actual plant is unknown, however, so the best one can do is to make the characteristic hold for a large enough set of plants. Finally, one fixed controller guarantees robustness. Consequently, it makes no sense to call a control system robust unless the particular characteristic and the set of plant models are specified.

The basis technique to model plant uncertainty is to model the plant as belonging to a set. Such a set can be either structured – for example, there is a finite number of uncertain parameters – or unstructured – the frequency response lies in a set in the complex plane for every frequency. The unstructured uncertainty model is more important for several reasons. Firstly, relying on the frequency response, it provides a good connection with the classical techniques and tools. Secondly, it is well suited to represent high-frequency modeling errors, which are generically present and caused by such effects as infinite-dimensional electromechanical resonance, transport delays, and diffusion processes. Finally, and most importantly, the unstructured model of uncertainty leads to a simple and useful design theory.

The unstructured set of plants is usually constructed as a neighborhood of the nominal plant, with the uncertainty represented by an additive, multiplicative, fractional, or feedback perturbation (Zhou and Doyle, 1998). The size of the neighborhood is measured by a suitable norm, most common being the H_∞ norm that is defined for any rational function analytic on the imaginary axis as

$$\|H\|_\infty = \sup_\omega |H(j\omega)| .$$

This norm has a simple control engineering interpretation. It is the distance in the complex plane from the origin to the farthest point on the Nyquist plot of the transfer function, and it appears as the peak value on the Bode magnitude plot.

This section will illustrate the design for *robust stability under unstructured norm-bounded multiplicative perturbations*. Consider a nominal plant with transfer function S and its neighborhood S_Δ defined by

$$S_\Delta(s) = [(1 + \Delta(s)M(s)]S(s) ,$$

where M is a fixed proper stable rational function and Δ is a variable proper stable rational function such that $\|\Delta\|_\infty \le 1$. The idea behind this uncertainty model is that ΔM is the normalized plant perturbation away from 1:

$$\frac{S_\Delta}{S} - 1 = \Delta M\ .$$

Hence if $\|\Delta\|_\infty \le 1$, then for all frequencies

$$\left| \frac{S_\Delta(j\omega)}{S(j\omega)} - 1 \right| \le |M(j\omega)|$$

so that $|M(j\omega)|$ provides the uncertainty profile while Δ accounts for phase uncertainty.

Now suppose that R is a controller that internally stabilizes the nominal plant S. It follows from the Nyquist diagram that

$$|S(j\omega)R(j\omega)M(j\omega)| < |1 + S(j\omega)R(j\omega)|$$

for all ω. Consequently, the controller R will internally stabilize the entire family of plants S_Δ if and only if

$$\left\| \frac{SR}{1+SR} M \right\|_\infty < 1\,.$$

This is a necessary and sufficient condition for robust stabilization of the nominal plant S.

The set of all internally stabilizing controllers for $S = B/A$ is described by the formula

$$R = \frac{Y + AW}{X - BW}\,,$$

where $AX + BY = 1$ and W is a proper and stable rational parameter. The robust stability condition then reads

$$\|P - QW\|_\infty < 1\,,$$

where $P := BYM$ and $Q := -\,BAM$ are proper stable rational functions. Any proper and stable rational W that satisfies this inequality then defines a robustly stabilizing controller R for S. In case W actually minimizes the norm one obtains the best robustly stabilizing controller.

As an example, consider a plant with the transfer function

$$S_\tau(s) = \frac{s+1}{s-1}\mathrm{e}^{-\tau s}\,,$$

where the time delay τ is known only to the extent that it lies in the interval $0 \le \tau \le 0.2$. The task is to find a controller that stabilizes the uncertain plant S_τ. The time-delay factor $e^{-\tau s}$ can be treated as a multiplicative perturbation of the nominal plant

$$S(s) = \frac{s+1}{s-1}$$

by embedding S_τ in the family

$$S_\Delta(s) = [(1 + M(s)\Delta(s)]\,S(s),$$

where Δ ranges over the set of proper and stable rational functions such that $\|\Delta\|_\infty \le 1$. To do this, M should be chosen so that the normalized perturbation satisfies

$$\left|\frac{S_\Delta(j\omega)}{S(j\omega)} - 1\right| = \left|e^{-j\omega\tau} - 1\right| \le |M(j\omega)|$$

for all ω and τ. A little time with the Bode magnitude plot shows that a suitable uncertainty profile is

$$M(s) = \frac{3s+1}{s+9}.$$

Figure 2 is the Bode magnitude plot of this M and $e^{-\tau s} - 1$ for $\tau = 0.2$, the worst value.

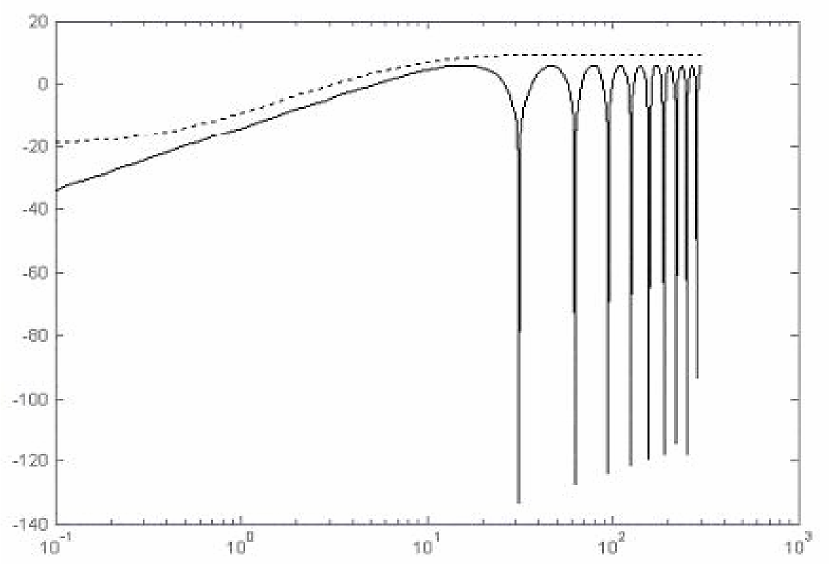


Figure 2. Bode plots of M (dotted) and $e^{-0.2s} - 1$ (solid)

The task of stabilizing the uncertain plant S_τ is thus replaced by that of stabilizing every element in the set S_Δ, that is to say, by robustly stabilizing the nominal plant S with respect to the multiplicative perturbations defined by M.

Take

$$A(s) = \frac{s-1}{s+1}, \quad B(s) = 1$$

in the fractional representation of the nominal plant S. The set of all stabilizing controllers for S is then given by

$$R(s) = -\frac{1 + \frac{s-1}{s+1} W(s)}{W(s)},$$

where W is a non-zero proper and stable rational parameter. The robust stability condition reads

$$\| P - QW \|_\infty < 1,$$

where

$$P(s) = \frac{3s+1}{s+9}, \quad Q(s) = -\frac{s-1}{s+1}\,\frac{3s+1}{s+9}.$$

Since Q has one unstable zero at $s = 1$, it follows from the maximum modulus theorem (Doyle, *et al*., 1992) that the minimum of the H_∞ norm taken over all proper and stable rational functions W is $P(1) = 2/5$ and this minimum is achieved for

$$W(s) = \frac{P(s) - P(1)}{Q(s)} = -\frac{13}{5}\,\frac{s+1}{3s+1}.$$

Thus the robust stability condition is satisfied and the corresponding best robustly stabilizing controller is

$$R(s) = \frac{2}{13}\,\frac{s+9}{s+1}.$$

3.4. Deadbeat control

The following application illustrates the design of discrete-time *deadbeat* controllers. Given a plant with discrete-time transfer function $S(z)$, written in the form of a coprime fraction of two proper and stable rational functions, $S = B/A$. The task is to determine a controller R that internally stabilizes the control system of Fig. 1 while rendering the output y to follow any reference r exactly in a minimum time. Consequently (Kučera and Kraus, 1995), the control system can have poles only at the point $z = 0$ and the reference-to-error transfer function

$$H = \frac{1}{1 + SR}$$

must be a polynomial of least degree possible in z^{-1}.

In the control system so designed, the polynomial H is given by the formula

$$H = A(X - BW) := P - QW$$

and its degree can be minimized by a choice of W: it suffices to identify W with the quotient of P/Q so that H becomes the remainder.

To illustrate, consider a discrete-time integrator plant

$$S(z) = \frac{\varepsilon z + (1-\varepsilon)}{z - 1} = \frac{\varepsilon + (1-\varepsilon)z^{-1}}{1 - z^{-1}}$$

sampled at the unit rate and displacement ε of input and output sampling instants, with $0 < \varepsilon \leq 1$. The Bézout equation admits a solution

$$X(z) = 1 + (1-\varepsilon)z^{-1}, \qquad Y(z) = z^{-1}.$$

The set of all stabilizing controllers that allocate the closed loop poles to the point $z = 0$ is

$$R(z) = \frac{z^{-1} - (1 - z^{-1})z^{-1}W}{1 + (1-\varepsilon)z^{-1} + [\varepsilon + (1-\varepsilon)z^{-1}]z^{-1}W}$$

for an arbitrary polynomial $W(z^{-1})$. The resultant transfer function from r to e is

$$H(z^{-1}) = (1 - z^{-1})[1 + (1-\varepsilon)z^{-1}] - (1 - z^{-1})[\varepsilon + (1-\varepsilon)z^{-1}]z^{-1}W.$$

Taking the quotient of the polynomial division of $1 + (1-\varepsilon)z^{-1}$ by $\varepsilon z^{-1} + (1-\varepsilon)z^{-2}$ gives the parameter $W = 0$; hence the optimal controller

$$R(z) = \frac{z^{-1}}{1 + (1-\varepsilon)z^{-1}}$$

and the polynomial

$$H(z) = (1 - z^{-1})[1 + (1-\varepsilon)z^{-1}].$$

The tracking error will vanish in three sampling periods.

3.5. Stabilization subject to input constraints

Most plants have inputs that are subject to hard limits on the range of variations that can be achieved. The effects of actuator saturation on a control system are poor performance and/or instability. Stabilization subject to input constraints can be formulated either as a local stabilization, when saturation is avoided for a set of initial states and the control system behaves as a linear one, or as a global stabilization, when saturation is allowed to occur and the control system is nonlinear.

Consider the saturation avoidance approach. Given a discrete-time plant

$$y(z) = S(z)u(z) + T(z)x_0$$

with the input

$$u(z) = u_0 + u_1 z^{-1} + u_2 z^{-2} + \ldots$$

subject to the constraints

$$-u^- \le u_k \le u^+, \quad k = 0, 1, 2, \ldots,$$

where u^+ and u^- are positive constants and x_0 is the initial state. The task is to find a controller (zero initial state assumed) of the form

$$u(z) = -R(z)y(z)$$

such that the control system is locally asymptotically stable for any initial state x_0 of the plant within a given polyhedron $P_N = \{x : Nx \le n\}$, where N is a matrix and n is a vector.

Denote $S = B/A$ and $T = C/A$ the fractional representation of the plant. The control sequence in a stable feedback system is

$$u = C(Y - AW)x_0 .$$

Taking W in the form of a power series around the point $z = \infty$

$$W(z) = p_1 z^{-1} + p_2 z^{-2} + \ldots$$

shows that the control sequence is an affine function of the parameters p_1, p_2, … of the form

$$u_k = M_k(p_1, p_2, \ldots)x_0, \quad k = 1, 2, \ldots,$$

and satisfies the given constraint if x_0 belongs to the polyhedron $P_M = \{x : M(p_1, p_2, \ldots)x \le m\}$, where

$$M(p_1, p_2, \ldots) = \begin{bmatrix} M_1(p_1, p_2, \ldots) \\ -M_1(p_1, p_2, \ldots) \\ M_2(p_1, p_2, \ldots) \\ -M_2(p_1, p_2, \ldots) \\ \vdots \end{bmatrix}, \quad m = \begin{bmatrix} u^+ \\ u^- \\ u^+ \\ u^- \\ \vdots \end{bmatrix}.$$

Now x_0 is in P_N, so that P_N must be contained in P_M. Applying the Farkas lemma (Henrion, *et al*., 2001), one concludes that the stabilization problem has a solution if and only if there exists a matrix P with non-negative entries and real numbers $p_1, p_2, \ldots$ such that

$$PN = M(p_1, p_2, \ldots), \quad Pn \le m .$$

This is a linear program for P and $p_1, p_2, \ldots$. The stabilizing controller is then obtained by putting

$$W(z) = p_1 z^{-1} + p_2 z^{-2} + \ldots .$$

If the power series W is approximated by a polynomial, then the program has a finite dimension.

To illustrate, consider the plant described by the input-output and state-output transfer functions

$$S(z) = \frac{1}{1-2z^{-1}}, \quad T(z) = \frac{2}{1-2z^{-1}} .$$

The plant input is constrained as

$$-1 \le u_k \le 1, \quad k = 1, 2, \ldots \ .$$

and the initial state x_0 belongs to the polyhedron

$$P_N : \begin{bmatrix} 1 \\ -1 \end{bmatrix} x \le \begin{bmatrix} 1/3 \\ 1/3 \end{bmatrix} \text{ (that is, } |x_0| \le 1/3).$$

The set of stabilizing controllers is found to be

$$R(z) = \frac{2z^{-1} - (1-2z^{-1})W(z)}{1+W(z)}$$

and the corresponding control sequence is

$$u(z) = [-4z^{-1} - 2(1-2z^{-1})W(z)]x_0 .$$

Now start with $W(z) = 0$ and check whether the resulting linear program for P is feasible:

$$P\begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -4 \\ 4 \end{bmatrix}, \qquad P\begin{bmatrix} 1/3 \\ 1/3 \end{bmatrix} \le \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

It is not, hence no controller of order 1 stabilizes the plant.

Proceed by choosing $W(z) = p_1 z^{-1}$ and check whether the resulting linear program for p_1 and P is feasible:

$$P\begin{bmatrix}1\\-1\end{bmatrix}=\begin{bmatrix}-4-2p_1 & 4+2p_1 & 4p_1 & -4p_1\end{bmatrix}^T,\ P\begin{bmatrix}1/3\\1/3\end{bmatrix}\le\begin{bmatrix}1 & 1 & 1 & 1\end{bmatrix}^T.$$

It is, and the solution

$$p_1=-\frac{2}{3},\ P=\frac{1}{3}\begin{bmatrix}8 & 0 & 8 & 0\\0 & 8 & 0 & 8\end{bmatrix}^T$$

furnishes the stabilizing controller

$$R(z)=z^{-1}\frac{2-2z^{-1}}{1.5+z^{-1}}.$$

The actual polyhedron of stabilizable initial states is

$$P_M:\frac{1}{3}\begin{bmatrix}-8 & 8 & -8 & 8\end{bmatrix}^T x_0\le\begin{bmatrix}1 & 1 & 1 & 1\end{bmatrix}^T \qquad \text{(that is, } |x_0|\le 3/8)$$

and it includes P_N as a proper subset.

The successive selection of a feasible parameter results in the increase of the order of the stabilizing controller. This points out to a potential weakness of the design procedure based on parametrization: each time an additional design specification is achieved, the order of the controller is increased.

4. Conclusions

The parametrization of internally stabilizing controllers can easily be extended to *multi-input multi-output* systems (Vidyasagar, 1985). Rational matrices are represented as „matrix fractions“, that is to say, as the left and right factorizations

$$S=B_P A_P^{-1}=A_L^{-1}B_L$$

of two proper and stable rational matrices, where A_P and B_P are right coprime and A_L and B_L are left coprime. The set of all internally stabilizing controllers is given by

$$R=(Y_P+A_P W)(X_P-B_P W)^{-1}=(X_L-WB_L)^{-1}(Y_L+WA_L),$$

where the proper and stable rational matrices X_L, Y_L and X_P, Y_P satisfy the Bézout identity

$$\begin{bmatrix}A_L & -B_L\\Y_L & X_L\end{bmatrix}\begin{bmatrix}X_P & B_P\\-Y_P & A_P\end{bmatrix}=I$$

and W is a proper and stable rational matrix parameter (Kučera, 1975; Youla, *et al*., 1976b).

It is interesting to note that the set of internally stabilizing controllers can be parametrized also for plants with *irrational* transfer functions. This is possible whenever such a transfer function is expressed in the form of a fraction of two *coprime* proper and stable rational functions. This property is by no means evident (Vidyasagar, 1985) and it holds, for instance, for transfer functions having a finite number of singularities in $\mathrm{Re}\, s \geq 0$, each of which is a pole.

Even more striking is the observation that internally stabilizing controllers can be parametrized for *nonlinear* plants, where transfer functions no longer exist. The key condition is again the possibility of factorizing the nonlinear mapping that defines the plant into two „coprime" mappings, one of them representing a stable system while the other one representing the inverse of a stable system (Hammer, 1985). Technical assumptions may prevent one from parametrizing the *entire* set of internally stabilizing controllers; still, the subset may be large enough for practical purposes.

The parametrization of all stabilizing controllers is a result that launched an entire new area of research and that has ultimately become a new paradigm for the design of optimal and robust control systems. Being of algebraic nature (Kučera, 1993), it is a result of high generality and elegance. The stabilizing controllers are obtained by solving a linear equation. This is not because the plant to be controlled is linear but because it is an element of the *ring of fractions* defined over the ring of stable plants (Vidyasagar, 1985). The requirement of stability is thus expressed as one of divisibility in a ring: an element a divides an element b if there exists an element x satisfying $ax = b$. That is why x is the solution of a *linear* equation.

There is a dual result: the parametrization of all plants that can be stabilized by a fixed controller. This result is useful in system identification. In fact, the (difficult) problem of closed-loop identification of the plant becomes a (simple) problem of open-loop identification of the parameter, as discussed by Anderson (1998). Consequently, the parametrization may facilitate the study of dual control.

References

Anderson, B.D.O. (1998). From Youla-Kucera to identification, adaptive and nonlinear control. Automatica, 34, 1485-1506.

Desoer, C.A., R.W. Liu, J. Murray and R. Saeks (1980). Feedback system design: The fractional representation approach to analysis and synthesis. IEEE Trans. Auto. Control, 25, 399-412.

Doyle, J.C., B.A. Francis and A.R. Tannenbaum (1992). Feedback Control Theory. Macmillan, New York.

Hammer, J. (1985). Nonlinear system stabilization and coprimeness. Int. J. Control, 44, 1349-1381.

Henrion, D., S. Tarbouriech and V. Kučera (2001). Control of linear systems subject to input constraints: a polynomial approach. Automatica, 37, 597-604.

Kučera, V. (1974). Closed-loop stability of discrete linear single variable systems. Kybernetika, 10, 146-171.

Kučera, V. (1975). Stability of discrete linear feedback systems. In: Proc. 6 th IFAC World Congress, Vol. 1, Paper 44.1. Pergamon, Oxford.

Kučera, V. (1979). Discrete Linear Control: The Polynomial Equation Approach. Wiley, Chichester.

Kučera, V. (1993). Diophantine equations in control – a survey. Automatica, 29, 1361-1375.

Kučera, V. (2002). Parametrization of stabilizing controllers. In: Proc. 8th IEEE International Conference on Methods and Models in Automation and Robotics, 87-94. Politechnika Szczecińska, Szczecin.

Kučera V. and F.J. Kraus F. J. (1995). FIFO stable control systems. Automatica, 31, 605-609.

Newton, G.C., L.A. Gould and J.F. Kaiser (1957). Analytic Design of Linear Feedback Controls. Wiley, New York.

Strejc, V. (1967). Synthese von Regelungssystemen mit Prozessrechner. Akademie-Verlag, Berlin.

Vidyasagar, M. (1985). Control System Synthesis: A Factorization Approach. MIT Press, Cambridge.

Youla, D.C., J.J. Bongiorno and H.A. Jabr (1976a). Modern Wiener-Hopf design of optimal controllers, part I: the single-input case. IEEE Trans. Auto. Control, 21, 3-14.

Youla, D.C., H.A. Jabr and J.J. Bongiorno (1976b). Modern Wiener-Hopf design of optimal controllers, part II: the multivariable case. IEEE Trans. Auto. Control, 21, 319-338.

Zhou, K. and J.C. Doyle (1998). Essentials of Robust Control. Prentice-Hall, Upper Saddle River.

METHODOLOGY FOR THE DESIGN OF FEEDBACK ACTIVE VIBRATION CONTROL SYSTEMS

Ioan Doré Landau, Aurelian Constantinescu,
Daniel Rey, Alphonse Franco
Laboratoire d'Automatique de Grenoble
ENSIEG, BP 46, 38402 Saint Martin d'Hères, France
fax: + 33 (0)4 76 82 62 44
e-mail: landau@lag.ensieg.inpg.fr

Patrice Loubat
Hutchinson - Centre de Recherche
Rue Gustave Nourry, BP 31, 45120 Châlette-sur-Loing, France

Abstract This paper presents an integrated methodology for feedback control of active vibration attenuation systems. The basic steps of the methodology are: open loop identification of the secondary path, design of a robust digital controller, identification in closed loop of a "control oriented" model, redesign of the controller based on the closed loop identified model and controller reduction. The feasability of this methodology is illustrated by its application on the Hutchinson active suspension.

Keywords: active control, active suspension, feedback control, closed loop identification, controller order reduction.

1. Introduction

Feedback is used in active vibration control mainly for three reasons:

- Absence of a measurement correlated with the vibration source (which is necessary for feedforward control).
- Wide band vibration attenuation.
- Potential robustness of performances with respect to system model variations.

A number of techniques has been proposed for the design of feedback controllers dedicated to active vibration attenuation [2, 3, 4]. Often an adaptation loop is added for the tuning of the controller because either the identified model is not enough accurate or because the designed controller is not enough robust. For more details on feedback used in active vibration control see [4]-Chapter 3.

The key contributions of the present paper are related to:

- the identification of a good "design model" by using up to date open loop and closed loop identification methods and model validation tests;
- the design of a robust controller allowing to achieve severe performance constraints in terms of the frequency attenuation characteristics;
- the use of a recent developed efficient controller reduction method preserving the desirable properties of the nominal closed loop system.

For more details on this methodology see [1].

To be specific we will start by presenting the system under consideration for the experimental verification of the methodology.

The structure of the system is presented in Fig. 1, a photo of the system being presented in Fig. 2. The controller will act upon the piston (through a power amplifier) in order to reduce the residual force. The system is controlled by a PC via an I/O card, the sampling frequency being $800Hz$.

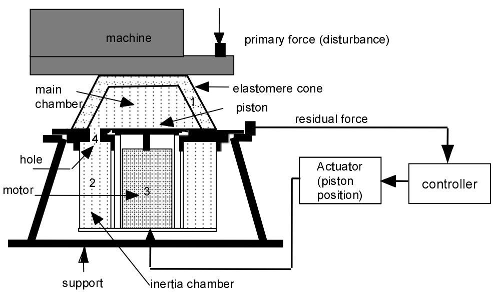


Figure 1. The active suspension system

The equivalent scheme is shown in Fig. 3.

The system input, $u(t)$ is the position of the piston, the output $y(t)$ being the residual force measured by a force sensor (see figs. 1, 3).

The principle of the active suspension is to vary the system's stiffness in order to attenuate the vibrations generated by the part that we want to isolate (primary force, disturbance). In our case, the primary force has been generated using a shaker controlled by a signal given by the computer.

Figure 2. Active suspension system (photo)

We call primary path transfer function ($q^{-d_1}\frac{C}{D}$) the transfer function between the signal sent to the shaker, p and the residual force $y(t)$. We call secondary path transfer function ($q^{-d}\frac{B}{A}$) the transfer function between the input of the system, $u(t)$ and the residual force. The input of the system being a position and the output a force, the secondary path transfer f

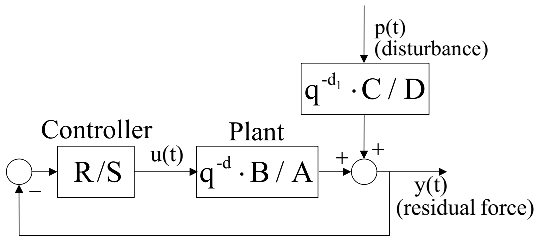


Figure 3. Block diagram of the active suspension system

The frequency characteristic of the identified primary path model (the effect of the disturbances on the output), between the excitation of the shaker and the residual force is shown in Fig. 4. The control objective is to minimize the modulus of the transfer function between the input signal of the shaker and the residual force at low frequencies, using a feedback control. In other words, to attenuate the first vibration mode, without amplifying the disturbance effect in low frequencies (below 31 Hz) and minimizing the maximum amplification of the disturbances over 35 Hz by distributing it through the high frequencies up to 200 Hz.

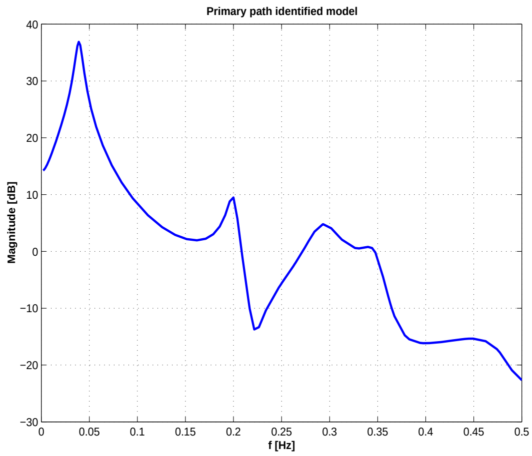


Figure 4. The frequency characteristic of the primary path model (input: shaker's signal, output: residual force)

Assuming that a good model of the system is available, the difficulty in controller design comes from the constraints in low frequencies (no amplification is allowed below $31Hz$). In our case the allowed amplification in the frequency region over $35Hz$ is $\leq 3dB$. See for example [2] for less stringent frequency specifications.

The methodology proposed for the design of feedback active vibration control is illustrated in Fig.5.

The first stage is the open loop identification and validation of a discrete time model for the secondary path (between the piston's position and the residual force).

Then a controller based on this open loop identified model is designed and implemented (open loop based controller). The pole placement with shaping of the sensitivity functions by convex optimization is used for the design [11].

Once the open loop based controller is implemented, an identification in closed loop is carried out [8]. This allows to get a better design model, since identification in closed loop (using appropriate algorithms) will enhance the precision of the estimated model in the critical frequency regions for control.

Then a re-design of the controller is done based on the closed loop identified model (nominal closed loop based controller).

The nominal controller is then implemented and tested.

The next stage is the reduction of the controller complexity, which can be done using simulated or real data. The algorithms used for controller reduction will preserve the desirable properties of the nominal closed loop system [9, 6].

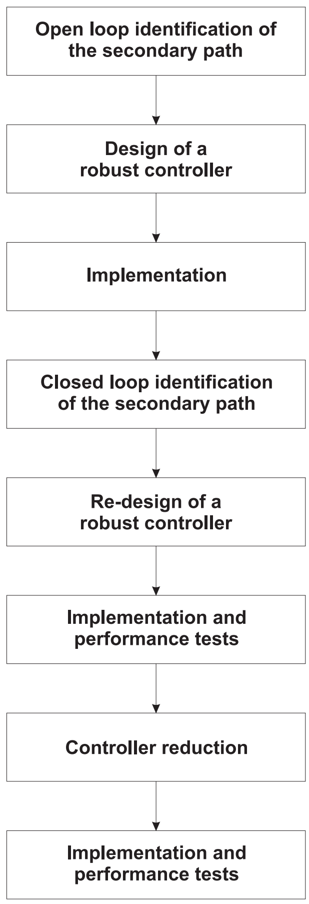


Figure 5. Design methodology

The last stage is a comparison in real time between the performances of the nominal and reduced order controllers.

The paper is organized as follows: Section 2 will discuss the model identification (in open and in closed loop). Section 3 will present the controller design methodology. Section 4 will present the controller reduction technique. Section 5 will illustrate the application of the design methodology to the Hutchinson active vibration attenuation system.

2. Open and closed loop identification

From a practical point of view, the identification of a plant model is the first thing to do for the design of a controller.

The identification of a system is an experimental approach for estimating a model of the real system. The identification procedure can be divided in four different steps:

- I/O data acquisition (under an experimental protocol).
- Estimation of the model complexity.
- Estimation of the model parameters and choice of the noise model.
- Validation of the identified model.

An important point in the identification of a system is the excitation signal. In this paper we use as excitation signal a PRBS (Pseudo Random Binary Sequence). See [10] for details.

2.1. Open loop identification

The open loop identification algorithms minimize the error between the output of the real system and the output of the estimated model. In other words, they try to estimate a model whose output fits as much as possible the part of the output of the real system generated by the excitation signal. For details on open loop identification and validation see [10].

2.2. Closed loop identification

Closed loop identification can be used when a controller (i.e. based on an open loop identified model) exists, in order to obtain a better model of the real plant since the precision of the estimated model is improved in the critical frequency regions for control. Of course, this requires to use specific algorithms.

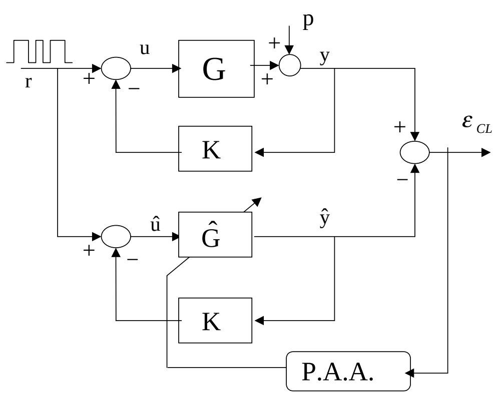


Figure 6. Identification in closed loop

The closed loop identification algorithms minimize the error between the true closed loop system and the adjustable predictor of the closed loop. The objective of the closed loop identification is to obtain a better predictor for the closed loop, using a better estimation of the plant model. The predictor of the closed loop is formed by the controller and the estimated plant model (see Fig. 6). For more details on the closed loop identification see [7, 8].

3. Controller design

The method used in this paper for the computation of a robust digital controller is the pole placement method with shaping of the sensitivity functions by convex optimization.

The combined pole placement/sensitivity function shaping method consists of placing the dominant closed loop poles, specifying some fixed parts of the controller and then adding auxiliary poles and controller parts to fulfill specifications on the output and input sensitivity functions by a convex optimization procedure [11].

3.1. Plant representation and controller structure

The structure of a linear time invariant discrete time model of the plant (on which is based the design of the controller) is

$$G(z^{-1}) = \frac{z^{-d}B(z^{-1})}{A(z^{-1})}, \tag{1}$$

where:

$$\begin{aligned} d &= \text{number of sampling periods on the plant pure time delay;} \\ A &= 1 + a_1 z^{-1} + \ldots + a_{n_A} z^{-n_A}; \\ B &= b_1 z^{-1} + \ldots + b_{n_B} z^{-n_B}. \end{aligned}$$

The controller to design is a RS-type controller (see Fig. 7). The sensi-

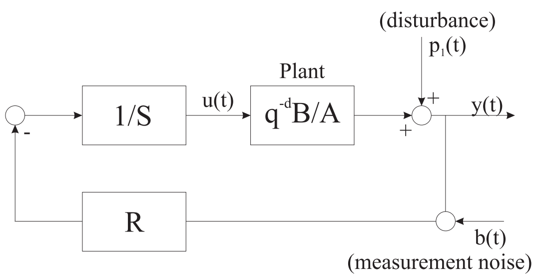


Figure 7. Structure of RS-controller

tivity functions for the closed loop are:

- the output sensitivity function (the transfer function between the perturbation $p_1(t)$ and the output $y(t)$):

$$S_{yp}(z^{-1}) = -\frac{A(z^{-1})S(z^{-1})}{P(z^{-1})}; \tag{2}$$

- the input sensitivity function (the transfer function between the input of the plant $u(t)$ and the output $y(t)$):

$$S_{up}(z^{-1}) = -\frac{A(z^{-1})R(z^{-1})}{P(z^{-1})}, \tag{3}$$

where $P(z^{-1})$ are the poles of the closed loop.

3.2. Design problem formulation

The problem may be formulated as follows: Given a nominal plant model $G_{nom} = \dfrac{z^{-d}B_{nom}}{A_{nom}}$ obtained by identification, define:

- the fixed parts of the controller (H_R and H_S);
- the desired closed loop behaviour (the dominant closed loop poles P_D and the acceptable region for the optimized ones);
- the desired upper bounds $W_x(\omega)$ for the modulus of the sensitivity functions (performance and robustness objectives);
- an objective to be minimized.

3.3. Controller parameterization

The parameterization used for the controller is the Youla-Kucera Parameterization:

$$\frac{R}{S} = \frac{H_R(R_0 + A_{nom}H_SQ)}{H_S(S_0 - z^{-d}B_{nom}H_RQ)}, \tag{4}$$

where the fixed parts of the controller, H_R, H_S and A_{nom}, B_{nom} are polynomials of z^{-1}.

The central controller ($Q = 0$) can be obtained by solving the Bezout equation for R_0 and S_0:

$$A_{nom}H_SS_0 + z^{-d}B_{nom}H_RR_0 = P_D, \tag{5}$$

where P_D is a stable polynomial containing the desired dominant closed loop poles. Expressing Q as a fraction of polynomials β and α (with α stable), we obtain:

$$\frac{R}{S} = \frac{H_R(R_0\alpha + A_{nom}H_S\beta)}{H_S(S_0\alpha - z^{-d}B_{nom}H_R\beta)} \tag{6}$$

$$P_{nom} = A_{nom}S + z^{-d}B_{nom}R = P_D\alpha, \tag{7}$$

where the zeros of P_D are the fixed closed loop poles and the zeros of α are the additional (optimized) ones.

Using the parameterization and constraint formulation presented above, a controller (R and S) with the required properties may be obtained by convex optimization. For more details on the optimization procedure see [11].

4. Controller reduction

The design of a robust controller (see Section 3) leads normally to high order controllers. There exist two main approaches to obtain a reduced order controller:

- to reduce the order of the plant model and then to compute a low order controller based on the reduced model;
- to compute a high order controller based on the nominal plant model, and then to reduce the order of the obtained controller.

The second approach seems more appropriate because the approximation is done in the final step of the controller design. In addition, the first approach in this case did not allowed to obtain enough simple controllers.

Identification in closed loop offers an efficient methodology for the controller order reduction. The most important aspect of the controller reduction is to preserve as much as possible the desirable closed loop properties.

One block diagram for reduced order controller identification is presented in Fig. 8. The simulated nominal closed loop system (the upper part of Fig. 8) is constituted by the nominal designed controller, K and the best identified plant model, $\widehat{G}$. The lower part is constituted by the estimated reduced order controller, $\widehat{K}$ and the plant model, $\widehat{G}$. It is assumed that the nominal controller stabilizes the real plant and the identified plant model.

The parametric adaptation algorithm will try find the best reduced order controller of a given order which will minimize the closed loop input error (the difference between the input of the plant model generated in the nominal simulated closed loop, u, and the input of the plant model generated by the closed loop using the reduced order controller, $\widehat{u}$).

Identification of a reduced order controller can also be done using real data [9].

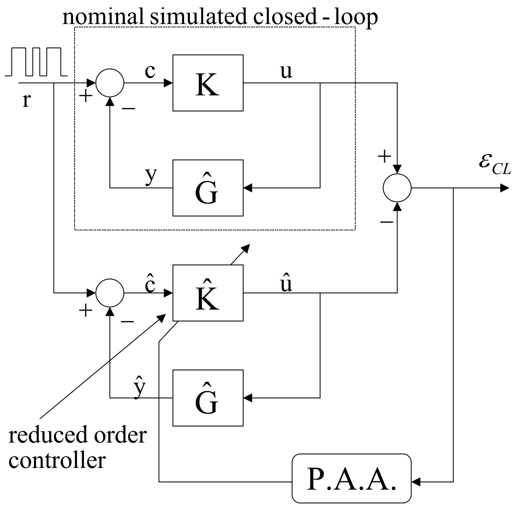


Figure 8. Closed-loop identification of reduced order controllers using simulated data (input matching: CLIM)

One can also consider as an objective for controller order reduction to minimize the closed loop error between the plant output generated in the nominal simulated closed loop and the plant output generated in the closed loop using the reduced order controller (CLOM algorithms) [5].

For more details on the algorithms see [9, 5, 6].

An important aspect for the reduction procedure, from the validation point of view, is that the reduced order controllers should stabilize the nominal model, $\widehat{G}$, and they should give sensitivity functions which are close to those obtained with the nominal controller in the critical frequency regions for performance and robustness [9].

5. Application of the designed methodology to an active suspension

The design methodology proposed in the previous sections will be illustrated on an active suspension. The active suspension has been presented in the Section 1 of this paper.

The primary path transfer function has been identified in open loop. The excitation signal sent at the input of the shaker is a PRBS with 10 cells shift register and the frequency divider $p = 2$. For the identification we used 2048 data points.

The primary path identified model has the following orders: $n_C = 12$, $n_D = 9$, delay $d_1 = 2$. The frequency characteristic of the identified model is presented in Fig. 4.

One can see that there exist several vibration modes, the first one (that we desire to attenuate) being at $31.5Hz$ with a damping factor 0.093.

We shall present below the results obtained, step by step.

5.1. Step 1: Open loop identification of a discrete model for the secondary path

The excitation signal sent at the input of the system (the piston) is a PRBS with 9 cells shift register and the frequency divider $p = 4$. For the identification we used 2048 data points.

The reason of using a frequency division $p = 4$ is that we are interested to obtain a good model in low frequencies, a PRBS with $p = 4$ having a higher energy in low frequencies. See [10] for details.

The identified model of the secondary path has the following orders: $nA = 12$, $nB = 11$, delay $d = 2$, and it has been identified using the Recursive Maximum Likelihood method with a variable forgetting factor.

The frequency characteristic of the identified model is presented in Fig. 9 (thin line), the first vibration mode being at $31.98Hz$ with a damping factor 0.078. As we can see, there are 6 vibration modes from which 5 are very low damped (< 0.078). We have chosen the best model from the point of view of the open loop validation techniques.

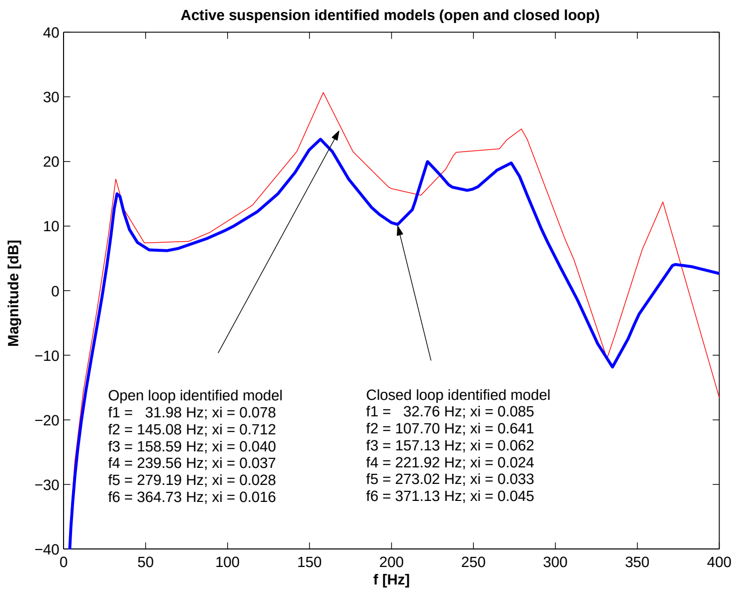


Figure 9. The frequency characteristics of the secondary path model (input: piston displacement, output: residual force)

5.2. Step 2: Design of a robust controller

Having now an open loop identified model, identified in Step 1, one can pass to the computation of a robust digital controller. We used the methodology presented in Section 3.

In order to compute the controller, we fixed a pair of dominant poles at the frequency of the first vibration mode of the open loop identified model with a damping factor 0.8. We introduced also some fixed parts in the controller (in H_R):

- $H_R = 1+q^{-1}$ (assures the opening of the loop at $0.5f_s$ - robustness reason);
- a pair of zeros at $215Hz$ with a damping factor 0.01 (assures a very low gain of the controller at the frequency where the energy of the PRBS used for identification is low ($0.25f_s$) and where therefore one has an uncertainty on the model);
- a pair of zeros at $20.5Hz$ with a damping factor 0.01 (assures the opening of the loop in low frequencies, because of the fact that attenuating the first vibration mode may other way produce amplification in low frequencies (below $31Hz$)).

The constraints on the sensitivity functions are the templates presented in Fig. 10. The template on the S_{yp} function has been established as a function of the desired disturbance rejection, the one on S_{up} is a function of the saturation problems of the actuator. The template on S_{yp} is at $-12dB$ at the frequency corresponding to the first vibration mode and at $0dB$ at the frequencies over $35Hz$ (up to $150Hz$), because we should like a very little amplification in this frequency region. We ask a little value of S_{up} at $0.25f_s$ because of the uncertainties of the identified model in this frequency region. The objective is to minimize the S_{yp} sensitivity function in the regions of interest (see the template) without sticking out from the template on S_{up}.

The resulting controller (the nominal one) has the following complexity (the orders of polynomials R and S): $n_R = 27$, $n_S = 28$. The output sensitivity function S_{yp}, respectively the input one, S_{up}, are presented in Fig. 10. From Fig. 10 one can see that we obtain a good disturbance rejection (low magnitude of S_{yp}) at the frequency corresponding to the first vibration mode ($\approx 31Hz$) and that we do not amplify at all in low frequencies. The maximum amplification over $31Hz$ is below $3dB$.

5.3. Step 3 and 4: Controller implementation and closed loop identification

Having now the nominal controller (which stabilizes also the real plant), computed at Step 2, we can proceed to the closed loop iden-

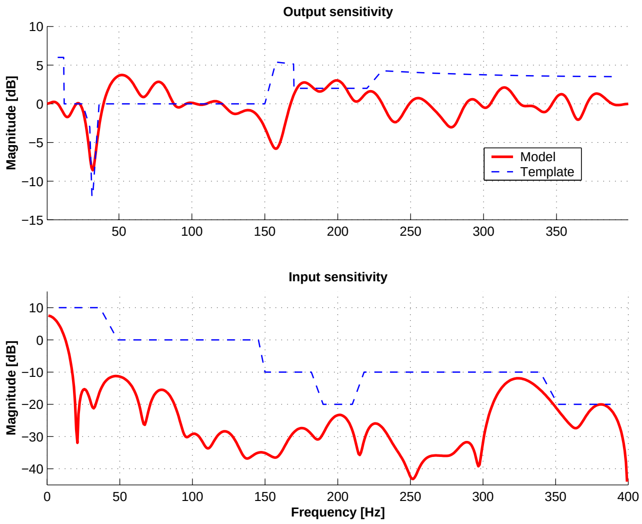


Figure 10. Sensitivity functions

Table 1. Closed loop validation results

Plant model	Open loop	Closed loop
C.L. Err. var.	1.5368	0.1477
$\|RN(i)\|_{max}$	0.4802	0.0340
Vinnicombe	0.7624	0.3587

tification of the secondary path in order to improve the quality of the open loop identified model (see Step1).

We use the same excitation signal from the open loop identification (Step 1) and we add it on the input of the system (the piston), which is now in feedback with the nominal controller (see Fig. 3)

We identify a model having the same complexity as the open loop one. We use the F-CLOE (Filtered Closed Loop Output Error) method [8].

The closed loop validation results for the open loop, respectively closed loop identified models, are presented in Table 1. The first two lines give the variance of the residual closed loop error and the maximum of the normalized cross correlations between estimated output and residual closed loop error. The third line gives the Vinnicombe distance between the identified transfer function of the real closed loop system and the closed loop transfer function of the simulated closed loop (nominal controller + model to validate). For details on closed loop validation and Vinnicombe distance see [8, 10], and [12], respectively.

One can see that the closed loop identified model validates better than the open loop one.

5.4. Step 5: Re-design of a controller

Having now a better model of the real system, we compute a robust digital controller based on this model, in order to improve the performances of the controller on the real system.

The controller based on the closed loop identified model is obtained using the same methodology and the same constraints (templates on the sensitivity functions) imposed in Step 2.

5.5. Step 6: Performance tests

In order to test the performances of the nominal controller on the real system, we present the spectral density of the residual force in open and in closed loop (see Fig. 11). One can see an attenuation of about $7dB$

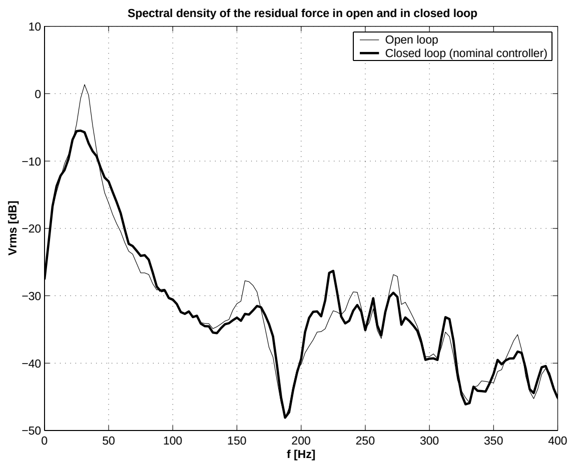


Figure 11. Spectral density of the residual force in open and in closed loop

at the frequency corresponding to the first vibration mode ($\approx 31Hz$), without any amplification in low frequencies. The tolerated amplification (3 dB) over 35 Hz has been verified, so the controller obtained accomplished the desired performances.

5.6. Step 7: Controller reduction

In order to do the order reduction of the nominal controller, we shall give the results obtained using the CLIM direct reduction method presented in Section 4, based on simulated data. The plant model used is the closed loop identified model. We use as external input a PRBS generated by 10 cells register and with a frequency divider of $p = 2$. We use 4096 data points and a variable forgetting factor. For more details, see [9, 6].

We present the reduction results by showing the sensitivity functions (S_{yp} and S_{up} in figs. 12, 13) for the nominal controller K_n with $n_R = 27, n_S = 28$ and for three reduced order controllers: K_1 with $n_R = 19$, $n_S = 20$, K_2 with $n_R = 12$, $n_S = 13$ and K_3 with $n_R = 9$, $n_S = 10$ respectively (a fixed part $H_R = 1+q^{-1}$ has been imposed in the reduced order controllers). K_3 controller has a lower complexity than the pole placement one.

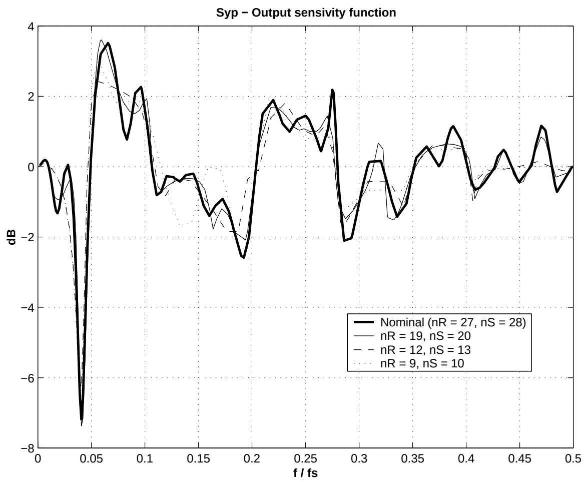


Figure 12. Output sensitivity for the active suspension (controller reduction using CLIM algorithm on simulated data)

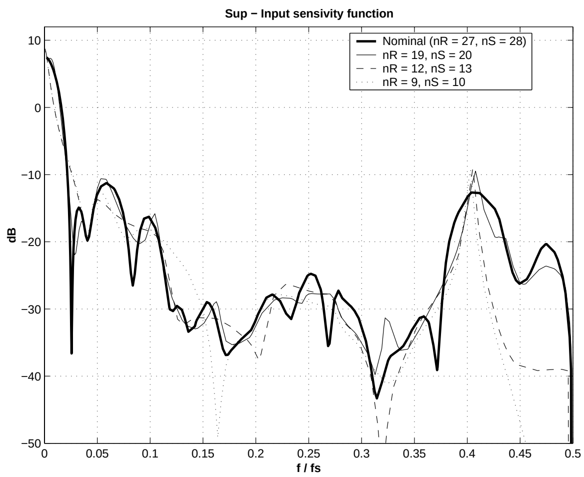


Figure 13. Input sensitivity for the active suspension (controller reduction using CLIM algorithm on simulated data)

Note: The reduced controller K_2 corresponds to the complexity of the pole placement controller.

One can see the closeness of the sensitivity functions.

Similar results are obtained by using real data for the reduction of the controller complexity.

5.7. Step 8: Performance tests

To illustrate the performances of the resulting controllers (K_n, K_1, K_2 and K_3) on the real system, the spectral density of the residual acceleration in open and in closed loop is shown in Fig. 14. The spectral densities obtained in closed loop operation are compared to those corresponding to the open loop operation.

The performances of the reduced order controllers are close to that of the nominal controller and all achieve a significant reduction of the vibrations around the first vibration mode of the plant model.

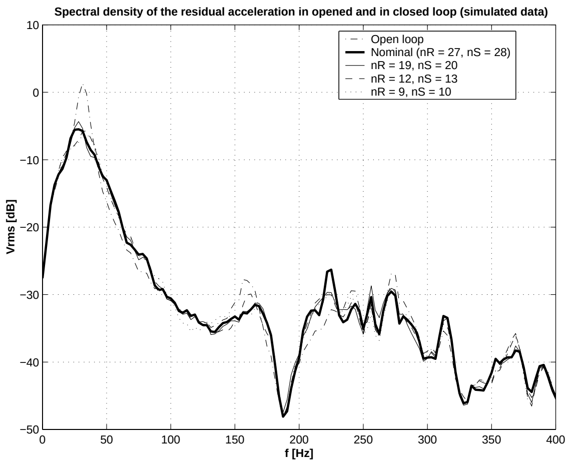


Figure 14. Spectral density of the residual acceleration in open and in closed loop (controller reduction using CLIM algorithm on simulated data)

6. Conclusions

The methodology presented in this paper allowed to design a feedback control for an active vibration system and has been successfully tested in real time.

The real system has been identified in open and in closed loop. The resulting controller accomplished the desired performances. The nominal controller has been simplified by a procedure preserving the closed loop properties. The reduced order controller (19 parameters instead of 55) gives very close results to those of the nominal controller.

References

[1] Constantinescu, A. (2001). Commande robuste et adaptative d'une suspension active. Thèse de doctorat. Institut National Polytechnique de Grenoble.

[2] Curtis, A. R. D. (1997). A methodology for the design of feedback control systems for the active control of sound and vibration. *Active 97* pp. 851–860. Budapest - Hungary.

[3] Friot, E. (1997). Optimal feedback control of a radiating plate under broadband acoustic loading. *Active 97* pp. 873–884. Budapest - Hungary.

[4] Fuller, C. R., S. J. Elliott and P. A. Nelson (n.d.). *Active control of vibration*. Academic Press. London.

[5] Karimi, A. and I. D. Landau (2000). Controller order reduction by direct closed loop identification (output matching). *Proceedings IFAC Symp. ROCOND 2000, Prague.*

[6] Landau, I. D., A. Karimi and A. Constantinescu (2001). Direct controller order reduction by identification in closed loop. *Automatica* (37), 1689–1702.

[7] Landau, I. D. and A. Karimi (1997*a*). An output error recursive algorithm for unbiased identification in closed loop. *Automatica* **33**(5), 933–938.

[8] Landau, I. D. and A. Karimi (1997*b*). Recursive algorithms for identification in closed loop: a unified approach and evaluation. *Automatica* **33**(8), 1499–1523.

[9] Landau, I. D. and A. Karimi (2000). Direct closed loop identification of reduced order controllers (input matching). *Proceedings of IFAC Symp. SYSID 2000, S. Barbara.*

[10] Landau, I. D., R. Lozano and M. M'Saad (1997). *Adaptive control*. Springer. London.

[11] Langer, J. and I. D. Landau (1999). Combined pole placement/sensitivity function shaping method using convex optimization criteria. *Automatica* **35**(6), 1111–1120.

[12] Vinnicombe, G. (1993). Frequency domain uncertainty and the graph topology. *IEEE Trans. on Automatic Control* **38**, 1371–1383.

FUTURE TRENDS IN MODEL PREDICTIVE CONTROL

Corneliu Lazar

"Gh. Asachi" Technical University of Iaşi
Department of Automatic Control and Industrial Informatics
Bd. D. Mangeron nr. 53A, 700050 Iaşi, Romania
E-mail:clazar@ac.tuiasi.ro

Abstract Model predictive control is a control technique in which a finite horizon optimal control problem is solved at each sampling instant to obtain the control input. The measured state is used as initial state and only the first control of the calculated optimal sequence of controls is applied to the plant. A key advantage of this form of control consists in its ability to cope with complex systems and hard constraints on controls and states. This resulted in a wide range of applications in industry, most of them in the petro-chemical branch. In this survey, a selected history of model predictive control is presented, with the purpose to outline the principles of this control methodology and to analyze the progress that has been made. The *initial* predictive control algorithms, mainly based on input/output models, are recalled in the introduction and then we focus on the more recent work done in nonlinear model and hybrid model predictive control. The stability problem and the computational aspects are discussed to formulate some fruitful ideas for the future research.

Keywords: model predictive control, discrete-time systems, nonlinear systems, hybrid systems, constraints, stability

1. Introduction

Model Predictive Control (MPC) is a control strategy that offers attractive solutions, already successfully implemented in industry, for the regulation of constrained linear or nonlinear systems and, more recently, also for the regulation of hybrid systems. Within a relatively short time, MPC has reached a certain maturity due to the continuously increasing interest shown for this distinctive part of control theory, and this has been illustrated by the prolific literature on this subject, including (Garcia *et al*., 1989; Qin and Badgwell, 1997; Allgöwer *et al*., 1999; Camacho and Bordons, 1999; Morari and Lee, 1999; Mayne *et al*., 2000; Rawlings, 2000; Maciejowski, 2002).

The reason for the rapid development of MPC algorithms mainly consists in the *intuitive* way of addressing the control problem. In comparison with conventional control, which uses a pre-computed control law, predictive control is built around the following *key principles*: the explicit use of a process model for calculating the future behavior of the plant, the optimization of an objective function subject to constraints (which yields an optimal control sequence) and the receding horizon control strategy. The MPC methodology involves solving on-line an open-loop optimal control problem subject to input, state and/or output constraints. The graphical interpretation of this concept is depicted in Fig. 1.

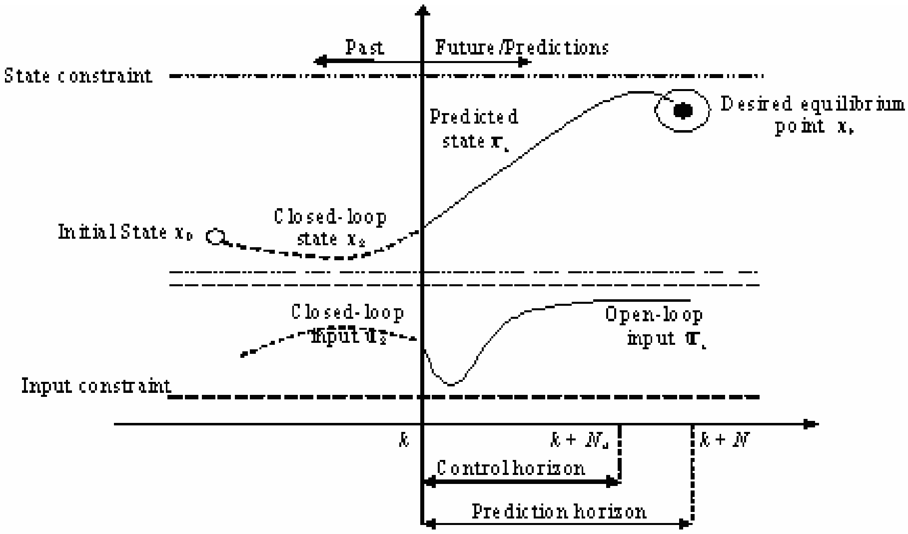


Figure 1. Graphical interpretation of model predictive control

At each sampling instant *k*, the measured variables and the *process model* (linear, nonlinear or hybrid) are used to predict the future behavior of the controlled plant over a specified discrete-time horizon called *prediction horizon (N)*. This is achieved considering a future *control scenario* as the input sequence applied to the process model, which must be calculated such that certain desired (imposed) objectives are fulfilled. To do that, a *cost function* is minimized subject to constraints, yielding an optimal control sequence over a discrete-time horizon named *control horizon* (N_u). Note that $N_u \leq N$ and if the control horizon is strictly smaller than the prediction horizon, the control input will be kept constant after N_u sampling time instants. According to the receding horizon control strategy, only the first element of the computed optimal control sequence is then applied to the plant.

The original MPC algorithms, addressing linear systems exclusively, utilized only input/output models. In this framework, several solutions have been proposed both in the industrial world, IDCOM – Identification and Command (later MAC – Model Algorithmic Control) at ADERSA (Richalet *et al*., 1978) and DMC – Dynamic Matrix Control at Shell (Cutler and Ramaker, 1980), which use step and impulse response models, and in the academic world (the adaptive control branch) MUSMAR – Multistep multivariable adaptive regulator (Mosca *et al*., 1984), predictor-based self tuning control (Peterka, 1984), EHAC – Extended Horizon Adaptive Control (Ydstie, 1984), EPSAC – Extend Predictive Self-Adaptive Control (De Keyser and Van Cauwenberghe, 1985). Other MPC algorithms were also developed later on, from which the most significant ones are GPC – Generalized Predictive Control (Clarke *et al*., 1987) and UPC – Unified Predictive Control (Soeterboek, 1992).

Next, the MPC algorithms have been designed for state-space models and extensions to nonlinear models followed shortly. In the framework of Nonlinear Model Predictive Control (NMPC), several alternatives have been studied and implemented with good results, such as dual-mode NMPC (Michalska and Mayne, 1993), quasi-infinite horizon NMPC (Chen and Allgöwer, 1996; Chen and Allgöwer, 1998) or contractive NMPC (de Oliveira Kothare and Morari, 2000). Also, a more recent stabilizing NMPC algorithm has been presented in (Magni *et al*., 2001).

The first MPC approach to the control of hybrid systems has been reported quite recently in (Bemporad and Morari, 1999). Since then, several MPC *schemes* have been proposed for particular relevant classes of hybrid systems, such as the ones in (Bemporad *et al.*, 2000; De Schutter and Van den Boom, 2001; De Schutter *et al*., 2002; Mayne and Rakovic, 2002; Lazar and Heemels, 2003).

By now, the linear MPC theory is quite mature. Important issues such as on line computations, the interplay between modeling-identification and control, and system theory subjects like stability and robustness are well defined. In the sequel, we will only focus on the NMPC and the MPC framework for hybrid systems and the associated problems.

2. NMPC - Basic concepts and problem formulation

When nonlinear systems (models) are employed, despite the slightly different problem formulation, the basic concepts are still the key principles of predictive control. The use of a nonlinear model only complicates the finite horizon optimal control problem that has to be solved on-line.

Consider the general discrete-time nonlinear systems described by the difference equation:

$$x_{k+1} = f(x_k, u_k) \tag{1}$$

subject to input and state constraints of the form:

$$u_k \in U,\ \forall k \geq 0 \quad x_k \in X,\ \forall k \geq 0, \tag{2}$$

where $x_k \in \Re^n$ is the state vector, $u_k \in \Re^m$ is the control input vector and $f(.) = [f_1(.), f_2(.), \ldots , f_n(.)]^T$ is a vector containing smooth nonlinear functions of their arguments, which are zero at zero. In the simplest form, the sets U and X are defined by:

$$U = \{u_k \in \Re^m \mid u_{\min} \leq u_k \leq u_{\max}\} \quad X = \{x_k \in \Re^n \mid x_{\min} \leq x_k \leq x_{\max}\}, \tag{3}$$

where u_{min}, u_{max}, x_{min}, x_{max} are given constant vectors.

The control objective is to regulate the state x to a desired equilibrium point x_r. As any equilibrium point x_r can be reduced to the origin via a suitable change of coordinates, we consider for the rest of the paper that the goal is to steer system (1) to the origin, while fulfilling the constraints (2).

The predictive control approach to the above stated control problem leads to the minimization of the cost function:

$$J(x_k, \mathbf{u}_k) = \sum_{i=0}^{N-1} F\left(x_{i|k}, u_{i|k}\right) + L\left(x_{N|k}\right) \tag{4}$$

subject to:

$$\begin{aligned} &x_{k+1|k} = f(x_{k|k}, u_{k|k}), \quad x_{k|k} = x_k, \\ &u_{i|k} \in U,\ \ i \in [0, N_u], \qquad u_{i|k} = u_{N_u},\ i \in [N_u, N], \\ &x_{i|k} \in X,\ \ i \in [0, N], \end{aligned} \tag{5}$$

where $\mathbf{u}_k = [u_{0|k}\ u_{1|k}\ \ldots\ u_{N-1|k}]$ are the manipulated controls, *F(.)* is the stage cost and *L(.)* is a suitable terminal state penalty term. The stage cost specifies the desired control performances and it is usually a quadratic function in x and u:

$$F\left(x_i, u_i\right) = x_i^T Q x_i + u_i^T R u_i \tag{6}$$

with Q and R positive definite and symmetric weighting matrices.

Note that, for simplicity, it is assumed that the prediction horizon N is equal to the control horizon N_u. Thus, the NMPC control problem can be stated as:

Problem 1 Solve:

$$\min_{\mathbf{u}_k}\{J(x_k,\mathbf{u}_k)\,|\,u_{i/k}\in U,\ i\in[0,N\text{-}1];\ x_{i/k}\in X,\ \ i\in[0,N];\ x_{0/k}=x_k\},\qquad(7)$$

which yields the optimal control sequence

$$\overline{\mathbf{u}}_k=[\overline{u}_{0/k}\ \overline{u}_{1/k}\ \ldots\ \overline{u}_{N-1/k}\,],\qquad(8)$$

and apply to plant (1) only the first element of (8).

The initial condition from (5) shows that the system model used to predict the future in the controller is initialized by the actual system state. In general, at each time instant *k* the full state x_k is assumed to be measured or must be estimated. Model-plant mismatch and disturbances are not represented in the optimization problem.

In general it would be desirable to use an infinite prediction horizon, i.e. $N=\infty$, and to minimize the cost function (4) with $L = 0$, in order to achieve stability of the closed-loop system. However, the open-loop optimal control problem that must be solved on-line is often formulated using a finite prediction horizon, resulting thus in a finite parameterized problem which allows a (real-time) numerical solution of the nonlinear programming Problem 1. It is obvious that the shorter the prediction horizon is, the less time consuming are the calculations involved, so it is advantageous from a computational point of view to implement MPC schemes using short horizons.

Still, the problem now consists in the fact that the actual closed-loop input and state trajectories will differ from the open-loop trajectories, even if no model mismatch and no disturbances are present. Moreover, it is by no means true that a repeated minimization of a finite horizon cost function in a receding horizon manner leads to an optimal solution also for the infinite horizon problem with the same stage cost *F* (Bitmead *et al*., 1990). In fact the two solutions differ significantly if a short horizon is utilized, which implies that there is no guarantee that the NMPC closed-loop system will be stable. Hence, when using finite horizons in standard NMPC, the employed cost function cannot be simply developed from the desired physical objectives.

2.1. NMPC algorithms with guaranteed stability

The most perceptive way to achieve stability, when an NMPC algorithm is utilized to calculate the control law, is to choose an infinite prediction horizon. In this case, with the state available for measurement, no model mismatch and no disturbances it follows directly from Bellmann's principle of optimality that the open-loop input and state trajectories calculated at a specific time instant *k* as a solution of the NMPC Problem 1, are in fact identical with the closed-loop trajectories of the nonlinear plant. This implies closed-loop stability because any feasible predicted trajectory goes to the origin.

Since nearly all stability proofs for NMPC follow along the same basic steps as for the infinite horizon proof, at this point, it is worth mentioning the key ideas: considering the cost function as a Lyapunov function, it is *first* shown that feasibility at one time instance does imply feasibility at the next time instance for the nominal case, in a *second* step, it is established that the cost function is strictly decreasing and thus the state and the input are converging to the origin, and finally, in a *third* step, asymptotic stability is established using the continuity of the cost function at the origin and its monotonicity properties.

Unfortunately, an infinite horizon for the NMPC Problem 1 is only useful as a theoretical concept, because the solution of such a high dimension optimization problem is extremely difficult, if not impossible to obtain. Due to this reason, finite horizon approaches are preferred for NMPC, despite the inconsistencies between the open-loop *predicted* trajectories and the closed-loop *actual* trajectories mentioned above. Instead of using an infinite prediction horizon, stability is achieved / guaranteed by adding suitable constraints (not connected with physical restrictions or desired performance/ requirements) and penalty terms to the *original cost function* (4). Therefore, these extra conditions are referred as *stability constraints*. In the following, two representative finite horizons NMPC schemes with guaranteed stability are presented.

Dual-mode NMPC This NMPC approach was introduced in (Michalska and Mayne, 1993) and consists in the use of two different controllers that are applied in different regions of the state space depending on the state being inside or outside of some terminal region that contains the origin. For the case in which the state is outside the terminal region an NMPC controller with a variable finite horizon is applied and, when the current state has entered the terminal region, a linear state feedback control law is employed. Thus, the proposed NMPC algorithm utilizes the following twofold control strategy:

$$u_k = \begin{cases} \overline{u}_{0/k}, & \textit{if } x_k \notin \Omega \\ Kx_k, & \textit{if } x_k \in \Omega\,. \end{cases} \tag{9}$$

The terminal region Ω and the state feedback are calculated off-line such that the terminal region is a positive invariant region of attraction for the nonlinear system controlled by the linear state feedback algorithm and thus, the input and state constraints are satisfied with this linear controller in Ω. According to the dual-mode approach, when the state is outside Ω, the length N of the horizon is considered as an additional minimizer and Problem 1 to be solved, becomes:

$$\min_{u_k, N} \left\{ J(x_k, u_k, N) \mid u_{i/k} \in U,\ i \in [0,\, N-1];\, x_{i/k} \in X,\, i \in [0,\, N];\, x_{0/k} = x_k \right\} \tag{10}$$

with the additional terminal inequality constraint

$$x_N \in \Omega_B\,, \tag{11}$$

which ensures that at the end of the horizon, the state has to lie on the boundary Ω_B of the terminal region Ω.

Starting from the outside of terminal region, the dual-mode NMPC controller guarantees the reaching of the terminal region boundary in a finite time. The close-loop stability is attained due to the use of a stabilizing local linear feedback control law. From the computational point of view, the dual-mode NMPC solution is more attractive because a inequality constraint is used, rather than a terminal equality constraint. The main drawback consists in the requirement to switch between control strategies and in determining the terminal region Ω.

Quasi-infinite horizon NMPC The quasi-infinite horizon NMPC strategy was presented in (Chen and Allgöwer, 1996) and then further developed in (Chen and Allgöwer, 1998), where the inequality stability constraint

$$x_N \in \Omega \tag{12}$$

and the quadratic terminal penalty term

$$L(x_N) = x_N^T P x_N \tag{13}$$

have been added to the standard NMPC Problem 1. The authors started from an infinite horizon cost function described by

$$J_\infty(x_k, \mathbf{u}_k^\infty) = \sum_{i=0}^{\infty} F(x_i, u_i), \tag{14}$$

where $\mathbf{u}_k^\infty$ is an infinite length control sequence. Splitting (14) in two parts, Problem 1 can be recast as:

$$\min_{\mathbf{u}_k^\infty} J_\infty(x_k, \mathbf{u}_k^\infty) = \min_{\mathbf{u}_k^\infty} \left\{ \sum_{i=0}^{N-1} F(x_i, u_i) + \sum_{i=N}^{\infty} F(x_i, u_i) \right\}. \tag{15}$$

The basic idea is that the final cost L is not a performance specification that can be chosen freely, but rather that the P matrix must be pre-computed such that the penalty term (13) is a good approximation of the second term in (15) (the infinite stage cost). Unfortunately, this is not usually feasible for general nonlinear systems, without introducing further restrictions. In particular, if the case is that the trajectories of the closed-loop system remain within some neighborhood of the origin (Ω) from the time instant $k + N$ towards infinity, then, an upper bound on the second term of (15) exists. The terminal region Ω is built such that a local state feedback law similar to the one employed in dual-mode NMPC asymptotically stabilizes the nonlinear system in Ω. Moreover, it is shown in (Chen and Allgöwer, 1998) that if the terminal region Ω and the terminal penalty matrix P is chosen according to

Procedure 1 of (Chen and Allgöwer, 1998), then

$$\sum_{i=N}^{\infty} F(x_i, Kx_i) \le x_N^T P x_N \ , \qquad (16)$$

holds and the following equality is obtained

$$\min_{\mathbf{u}_k^{\infty}} J_{\infty}(x_k, \mathbf{u}_k^{\infty}) = \min_{\mathbf{u}_k^N} J(x_k, \mathbf{u}_k), \qquad (17)$$

with J being the cost function as in the standard NMPC Problem 1. This implies that the optimal value of the finite horizon optimization problem bounds the value of the corresponding infinite horizon optimization problem, and thus asymptotic stability of the closed-loop system can be achieved, irrespectively of the control performance specifications.

Although the idea of using a local state feedback and a terminal inequality constraint are inspired by dual-mode NMPC, the main advantage of quasi-infinite NMPC comes from the fact that the control law is calculated solving the same NMPC problem, not depending on the state being inside or outside the terminal region, so no switching is involved in this case.

3. MPC for hybrid systems

A general model of hybrid systems leads to an extremely high complexity approach for the synthesis, analysis and computation of the controller. So it is necessary to focus on particular subclasses of hybrid systems that allow efficient computational methods for MPC and capture a wide range of industrially relevant processes. In the following, we present two subclasses to which MPC has already been applied successfully.

3.1. MPC of mixed logical dynamical systems

Mixed logical dynamical (MLD) systems are a subclass of hybrid systems described by interacting physical laws, logical rules and operating constraints. For this class of systems there are both continuous and binary inputs, states, outputs and auxiliary variables. MLD systems include a wide set of models, among which linear hybrid systems, finite state machines, some classes of discrete event systems, systems with discrete or qualitative inputs, constrained linear systems, bilinear systems, piecewise linear output functions, nonlinear dynamic systems where the nonlinearity can be expressed through combinational logic. The main motivation for the MLD framework is that in many applications the system to be controlled should be described by means of differential equations as well as by means of logic (i.e. due to on/off switches, gears or valves). In the MLD setting the logic is converted into integer variable linear inequalities to make the model mathematically tractable. This leads to a description of MLD systems with linear dynamic equations subject to linear mixed-integer inequalities (i.e. inequalities involving continuous and logical variables).

The MLD systems are described by equations of the form (Bemporad and Morari, 1999):

$$\begin{aligned} x_{k+1} &= Ax_k + B_1 u_k + B_2 \delta_k + B_3 z_k \\ y_k &= Cx_k + D_1 u_k + D_2 \delta_k + D_3 z_k \\ E_2 \delta_k + E_3 z_k &\le E_1 u_k + E_4 x_k + E_5 \ , \end{aligned} \tag{18}$$

where $x_k = [x_r^T(k)\ x_b^T(k)]^T$ is the state vector with $x_r(k) \in \Re^{n_r}$ denoting the continuous part of the state and with $x_b(k) \in \{0,1\}^{n_b}$ denoting the logical (discrete) part of the state, and k is sampling time instant. The output y_k and the input u_k have a similar structure as the state vector, and $z_k \in \Re^{r_r}$, $\delta_k \in \{0,1\}^{r_b}$ are auxiliary variables. MPC proved to be a successful tool for stabilizing MLD systems to a desired reference point, or for solving the (reference trajectory) tracking problem. The first MPC algorithm formulated for MLD systems (Bemporad and Morari, 1999) performs, at each sampling time instant k, the following operations:

Solve:

$$\begin{aligned} \min_{\{v_k)\}} J(v_k, x_k) = \sum_{i=1}^{N} \left\| x_{k+i|k} - x_e \right\|_{Q_1}^2 + \left\| u_{k+i-1|k} - u_e \right\|_R^2 + \\ + \left\| z_{k+i-1|k} - z_e \right\|_{Q_2}^2 + \left\| \delta_{k+i-1|k} - \delta_e \right\|_{Q_3}^2 + \\ + \left\| y_{k+i-1|k} - y_e \right\|_{Q_4}^2 \end{aligned} \tag{19}$$

subject to:

$$\begin{aligned} x_{k+1} &= Ax_k + B_1 u_k + B_2 \delta_k + B_3 z_k \\ y_k &= Cx_k + D_1 u_k + D_2 \delta_k + D_3 z_k \end{aligned} \tag{20.1}$$

$$E_2 \delta_k + E_3 z_k \le E_1 u_k + E_4 x_k + E_5 \tag{20.2}$$

$$x_{k+N|k} = x_e \ . \tag{20.3}$$

Apply to plant (18) only the first element of the optimal control sequence $v_k^* = [u_{k|k}^*, u_{k+1|k}^*, \ldots, u_{k+N-1|k}^*]$, namely $u_{k|k}^*$, accordingly to the receding horizon control strategy.

In the above equations, the standard MPC notation has been used, where N is the prediction horizon and $x_{k+i|k}$ represents the value of the state vector after *i - steps* ahead in the future, calculated at the sampling time instant *k* using the MLD model (18), the measured state x_k and the corresponding inputs from the control sequence v_k. Also, in equation (19), the quintuple $(x_e, u_e, z_e, \delta_e, y_e)$ is an equilibrium quintuple for the MLD system (18) and $\|.\|_Q$ denotes the Euclidean norm weighted by matrix Q. Note that the optimization of the cost function (19) must be fulfilled subject to the feasibility constraints (20.1) (which ensure that feasibility at one sampling instant implies feasibility at the next time instant), the constraints (20.2), imposed by the nature of the MLD system, and the terminal state equality constraint (20.3) that guarantees stability. Provided that the matrices used in the objective function (19) are positive definite and that the constrained minimization problem is initially feasible, it has been proven in (Bemporad and Morari, 1999) that the MPC control law (19)-(20) stabilizes the closed loop system (18)-(19)-(20). The potentialities of the method and its impact in process control have been demonstrated through simulation case studies on a Kawasaki Steel gas supply system in (Bemporad and Morari, 1999).

However, this MPC approach has a drawback, which consists of a *mixed integer quadratic programming* problem (MIQP), *NP* hard, which must be solved on-line at each sampling time instant subject to constraints. Although there are several algorithms for solving the MIQP problem, such as *cutting plane methods*, *decomposition methods*, *logic-based methods* and *branch and bound methods*, a high computational effort is required, which restricts this control scheme to slow processes. Also, condition (20.3) will not be satisfied for any *N* so if the prediction horizon is chosen too small, a solution to the constrained optimization problem may no longer exist, hence feasibility is lost. These problems were addressed in (Bemporad *et al*., 2000), where infinity norms have been used in (19) instead of 2-norms and an explicit solution has been obtained for the MPC problem for MLD systems. This can be achieved by reformulating the *original* MPC problem as a *multiparametric mixed integer linear program* (mp-MILP) and obtaining the explicit piecewise linear control law off-line. The equality constraint (20.3) has been removed, and a terminal cost has been added to the cost function (19), such that stability is guaranteed irrespective of the length of the prediction horizon. Thus, the computational effort is reduced and feasibility is increased.

However, due to the fact that the explicit piecewise linear control law calculated off-line consists in a set of state feedbacks, which are used in certain regions of the state space, a finite number of linear inequalities has to checked at each sampling time to determine in which region the current state resides. Hence, the *explicit* solution still requires some on-line computations, which increase in complexity with the length of the prediction horizon (the number of control regions increases).

3.2. MPC of Piecewise Affine (PWA) systems

Another relevant class of hybrid systems is the class of PWA systems, described by equations of the form (Sontag 1981):

$$x(k+1) = A_j x(k) + B_j u(k) + f_j \quad \text{for} \quad \begin{bmatrix} x(k) \\ u(k) \end{bmatrix} \in \Omega_j, \tag{21}$$

where $A_j \in \Re^{nxn}$, $B_j \in \Re^{nxm}$, $f_j \in \Re^n$, $\forall j \in S = \{1, 2, ..., s\}$, $k \geq 0$ denotes the discrete-time instant and $\{\Omega_j \mid j \in S\}$ is a finite set of polytopes (i.e. compact and convex polyhedrons) with mutually disjoint interiors. The PWA systems are particularly important due to the fact that equivalences exist between piecewise affine systems and several other relevant classes of hybrid systems (Heemels *et al*., 2001).

Recently an optimal control solution and an optimal control and receding horizon control solution have been presented in (Kerrigan and Mayne, 2002) for constrained PWA systems with bounded disturbances. The optimal control is determined in this case by comparing the solutions of a finite number of multiparametric linear programming problems, instead of solving on-line a multiparametric mixed integer linear programming problem. Although this approach might not be realizable for large or complex systems (due to the computations required for calculating the robust controllable sets), the controllable sets theory can be used for studying the feasibility of the MPC problem for MLD systems (i.e. a lower bound $\tilde{N}$ on the prediction horizon can be estimated such that feasibility is guaranteed for any $N \geq \tilde{N}$). Another MPC algorithm for piecewise affine systems, that uses reverse transformation, has been presented in (Mayne and Rakovic, 2002).

An approach for reducing the on-line computational load encountered in MPC algorithms for piecewise affine systems has been presented in (Lazar and Heemels, 2003). This method is based on an algorithm that solves off-line the controllability problem with respect to an invariant target set. The algorithm calculates a minimum of discrete events controllable path to the target set and organizes the resulting state space regions (sets) in a *tree-like* structure. For an initial state (or a measured state), a controllable path to the target set with a minimal number discrete events is easily obtained and a resulting ordered sequence of state space regions is pre-computed; each region corresponds to a single sub-model, part of the piecewise affine system. Then, it has been shown that under suitable assumptions the minimal discrete events controllable path can be used to develop a semi-explicit (sub-optimal) computationally more friendly MPC algorithm for piecewise affine systems.

4. Conclusions

The research on model predictive control has now reached a relatively mature stage. This rapidly evolving control methodology has proved to be a successful solution for the control of industrial applications where hard constraints are presents. The research in the academic world has been focused on the stability and the robustness of model predictive control. These problems have already been thoroughly investigated for linear and nonlinear systems, leading to a (all most) complete framework. Recently, model predictive control has also been extended to some relevant classes of hybrid systems and hybrid MPC tends to become as trendy as nonlinear MPC.

References

Allgöwer, F., T.A. Badgwell, S.J. Qin, J.B. Rawlings and S.J. Wright, (1999). Nonlinear predictive control and moving horizon estimation – An introducing overview. In *Advances in Control: Highlights of ECC'99*, (P.M. Frank, Ed. London: Springer), pp. 391-449.

Bemporad, A. and M. Morari (1999). Control of systems integrating logic, dynamics, and constraints, *Automatica*, Vol. 35, pp. 407-427.

Bemporad, A., F. Borrelli and M. Morari, (2000). Optimal controllers for hybrid systems: Stability and piecewise linear explicit form. In *Proceedings of the 39th IEEE Conference on Decision and Control*, Sydney, Australia.

Bitmead, R.R., M. Gevers, and V. Wertz, (1990). *Adaptive optimal control – The thinking man's GPC*, Englewood Cliffs, NJ: Prentice-Hall.

Camacho, E.F. and C. Bordons, (1999). *Model Predictive Control*, Springer-Verlag, London.

Chen, H. and F. Allgöwer, (1996). A quasi-infinite horizon predictive control scheme for constrained nonlinear systems. *In Proc. 16th Chinese Control Conference*, pp. 309-316, Qindao.

Chen, H. and F. Allgöwer, (1998). Nonlinear model predictive control schemes with guaranteed stability. In C. Berber, R. und Kravaris, editor, *Nonlinear Model Based Process Control*, pp.465-4494, Kluwer Academic Publishers, Dodrecht.

Clarke, D.W., C. Mohtadi and P.S. Tuffs, (1987). Generalized predictive control: I – the basic algorithm and II – Extensions and interpretations, *Automatica*, Vol. 23, pp. 137-160.

Cutler, C.R. and B.L. Ramaker, (1980). Dynamic matrix control – A computer control algorithm, *Proc. of Joint Automatic Control Conference*, San Francisco, USA.

De Keyser, R.M.C., and A.R. Van Cauwenberghe, (1985). Extended prediction self-adaptive control, *IFAC Symp. on Identification and System Parameter Estimation*, York, U.K., pp. 1255-1260.

De Oliveira Kothare, S.L. and M. Morari., (2000). Contractive Model Predictive Control for Constrained Nonlinear Systems, *IEEE Transactions on Automatic Control*, Vol. 45(6), pp. 1053-1071.

De Schutter, B. and T.J.J. van den Boom, (2001). Model predictive control for max-plus-linear discrete events systems, *Automatica*, Vol. 37(7), pp.1049-1056.

De Schutter, B., T.J.J. van den Boom and G.J. Benschop, (2002). MPC for continuous piecewise-affine systems. In *Proceedings of the 15th IFAC World Congress*, Barcelona, Spain.

Garcia, C.E., D.M. Prett and M. Morari, (1989). Model Predictive Control: Theory and Practice – a Survey, *Automatica*, Vol. 25(3), pp. 1753-1758.

Heemels, W.P.M.H., B. De Schutter and A. Bemporad, (2001). Equivalence of hybrid dynamical models, *Automatica*, Vol. 37, pp. 1085-1091.

Kerrigan, E.C. and D.Q. Mayne, (2002). Optimal control of constrained, piecewise affine systems with bounded disturbances. In *Proceedings of the 41st Conference on Decision and Control*, Las Vegas.

Lazar, M. and W.P.M.H. Heemels, (2003). A semi-explicit MPC set-up for constrained piecewise affine systems. In *Proceedings of the European Control Conference*, Cambridge.

Maciejowski, J.M., (2002). *Predictive Control with Constraints*. Prentice Hall. Harlow, England.

Magni, L., G. De Nicolao, L. Magnani and R. Scattolini, (2001). A stabilizing model-based predictive control algorithm for nonlinear systems, *Automatica*, Vol. 37(9), pp. 1351–1362.

Mayne, D.Q., J.B. Rawlings, C.V. Rao and P.O.M. Scokaert, (2000). Constrained model predictive control, *Automatica*, Vol. 36, pp. 789-814.

Mayne, D.Q. and S. Rakovic, (2002). Optimal control of constrained piecewise affine discrete time systems using reverse transformation. In *Proceedings of the 41st Conference on Decision and Control*, Las Vegas.

Michalska, H. and D. Q. Mayne, (1993). Robust receding horizon control of constrained nonlinear systems, *IEEE Transactions on Automatic Control*, AC-Vol. 38(11), pp. 1623-1633.

Morari, M. and J.H. Lee, (1999). Model predictive control: Past, present and future. *Computers and Chemical Engineering*, 23, 667-682.

Mosca, E., G. Zappa and C. Manfredi, (1984). Multistep horizon self-tuning controllers. In *Proceedings of the 9th IFAC World Congress*, Budapest.

Qin, S.J. and T.A. Badgwell, (1997). An overview of industrial model predictive control technology. In *Chemical Process Control-AIChE Symposium Series* (J. Kantor, C. Garcia and B. Carnahan, Eds. New York: AIChE), pp. 232-256.

Peterka, V., (1984). Predictor-based self tuning control. *Automatica*, Vol. 20, pp. 39-50.

Rawlings, J.B., (2000). Tutorial Overview of Model Predictive Control, *IEEE Control Systems Magazine,* Vol. 20(3), pp. 38-52.

Richalet, J.A., A. Rault, J.L. Testud and J. Papon, (1978). Model predictive heuristic control: applications to an industrial process, *Automatica*, Vol. 14, pp.413-428.

Soeterboek, R., (1992). *Predictive Control - A unified approach*, Englewood Cliffs, NJ:

Prentice-Hall.

Sontag, E.D., (1981). Nonlinear regulation: the piecewise linear approach, *IEEE Transactions on Automatic Control*, Vol. 26(2), pp. 346-357.

Ydstie, B.E., (1984). Extended horizon adaptive control. *Proc. of the 9th IFAC World Congress,* Budapest, Hungary.

BLOCKING PHENOMENA ANALYSIS FOR DISCRETE EVENT SYSTEMS WITH FAILURES AND/OR PREVENTIVE MAINTENANCE SCHEDULES*

Jose Mireles Jr.

Instituto de Ingeniería y Tecnología de la
Universidad Autónoma de Ciudad Juárez
Ave. del Charro 450 Nte., Cd. Juárez, Chih. MÉXICO
E-mail: jmireles@arri.uta.edu, jmireles@uacj.mx

Frank L. Lewis

Automation & Robotics Research Institute
University of Texas at Arlington (UTA)
7300 Jack Newell Blvd. S., Fort Worth, TX 76118-7115, USA
E-mails: flewis@controls.uta.edu

Abstract We present an analysis of possible blocking phenomena, deadlock, in Discrete Event Systems (DES) having corrective and/or Preventive Maintenance Schedules (PMS). Although deadlock avoidance analysis for several classes of DES systems has been widely published, and although different approaches for PMS exist, it is not obvious how to mix deadlock avoidance and maintenance theories to improve throughput. In this paper we show that for some DES structures having reentrant flow lines, it is not necessary to stop activities in the DES, for the case one or more machines in production lines are in PMS. However, PMS may cause deadlock to occur if activities continue in some machines. We propose deadlock-free dispatching rules derived by performing circular wait analysis for possible deadlock situations in systems with PMS. This is accomplished by integrating the PMS structure and failure dynamics into

* Research supported by ARO Grants DAAD19-00-1-0037, NSF-CONACyT DMI-0219195, and PROMET from Mexico

a separate DES system that acts as a disturbance in the primary Reentrant Flow-line DES system. We propose a matrix formulation and a Finite State Machine to synchronize both subsystems.

Keywords: deadlock avoidance, Petri nets, discrete event systems, reentrant flow lines, maintenance

1. Introduction

In this paper we address the problem of avoiding possible deadlock situations on Flexible Manufacturing Systems or Discrete Event Systems (DES) having shared resources in Reentrant Flow-lines [*Kumar 93*], with scheduled maintenance jobs. It is no doubt Preventive Maintenance (PM) is a vital activity for improving machines availability in DES. This improving of availability is due to the decreased number of corrective maintenance jobs in machines, which lead to a much more costly production times. PM methods, like the Reliability-Centered Maintenance method has been used for years, and is still a recommended approach [*Smith 1992*]. Recent studies have proven advantages of using PM techniques. For example, [*Hicks 1990*] has shown improvements in cost-reduction in different Army sites in the state of Texas. In Hicks' work, recommendations are given to keep improving PM schedules. One important recommendation is the search for automated expert systems for optimal use of machines in systems with PM schedules. In this paper, we present one expert system with PM schedules based on matrices that avoid blocking phenomena in reentrant flow-lines. If DES contains Multipart Reentrant flow-lines (MRF), i.e. shared resources perform more than one job for same product, in a system producing several products, and if it is possible not to stop processes, even if one or more machines are in PM, then blocking phenomena can occur if jobs are not correctly sequenced in the remaining non-in-maintenance resources. This blocking phenomenon is known as system ***deadlock*** [*Banaszak et al. 90, Hsieh et al. 94, Ezpeleta et al. 95, Fanti et al. 97, Lewis et al. 98*]. Therefore, it is very important that the Discrete Event (DE) controller, after knowing which resources are in PM or corrective maintenance, properly sequences jobs and assigns available resources.

In this paper we restrict our analysis to systems lacking ***key resources*** [*Gurel et al. 00*]. These key resources are critical structured resources that might lead to possible *Second Level Deadlock* (SLD) [*Fanti et al. 00*]. Systems lacking SLD are called regular. In [*Mireles et al. 02*], we provide a matrix tests for system regularity. Based on the decision-making matrix formulation introduced in [*Lewis 92-93*], this paper presents the development of a deadlock-free ***augmented*** discrete event controller for regular MRF systems with failures and PMS. This augmented controller contains a framework capable of handling failures and maintenance-capabilities in the DES structure. We describe the DE controller (DEC) formulation, and show how to analyze and compute in matrix notation the structures needed for deadlock-free dispatching algorithms. Based on these matrix constructions, we integrate PM systems' information for deadlock-free dispatching rules in our augmented DEC matrix formulation by limiting the work-in-progress (WIP) in some critical subsystems, which we define later. This is accomplished by integrating a Finite State Automata system composed of the primary Reentrant Flow-line DES system, and the disturbance-acting PMS structure containing failure dynamics.

2. Matrix-based discrete event controller

A novel Discrete Event Controller (DEC) for manufacturing workcells was described in [*Lewis et al. 93, Mireles et al. 01a-b*]. This DEC is based on matrices, and it was shown to have important advantages in design, flexibility and computer simulation. The definition of the variables of the Discrete Event Controller is as follows. Let *v* be the set of tasks or jobs used in the system, *r* the set of resources that implement/perform the tasks, *u* the set of inputs or parts entering the DES. The DEC Model State Equation is described as

$$\bar{x} = F_v \otimes \bar{v} \oplus F_r \otimes \bar{r} \oplus F_u \otimes \bar{u} \oplus F_{uc} \otimes \bar{u}_C \ , \tag{1}$$

where: $\bar{x}$ is the task or state logical vector, F_v is the job sequencing matrix, F_r is the resource requirements matrix, F_u is the input matrix, F_{uc} is the conflict resolution matrix, and u_c is a conflict resolution vector.

This DEC equation is performed in the AND/OR algebra. That is, multiplication $\otimes$ represents logical "AND," addition $\oplus$ represents logical "OR," and the over-bar means logical negation. From the model state equation, the following four interpretations are obtained. The job sequencing matrix $\boldsymbol{F_v}$ reflects the states to be launched based on the current finished jobs. It is the matrix used by [*Steward 81*] and others and can be written down from the manufacturing Bill of Materials. The resource requirement matrix $\boldsymbol{F_r}$ represents the set of resources needed to fire possible job states this is the matrix used by [*Kusiak et al. 92*]. The input matrix $\boldsymbol{F_u}$ determines initial states fired from the input parts. The conflict resolution matrix $\boldsymbol{F_{uc}}$ prioritizes states launched from the external dispatching input u_C, which has to be derived via some decision making algorithm [*Graves 81*]. The importance of this equation is that it incorporates matrices F_v and F_r, previously used in heuristic manufacturing systems analysis, into a rigorous mathematical framework for DE system computation.

For a complete DEC formulation, one must introduce additional matrices, S_r and S_v, as described next. The state logic obtained from the state equation is used to calculate the jobs to be fired (or task commands), to release resources, and to inform about the final products produced by the system. These three important features are obtained by using the three equations:

Start Equation (task commands) $v_S = S_V \otimes x$ (2)

Resource Release Equation $r_S = S_r \otimes x$ (3)

Product Output Equation $y = S_y \otimes x$. (4)

3. Matrix analysis of MRF systems

In these sections we present a technique for deadlock-free dispatching for MRF systems with maintenance schedules, and show how to implement some notions from other papers using matrices. First, we integrate PM systems in MRF structures using our matrix approach, and then, we determine the deadlock constructions needed for free dispatching. This yields computationally efficient algorithms for analyzing the structure of MRF and deadlock-free dispatching.

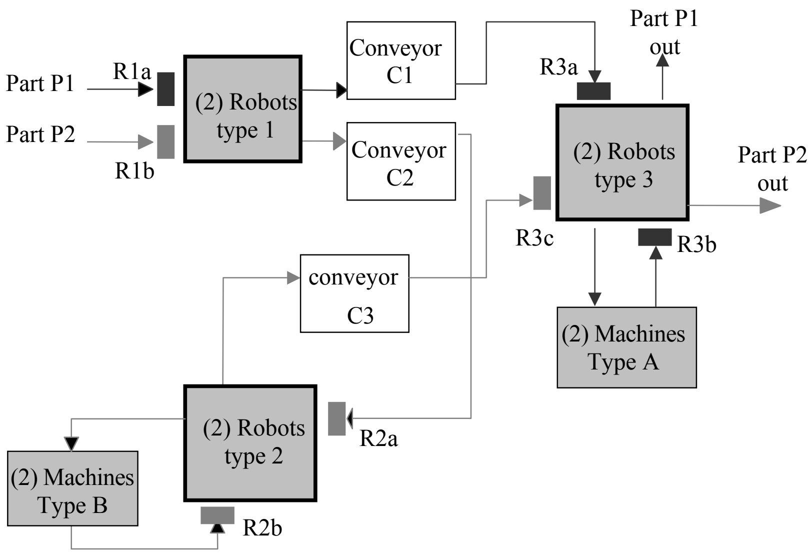


Figure 1. Multipart Reentrant Flow Line Problem

Consider the following definition of Multiple Reentrant Flow-lines, which define the sort of discrete-part manufacturing systems that can be described by a Petri net. The characteristics of MRF systems are:

- No preemption. A resource cannot be removed from a job until this job is completed.
- Mutual exclusion. A single resource can be used for only one job at a time.

- Hold while waiting. A process holds the resources already allocated to it until it has all resources required to perform a job.
- For the DE systems we consider in our analysis, the following are their particularities:
- Each job uses only one resource.
- After each resource executes one job, it is released immediately for its availability.
- In this paper we also consider handling scheduled preventive maintenance, as well as machine failures.

An example of a class of MRF system is given next. Consider the Multipart Reentrant Flow-line problem shown in Figure 1. This system uses two types of machining resources and three types of robotic resources, machine types A and B, and robots type 1, 2 and 3. Any of the (two) robotic resources type 1 moves incoming parts P1 and P2 to conveyors C1 and C2 respectively. Any of the (two) robotic resources type 2 can accomplish two jobs, jobs R2a and R2b. Job type R2a moves part type P2 from conveyor C2 to buffer of (any of the two) machines type B. Job type R2b moves machined part type P2 from (any of the two) machines type B to conveyor C3. Any of the (two) robotic resources type 3 can accomplish three jobs, jobs R3a, R3b, and R3c. Job type R3a moves part type P1 from conveyor C1 to buffer of (any of the two) machines type A. Job type R3b moves machined part type P1 from (any of the two) machines type A to parts out P1. Job type R3c moves machined part type P2 from conveyor C3 to parts out P2.

In this example, for simplicity, we are assuming buffer sizes on conveyors and machines equal to one. This assumption will help us emphasize possible deadlock situations when resources are been in failure or scheduled for maintenance. Also, if we consider larger buffers, we will reach a practical point where the buffer might be full and so our same deadlock situation will appear.

3.1. Failure/Maintenance DES structure

In this section we present an extension of the matrix framework presented in section 2 to incorporate DES systems with Failure and/or PMS. When human operators proceed to fix failures in machines/resources or proceed to perform a preventive maintenance, their jobs can be seen as specific jobs holding such machines/resources. The problem is that holding such resources being in Failure or PMS can lead to system deadlock. Therefore, in order to be able to control a DES with failures and/or maintenance schedules, one has to consider that each of such machine/robotic resources is in one of three possible states: In-Service state, Failure state, or in PM state. Then, for each resource in a PN representation, has to illustrate the Failure and PM states, as in the PN addition system in figure 2. We call this PN system the Failure-

Maintenance (*FM*) system. In this figure, the places and transitions highlighted as "In-Service Status" belong to the FMR system, where t_x and t_y represent transitions $\bullet Job_{ij}$ and $Job_{ij}\bullet$, for the j number of jobs from resource R_i. Notice that transition t_{fij} fires when a failure occurs in resource R_i (for i=1,2…n=number of resources) while performing operation Job_{ij}, after finishing this repair job, t_{frij} should be fired (in the PN from figure 2, this can be easily ensured by adding a virtual place between each t_{fij} and t_{frij} transition pairs). Transition t_{mi} will fire when a preventive maintenance M_{pi} for resource R_i is requested. When a transition t_{fij} fires, a failure repair job, F_{repi}, is requested for execution. Maintenance times for jobs type M_{pi} are deterministic times. However, repair time jobs, type F_{repi}, are stochastic and not deterministic, and usually F_{repi} job times are larger than M_{pi} job times. Note that in order to improve throughout, transitions t_{fij} are preferred over all others. However, transition t_{mi} is not always an 'urgent' transition to fire due to a scheduled PM, by presence of a new token in place M_{anti}. This is, the supervisor can decide whether it is more important to finish pending jobs, or proceed to maintenance of corresponding resource R_i.

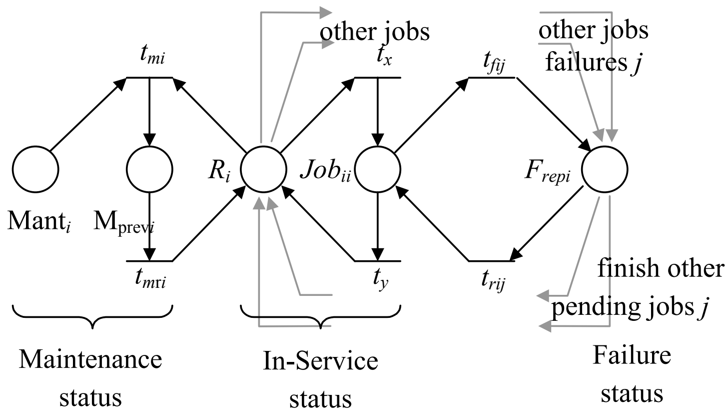


Figure 2. Corrective, In-Service and Preventive status of FM

The definition of the Failure-Maintenance system structure follows. Since the structure discussed in section 2 is now augmented by the addition of corrective and PMS, the *FM* structures, we need to re-define the formulation from section 2. For this, we need to include jobs type F_{repi} and M_{previ} (the repair and the maintenance jobs, respectively), and the control transitions that activate these jobs for every type of resources R_i. We include these sets in our now augmented matrix form. We integrate these FM structures by incorporating in matrices F and S the transitions and places shown in figure 3. This figure shows black and gray dots, representing ones and zeros in the rows & columns shown. To properly maintain FM structures, we supervise the maintenance-integrated system, and keep track of job markings that belong to this system. That is, the number of tokens in the *FM* addition system for each resource plus the number of tokens in the job set of same resource is always constant and equal to the initial marking in that resource (assuming no maintenance is in schedule at the time the initial marking is calculated.)

Resource places are he only places shared between PM structures and the original PN (with no PMs). Notice also that for any of these two options, the 'travel' of tokens between one system to the other is through resource places *R*. Unless, of course, if a failure happens at the moment a machine is performing a job, a token passes from that job to failure status job place F_{repi} (by firing corresponding t_{fij}). For this case, we consider the part was not finished, and stays in standby as a damaged part or for to be re-machined. Then, when failure happens, t_{fij} is fired with high priority and start maintenance failure job type F_{repi}.

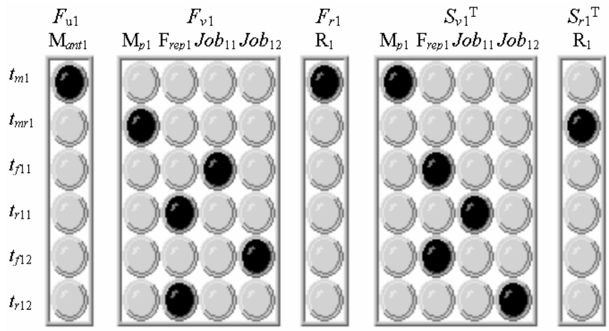


Figure 3. F_u, F_v, F_r, S_v^T, S_r^T matrices for resource R_1

This separation of systems MRF and FM is practical for the following reasons:

1) Since *FM* system does not have resource loops and does not generate extra resource loops if exist any in the general existing system, this facilitates deadlock analysis on the MRF system without worrying about dynamics on *FM* systems.
2) It is possible to maintain and control an independent *FM* subsystem with its appropriate PMS, and the existing general system by properly handling the marking vectors from both systems. It is clear that at any given time, the total number of tokens in a job set from a specific resource set, plus the available set of resources from that set is maintained equal to the initial marking of that resource set. This total number of tokens is diminished by one, for every job been in maintenance, i.e. been in its corresponding *FM* system's job set. Then, by maintaining for each resource this number of tokens equal always to the sum of tokens from both systems, it is possible to maintain control of the MRF and *FM* systems.

Figure 2 shows the FM Petri net system structure that has to be added for each resource in the FMR system to supervise preventive and corrective jobs. Figure 3 shows the matrix representation section representing only the FM system of resource R_1. For the class of MRF systems we are considering including *FM*, deadlock can occur only if there is a ***circular wait relation*** among resources [*Deitel 84, Gurel et.al 00*]. Circular wait relations are ubiquitous in reentrant flow-lines and in themselves do not present a problem. However, if a circular wait relation develops into circular blocking, then one has deadlock. But, as long as dispatching is carefully performed, the existence of circular wait relations presents no problem for ***regular systems*** [*Gurel et.al 00*]. In this paper we restrict our analysis to regular systems. This systems lack ***key resources***. These key resources are critical structured resources that might lead to possible *Second Level Deadlock* (SLD) [*Fanti et al. 00*] situations in MRF systems. In [*Mireles et al. 02a-b*], we provide a matrix tests for system regularity.

3.2. Circular waits: simple circular waits and their unions

In this section we present a matrix procedure to identify all circular waits (CW) in MRF systems. CWs are special wait relationships among resources described as follows. Given a set of resources R, for any two resources $r_{i,}$ r_j $\subset R$, r_i is said to wait for r_j, denoted $r_i \rightarrow r_j$, if the availability of r_i is an immediate requirement to release r_j, or equivalently, if there exists at least one transition $x \in \bullet r_i \cap r_j \bullet$. Circular waits among resources are a set of resources r_a, r_b,…r_w, which wait relationships among them are $r_a \rightarrow r_b \rightarrow \ldots \rightarrow r_w$, and $r_w \rightarrow r_a$. The simple Circular Waits (sCW), are primitive CWs which do not contain other CWs. If sCW are present in the PN system structure, these are identified by constructing a ***digraph*** of resources. [*Hyenbo 95*] demonstrated a technique to identify such sCW. We used his approach to construct digraphs in matrix form. The entire sets of CWs are the sCW plus the circular waits composed of unions of non-disjoint sCW (unions through shared resources among sCW.)

In [*Mireles et al. 01*], we obtained two matrices, C_{out} and G, using digraph theory and string algebra. C_{out} provides the set of resources which compose every CW (in rows), that is, an entry of 'one' on every (*i*,*j*) position means that resource *j* is included in the i^{th} CW. G provides the set of composed CWs (rows) from unions of sCW (columns), that is, an entry of 'one' on every (*i*,*j*) position means that j^{th} sCW is included in the i^{th} composed CW.

3.3. Deadlock analysis: identifying critical siphons and critical subsystems

Three important sets associated with the CWs C are the ***siphon-job*** sets $J_s(C)$, the ***critical siphons***, $S_c(C)$, and ***critical subsystems***, $J_o(C)$. The critical siphon of a CW is the smallest siphon containing the CW. Note that if the critical siphon ever becomes empty, the CW can never again receive any tokens. This is, the CW has become a circular blocking. The siphon-job set, $J_s(C)$, is the set of jobs which, when added to the set of resources contained in CW C, yields the critical siphon. The critical siphons of that CW C are the conjunction of sets $J_s(C)$ and C. The critical subsystems of the CW C, are the ***job sets*** J(C) from that C not contained in the siphon-job set $J_s(C)$ of C. That is $J_o(C) = J(C) \setminus J_s(C)$. The job sets of CW C are defined by $J(C) = \cup_{r \in C} J(r)$, for $J(r) = r^{\bullet\bullet} \cap J$, where J is the set of all jobs.

In order to implement efficient real-time control of the DES, we need to compute these sets in matrix form. We need intermediate quantities ${}^{\bullet}C$ and $C^{\bullet}$, ***input*** and ***output*** transitions from *C*, and which in matrix form for each CW are denoted ${}_{d}C$ and C_d respectively, computed as,

$$ {}_{d}C = C_{out} S_r, \text{ and} \tag{5}$$

$$ C_d = C_{out} F_r^T . \tag{6}$$

In terms of these constructions, matrix form sets are described next, indicating 'one' on every entry (*i,j*) for places that belong to that set existing in every i^{th} CW. The job sets described earlier for each CW C, J(C), in matrix form (for all CWs arranged in rows) are described by

$$ J_C = {}_{d}C F_v = C_d \; S_v^T . \tag{7}$$

The ***siphon-job*** sets are defined for each i^{th} CW C_i as $J_s(C_i) := J(C_i) \cap ({}^{\bullet}C \setminus C^{\bullet})$. In matrix notation, we can obtain them for all CWs by

$$ J_s = J_C \wedge (\overline{C_d F_v}) . \tag{8}$$

The ***critical subsystems***, $J_o(C_i) = J(C_i) \setminus J_s(C_i)$, in matrix form for all CWs C_i are obtained by

$$ J_o = J_C \wedge (C_d \, Fv) . \tag{9}$$

4. Deadlock avoidance

In terms of the constructions just given, we now present a minimally restrictive resource dispatching policy that guaranties absence of deadlock for multi-part reentrant flow lines. To efficiently implement in real time a DE controller with this dispatching policy we use matrices for all computations. We consider the case where the ***system is regular***, that is, it cannot contain the Critical Resources (CR) (so-called structured bottleneck resources or 'key resources' [*Gurel et al. 00*] existing in Second Level

Deadlock (SLD) structures [*Fanti et al. 97, 00*].) For this case, we described in [*Mireles et al. 02*], a mathematical test to verify that MRF systems are regular. If that is not the case, we can still use this matrix formulation, but with a different dispatching policy designed for systems containing second level deadlock structures. We will present such dispatching policy for FMRF systems having CR in a forthcoming work.

4.1. Dispatching policy

In this section we consider dispatching for regular systems. In [*Lewis et al. 98*] was given a minimally restrictive dispatching policy for regular systems that avoids deadlock for the class of MRF systems considered in this paper, but without the failures or PMS. To understand this policy, note that, for this class of systems, a deadlock is equivalent to a circular blocking (CB). There is a CB if and only if there is an empty circular wait (CW). However, CB is possible (for regular systems) if and only if (iff) the corresponding critical siphon from any CW is empty. This is, there is a deadlock iff all tokens of the CW are in the Critical Subsystem.

Therefore, the key to deadlock avoidance is to ensure that the WIP in the Critical Subsystems is limited to one less job than the total number of initial tokens in the CW (i.e. the total number of resources available in the CW). Preliminary off-line computations using matrices are used to compute the Critical Systems. A supervisor is assigned to each Critical Subsystem (CS) who is responsible for *dynamic dispatching* by counting the jobs in that CS and ensuring that they do not violate the following condition, for each CW C_i,

$$m(J_o(C_i)) < m_o(C_i). \tag{10}$$

That is, the number of enabled places contained in the CS for each C_i must not reach the total number of resources contained in that C_i. In (10), $m_o(C_i)$, is the initial marking of C_i,. However, having failures and PM jobs, the total number of available resources will be diminished. So that $m_o(C_i)$ does not represent anymore the actual available resources contained for that C_i. To be able to keep track of such available resources, we need to define the total number of job places from systems *FM* corresponding to resources contained in a CW C_i, by $J_{MF}(C_i)$. Then, if we diminish $m_o(C_i)$ by jobs currently in failure and/or PM in $J_{MF}(C_i)$, our CB supervision test (10), we will be able to ensure actual available resources which will ensure deadlock-free dispatching. This is, our new CB supervision test is

$$m(J_o(C_i)) < \{m_o(C_i) - J_{MF}(C_i)\}. \tag{11}$$

A graphical example of using (11) is pictured in Figure 4. This system has two circular waits, C_1={M1, R_3}, and C2={M2, R_2}. This system contains five *FM* systems split as separate subsystems. Notice that initial $m_o(C_i)=4$ for $i=1,2$.

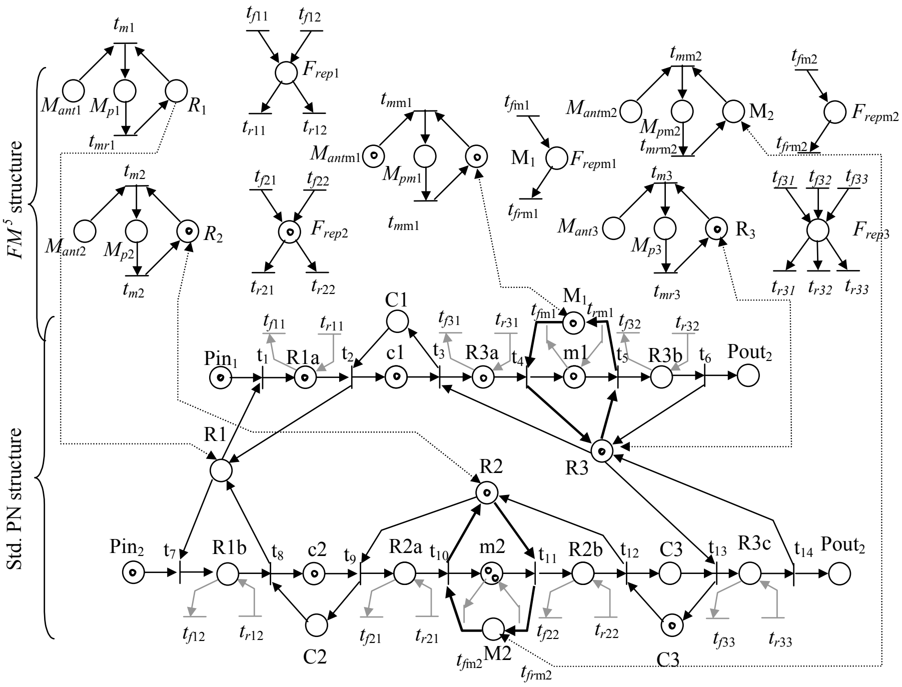


Figure 4. Complete Petri Net and FM^5 system structures

The current status shown in Figure 4, is that CW C_1 has two jobs pending in $m(J_o(C_1))=2$, jobs R3a and m1. Then, since $J_{MF}(C_1)=0$ (no jobs in places Fm_1, Mm_{p1}, F_{r3}, and M_{p3}), and $m_o(C_1)=4$, we are able to fire transition t_3 to have a total of three tokens allowed by (11). However, since a new attempt to start a PM job at place M_{pm1} is in place, and if we fire transition t_{mm1}, $J_{MF}(C_1)$ will become one, then we should not fire t_3 since C1 would be in deadlock, due to (11). For CW C_2, the allowable number of resources should be $<\{m_o(C_2) - J_{MF}(C_2)\}$. This is, should be smaller that 3. Then, we can not fire transition t_9, since C_2 will get into CB until failure maintenance $Frep_2$ is finished. Therefore, it is better not to get into blocking and wait till one of the jobs m2 is finished to diminish $m(J_o(C_2))$ by firing t_{11}.

The appropriate way to keep the markings of resources equal in both systems is to use Finite State Automata techniques to supervise both subsystems alternatively. This is, run one (several) discrete event(s) in any one of these subsystems, then holds its markings and passes the new marking of resources *R, m*(*R*), before one run event(s) in the other subsystem. This Finite Element Machine interaction between subsystems is shown in Figure 5.

For implementation of the DEC, in every DE iteration, we can use any desired dispatching policy. For example, FBFS, which maximizes WIP and machine percent utilization. However, it is known that FBFS often results in deadlock. Therefore, we combine FBFS with our new deadlock avoidance test (11). Thus, before we dispatch the FBFS resolution, we must examine the marking outcome with our deadlock policy. If this resulting outcome does not satisfy (11), then the algorithm denies or ***pre-filters*** in real time the firing and we apply again the FBFS conflict resolution strategy for the next possible allowable firing sequence. Then, using FBFS while permitted, we will try to satisfy in most of the current status of the cell the case $m(J_o(Ci)) = \{m_o(C_i) - J_{MF}(C_i)\}$-1. The later condition is an extended policy from that called MAXWIP policy, defined in [*Huang et al. 96*].

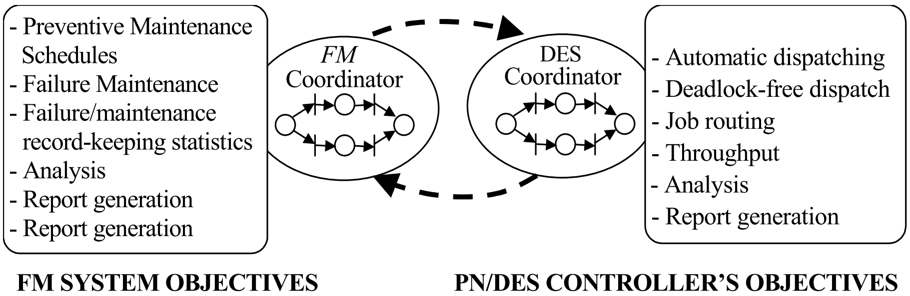


Figure 5. Finite State Automata interactions between the *FM* subsystem and DES controller structure

5. Conclusions

We show an analysis of blocking phenomena in Discrete Event Systems (DES) having corrective and/or Preventive Maintenance Schedules (PMS). We show that for some DES structures having reentrant flow-lines, it is not necessary to stop all activities in the DES, for the case one or more machines are in corrective and/or PMS. We proposed deadlock-free dispatching rules derived by performing circular wait analysis for possible deadlock situations. We analyzed the so-called critical siphons, certain critical subsystems and resources to develop a DE controller that guarantees deadlock-free dispatching with PMS by limiting the work-in-progress in the critical subsystems associated with each CW. This is accomplished by integrating a Finite State Automata supervision between two subsystems. One system is the Reentrant Flow-line system structure controlled by the DES matrix formulation, and an extra DES system contains the failure and preventive maintenance dynamics, called *FM* system structure. Deadlock-free dispatching is possible by passing the markings of available resources between these two subsystems. The extra

FM DES system acts as a disturbance in the primary Reentrant Flow-line DES system.

References

Banaszak Z. A. and B. H. Krogh. "Deadlock Avoidance in Flexible Manufacturing Systems with Concurrently Competing Process Flows." *IEEE Trans. Robotics and Automation*, RA-6, pp. 724-734 (1990).

Ezpeleta S. D., J. M. Colom and J. Martinez. "A Petri Net Based Deadlock Prevention Policy for Flexible Manufacturing Systems." *IEEE Trans. Robotics and Automatio*n, RA-11, pp. 173-184 (1995).

Fanti M.P., B. Maione, S. Mascolo, and B. Turchiano. "Event-Based Feedback Control for Deadlock Avoidance in Flexible Production Systems." IEEE Transactions on Robotics and Automation, Vol. 13, No. 3, June 1997.

Graves S.C. "A Review of Production Scheduling." *Operations Research*, vol. 29, no. 4 (1981).

Gurel A., S. Bogdan, and F.L. Lewis. "Matrix Approach to Deadlock-Free Dispatching in Multi-Class Finite Buffer Flowlines." IEEE Transactions on Automatic Control. Vol. 45, no. 11, Nov. 2000, pp. 2086-2090.

Hicks D.K. "Preventive Maintenance Program: Evaluation and Recommendations for Improvements." U.S. Army Construction Engineering Research Laboratory (USACERL). Report OMB No. 0704-0188. June 1990.

Hsieh F.-S. and S.-C. Chang. "Dispatching-Driven Deadlock avoidance controller Synthesis for Flexible Manufacturing Systems." *IEEE Trans. Robotics and Automation*, RA-11, pp. 196-209 (1994).

Hyuenbo C., T. K. Kumaran, and R. A. Wysk. "Graph-Theoretic Deadlock Detection and Resolution for Flexible Manufacturing Systems." *IEEE Transactions on Robotics and Automation*, vol. 11, no. 3, pp. 413-421 (1995).

Kumar, P.R. "Re-entrant lines." *Queueing Systems: Theory and Applications*. vol. 13, pp. 87-110, SW, (1993).

Kusiak A. and J. Ahn. "Intelligent Scheduling of Automated Machining Systems." *Computer Integrated Manufacturing Systems*, vol.5, no.1, Feb. 1992, pp. 3-14. UK (1992).

Lewis F. L.. "A Control System Design Philosophy for Discrete Event Manufacturing Systems." *Proc. Int. Symp. Implicit and Nonlinear Systems*, pp. 42-50, TX (1992).

Lewis, F.L., H.-H. Huang and S. Jagannathan. "A systems approach to discrete event controller design for manufacturing systems control." *Proceedings of the 1993 American Control Conference* (IEEE Cat. No.93CH3225-0). American Autom. Control Council. pp.1525-31 vol.2. Evanston, IL, USA (1993).

Lewis F.L., Gurel A, Bogdan S, Docanalp A, Pastravanu OC. "Analysis of Deadlock and Circular Waits using a Matrix Model for Flexible Manufacturing Systems*.*" *Automatica,* vol.34, no.9, Sept. 1998, pp.1083-100. Publisher: Elsevier, UK (1998).

Mireles, J. and F.L. Lewis, "On the Development and Implementation of a Matrix-Based Discrete Event Controller." *MED01, Proceedings of the 9th Mediterranean Conference on Control and Automation.* Pub. on CD, ref MED01-012. June 27-29 2001. Dubrovnik, Croatia 2001, a).

Mireles, J. and F.L. Lewis. "Intelligent Material Handling: Development and Implementation of a Matrix-Based Discrete Event Controller." *IEEE Transactions on Industrial Electronics*. Vol. 48, No. 6, December 2001, b).

Mireles, J. and F.L. Lewis, A. Gurel. "Deadlock Avoidance for Manufacturing Multipart Reentrant Flow Lines Using a Matrix-Based Discrete Event Controller." *Int. Journal of Production Research*. 2002.

Smith A.M. "Preventive Impact on Plant Availability." *Proceedings 1992 Annual Reliability and Maintainability Symposium*. pp. 177-180. 1992.

Steward, D. V. "The Design Structure System: A Method for Managing the Design of Complex Systems." *IEEE Trans. On Engineering Management*, vol. EM-28, no. 3, pp. 71-74 (1981).

INTELLIGENT PLANNING AND CONTROL IN A CIM SYSTEM

Doru Pănescu and Ştefan Dumbravă
"Gh. Asachi" Technical University of Iaşi
Department of Automatic Control and Industrial Informatics
Bd. D. Mangeron nr. 53A, 700050 Iaşi, Romania

Abstract A CIM planning and control architecture is presented, mainly regarding its software aspects. The system was developed and tested on a flexible manufacturing cell endowed with industrial equipment. Artificial intelligence methods and tools were used, namely expert systems, multiagent systems and rule based programming. The approach that combines centralized planning and monitoring with de-centralized and distributed decision making and control sub-systems aims at a high flexibility and autonomy.

Keywords: computer-integrated manufacturing, artificial intelligence, agents, expert systems, planning

1. Introduction

The use of new Artificial Intelligence (AI) techniques for planning and controlling in Robotics and CIM (Computer-Integrated Manufacturing) systems is a topical approach (Murphy, 2000; Parunak, 1999). One problem in applying the new methods is the difficulty of testing them on industrial systems. With respect to this, the presented research has the advantage of being based on a benchmark system that is used for both research and education. As depicted in Fig. 1, this system is exploiting a Flexible Manufacturing Cell (FMC), mainly endowed with real industrial equipment. In order to test a new planning and control architecture it was considered that the respective cell is part of a CIM system, so that various FMCs may exchange resources among them and the planning and control approach must be able to handle this. The goal for the whole system is to assemble the desired type and number of products, in the shortest time. A classical planning and control solution was carried out first (Pănescu, 2001); now a new one is under development, so that an optimal operation, greater autonomy and flexibility should be obtained.

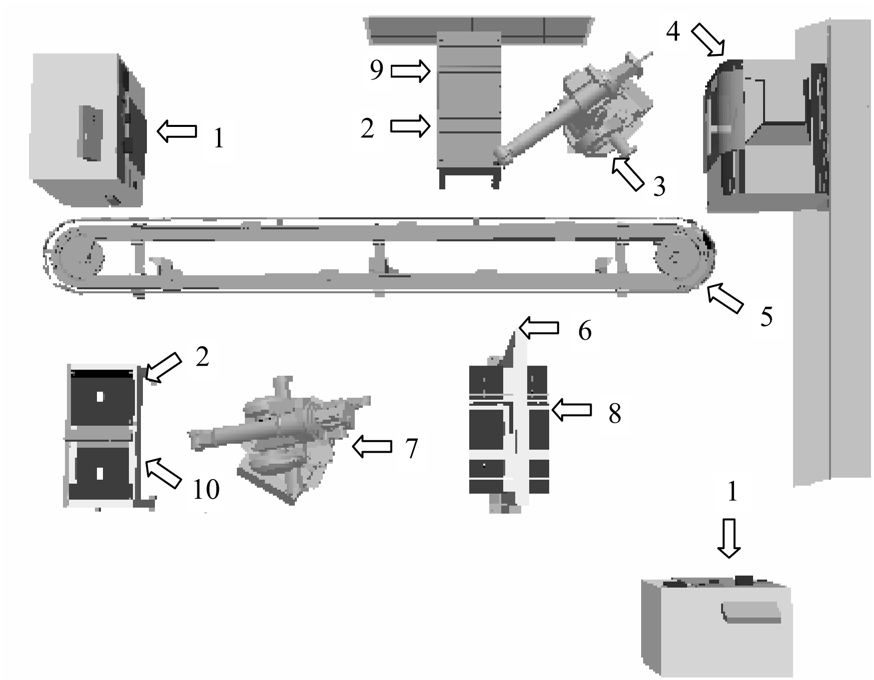


Figure 1. The layout of the FMC; 1 – robot controllers; 2 – storage devices for raw parts; 3 – IRB 1400 robot; 4 – PC Mill 55 machine tool; 5 – FlexLink conveyor; 6 – OptiMaster vision control station; 7 – IRB 2400 robot; 8 – assembly table; 9 – storage device for processed parts; 10 – storage device for final products

Though in the considered manufacturing system there is a certain sequence of main operations, namely part processing, quality control and assembly, certain issues make the planning and control processes difficult. These are as follows:

- The main operations are accompanied by auxiliary operations, which must be properly planned to assure no interruption for the system; these refer to part transfer, feeding/unfeeding of the machine tools and storage devices.
- The time parameters of the FMCs are variable.
- There are certain resources that act as bottlenecks. For example, in the FMC where the experiments were conducted these are the two robots, the conveyor and the storage devices. Concerning them, a wrong planning may cause deadlocks.
- To increase the CIM system autonomy and flexibility, the production goals were considered to appear randomly, and the planning system

should be able to deal with them. It means that the planning process is to be interleaved with execution phases and re-planning when needed.

To handle all these situations a new architecture was developed, making use of expert systems and agents techniques, as it is shown in the next paragraphs.

2. The architecture of the planning and control system

The control of the CIM system is achieved in a hierarchical manner. The decision part is placed on the higher level sending main commands towards the classical control sub-systems (robots and machine tools controllers, etc). This level is composed of a multi-expert system in conjunction with some multiagent systems, as shown in Fig. 2. An expert system deals with CIM planning and another one is in charge with production monitoring. The first expert system is continually receiving the goals and is also considering the tasks already fulfilled by the CIM system, based on the results from the monitoring expert system. Thus an adaptable operation is possible; as soon as some outputs (e.g. processed parts) are available from any FMC, they will be considered in the planning process. Such an operation is supported by the chosen implementation that is a rule based one; the monitoring expert system sends facts towards the planning one, and these activate in an opportunistic manner rules that plan new activities based on newly available resources (e.g. certain products' assembly using the recently processed parts).

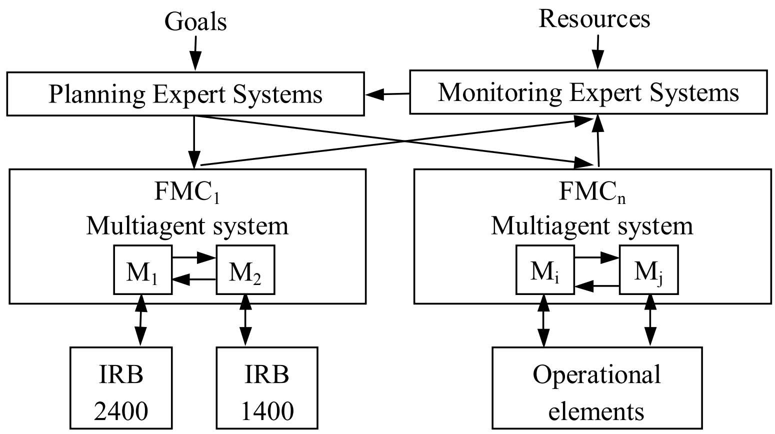


Figure 2. Planning and control architecture

It is to note that the planning and control architecture possesses both centralized and decentralized features. The two expert systems provide a centralized management of the main CIM systems goals – assembling of the products. Otherwise a distributed architecture is achieved through the use of several FMCs with specific planning and control systems. As already mentioned, for each FMC the control architecture is a hierarchical one. On the top level there is a software multiagent system that gets the action decisions. These are transmitted to the control systems of the real equipment, namely to the operational elements. For example, in the FMC_1, the one used for experiments (see Fig. 1), the two robots receive the action decisions. In this way, in an indirect manner, the robots have been transformed in the two agents of a multiagent system. This solution was adopted as the robot controllers can run only programs in their specific programming environment (the RAPID language). Thus, the CLIPS programs (Giarratano and Riley, 1989) that implement two software agents, representing action planning for each robot, are running on two distinct computers, being connected (by an Ethernet link) with the robot controllers.

3. On flexible manufacturing planning based on multiagent systems

In the considered FMC and generally in a manufacturing process, the robots may appear as actors establishing the main events. Many of the features required for such a role, like autonomous, flexible operation and continuous interaction with the environment, are naturally conducting towards the agent systems technique (Wooldridge, 2002). According to the architecture of Fig. 2, two software agents have been considered, named M_1 and M_2, corresponding to and guiding the two robots of the FMC_1. By the interaction with the planning expert system and between themselves, both agents are goal based ones (Russel and Norvig, 1995). As FMC_1 is a cell that can provide both part processing and assembly, for M_1 the main goals refer to assembling certain products, and for M_2 these correspond to the auxiliary operations related to part processing on the machine tool. Besides these, secondary goals for M_1 are the ones received from M_2 (when this one needs its help) and also those arising from its own motivation (e.g. the conveyor discharge). For M_2, when there is no main goal, it may also take into account secondary goals received from M_1 or its own goals, which refer to liberating some resources. For example, if there is no primary goal from M_1 (no goal to supply the machine tool), and some raw parts are on conveyor, M_2 will decide to transfer a raw part from conveyor to the storage device; thus the conveyor that is a shared resource is freed.

A hierarchy exists between the two agents: M_1 is ranked higher than M_2; this means that M_2 must accomplish the goals on part processing received from M_1 first (these are its primary goals) and only then consider its own goals. Even so, the dependence between the two agents is not a unilateral one, but a reciprocal one (Wooldridge, 2002). Indeed M_1 depends on M_2 in satisfying certain production goals. This is the case when a certain part is needed for assembling and is not available from another production cell, conducting to the necessity of being produced on the machine tool of FMC_1.

In such a case M_2 has to consider the goal received from M_1 and decide about a sequence of actions for machine tool feeding, starting of the processing, and finally transferring of the processed part, via conveyor, to the robot represented by the agent M_1. Meanwhile M_2 depends on M_1, as in certain conditions it cannot fulfill some tasks only by itself. This is the case when the machine tool must be fed and there is no raw part in the storage device near the IRB 1400 robot. Then the agent M_2 will ask the help of M_1, which may pass a raw part from the storage device near it, by using the conveyor.

A main property for a multiagent system is its communication ability. In the designed scheme there are several communication channels. First, there is a transfer of information from the planning expert system to the multiagent systems dedicated to various FMCs, representing the goals these must achieve. Then, the agents of the same system can exchange messages between them. Semantically these refer to assertions, requests, acceptances and refusals (Wooldridge, 2002), while syntactically they are all under the form of the facts in a rule based system (the agents are implemented in CLIPS). Even for the information received from environment, namely from the robot controllers, there is a C interface, which converts it into facts that are included in the agents' knowledge bases.

As an example on how planning and control of the FMC_1 are managed, starting from the goal of assembling a certain product, M_1 will make a request towards M_2 when a certain part is needed for assembling and it is not already available from another FMC. When the message is received – the formalism is closed to that of KQML (Wooldridge, 2002) – in the knowledge base of M_2 the following fact appears:

(goal of M1 part type D)

The above fact is under the CLIPS appearance and if it activates a chain of rules that find a plan for the respective goal fulfillment, then a message of acceptance will be sent to M_1. When the agent M_2 cannot manage by itself, but there is a plan of fulfilling the goal by cooperation, a conditional acceptance is sent back towards M_1. This may be the case in the considered example when the raw part necessary for processing a part of type D is not available in the storage device near the IRB 1400 robot. After receiving the message, the following fact appears in the knowledge base of M_1:

(goal of M_2 raw_part for type D)

This one activates a chain of rules, which search for an adequate raw part in the storage near the IRB 2400 robot and the possibility of transferring it to the IRB 1400 robot. In the peculiar case when the conveyor and the storages near the IRB 2400 robot are full, the agent M_1 further asks the help of M_2 that will be requested to discharge a position from the conveyor.

The above example illustrates how a multiagent cooperation strategy can conduct to solving, in an autonomous manner, manufacturing problems with multiple robotic actions. The mechanisms of agents' goal based motivation and communication proved to be useful in some other cases, as follows.

- When a storage device or the conveyor is full the robots will take the initiative (there is a self-motivation with respect to this) to discharge it, even if there is no production goal justifying this by that time. After such an action, the agent responsible for it will inform the other ones, by sending a message of assertion type. Such behavior was chosen when the Petri net of the FMC_1 was studied and the necessity of deadlock avoidance was considered.
- As already mentioned, there is a continuous exchange of information between the multiagent systems and the two expert systems from the top level. This allows the planning and control system to manage the whole CIM system operation, even if the duration of various operations is variable. For example, in the FMC_1, when a part is processed on the machine tool and one of the same kind arrives from another FMC, the IRB 2400 robot will immediately use this in the assembling process, in order to minimize the time of product delivery. In such a case the agent M_1 informs M_2 about the event and so M_2 will make a plan to store the processed part in its storage. If such a plan fails (e.g. its storage of processed parts is full), M_2 will further ask for the cooperation of M_1. In the same time, as the monitoring expert system is informed about such events by the multiagent systems of the corresponding FMCs, a feedback is sent towards the planning expert system and re-planning is started.
- Besides messages of acceptance, illustrated in the previous cases, refusals are also possible. As an example, when M_2 receives a goal to process a certain part on the machine tool and there is no corresponding raw part in its storage device it will ask the help of M_1. In the case that the storage device near the IRB 2400 robot does not contain any corresponding raw part, the answer of M_1 to the request of M_2 will be a refusal. In this case, the monitoring expert system is informed about the failure of the processing operation, and again the planning expert system will have to re-plan the goal towards another cell.

4. Conclusion

A few conclusions resulted from the research developed so far on AI based planning and control in CIM. As with other new techniques, there is a gap between research and industrial application of multiagent systems; the presented architecture is to be regarded as reducing this gap, because real equipment and problems were used. Agents and expert systems, through the way they consider the environment, the communication and the opportunist management of events, are able to deal with ill-structured problems. This is the case for many CIM systems, since it can be difficult to possess, from the beginning, all necessary information for planning and control, but this may appear randomly, as it is the case of the market's requests. Moreover, when the CIM system is a complex one, with several manufacturing cells, the combination of centralized/de-centralized planning and control provided by the proposed architecture, together with the modularity of the respective AI methods, conduct to a greater adaptability and autonomy, in comparison with the classical solutions.

Through a hierarchical structure, with the AI sub-systems on higher levels, robotic applications get new enhancements, as this contribution showed. Even using current industrial robots, which possess little intelligence, when connected with software agents these can become more flexible tools, and CIM systems that have such robots as central actors can be much easier deployed.

References

Giarratano, J. and G. Riley (1989). *Expert Systems: Principles and Programming*, pp. 373 – 497. PWS-KENT, Boston.

Murphy, R. (2000). *Introduction to AI Robotics*, pp. 15 – 36, 293 – 309. The MIT Press, Cambridge, Massachusetts.

Parunak, V. D. H. (1999). Industrial and Practical Application of DAI. In: *Multiagent Systems: a modern approach to distributed artificial intelligence* (Weiss, G. (Ed)), pp. 377 – 414. The MIT Press, Cambridge, Massachusetts.

Pănescu, D., M. Voicu, Şt. Dumbravă, R. Roşioru, G. Porumb, C. Brăescu and V. Dorin (2001). The Development of a Flexible Manufacturing System for CIM Education. *Bulletin of the Polytechnic Institute of Iaşi*, **XLVII**, pp. 75 – 83.

Russell, S. and P. Norvig (1995). *Artificial Intelligence: A Modern Approach*, pp. 42 – 44. Prentice Hall, Upper Saddle River, New Jersey.

Wooldridge, M. (2002). *An Introduction to Multiagent Systems*, pp. 17 – 43, 125, 163 - 183. John Wiley & Sons, Baffins Lane, Chichester.

PETRI NET TOOLBOX – TEACHING DISCRETE EVENT SYSTEMS UNDER MATLAB

Octavian Pastravanu, Mihaela-Hanako Matcovschi and Cristian Mahulea
Department of Automatic Control and Industrial Informatics
Technical University "Gh. Asachi" of Iasi
Blvd. Mangeron 53A, 700050 Iasi, Romania, Tel./Fax: +40-232-230.751
E-Mail: {opastrav, mhanako, cmahulea}@delta.ac.tuiasi.ro

Abstract A MATLAB toolbox has been developed to handle the basic problems of discrete event dynamical systems that are modeled by Petri nets. In the ***Petri Net Toolbox*** five types of Petri nets (untimed, transition-timed, place-timed, stochastic and generalized stochastic), with finite or infinite capacity, can be used. A user-friendly graphical interface allows activating three simulation modes (accompanied or not by animation) and running specific functions that cover the key topics of analysis such as coverability tree, structural properties (including invariants), time-dependent performance indices, max-plus state-space representations. A design procedure is also available, based on parameterized models. By incorporating instruments to explore the dynamics of Petri net models, as well as animation facilities to support the intuitive understanding and to guide the users in the exploitation of the software, the ***Petri Net Toolbox*** proves to be a valuable aid for Control Engineering education.

Keywords Control Engineering education, discrete event systems, Petri nets, MATLAB software

1. Motivation and objectives

The ***Petri Net Toolbox*** (***PN Toolbox***) was designed, implemented and tested at the Department of Automatic Control and Industrial Informatics of the Technical University „Gh. Asachi" of Iaşi. It is software for simulation, analysis and design of *discrete event systems* (DES), based on *Petri net* (PN) models, and embedded in the MATLAB environment. Our initiative brought remarkable benefits for training and research because Control Engineering students are familiar with the exploitation of *Graphical User Interfaces* (GUIs) (The MathWorks, 2001a) based on this popular software. The integration of the ***PN Toolbox*** with the MATLAB philosophy has the incontestable merit of broadening the MATLAB's utilization domain towards the area of discrete-event systems, which is now covered only by the State-Flow package. The orientation of the ***PN Toolbox*** was also meant to permit further development in the sense of hybrid systems, since MATLAB incorporates comprehensive libraries for studying continuous and discontinuous dynamics.

In the current version (namely 2.0) of the ***PN Toolbox***, five types of classic PN models are accepted, namely: untimed, transition-timed, place-timed, stochastic and generalized stochastic. The timed nets can be deterministic or stochastic, and the stochastic case allows using appropriate functions to generate random sequences corresponding to probability distributions with positive support. The default type of an arc is regular, but the user is allowed to change it into double or inhibitor, if necessary. Unlike other PN software, where places are meant as having finite capacity, our toolbox is able to operate with infinite-capacity places. In addition, the ***PN Toolbox*** allows the assignment of priorities and/or probabilities to conflicting transitions. As an educational aid, this software is suitable for applications illustrating the theoretical concepts provided by courses on PNs with different levels of difficulty, e.g. (Pastravanu, 1997), (Pastravanu *et al.* 2002), (Matcovschi, 2003), allowing relevant experiments for studying the event-driven dynamics of physical systems encountered in many technical fields (such as flexible manufacturing systems (FMSs), computer systems, communication protocols, power plants, power electronics).

The main goal envisaged by the designers of the ***PN Toolbox*** was to provide a collection of instruments for education and training at a graduate level, exploitable under MATLAB. Therefore, the focus was placed on developing students' skills in mastering PN models as a generous framework for dealing with discrete-event systems. Although a large number of tools are advertised for various types of PN problems (Mortensen, 2003), the unified treatment permitted by the ***PN Toolbox*** for untimed, deterministic/stochastic *P*- and *T*-timed PNs, stochastic and generalized stochastic PNs, ensures the premises for an efficient instruction. Thus, the user needs a short time to learn how to handle the ***PN Toolbox*** and his major intellectual effort is invested in the construction and careful analysis of the PN models. The interest shown by the authors for the convenient usage of the ***PN Toolbox*** is reflected by the numerous improvements brought to its previous versions (Mahulea *et al.*, 2001), (Matcovschi *et al.*, 2001), (Matcovschi *et al.*, 2002).

For attaining the proposed teaching goal, we preferred to orient our work towards enhancing the quality and reliability of the procedures devoted to standard topics rather than developing new algorithms. Consequently, the authors' attention focused on the following targets: (i) implementation of efficient algorithms for simulation, analysis and synthesis, (ii) creation of powerful visual support for the intuitive understanding of PN model usage, and (iii) elaboration of a comprehensive online help including animated demonstrative examples of handling the software.

2. Simulation, analysis and design

The ***PN Toolbox*** has an easy to exploit GUI (Matcovschi *et al.*, 2003) that gives the possibility to draw PNs in a natural fashion and allows a straightforward access to various commands starting adequate procedures for exploiting the PN models.

The simulation mechanism is based on the rule for enabling and firing of transitions specific to the type of the current PN model. Consequently, the simulation is driven by an asynchronous clock corresponding to the occurrence of events (Cassandras, 1993). In the untimed case, the sequencing of the events is reduced to simply ordering their occurrence, without any temporal significance, unlike the timed case when simulation requires a continuous correlation with physical time.

Three modes of simulation are implemented in the ***PN Toolbox***, namely: *Step*, *Run Slow* and *Run Fast*. The *Step* and *Run Slow* simulation modes are accompanied by animation; the user can record the progress of the simulation in a log file with HTML format. After ending a simulation (run in any of the three modes) a number of *Performance Indices* are available to globally characterize the simulated dynamics. They refer to: (i) transitions: *Service Sum* (the total number of firings during the simulation), *Service Rate* (the mean frequency of firings), *Service Distance* (the mean time between two successive firings), *Utilization* (the fraction of time when server is busy); and (ii) places: *Arrival Sum*, *Throughput Sum* (the total number of arrived/departed tokens), *Arrival Distance*, *Throughput Distance* (the mean time between two successive instants when tokens arrive in/depart from the place), *Waiting Time* (the mean time a token spends in a place), *Queue Length* (the average number of tokens weighted by time). For timed or (generalized) stochastic PNs, while in the *Step* and *Run Slow* simulation modes, the *Scope* facility opens a new MATLAB window that displays (dynamically) the evolution of a selected performance index versus time.

For untimed PN models, the *behavioral properties* (e.g. boundedness, liveness, reversibility, etc.) may be studied based on the *coverability tree* of the net. The coverability tree is built with or without the ω-convention. The ω-convention means the usage of a generic symbol (herein denoted by "ω") for referring to unbounded markings (Murata, 1989). The *structural* *properties* are approached as integer programming problems (Matcovschi *et al.*, 2001); the minimal-support P- and T-invariants (Martinez and Silva, 1982), (David and Alla, 1992) are displayed, on request, in separate windows.

A facility for the synthesis of timed or (generalized) stochastic PN models is *Design*, which allows exploring the dependence of a *Design Index* on one or two *Design Parameters* that vary within intervals defined by the user. For each test-point belonging to this (these) interval(s) a simulation-experiment is performed in the *Run Fast* mode. The results of all these simulation-experiments yield a graphical plot (2-D or 3-D, respectively) defining the dependence of the selected *Design Index* on the *Design Parameter*(*s*); the extreme values of the *Design Index* are numerically displayed.

The ***PN Toolbox*** is able to derive, directly from the topology and initial marking of a place-timed event graph, the max-plus state-space representation (Bacelli *et al.*, 1992). The following facilities are available for the max-plus analysis (Matcovschi *et al.*, 2002): ● displaying the matrix-form of the equations; ● max-plus simulation; ● graphical plots of the simulation results.

3. Visual information and animated demos

To enlarge the addressability of the ***PN Toolbox***, it includes a series of animation facilities aiming either to support the intuitive understanding or to guide the users in the exploitation of the software.

In the simulation modes *Step* and *Run Slow*, numerical computation is accompanied by animation whose role consists in feeding the user with visual information (current token contents of the places, currently firing transition), complementary to the numerical data available at the end of a simulation experiment. The animation technique is based on the general philosophy of the *object-oriented graphics system*, called **Handle Graphics** (The MathWorks Inc., 2001b). The nodes and arcs of a model are uniquely identified as MATLAB objects whose properties define (i) the characteristics of the PN, (ii) the graphical representation of the objects in the special area reserved for model drawing and (iii) the simulation status. The animation effects are obtained by automatically calling the **set** function for the properties referring to the appropriate instance of an object.

At the same time, the ***PN Toolbox*** was meant to illustrate, by short movies, behaviors that are typical for discrete event systems, for example sequential/parallel sharing of resources, routing policies, services in queuing networks, etc. The implementation combines, by means of the ActionScript Toolbox for Macromedia Flash (Macromedia, 2003), various techniques such as 2D and 3D graphics developed in Adobe Photoshop 7 (Adobe Systems Inc, 2003) and Maya 4.5 (Alias|Wavefront Inc, 2003), respectively. Each movie shows the physical motion of a real-life system synchronized with the token dynamics in the associated PN model, as resulting from the tutorial examples commented on in the following section. The movies are accessible on the web site we have created for the ***PN Toolbox*** (Mahulea *et al.*, 2003).

On the ***PN Toolbox*** site, the user can also find the online help of our software as well as some animated demos whose purpose is to present specific sequences of operations in handling the GUI and the interpretation of numerical results. Watching these demos, the user learns how to handle the key problems of discrete event systems within a PN framework: usage of adequate PN type (untimed, P/T-timed, stochastic or generalized stochastic) in model construction, study of behavioral/structural properties, analysis of max-plus representation, simulation and interpretation of the results, parameterized design, etc.

4. Tutorial examples

The four tutorial examples briefly described below were designed to prove the effectiveness of the ***PN Toolbox*** in assisting the DES training based on the Petri net theory. These examples cover a large area of classical topics and the incorporated animation is extremely profitable especially for the beginners along the lines detailed in the previous section.

Demo 1 refers to a computer system with two processors sharing two disks (in parallel) which is a version of the "Two Dinning Philosophers" well-known problem (Dijkstra, 1968), illustrated by a movie from which a frame is captured in fig. 1.

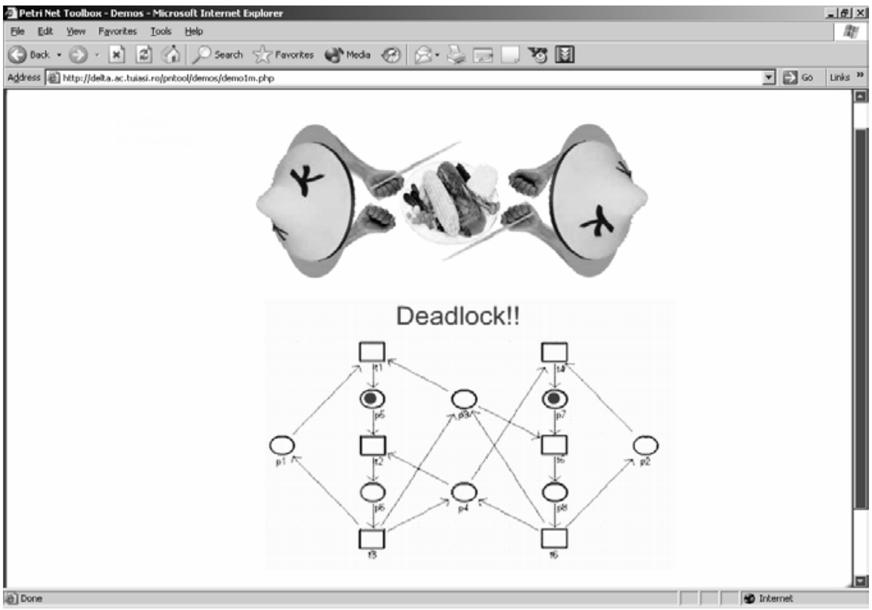


Figure 1. Frame in the *Demo 1* movie illustrating the "Two Dinning Philosophers" problem together with the dynamics of the associated PN model

The addressed problems are: ● construction of an untimed Petri net model; ● analysis of deadlock (via the coverability tree); ● prevention of deadlock through lookahead feedback (Lewis *et al.*, 1995); ● access to the following information about the Petri net model: incidence matrix, minimal-support P- and T-invariants, structural properties.

Demo 2 refers to a manufacturing system with a sequentially shared robot (Desrocheres and Al-Jaar, 1993), (Zhou and DiCesare, 1993), illustrated by a movie from which a frame is captured in fig. 2. The addressed problems are: ● construction of a P-timed Petri net model; ● analysis of deadlock (via simulation); ● prevention of deadlock by limiting the number of pallets; ● analysis of time-dependent performance indices and ● study of a performance index depending on two design parameters (see fig. 3).

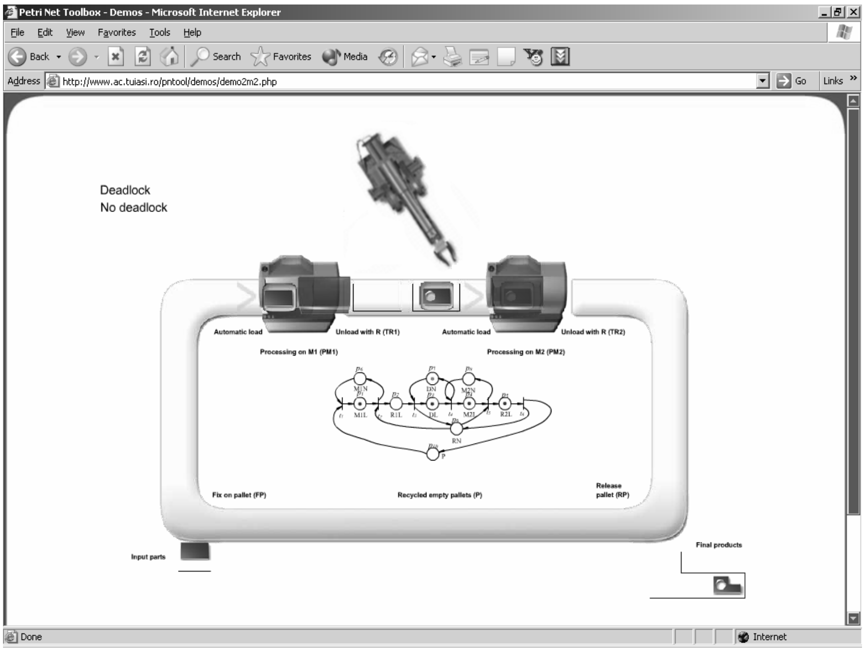


Figure 2. Frame in the *Demo 2* movie illustrating the functioning of a manufacturing system concomitantly with the dynamics of the associated PN model

Demo 3 refers to a flow-shop system with three machines, adapted from (Bacelli *et al.*, 1992). The addressed problems are: ● simulation and animation in the Run Slow mode; ● record of the simulation results in a log file; ● computation of the cycle time, ● max-plus analysis of a place-timed event graph: max-plus state-space representation, setting of the values for the input vectors, max-plus based simulation and plots of the components for the input, state or output vectors (see fig. 4).

Demo 4 refers to an open markovian queuing network (Cassandras, 1993). This demo illustrates: ● construction of a generalized stochastic Petri net model; ● usage of the Scope and Diary facilities; ● analysis of time-dependent performance indices.

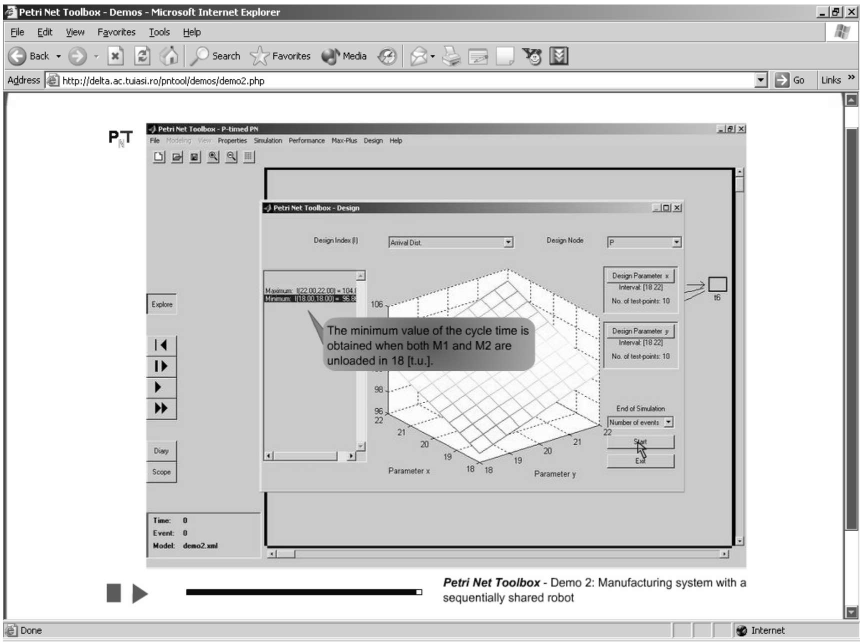


Figure 3. Frame in *Demo 2* introducing, by animation, the usage of parameterized design

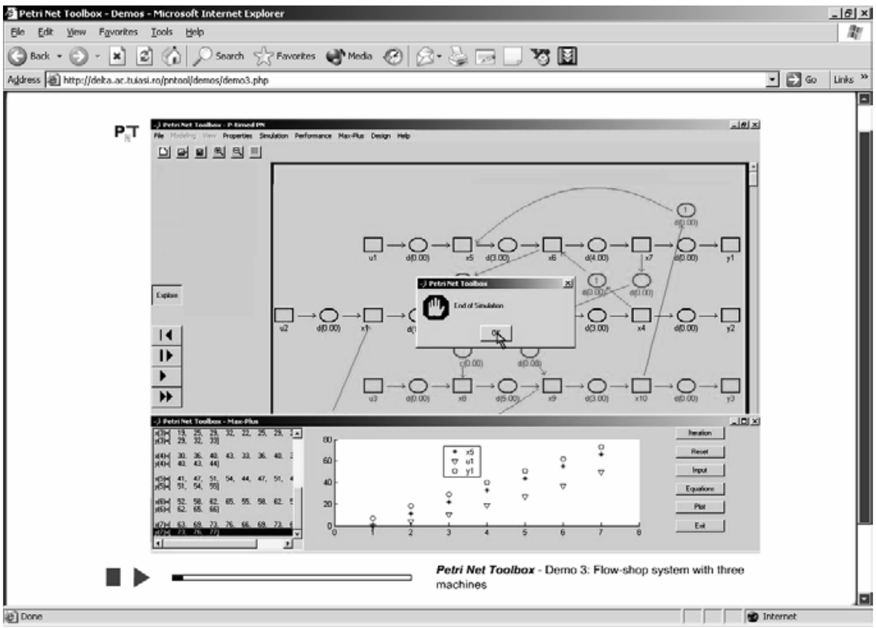


Figure 4. Frame in *Demo 3* presenting, by animation, the usage of max-plus analysis

5. Conclusions

Despite the large offer of software products available for MATLAB, none of its toolboxes provides instruments able to handle Petri net models. This fact has motivated the development of the ***PN Toolbox*** based on a user-friendly graphical interface that makes it very attractive for students because they don't have to spend time for code writing and their attention can exclusively focus on the topics of Control Engineering. The facilities created for simulation, analysis and design prove useful in many types of applications including a wide range of event-driven dynamics, as illustrated by the four tutorial examples briefly presented in the text.

References

Adobe Systems Inc (2003). *Home Page*, http://www.adobe.com.

Alias|Wavefront Inc. (2003). *Home Page*, http://www.aliaswavefront.com.

Bacelli, F., G. Cohen, G.J. Olsder and J.P. Quadrat (1992). *Synchronization and Linearity, An Algebra for Discrete Event Systems*, Wiley, New York.

Cassandras, C.G. (1993). *Discrete Event Systems: Modeling and Performance Analysis*, Irwin.

David, R. and H. Alla (1992). *Du Grafcet aux Réseaux de Petri* (2e édition), Hermes, Paris.

Desrocheres, A.A. and R.Y. Al-Jaar (1993). *Modeling and Control of Automated Manufacturing Systems*. IEEE Computer Society Press, Rensselaer, Troy, New-York.

Djikstra, E. W. (1968). "Co-operating sequential processes", in F. Genyus (Ed.), *Programming Languages*, New York Academic, pp. 43-112.

Lewis, F., H.H. Huang, O. Pastravanu and A. Gurel (1995). "Control Systems Design for Flexible Manufacturing Systems", in A. Raouf, and M. Ben-Daya (Eds.), *Flexible Manufacturing Systems: Recent Developments*, Elsevier Science, pp. 259-290.

Macromedia Inc. (2003). *Home Page*, http://www.macromedia.com.

Mahulea, C., L. Bârsan and O. Pastravanu (2001). "Matlab Tools for Petri-Net-Based Approaches to Flexible Manufacturing Systems", in F.G. Filip, I. Dumitrache and S. Iliescu (Eds.), *9th IFAC Symposium on Large Scale Systems LSS 2001*, Bucharest, Romania, pp. 184-189.

Mahulea, C., M.H. Matcovschi and O. Pastravanu (2003). *Home Page of the Petri Net Toolbox*, http://www.ac.tuiasi.ro/pntool.

Martinez, J. and M. Silva (1982). "A Simple and Fast Algorithm to Obtain All Invariants of a Generalized Petri Net", in C. Girault and W. Reisig (Eds), *Application and Theory of Petri Nets*, Informatik Fachberichte 52, Springer, pp. 301-310.

Matcovschi, M.H., C. Mahulea and O. Pastravanu (2001). "Exploring Structural Properties of Petri Nets in MATLAB", *7th International Symposium on Automatic Control and Computer Science SACCS 2001*, Iasi, Romania.

Matcovschi, M.H., C. Mahulea and O. Pastravanu (2002). "Computer Tools For Linear Systems Over Max-Plus Algebra", *5th International Conference on Technical Informatics CONTI'2002*, Timisoara, Romania.

Matcovschi, M.H. (2003). *Markov Chains and Markovian Queueing Systems* (in Romanian), Ed. Gh. Asachi.

Matcovschi, M.H., C. Mahulea and O. Pastravanu (2003). "Petri Net Toolbox for MATLAB", *11th IEEE Mediterranean Conference on Control and Automation MED'03*, Rhodes, Greece.

The MathWorks Inc. (2001a). *Building GUIs with MATLAB*. Natick, Massachusetts.

The MathWorks Inc. (2001b). *Using MATLAB Graphics*. Natick, Massachusetts.

Mortensen, K.H. (2003). *Petri Nets Tools and Software*, http://www.daimi.au.dk/PetriNets/tools.

Murata, T. (1989). "Petri Nets: Properties, Analysis and Applications", *Proc. of the IEEE*, vol. **77**, pp. 541-580.

Pastravanu, O. (1997). *Discrete Event Systems. Qualitative techniques in a Petri Net Framework* (in Romanian), Ed. Matrix Rom.

Pastravanu, O., M.H. Matcovschi and C. Mahulea (2002). *Applications of Petri Nets in Studying Discrete Event Systems* (in Romanian), Ed. Gh. Asachi.

Zhou, M.C. and F. DiCesare (1993). *Petri Net Synthesis for Discrete Event Control of Manufacturing Systems*, Kluwer, Boston.

COMPONENTWISE ASYMPTOTIC STABILITY - FROM FLOW-INVARIANCE TO LYAPUNOV FUNCTIONS

Octavian Pastravanu and Mihail Voicu

Department of Automatic Control and Industrial Informatics
Technical University "Gh. Asachi" of Iasi
Blvd. Mangeron 53A, 700050 Iasi, Romania
Phone/Fax: +40-232-230751
E-mail: opastrav@delta.ac.tuiasi.ro, mvoicu@delta.ac.tuiasi.ro

Abstract In previous works on componentwise asymptotic stability (CWAS), the analysis of CWAS for a given linear system requested the investigation of an auxiliary system of difference (in the discrete-time case) or differential (in the continuous-time case) inequalities, built from the state equation of the studied system. Our paper shows that, by the adequate usage of the infinity norm, the analysis of CWAS can circumvent the construction of such inequalities and can apply the standard tools of asymptotic stability (ε - δ formalism, properties of the operator describing the system dynamics, Lyapunov functions) directly to the studied system. These novel results reveal the complete meaning of CWAS as a special type of asymptotic stability.

Keywords: componentwise asymptotic stability, stability analysis, flow-invariant sets, linear systems

1. Introduction

The concepts of *componentwise asymptotic stability* (CWAS) and *componentwise exponential asymptotic stability* (CWEAS) were introduced and characterized for continuous-time dynamical systems by Voicu, who explored the linear dynamics in (Voicu, 1984a; b) and the nonlinear dynamics in (Voicu, 1987). Voicu's works relied on the theory of time-dependent flow-invariant sets (Pavel, 1984) which allowed a refinement of the standard stability notions, by the individual monitoring of the state-space trajectories approaching an equilibrium point. Later on, CWAS and CWEAS were extended by Hmamed to continuous-time delay linear systems (Hmamed, 1996) and to 1-D and 2-D linear discrete systems (Hmamed, 1997). Recently, Pastravanu and Voicu dealt with CWAS and CWEAS of interval matrix systems in both discrete-time and continuous-time cases (Pastravanu and Voicu, 1999; 2002). For a survey of some results based on time-dependent flow-invariant sets see (Voicu and Pastravanu, 2003).

All the researches mentioned above focused on the characterization of CWAS / CWEAS via difference inequalities (in the discrete-time case) and differential inequalities (in the continuous-time case). Consequently, emphasis was placed on studying the properties of the operators defining such inequalities, which were different from the operators describing the system dynamics.

The purpose of the current paper is to point out the existence of direct links between the dynamics of the studied system and CWAS / CWEAS as a special type of asymptotic stability. It is shown that such links are ensured by the usage of infinity norm and operate as particular forms of well-known results in the classical theory of stability. Thus, the analysis of CWAS / CWEAS can circumvent the construction of the inequalities mentioned above and can apply standard tools in stability theory directly to the investigated system.

During the last decade, the infinity norm has been used in several works devoted to the study of polyhedral invariant sets and their application in control – see, for instance, the remarkable survey paper (Blanchini, 1999) and the papers cited therein. For most of these researches, the polyhedral invariant sets do not depend on time, or if they do, the time-dependence is understood as a contraction of exponential type, operating uniformly on the constraints of the initial conditions (which is actually induced by the exponential-type decreasing of a non-quadratic Lyapunov function associated with linear systems). Therefore, such researches (focusing on the generality of the polyhedrons, but neglecting the generality of the time dependence) do not realize that the studied invariance is strongly related to a special type of asymptotic stability (actually meaning CWAS / CWEAS).

Besides the intrinsic value of the stability analysis tools developed by our paper, we are also able to bridge the gap between the research trend commented above and the CWAS / CWEAS framework. Thus, CWAS / CWEAS as special type of AS, reveal the complete meaning of the invariance for symmetrical rectangular sets, whose dependence of time is *a priori* stated and explicitly defined.

2. CWAS and CWEAS derived from flow invariance

This short presentation of the key concepts and results on CWAS and CWEAS is based on the initial formulation proposed for the continuous-time case in (Voicu, 1984a; b) and, later on, unified for discrete-time and continuous-time cases in (Pastravanu and Voicu, 1999, 2002).

Consider the linear system:

$$\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t), \ \mathbf{A} \in \mathbf{R}^{n \times n}, \tag{1}$$

where $t \in \mathbf{T}$ denotes the independent variable with discrete-time meaning $\mathbf{T} = \mathbf{Z}_+$, or continuous-time meaning $\mathbf{T} = \mathbf{R}_+$, and the action of the operator ( )' is defined by:

$$\mathbf{x}'(t) = \begin{cases} \mathbf{x}(t+1) \ \text{ for the discrete-time case } \ t \in \mathbf{T} = \mathbf{Z}_+; \\ \dot{\mathbf{x}}(t) \ \text{ for the continuous-time case } \ t \in \mathbf{T} = \mathbf{R}_+. \end{cases} \tag{2}$$

Definition 1. Given the vector function $\mathbf{h}(t) : \mathbf{T} \to \mathbf{R}^n$, which fulfils the following conditions:

(a) in the discrete-time case ($\mathbf{T} = \mathbf{Z}_+$), **h**(t) has positive components $h_i(t) > 0$, $i = 1, \ldots, n$, and $\lim_{t \to \infty} \mathbf{h}(t) = 0$,

(b) in the continuous-time case ($\mathbf{T} = \mathbf{R}_+$), **h**(t) is differentiable, has positive components $h_i(t) > 0$, $i = 1, \ldots, n$, and $\lim_{t \to \infty} \mathbf{h}(t) = 0$, system (1) is called componentwise asymptotically stable (CWAS) with respect to $\mathbf{h}(t)$ if

$$\forall t_0, t \in \mathbf{T}, \ \ t_0 \le t : \ |x_i(t_0)| \ \le \ h_i(t_0) \ \Rightarrow |x_i(t)| \ \le \ h_i(t), \quad i = 1, \ldots, n, \tag{3}$$

where $x_i(t), \ \ i = 1, \ldots, n$, denote the state variables of system (1).□

CWAS allows the individual monitoring of each state variable and therefore it represents a refinement of the standard concept of asymptotic stability where the evolution is characterized in the global terms of a vector norm.

Theorem 1 *All the functions* $\mathbf{h}(t)$ *that fulfill the conditions in Definition 1 are solutions of the difference inequality (in the discrete-time case) or differential inequality (in the continuous-time case):*

$$\mathbf{h}'(t) \ \ \ge \ \ \overline{\mathbf{A}} \quad \mathbf{h}(t), \tag{4}$$

where the matrix $\overline{\mathbf{A}} \in \mathbf{R}^{n \times n}$ *is built from matrix **A** in equation (1), as follows:*

(a) for the discrete-time case:

$$\overline{a}_{ij} = | a_{ij} |, \qquad i, j = 1, \cdots, n; \tag{5a}$$

(b) for the continuous-time case:

$$\begin{aligned} &\overline{a}_{ii} = a_{ii}, \ i = 1, \cdots, n, \\ &\overline{a}_{ij} = | a_{ij} |, \ \ i \ne j, \ i, j = 1, \cdots, n. \end{aligned} \tag{5b}$$

System (4) confers a consistent dynamical signification to the operator $\bar{\mathbf{A}}$, pointing out the origin of the CWAS concept in the theory of flow-invariant sets. Within this context, it is worth saying that system (4) might have solutions $\mathbf{h}(t)$ that do not fulfill the condition $\lim_{t\to\infty}\mathbf{h}(t)=0$ in Definition 1, but such solutions are able to define time-dependent sets, which are flow-invariant with respect system (1).

Theorem 2 *System (1) is CWAS with respect to an arbitrary* $\mathbf{h}(t)$ *which fulfils the conditions in Definition 1, if and only if the matrix* $\bar{\mathbf{A}}$ *built according to (5a) or (5b) is stable in Schur or Hurwitz sense, respectively.*

The usage of CWAS with respect to a particular vector function $\mathbf{h}(t)$ of exponential type yields:

Definition 2. (a) In the discrete-time case, system (1) is called componentwise exponential asymptotically stable (CWEAS) if there exist a vector $\mathbf{d}\in\mathbf{R}^n$, with positive components $d_i>0,\quad i=1,\dots,n$, and a constant $0<r<1$ such that

$$\forall t_0, t\in\mathbf{T}=\mathbf{Z}_+,\quad t_0\le t:\ |x_i(t_0)|\le\ d_i r^{t_0}\ \Rightarrow\ |x_i(t)|\le\ d_i r^t,\ i=1,\dots,n. \quad (6a)$$

(b) In the continuous-time case, system (1) is called componentwise exponential asymptotically stable (CWEAS) if there exist a vector $\mathbf{d}\in\mathbf{R}^n$, with positive components $d_i>0,\quad i=1,\dots,n$, and a constant $r<0$ such that

$$\forall t_0, t\in\mathbf{T}=\mathbf{R}_+,\ t_0\le t:\ |x_i(t_0)|\le d_i e^{rt_0}\ \Rightarrow\ |x_i(t)|\le d_i e^{rt},\ i=1,\dots,n. \square \quad (6b)$$

The linearity of the dynamics of system (1) guarantees the equivalence between CWAS and CWEAS.

Theorem 3 *For both discrete-time and continuous-time cases, system (1) is CWAS with respect to an arbitrary* $\mathbf{h}(t)$ *which fulfils the conditions in Definition 1 if and only if system (1) is CWEAS.*

On the other hand, the exponential form of the vector function $\mathbf{h}(t)$ considered in Definition 2 results in an algebraic characterization of CWEAS, or, equivalently, CWAS.

Theorem 4 *System (1) is CWAS (or equivalently CWEAS), if and only if the system of inequalities constructed with the matrix* $\bar{\mathbf{A}}$ *(5a) or (5b):*

$$\bar{\mathbf{A}}\mathbf{d}\le r\mathbf{d},\quad \mathbf{d}\in\mathbf{R}^n,\ d_i>0,\quad i=1,\dots,n,\quad r\in\mathbf{R}, \quad (7)$$

has solutions $0<r<1$ *in the discrete-time case, or* $r<0$ *in the continuous-time case, respectively.*

The special structure of matrix $\bar{\mathbf{A}}$ built according to (5a) or (5b) induces a spectral property to $\bar{\mathbf{A}}$ of crucial importance for the compatibility of inequality (7):

Theorem 5 *Denote by* $\lambda_i(\bar{\mathbf{A}})$, $i = 1,\dots,n$, *the eigenvalues of the matrix* $\bar{\mathbf{A}}$.

i) *(a) If* $\bar{\mathbf{A}}$ *is defined according to (5a), then* $\bar{\mathbf{A}}$ *has a real nonnegative eigenvalue (simple or multiple) denoted by* $\lambda_{\max}(\bar{\mathbf{A}})$, *meaning the spectral radius, which fulfills the dominance condition*

$$|\lambda_i(\bar{\mathbf{A}})| \leq \lambda_{\max}(\bar{\mathbf{A}}), \quad i = 1,\dots,n. \tag{8a}$$

(b) If $\bar{\mathbf{A}}$ *is defined according to (5b), then* $\bar{\mathbf{A}}$ *has a real eigenvalue (simple or multiple), denoted by* $\lambda_{\max}(\bar{\mathbf{A}})$, *meaning the spectral abscissa, which fulfills the dominance condition*

$$\mathrm{Re}[\lambda_i(\bar{\mathbf{A}})] \leq \lambda_{\max}(\bar{\mathbf{A}}), \quad i = 1,\dots,n. \tag{8b}$$

ii) *The system of inequalities (7) is compatible if and only if*

$$\lambda_{\max}(\bar{\mathbf{A}}) \leq r\,. \tag{9}$$

3. CWAS / CWEAS and ε ~ δ formalism

Although it was eminently clear that CWAS, or, equivalently, CWEAS represented a stronger concept than the standard asymptotic stability, no proof has been constructed yet for this statement in terms of norms (which actually provide the classical tools for defining asymptotic stability). Let us show that the exponential asymptotic stability incorporates the concept of CWEAS as a special case, by using the well known ε ~ δ language. Therefore consider the following general condition which ensures the *exponential asymptotic stability* for the equilibrium point {0} of linear system (1) (e.g. (Michel and Wang, 1995), pp. 107):

(a) for the discrete-time case:

$$\forall\varepsilon > 0 \;\; \exists\delta(\varepsilon) > 0,\; 0 < \omega < 1 : \|\mathbf{x}(t_0)\| \leq \delta(\varepsilon) \Rightarrow \forall t \geq t_0 : \|\mathbf{x}(t)\| \leq \varepsilon\omega^{(t-t_0)}; \tag{10a}$$

(b) for the continuous-time case:

$$\forall\varepsilon > 0 \;\; \exists\delta(\varepsilon) > 0, \;\; \omega < 0 : \|\mathbf{x}(t_0)\| \leq \delta(\varepsilon) \Rightarrow \;\forall t \geq t_0 : \|\mathbf{x}(t)\| \leq \varepsilon e^{\omega(t-t_0)}, \tag{10b}$$

where $\| \; \|$ denotes an arbitrary vector norm in $\mathbf{R}^n$.

On the other hand, define the vector norm:

$$\|\mathbf{x}\|_{\mathbf{D}\infty} = \|\mathbf{D}^{-1}\mathbf{x}\|_{\infty}, \tag{11}$$

where the diagonal matrix

$$\mathbf{D} = \mathrm{diag}\{d_1, \cdots, d_n\} \tag{12}$$

is built with the positive constants $d_i > 0,\ i = 1, \cdots, n$.

Theorem 6 *System (1) is CWEAS if and only if condition (10) is met with* $\delta(\varepsilon) = \varepsilon$, $\omega = r$ *and for the vector norm* $\| \, \|_{\mathbf{D}\infty}$ *given by (11).*

Proof. The inequality $\| \mathbf{x}(t_0) \|_{\mathbf{D}\infty} \leq \varepsilon$ is equivalent to the componentwise inequality $| \mathbf{x}(t_0) | \leq \varepsilon \quad \mathbf{d}$ and

(a) for the discrete-time case, the inequality $\| \mathbf{x}(t) \|_{\mathbf{D}\infty} \leq \varepsilon \quad r^{(t-t_0)}$ is equivalent to the componentwise inequality $| \mathbf{x}(t) | \leq \varepsilon \quad \mathbf{d} \quad r^{(t-t_0)}$ for $t \geq t_0$;

(b) for the continuous-time case, the inequality $\| \mathbf{x}(t) \|_{\mathbf{D}\infty} \leq \varepsilon \quad e^{r(t-t_0)}$ is equivalent to the componentwise inequality $| \mathbf{x}(t) | \leq \varepsilon \quad \mathbf{d} \quad e^{r(t-t_0)}$ for $t \geq t_0$.■

Proving that the CWEAS property is obtainable from the general definition of the exponential asymptotic stability, this result motivates us to further explore the standard instruments used by the stability analysis of linear systems in order to characterize CWAS / CWEAS.

4. CWAS / CWEAS and properties of operator A

Theorems 4 and 5 are extremely valuable in characterizing the CWAS (CWEAS) of system (1), because they permit a complete exploration of the link between the scalar *r*, vector **d** and matrix $\bar{\mathbf{A}}$ constructed according to (5). Nevertheless, they are unable to link *r* and **d** directly to matrix **A** used in system (1). One can overcome this disadvantage, by introducing the *matrix norm* subordinate to the vector norm $\| \, \|_{\mathbf{D}\infty}$ defined in (11) with (12):

$$\| \mathbf{M} \|_{\mathbf{D}\infty} = \| \mathbf{D}^{-1}\mathbf{M}\mathbf{D} \|_{\infty},\ \mathbf{M} \in \mathbf{R}^{n \times n}. \tag{13}$$

Theorem 7 *Consider a square matrix **A** and the matrix* $\bar{\mathbf{A}}$ *built from it according to (5). A positive vector **d** and a constant r are a solution of inequality (7) if and only if*

$$\mu_{\mathbf{D}\infty}(\mathbf{A}) \leq r\,, \tag{14}$$

where $\mu_{\mathbf{D}\infty}(\mathbf{A})$ *denotes a matrix measure defined by:*

(a) for $\bar{\mathbf{A}}$ *built according to (5a):*

$$\mu_{\mathbf{D}\infty}(\mathbf{A}) = \| \mathbf{A} \|_{\mathbf{D}\infty}\,; \tag{15a}$$

(b) for $\bar{\mathbf{A}}$ *built according to (5b):*

$$\mu_{\mathbf{D}\infty}(\mathbf{A}) = \lim_{\tau \to 0+} \frac{\| \mathbf{I} + \tau \mathbf{A} \|_{\mathbf{D}\infty} - 1}{\tau} . \tag{15b}$$

Proof. Algebraic inequality (7) can be written as:

$$(1/d_i) \sum_{j=1}^{n} \bar{a}_{ij} d_j \le r \ , \ i = 1, \ldots, n \ , \tag{16}$$

or, equivalently:

$$\max_{i=1,\cdots,n} \left\{ (1/d_i) \sum_{j=1}^{n} \bar{a}_{ij} d_j \right\} \le r . \tag{17}$$

(a) For the discrete-time case, all the elements $\bar{a}_{ij}$ constructed in accordance with (5a) are nonnegative and, therefore, (17) is equivalent to:

$$\| \mathbf{D}^{-1} \mathbf{A} \mathbf{D} \|_{\infty} \le r , \tag{18a}$$

which, taking into account (15a), means inequality (14).

(b) For the continuous-time case, in accordance with (5b) all the elements $\bar{a}_{ij}$, $i \ne j$, are nonnegative. If the same big positive constant $1/\tau \ge \| \mathbf{A} \|_2$, is added to both sides of each inequality (16), then all the elements $\bar{a}_{ii} + 1/\tau$ become also nonnegative and, therefore, (16) is equivalent to:

$$\| \mathbf{D}^{-1} (\frac{1}{\tau} \mathbf{I} + \mathbf{A}) \mathbf{D} \|_{\infty} \le r + \frac{1}{\tau} , \tag{18b}$$

which, taking into account (15b), means inequality (14).■

Remark 1. The matrix measure defined by (15b) for $\mathbf{D} = \mathbf{I}$ the identity matrix is frequently referred to as the "logarithmic norm" (Deutsch, 1975)], although it does not meet all the properties of a norm.□

Remark 2. The n inequalities given by (16), which are equivalent to CWEAS, express the condition that the generalized Gershgorin disks of the matrix $\bar{\mathbf{A}}$ lay inside the unit circle or in the left half plane of the complex plane. In the continuous-time case these disks are identical to those of the matrix **A** (as pointed out in (Voicu, 1984b)), and in the discrete-time case, they can be identical to those of the matrix **A,** or symmetrical with respect to the imaginary axis of the complex plane. Therefore the usage, in the very recent paper (Polyak and Shcherbakov, 2002), of condition (16) for the particular case $d_i = 1$, $i = 1, \ldots, n$, as a parametric definition for a property called "superstability" has no reason and yields particular forms of the CWEAS results available from (Voicu, 1984a; b; 1987)], (Pastravanu and Voicu, 1999; 2002).□

Theorem 8 *The dominant eigenvalue* $\lambda_{\max}(\bar{\mathbf{A}})$ *introduced in Theorem 4 fulfills the condition:*

$$\lambda_{\max}(\bar{\mathbf{A}}) = \min_{\mathbf{D}=diag\{d_i\}} \mu_{\mathbf{D}\infty}(\mathbf{A}), \tag{19}$$

where $\mu_{\mathbf{D}\infty}(\mathbf{A})$ *is defined by (15a) or (15b), in accordance with the procedure for building* $\bar{\mathbf{A}}$ *(5a) or (5b), respectively.*

Proof. (a) In the discrete-time case, (19) results from the equality proven in (Theorem 2, (Stoer Witzgall, 1962)) for nonnegative matrices:

$$\lambda_{\max}(\bar{\mathbf{A}}) = \min_{\mathbf{D}=diag\{d_i\}} \| \mathbf{D}^{-1}\bar{\mathbf{A}}\mathbf{D} \|_{\infty}, \tag{20a}$$

together with:

$$\| \mathbf{D}^{-1}\bar{\mathbf{A}}\mathbf{D} \|_{\infty} = \| \mathbf{D}^{-1}\mathbf{A}\mathbf{D} \|_{\infty} = \mu_{\mathbf{D}\infty}(\mathbf{A}). \tag{21a}$$

(b) In the continuous-time case, (19) results along the same lines, by taking into consideration the nonnegativeness of the matrix $\mathbf{I}/\tau + \bar{\mathbf{A}}$, as well as the fact that for small $\tau > 0$ (i.e. $\tau \le 1/\|\bar{\mathbf{A}}\|_2$ satisfied) one can write:

$$\lambda_{\max}(\frac{1}{\tau}\mathbf{I} + \bar{\mathbf{A}}) = \min_{\mathbf{D}=diag\{d_i\}} \| \mathbf{D}^{-1}(\frac{1}{\tau}\mathbf{I} + \bar{\mathbf{A}})\mathbf{D} \|_{\infty} \tag{20b}$$

and

$$\lim_{\tau\to 0+}\left(\| \mathbf{D}^{-1}(\frac{1}{\tau}\mathbf{I} + \bar{\mathbf{A}})\mathbf{D} \|_{\infty} - \frac{1}{\tau} \right) = \lim_{\tau\to 0+}\left(\| \mathbf{D}^{-1}(\frac{1}{\tau}\mathbf{I} + \mathbf{A})\mathbf{D} \|_{\infty} - \frac{1}{\tau} \right) = \mu_{\mathbf{D}\infty}(\mathbf{A}). \blacksquare \tag{21b}$$

Theorem 9 *Linear system (1) is CWAS / CWEAS if and only if*

(a) for the discrete-time case, there exists a vector with positive entries $\mathbf{d} \in \mathbf{R}^n$*, such that*

$$\mu_{\mathbf{D}\infty}(\mathbf{A}) < 1, \tag{22a}$$

(b) for the continuous-time case, there exists a vector with positive entries $\mathbf{d} \in \mathbf{R}^n$*, such that*

$$\mu_{\mathbf{D}\infty}(\mathbf{A}) < 0, \tag{22b}$$

where $\mu_{\mathbf{D}\infty}(\mathbf{A})$ *is defined according to (15a) and (15b), respectively.*

Proof. It results directly from Theorems 2 and 5 combined with Theorem 8. ■

5. CWAS / CWEAS and Lyapunov functions

The previous results fully motivates the idea of investigating CWAS by special Lyapunov functions, whose expressions contain precise information about the vector functions $\mathbf{h}(t)$ used in Definition 1.

Theorem 10 *Consider a vector function* $\mathbf{h}(t)$ *that fulfills the conditions in Definition 1. System (1) is CWAS with respect to* $\mathbf{h}(t)$*, if and only if*

$$V(t, x(t)) = \| \mathbf{H}^{-1}(t) x(t) \|_{\infty}, \quad \mathbf{H}(t) = \mathrm{diag}\{h_1(t), \dots, h_2(t)\} \tag{23}$$

is a weak Lyapunov function for system (1).

Proof. Given the properties of the vector function $\mathbf{h}(t)$, in both discrete-time and continuous-time cases $V(t, \mathbf{x}(t)) > 0$ for any t and $\mathbf{x}(t) \neq 0$.

(a) In the discrete-time case, $V(t, \mathbf{x}(t))$ is a weak Lyapunov function for system (1) means:

$$\forall t \in \mathbf{T} = \mathbf{Z}_+,\ \forall \mathbf{x}(t) \in \mathbf{R}^n \setminus \{0\} : \frac{V(t+1, \mathbf{x}(t+1))}{V(t, \mathbf{x}(t))} \tag{24a}$$

which can be also written as:

$$\forall t \in \mathbf{T} = \mathbf{Z}_+,\quad \forall \mathbf{x}(t) \in \mathbf{R}^n \setminus \{0\} : \ \frac{\| (\mathbf{H}^{-1}(t+1)\mathbf{A}\mathbf{H}(t))(\mathbf{H}^{-1}(t)\mathbf{x}(t)) \|_{\infty}}{\| \mathbf{H}^{-1}(t)\mathbf{x}(t) \|_{\infty}} \leq 1. \tag{25a}$$

If (25a) is true, then we have:

$$\forall t \in \mathbf{T} = \mathbf{Z}_+ : \max_{\| \mathbf{H}^{-1}(t)\mathbf{x}(t) \|_{\infty} = 1} \| (\mathbf{H}^{-1}(t+1)\mathbf{A}\mathbf{H}(t))(\mathbf{H}^{-1}(t)\mathbf{x}(t)) \|_{\infty} \leq 1, \tag{26a}$$

that is equivalent to the boundedness of the operator norm:

$$\forall t \in \mathbf{T} = \mathbf{Z}_+ : \| \mathbf{H}^{-1}(t+1)\mathbf{A}\mathbf{H}(t) \|_{\infty} \leq 1. \tag{27a}$$

Now, taking into account the equality:

$$\forall t \in \mathbf{T} = \mathbf{Z}_+ : \ \| \mathbf{H}^{-1}(t+1)\mathbf{A}\mathbf{H}(t) \|_{\infty} = \| \mathbf{H}^{-1}(t+1)\bar{\mathbf{A}}\mathbf{H}(t) \|_{\infty}, \tag{28a}$$

relationship (27a) yields:

$$\forall t \in \mathbf{T} = \mathbf{Z}_+ : \| \mathbf{H}^{-1}(t+1)\bar{\mathbf{A}}\mathbf{H}(t) \|_{\infty} \leq 1, \tag{29a}$$

which means that inequality (4) is satisfied with $\mathbf{h}(t)$ meeting conditions in Definition 1, i.e. system (1) is CWAS with respect to $\mathbf{h}(t)$.

Conversely, if system (1) is CWAS with respect to $\mathbf{h}(t)$, then relationship (27a) holds and allows writing:

$$\begin{aligned} &\forall t \in \mathbf{T} = \mathbf{Z}_+,\ \forall \mathbf{x}(t) \in \mathbf{R}^n \setminus \{0\} : \ \frac{\| (\mathbf{H}^{-1}(t+1)\mathbf{A}\mathbf{H}(t))(\mathbf{H}^{-1}(t)\mathbf{x}(t)) \|_{\infty}}{\| \mathbf{H}^{-1}(t)\mathbf{x}(t) \|_{\infty}} \leq \\ &\leq \frac{\| \mathbf{H}^{-1}(t+1)\mathbf{A}\mathbf{H}(t) \|_{\infty} \| \mathbf{H}^{-1}(t)\mathbf{x}(t) \|_{\infty}}{\| \mathbf{H}^{-1}(t)\mathbf{x}(t) \|_{\infty}} = \| \mathbf{H}^{-1}(t+1)\mathbf{A}\mathbf{H}(t) \|_{\infty} \leq 1, \end{aligned} \tag{30a}$$

which shows that (25a) is true, i.e. $V(t,\mathbf{x}(t))$ defined by (23) is a weak Lyapunov function.

(b) In the continuous-time case, $V(t,\mathbf{x}(t))$ is a weak Lyapunov function for system (1) means that $V(t,\mathbf{x}(t))$ is nonincreasing along any trajectory of system (1), i.e.

$$\forall t\in\mathbf{T}=\mathbf{R}_+,\quad \forall\mathbf{x}(t)\in\mathbf{R}^n\setminus\{0\},\ \forall\tau>0: V(t+\tau,\mathbf{x}(t+\tau))-V(t,\mathbf{x}(t))\le 0, \quad (24b)$$

which, for small $\tau>0$ can be also written as:

$$\forall t\in\mathbf{T}=\mathbf{R}_+,\forall\mathbf{x}(t)\in\mathbf{R}^n\setminus\{0\}:\frac{\|(\mathbf{H}^{-1}(t+\tau)(I+\tau\mathbf{A})\mathbf{H}(t))(\mathbf{H}^{-1}(t)\mathbf{x}(t))\|_\infty}{\|\mathbf{H}^{-1}(t)\mathbf{x}(t)\|_\infty}\le 1$$
$$. \quad (25b)$$

If (25b) is true, then, for small $\tau>0$, we have:

$$\forall t\in\mathbf{T}=\mathbf{R}_+,\ \max_{\|\mathbf{H}^{-1}(t)\mathbf{x}(t)\|_\infty=1}\|(\mathbf{H}^{-1}(t+\tau)(\mathbf{I}+\tau\mathbf{A})\mathbf{H}(t))(\mathbf{H}^{-1}(t)\mathbf{x}(t))\|_\infty\le 1, (26b)$$

that is equivalent to boundedness of the operator norm:

$$\forall t\in\mathbf{T}=\mathbf{R}_+:\|\mathbf{H}^{-1}(t+\tau)(\mathbf{I}+\tau\mathbf{A})\mathbf{H}(t)\|_\infty\le 1. \quad (27b)$$

Now, taking into account the equality:

$$\forall t\in\mathbf{T}=\mathbf{R}_+:\ \|\mathbf{H}^{-1}(t+\tau)(\mathbf{I}+\tau\mathbf{A})\mathbf{H}(t)\|_\infty=\|\mathbf{H}^{-1}(t+\tau)(\mathbf{I}+\tau\bar{\mathbf{A}})\mathbf{H}(t)\|_\infty, (28b)$$

valid for small $\tau>0$, relationship (27b) yields:

$$\forall t\in\mathbf{T}=\mathbf{R}_+:\|\mathbf{H}^{-1}(t+\tau)(\mathbf{I}+\tau\bar{\mathbf{A}})\mathbf{H}(t)\|_\infty\le 1, \quad (29b)$$

which means that inequality (4) is satisfied with **h**(t) meeting conditions in Definition 1, i.e. system (1) is CWAS with respect to **h**(t).

Conversely, if system (1) is CWAS with respect to **h**(t), then relationship (27b) holds and allows writing:

$$\forall t\in\mathbf{T}=\mathbf{R}_+,\forall\mathbf{x}(t)\in\mathbf{R}^n\setminus\{0\}:\frac{\|(\mathbf{H}^{-1}(t+\tau)(\mathbf{I}+\tau\mathbf{A})\mathbf{H}(t))(\mathbf{H}^{-1}(t)\mathbf{x}(t))\|_\infty}{\|\mathbf{H}^{-1}(t)\mathbf{x}(t)\|_\infty}\le$$
$$\le\frac{\|\mathbf{H}^{-1}(t+\tau)(\mathbf{I}+\tau\mathbf{A})\mathbf{H}(t)\|_\infty\|\mathbf{H}^{-1}(t)\mathbf{x}(t)\|_\infty}{\|\mathbf{H}^{-1}(t)\mathbf{x}(t)\|_\infty}=\|\mathbf{H}^{-1}(t+\tau)(\mathbf{I}+\tau\mathbf{A})\mathbf{H}(t)\|_\infty\le 1,$$
$$(30b)$$

which shows that (25b) is true, i.e. $V(t,\mathbf{x}(t))$ defined by (23) is a weak Lyapunov function. ■

For the particular case when testing CWEAS and the vector function

$\mathbf{h}(t)$ considered in Definition 1 is of exponential type (see Definition 2), the explicit time-dependence of the Lyapunov function becomes redundant as shown below.

Theorem 11 *System (1) is CWEAS with* $d_i > 0, i = 1, \cdots, n$ *, if and only if*

$$V(\mathbf{x}(t)) = \| \mathbf{x}(t) \|_{\mathbf{D}\infty} \tag{31}$$

is a strong Lyapunov function.

Proof. Given the particular form of matrix D used in (31), $V(\mathbf{x}(t)) > 0$ for any t and $\mathbf{x}(t) \neq 0$), in both discrete-time and continuous-time cases.

(a) In the discrete-time case, $V(\mathbf{x}(t))$ is a strong Lyapunov function for system (1) means:

$$\forall t \in \mathbf{T} = \mathbf{Z}_+, \ \forall \mathbf{x}(t) \in \mathbf{R}^n \setminus \{0\} : \ \frac{V(\mathbf{x}(t+1))}{V(\mathbf{x}(t))} < 1, \tag{32a}$$

which can be also written as:

$$\forall t \in \mathbf{T} = \mathbf{Z}_+, \ \forall \mathbf{x}(t) \in \mathbf{R}^n \setminus \{0\} : \ \frac{\| (\mathbf{D}^{-1}\mathbf{A}\mathbf{D})(\mathbf{D}^{-1}\mathbf{x}(t)) \|_\infty}{\| \mathbf{D}^{-1}\mathbf{x}(t) \|_\infty} < 1 . \tag{33a}$$

If (33a) is true, then we have:

$$\forall t \in \mathbf{T} = \mathbf{Z}_+ : \max_{\|\mathbf{D}^{-1}(t)\mathbf{x}(t)\|_\infty = 1} \| (\mathbf{D}^{-1}\mathbf{A}\mathbf{D})(\mathbf{D}^{-1}(t)\mathbf{x}(t)) \|_\infty < 1, \tag{34a}$$

that is equivalent to boundedness of the operator norm:

$$\forall t \in \mathbf{T} = \mathbf{Z}_+ : \| \mathbf{D}^{-1}\mathbf{A}\mathbf{D} \|_\infty < 1 . \tag{35a}$$

Thus, we have shown that

$$\mu_{\mathbf{D}\infty}(\mathbf{A}) = \| \mathbf{A} \|_{\mathbf{D}\infty} = \| \mathbf{D}^{-1}\mathbf{A}\mathbf{D} \|_\infty < 1, \tag{36a}$$

which, in accordance with Theorem 9, ensures CWEAS of system (1) with $d_i > 0, \quad i = 1, \cdots, n$.

Conversely, CWEAS of system (1) with $d_i > 0, \quad i = 1, \cdots, n$, means CWAS with respect to $\mathbf{h}(t) = \mathbf{d}r^t$, $0 < r < 1$, which, according to Theorem 10, is equivalent to:

$$\forall t \in \mathbf{T} = \mathbf{Z}_+, \ \forall \mathbf{x}(t) \in \mathbf{R}^n \setminus \{0\} : \ \frac{\| (r^{-(t+1)}\mathbf{D}^{-1}\mathbf{A}r^t\mathbf{D})(r^{-t}\mathbf{D}^{-1}\mathbf{x}(t)) \|_\infty}{\| r^{-t}\mathbf{D}^{-1}\mathbf{x}(t) \|_\infty} \le 1 \tag{37a}$$

or, furthermore:

$$\forall t \in \mathbf{T} = \mathbf{Z}_+, \ \forall \mathbf{x}(t) \in \mathbf{R}^n \setminus \{0\}: \ \frac{\| (\mathbf{D}^{-1}\mathbf{A}\mathbf{D})(\mathbf{D}^{-1}\mathbf{x}(t)) \|_\infty}{\| \mathbf{D}^{-1}\mathbf{x}(t) \|_\infty} \le r < 1. \quad (38a)$$

Thus, we have proved the validity of (33a) and, consequently of (32a), i.e. $V(\mathbf{x}(t))$ is a strong Lyapunov function for system (1).

(b) In the continuous-time case, $V(\mathbf{x}(t))$ is a strong Lyapunov function for system (1) means:

$$\forall t \in \mathbf{T} = R_+, \ \forall \mathbf{x}(t) \in \mathbf{R}^n : \lim_{\tau \to 0+} \frac{V(\mathbf{x}(t+\tau)) - V(\mathbf{x}(t))}{\tau} < 0, \quad (32b)$$

which, for small $\tau > 0$ can be also written as:

$$\forall t \in \mathbf{T} = \mathbf{R}_+, \ \forall \mathbf{x}(t) \in \mathbf{R}^n \setminus \{0\}: \ \frac{\| \mathbf{D}^{-1}(\mathbf{I} + \tau\mathbf{A})\mathbf{D}\mathbf{D}^{-1}\mathbf{x}(t)) \|_\infty}{\| \mathbf{D}^{-1}\mathbf{x}(t) \|_\infty} < 1. \quad (33b)$$

If (33b) is true, then, for small $\tau > 0$, we have:

$$\forall t \in \mathbf{T} = \mathbf{R}_+, \ \max_{\|\mathbf{D}^{-1}\mathbf{x}(t)\|_\infty = 1} \| (\mathbf{D}^{-1}(\mathbf{I} + \tau\mathbf{A})\mathbf{D})(\mathbf{D}^{-1}\mathbf{x}(t)) \|_\infty < 1, \quad (34b)$$

that is equivalent to boundedness of the operator norm:

$$\forall t \in \mathbf{T} = \mathbf{R}_+ : \| \mathbf{D}^{-1}(\mathbf{I} + \tau\mathbf{A})\mathbf{D} \|_\infty < 1. \quad (35b)$$

Thus, we have shown that

$$\mu_{\mathbf{D}\infty}(\mathbf{A}) = \lim_{\tau \to 0+} \frac{\| \mathbf{I} + \tau\mathbf{A} \|_{\mathbf{D}\infty} - 1}{\tau} = \lim_{\tau \to 0+} \frac{\| \mathbf{D}^{-1}(\mathbf{I} + \tau\mathbf{A})\mathbf{D} \|_\infty - 1}{\tau} < 0, \quad (36b)$$

which, in accordance with Theorem 9, ensures CWEAS of system (1) with $d_i > 0, \quad i = 1, \cdots, n$.

Conversely, CWEAS of system (1) with $d_i > 0, \quad i = 1, \cdots, n$, means CWAS with respect to $\mathbf{h}(t) = \mathbf{d}e^{rt}$, $r < 0$, which, according to Theorem 10, is equivalent to:

$$\forall t \in \mathbf{T} = \mathbf{R}_+, \forall \mathbf{x}(t) \in \mathbf{R}^n \setminus \{0\}:$$
$$\frac{\| (e^{-r(t+\tau)}\mathbf{D}^{-1}(\mathbf{I} + \tau\mathbf{A})e^{rt}\mathbf{D})(e^{-rt}\mathbf{D}^{-1}\mathbf{x}(t)) \|_\infty}{\| e^{-rt}\mathbf{D}^{-1}\mathbf{x}(t) \|_\infty} \le 1 \quad (37b)$$

or, furthermore:

$$\forall t \in \mathbf{T} = \mathbf{R}_+, \ \forall \mathbf{x}(t) \in \mathbf{R}^n \setminus \{0\}: \ \frac{\| \mathbf{D}^{-1}(\mathbf{I} + \tau\mathbf{A})\mathbf{D}\mathbf{D}^{-1}\mathbf{x}(t)) \|_\infty}{\| \mathbf{D}^{-1}\mathbf{x}(t) \|_\infty} \le e^{r\tau} < 1. \quad (38b)$$

Thus, we have proved the validity of (33b) and, consequently, of (32b), i.e. $V(x(t))$ is a strong Lyapunov function for system (1). ■

Remark 3. In papers (Kiendl et al, 1992), (Polanski, 1995), (Loskot et al, 1998) the usage of Lyapunov function (31) is understood in the sense of standard AS, but pointing out the invariance of a time-independent polyhedral set. Papers (Blanchini, 1994; 1995) notice that Lyapunov function (31) induces a time-dependence of exponential type for the invariant polyhedral sets; however the stability analysis is addressed within the classical framework, without any interpretation of the componentwise meaning. Moreover, the case of invariant polyhedral sets with arbitrary time-dependence (not only exponential) remains completely ignored by these two papers. □

6. Conclusions

By using the infinity norm, well-known results from the classical theory of stability can be particularized so as to characterize CWAS / CWEAS as a special type of asymptotic stability. Thus, our approach allows developing connections between the dynamics of system (1) and CWAS / CWEAS, by circumventing the usage of auxiliary system (4) and applying standard tools in stability theory directly to system (1). The key results refer to the exploitation of the following instruments: ε - δ formalism (Theorem 6), properties of operator **A** (Theorem 9), time-dependent Lyapunov functions for testing CWAS with respect to an arbitrary vector function (Theorem 10) and time-independent Lyapunov functions for testing CWEAS (Theorem 11).

References

Blanchini, F. (1994) "Ultimate boundedness control for uncertain discrete-time systems via set-induced Lyapunov functions", *IEEE Trans. on Aut. Control,* vol. 39, pp. 428-433.

Blanchini, F. (1995). "Nonquadratic Lyapunov functions for robust control", *Automatica.,* vol. 31, pp.2061-2070.

Blanchini, F. (1999). "Set invariance in control - Survey paper", *Automatica*, vol. *35*, pp. 1747-1767.

Deutsch, E. (1975). "On matrix norms and logarithmic norms", *Numerische Mathematik*, vol. *24*, pp. 49-51.

Hmamed, A. (1996). "Componentwise stability of continuous-time delay linear systems", *Automatica*, vol. 32, pp. 651-653.

Hmamed, A. (1997). "Componentwise stability of 1-D and 2-D linear discrete systems", *Automatica,* vol. *33*, pp. 1759-1762.

Kiendl, H., J. Adamy and P. Stelzner (1992). "Vector norms as Lyapunov functions for linear systems", *IEEE Trans. on Aut. Control,* vol. 37, pp. 839-842.

Loskot, K., A. Polanski and A. Rudnicki (1998). "Further comments on vector norms as Lyapunov functions for linear systems", *IEEE Trans. on Aut. Control,* vol. 43, pp. 289-291.

Michel, A. N. and K. Wang (1995). *Qualitative Theory of Dynamical Systems*, Marcel Dekker, New York.

Pastravanu, O. and M. Voicu (1999). "Flow-invariant rectangular sets and componentwise asymptotic stability of interval matrix systems", in *Proc. 5th European Control Conference*, Karlsruhe, CDROM.

Pastravanu, O. and M. Voicu (2002). "Interval matrix systems - Flow-invariance and componentwise asymptotic stability", *Differential and Integral Equations* vol. 15, pp. 1377-1394.

Pavel, H. N. (1984). *Differential Equations, Flow-Invariance and Applications*. Pitman, Boston.

Polanski, A. (1995). "On infinity norms as Lyapunov functions for linear systems", *IEEE Trans. on Aut. Control,* vol. 40, pp. 1270-1273.

Polyak, B. T. and P. S. Shcherbakov (2002). "Superstable linear control systems. I. Analysis", Avtomat. i Telemekh, pp.37--53, (Russian).

Stoer, J. and C. Witzgall (1962)."Transformations by diagonal matrices in a normed space", *Numerische Mathematik*, vol. *4*, pp.158–171.

Voicu, M. (1984a). "Free response characterization via flow-invariance", in *Prep. 9th IFAC World Congress*, Budapest, vol. 5, pp. 12–17.

Voicu, M. (1984b). "Componentwise asymptotic stability of linear constant dynamical systems", *IEEE Trans. Automat. Control*, vol. 10, pp. 937–939.

Voicu, M. (1987). "On the application of the flow-invariance method in control theory and design", in *Prep. 10th IFAC World Congress*, Munich, vol. 8, pp. 364–369.

Voicu, M., Pastravanu O. (2003). "Flow-invariance method in control: a survey of some results", in Voicu M. (Ed.), *Advances in Automatic Control*, Kluwer, pp. 393–434 (in this volume).

INDEPENDENT COMPONENT ANALYSIS WITH APPLICATION TO DAMS DISPLACEMENTS MONITORING

Theodor D. Popescu
National Institute for Research and Development in Informatics
8-10 Maresal Averescu Avenue, 71316 Bucharest, Romania
E-mail: pope@u3.ici.ro

Abstract Independent Component Analysis (ICA) is an emerging field of fundamental research with a wide range of applications such as remote sensing, data communications, speech processing and medical diagnosis. It is motivated by practical scenarios that involve multisources and multisensors. The key objective of ICA is to retrieve the source signals without resorting to any a priori information about the source signals and the transmission channel. ICA using second-order statistics and high-order statistics based techniques and the corresponding algorithms will be presented to perform the blind separation of stationary or cyclostationary sources. In the last part of the paper, a case study with real data having as subject dams displacements monitoring will be presented.

Keywords: independent component analysis, blind source separation, second-order statistics, high-order statistics, large dams monitoring

1. Independent component analysis

1.1. Problem Formulation

Independent Component Analysis (ICA) is a statistical and computational technique, that can be seen as an extension to Principal Component Analysis (PCA) and Factor Analysis (FA)(Hyvärinen, *et al.*, 2001). ICA is a much more powerful technique, capable of finding the underlying factors or sources when these classic methods fail completely. The data analysed by ICA could originate from many different kinds of application fields, including digital images, economic indicators and psychometric measurements.

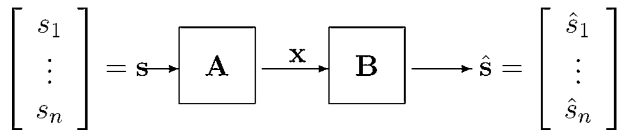


Figure 1. Mixing and separating. Unobserved signals: $\mathbf{s}$; observations: $\mathbf{x}$; estimated source signals: $\hat{\mathbf{s}}$

The simple ICA model assumes the existence of n independent signals $s_1(t), \dots, s_n(t)$ and the observation of as many mixtures $x_1(t), \dots, x_n(t)$, these mixtures being linear and instantaneous, i.e.

$$x_i(t) = \sum_{j=1}^{n} a_{ij} s_j(t) \tag{1}$$

for each $i = 1, n$. This is compactly represented by the mixing equation

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \tag{2}$$

where $\mathbf{s}(t) = [s_1(t), \dots, s_n(t)]^T$ is an $n \times 1$ column vector collecting the source signals, vector $\mathbf{x}(t)$ similary collects the n observed signals and the square $n \times n$ "mixing matrix" $\mathbf{A}$ contains the mixture coefficients. The ICA problem consists in recovering the source vector $\mathbf{s}(t)$ using only the observed data $\mathbf{x}(t)$, the assumption of independence between the entries of the input vector $\mathbf{s}(t)$ and possible some a priori information about the probability distribution of the inputs. It can be formulated as the computation of an $n \times n$ "separating matrix" $\mathbf{B}$ whose output $\hat{\mathbf{s}}(t)$

$$\hat{\mathbf{s}}(t) = \mathbf{B}\mathbf{x}(t) \tag{3}$$

is an estimate of the vector $\mathbf{s}(t)$ of the source signals (see Figure 1).

ICA is closely related to the method Blind Source Separation (BSS) or blind signal separation. A "source" means here an original signal, i.e. independent component. "Blind" means that we no very little, if anything, on the mixing matrix, and make little assumptions on the source signals. ICA is one method, perhaps the most widely used, for performing blind source separation.

In many applications, it would be more realistic to assume that there is some noise in the measurements, which would mean adding a noise term in the model:

$$\begin{aligned} \mathbf{y}(t) &= \mathbf{A}\mathbf{s}(t) \\ \mathbf{x}(t) &= \mathbf{y}(t) + \mathbf{n}(t). \end{aligned} \tag{4}$$

1.2. Identificability of the ICA model

The identificability of the noise-free ICA model has been treated in Comon (1994). By imposing the following fundamental restrictions (in addition to the basic assumption of statistical independence), the identifiability of the model can be assured:

1 All the independent components s_i with the possible exception of one component, must be non-Gaussian.
2 The number of the observed linear mixtures m must be at least as large as the number of independent components n, i.e. $m \geq n$.
3 The matrix $\mathbf{A}$ must be of full column rank.

Usually, it is also assumed that $\mathbf{x}$ and $\mathbf{s}$ are centered. If $\mathbf{x}$ and $\mathbf{s}$ are interpreted as stochastic processes instead of simply random variables, additional restrictions are necessary. At the minimum, one has to assume that the stochastic processes are stationary in the strict sense. Some restriction of ergodicity with respect to the quantities estimated are also necessary.

In the ICA model of eq. (2), it is easy to see that the following ambiguities will hold:

1 We cannot determine the variances (energies) of the independent components. The reason is that, both $\mathbf{s}$ and $\mathbf{A}$ being unknown, any scalar multiplier in one of the sources s_i could always be cancelled by dividing the corresponding column $\mathbf{a}_i$ in $\mathbf{A}$ by the same scalar. As a consequence we may quite as well fix the magnitudes of the independent components; as they are random variables, the most natural way to do this is to assume that each has unit variance: $E[s_i^2] = 1$. Then the matrix $\mathbf{A}$ will be adapted in the ICA solution methods to take into account this restriction.
2 We cannot determine the order of the independent components. The reason is that, again both $\mathbf{s}$ and $\mathbf{A}$ being unknown, we can freely change the order of the terms in the sum (1), and call any of the independent components the first one.

1.3. Algorithms for ICA

Independent Component Analysis is mainly performed using the information on signal statistics. When the signals are temporal coherent, it is possible to solve the problem using only the second-order statistics, but if the signals are temporal white or have identical normalized spectral densities, without any information on a priori source distributions, the solution will need using of order statistics higher then second order for the received signals. If the source signal distributions are known, the problem could be solved by maximum likelihood method. We underline that in the case of source signals temporal white and Gaussian, the blind source separation problem has not solution.

In the next two sections we present two approaches: the first supposes the signals temporal coherent and exploits the second-order statistics using intercovariance matrix of observations, and the second supposes the signals white temporal and exploits statistics of order higher than two, using non-linear functions.

2. ICA using second-order statistics

2.1. Second-Order Information

The first step of the ICA procedure (Belouchrani, *et al.*, 1997) consists of prewhitening the signal part $\mathbf{y}(t)$ of the observation. This is done via a whitening matrix $\mathbf{W}$, i.e. a $n \times m$ matrix (we consider n sources and m mixtures) such that $\mathbf{Wy}(t)$ is spatially white. The whiteness condition is

$$\mathbf{I}_n = \mathbf{W}\mathbf{R}_y\mathbf{W}^T = \mathbf{W}\mathbf{A}\mathbf{A}^T\mathbf{W}^T, \tag{5}$$

where $\mathbf{I}_n$ denotes the $n \times n$ identity matrix. Equations (5) implies that $\mathbf{WA}$ is a unitary matrix: for any whitening matrix $\mathbf{W}$, it then exists a unitary matrix $\mathbf{U}$ such that $\mathbf{WA} = \mathbf{U}$. As a consequence, matrix $\mathbf{A}$ can be factored as

$$\mathbf{A} = \mathbf{W}^{\#}\mathbf{U} = \mathbf{W}^{\#}[\mathbf{u}_1, \dots, \mathbf{u}_n], \tag{6}$$

where $\#$ denotes the pseudoinverse and $\mathbf{U}$ is unitary. The use of second-order information - in the form of an estimate of $\mathbf{R}_y(0)$ which is used to solve for $\mathbf{W}$ in (5) - reduces the determination of the $m \times n$ mixing matrix $\mathbf{A}$ to the determination of a unitary $n \times n$ matrix $\mathbf{U}$. The whitened process $\mathbf{x}_w(t) = \mathbf{Wx}(t)$ still obeys a linear model:

$$\mathbf{x}_w(t) \overset{def}{=} \mathbf{Wx}(t) = \mathbf{W}(\mathbf{As}(t) + \mathbf{n}(t)) = \mathbf{Us}(t) + \mathbf{Wn}(t). \tag{7}$$

The signal part of the whitened process now is a unitary mixture of the source signals. Note that all the information contained in the covariance is 'exhausted' after the whitening, in the sense that changing $\mathbf{U}$ in (7) to any other unitary matrix leaves unchanged the covariance of $\mathbf{x}_w(t)$.

2.2. Whitening Matrix Computation

This step is implemented via eigendecomposition of the sample covariance matrix $\hat{\mathbf{R}}_y(0)$. We consider here that the noise covariance is of the form $\mathbf{R}_n(0) = \sigma^2\mathbf{I}_n$. The whitening procedure is the following:

1 Estimate the covariance matrix $\hat{\mathbf{R}}_x(0)$ using T samples of the observations:

$$\hat{\mathbf{R}}_x(0) = \frac{1}{T}\sum_{t=1}^{T}\mathbf{x}(t)\mathbf{x}(t)^T. \tag{8}$$

2 Perform the eigendecomposition of the $\hat{\mathbf{R}}_x(0)$ covariance matrix

$$\hat{\mathbf{R}}_x(0) = \mathbf{H}\Delta\mathbf{H}^T, \tag{9}$$

where

$$\mathbf{H} = [\mathbf{h}_1, ..., \mathbf{h}_m]$$

and

$$\Delta = diag[\lambda_1, \ldots, \lambda_m]$$

with $\lambda_i \geq \lambda_j$ for $i < j$. The number of sources can be estimated starting from the spectrum Δ (Wax and Kailath, 1983; Yin and Krishnaiah, 1987).

3 Estimate noise variance $\hat{\sigma}^2$ as the average of the $m - n$ smallest eigenvalues of Δ

$$\hat{\sigma}^2 = \frac{1}{m-n}\sum_{i=n+1}^{m}\lambda_i. \tag{10}$$

4 Compute the whitening matrix $\hat{\mathbf{W}}$ as:

$$\hat{\mathbf{W}} = \Delta^{'}\mathbf{H}^{'T}, \tag{11}$$

where

$$\Delta^{'} = diag[(\lambda_1 - \hat{\sigma}^2)^{-1/2}, \ldots, (\lambda_n - \hat{\sigma}^2)^{-1/2}]$$

and

$$\mathbf{H}^{'} = [\mathbf{h}_1, \ldots, \mathbf{h}_n].$$

This resulted matrix is used to obtain the whitened process

$$\hat{\mathbf{x}}_w(t) = \hat{\mathbf{W}}\mathbf{x}(t), \quad t = 1, \ldots, T. \tag{12}$$

2.3. Intercovariance Matrix Estimation

Starting from the whitened process $\mathbf{x}_w(t)$, K intercovarince matrices of this process are computed:

$$\hat{\mathbf{R}}_w(k) = \frac{1}{T-k}\sum_{t=k+1}^{T}\mathbf{x}_w(t)\mathbf{x}_w(t-k)^T, \tag{13}$$

where $1 \geq k \geq K$. The resulted matrices are of $n \times n$ dimension, and the computation effort does not depend of number of sensors, m. The value of K will be selected to realize a trade off between the statistic efficiency and computation effort. The value of the delays used in computation depends also on the length of the signal correlations. If we have a priori information on spectral density of sources, the value of K can be optimal chosen.

2.4. Joint Diagonalization

Let $\mathbf{R}_w = \{\mathbf{R}_w(k) | 1 \leq k \leq K\}$ be a set of K matrices with common size $n \times n$. A joint diagonalizer of the set $\mathbf{R}_w$ is defined as a unitary maximizer of the criterion

$$C(\mathbf{U}) \stackrel{def}{=} \sum_{k=1}^{K} |diag(\mathbf{U}^T \mathbf{R}_w(k) \mathbf{U})|^2, \tag{14}$$

where $|diag(\cdot)|$ is the norm of the vector build from the diagonal of the matrix argument. The problem is solved by a generalization of Jacobi technique (Golub and Loan, 1989; Souloumiac and Cardoso, 1991; 1994).

2.5. Mixing Matrix and Source Signals Estimation

Let $\hat{\mathbf{U}} = [\hat{\mathbf{u}}_1, \ldots, \hat{\mathbf{u}}_n]$ be the unitary matrix resulted by joint diagonalization. If the objective of the blind identification is source separation, a brute estimation of these can be computed by:

$$\hat{\mathbf{s}}(t) = \hat{\mathbf{U}}^T \hat{\mathbf{x}}_w(t). \tag{15}$$

To estimate the mixing matrix need to inverse the effect of whitening, and the mixing matrix can be estimated by

$$\hat{\mathbf{A}} = \hat{\mathbf{W}}^{\#} \hat{\mathbf{U}}. \tag{16}$$

To obtain at the output of the separator a maximum signal/noise ratio the source signals are estimated by

$$\hat{\mathbf{s}}(t) = \hat{\mathbf{A}}^T \hat{\mathbf{R}}_x(0)^{-1} \mathbf{x}(t). \tag{17}$$

2.6. The Algorithm

The general scheme of the SOBI algorithm (Second Order Blind Identification) can now be described by the following steps:

Step 1. Form the sample covariance $\hat{\mathbf{R}}_x(0)$ and compute the whitening matrix $\hat{\mathbf{W}}$.

Step 2. Whitening the data provided by the sensors:

$$\hat{\mathbf{x}}_w(t) = \hat{\mathbf{W}}\mathbf{x}(t), \quad t = 1, \ldots, T.$$

Step 3. Estimate K intercovariance matrices $\hat{\mathbf{R}}_w(k)$ of $\hat{\mathbf{x}}_w(t)$ for different time delay $k = 1, \ldots, K$.
Step 4. Jointly diagonalize the set of intercovariace matrices in a base $\hat{\mathbf{U}} = [\hat{\mathbf{u}}_1, \ldots, \hat{\mathbf{u}}_n]$.
Step 5. Estimate the mixing matrix with

$$\hat{\mathbf{A}} = \hat{\mathbf{W}}^{\#}\hat{\mathbf{U}}.$$

Step 6. Estimate the source signals by

$$\hat{\mathbf{s}}(t) = \hat{\mathbf{A}}^T\hat{\mathbf{R}}_x(0)^{-1}\mathbf{x}(t).$$

Note that at the second step of the algorithm the observation dimension is reduced to n, the source number. It results that the intercovariance matrices estimation is performed in a space of reduced dimension.

3. ICA using high-order statistics

In the basic approach to solve ICA problem, the temporal structure of the received signals is in fact omitted and $\mathbf{s}(t)$ and $\mathbf{x}(t)$ are regarded as realizations of random vectors $\mathbf{s}$ and $\mathbf{x}$. We seek the solution of the form (3).

The problem for solving the separating matrix $\mathbf{B}$ is somewhat simplified if we consider only one of the source signals at a time. From equation (3) it follows:

$$\hat{s}_i = \mathbf{b}_i^T\mathbf{x} \tag{18}$$

with $\mathbf{b}_i^T$ the i-th row of $\mathbf{B}$.

The problem is further simplified by performing a prewhitening of the data $\mathbf{x}$: the observed vector $\mathbf{x}$ is firstly linearly transformed to another vector whose elements are mutually uncorrelated and all have unit variance. It can be shown that after this step, $\mathbf{B}$ will be an orthogonal matrix.

A recent review of various information theoretic contrast functions for solving $\mathbf{B}$, like mutual information, negentropy, maximum entropy, and infomax, as well as the maximum likelihood approach is given by Cardoso (1998).

As an example of contrast functions, consider the case of maximizing the kurtosis $E\{\hat{s}_i^4\} - 3[E\{\hat{s}_i^2\}]^2$ of the estimated signals $\hat{s}_i$. Because we assumed that the estimated signals have unit variance, this reduces to maximizing the fourth moment $E\{\hat{s}_i^4\}$. Its gradient with respect to $\mathbf{b}_i$

is $4E\{(\mathbf{b}_i^T\mathbf{x})^3\mathbf{x}\}$. In a gradient learning type rule, the row $\mathbf{b}_i^T$ of the separating $\mathbf{B}$ would be sought using a version of this gradient, in which the expectation is dropped and the gradient is computed separately for each input vector $\mathbf{x}$. In addition, a normalization term would be needed that keeps the norm of $\mathbf{b}_i$, equal to one - remember that the matrix $\mathbf{B}$ would be orthogonal due to the prewhitening of the data $\mathbf{x}$.

A much more efficient algorithm is the following fixed point iteration (Hyvärinen and Oja, 1997):

1 Take a random initial vector $\mathbf{b}(0)$ of norm 1. Let $k = 1$.

2 Let $\mathbf{b}(k) = E\{\mathbf{x}(\mathbf{b}(k-1)^T\mathbf{x})^3\} - 3\mathbf{b}(k-1)$. The expectation can be estimated using a large sample of $\mathbf{x}$ vectors.

3 Divide $\mathbf{b}(k)$ by its norm.

4 If $|\mathbf{b}(k)^T\mathbf{b}(k-1)|$ is not close enough to 1, let $k = k+1$ and go back to step 2. Otherwise, output the vector $\mathbf{b}(k)$.

The final vector $\mathbf{b}(k)$ given by the algorithm equals the transpose of one of the rows of the (orthogonal) separating matrix $\mathbf{B}$.

To estimate n independent components, we run this algorithm n times. To ensure that we estimate each time a different independent component, we use the deflation algorithm that adds a simple orthogonalizing projection inside the loop. Recall that the rows of the separating matrix $\mathbf{B}$ are orthogonal because of the prewhitening. Thus we can estimate the independent components one by one by projecting the current solution $\mathbf{b}(k)$ on the space orthogonal to the rows of the separating matrix $\mathbf{B}$ previously found.

This algorithm, with the whitening and several extensions, is implemented in Matlab in the FastICA package which is a public domain package. A remarkable property of the FastICA algorithm is that a very small number of iterations seems to be enough to obtain the maximal accuracy allowed by the sample data. This is due to the cubic convergence of the algorithm.

Another algorithm intensively used in practice is JADE (Joint Approximate Diagonalization of Eigen-matrices) (Cardoso and Souloumiac, 1993). It is a typically batch algorithm using tensorial techniques as eigenmatrix decomposition. The algorithm is quite complicated, requiring sophisticated matrix manipulation.

4. Case study - DAMS displacements monitoring

One of the main objectives for dams displacements monitoring is to detect any abnormal behaviour alteration as early as possible. Any change in a dam response under the same loads may be due to a structural

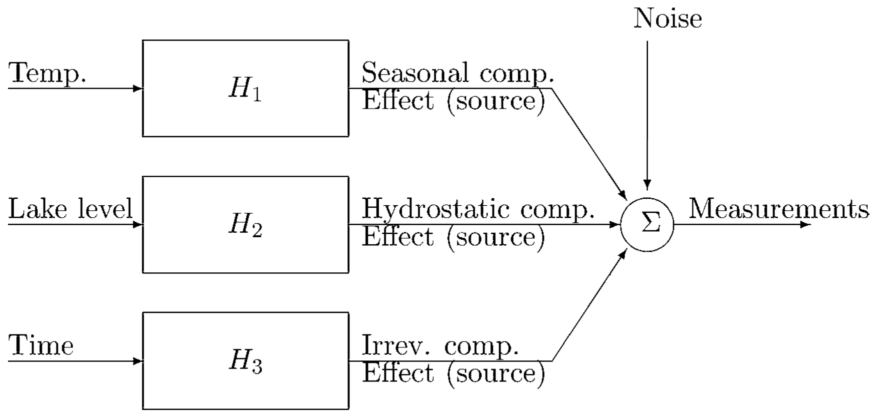


Figure 2. Arch dam physical model

deterioration culminating with the dam collapse. A change detected in real time can be decisive for the possible strengthening works.

Experience (Mazenot, 1971; Ispas, *et al.*, 2000) shows that the values of gross measurements recorded for dams point out the superposition of the following main three components: time, hydrostatic load, and temperature (see Figure 2).

The time or irreversible component corresponds to an evolution in time, those trend is that of being amortized (strengthened) or amplified (deteriorated); the reversible hydrostatic component corresponds to the hydrostatic pressure effect of the lake, while the reversible seasonal component depends on the distribution of temperatures and precipitation.

The objective of the application was to separate the components (sources) mentioned above starting from the displacements of the dam, without a priori knowledge of the generator phenomena or of the propagation environment, and by using only of the raw displacement measures. The application was dedicated to Vidraru dam, Romania, for a period of 1200 days.

The evolution of the dam displacements for x and y directions are given in Figure 3 and Figure 4, respectively, at different levels.

For these displacements, when SOBI algorithm has been used, resulted 3 independent sources which can be assimilated with the hydrostatic pressure component (lake level), seasonal component (temperature) and irreversible component. These are represented together with the lake level and temperature in Figure 4. It can be noted that there are strong similarities between the estimated sources, representing seasonal and hydrostatic components, and temperature and lake level evolutions. The irreversible component, last represented, does not create special problems concerning dam safety.

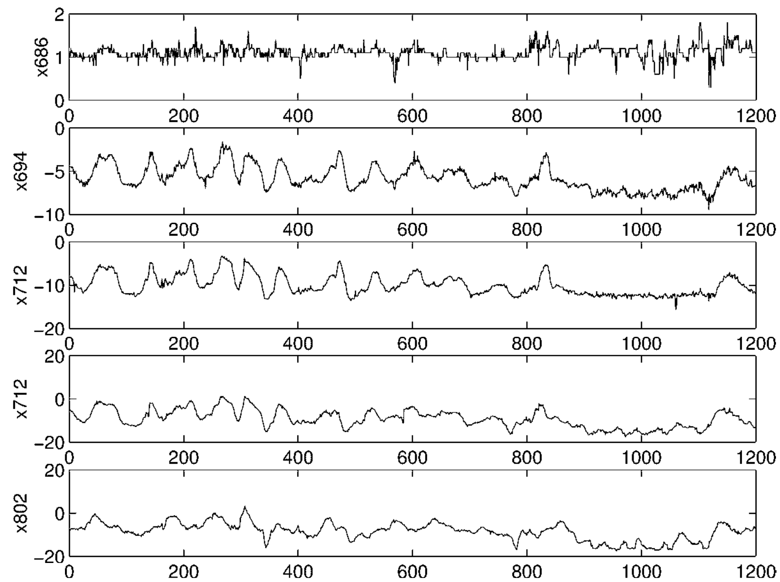


Figure 3. Displacements for x axis at different levels

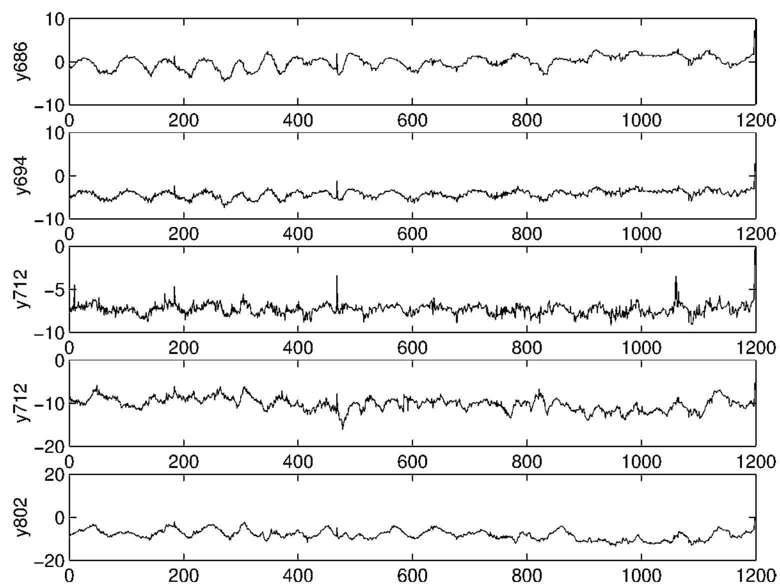


Figure 4. Displacements for y axis at different levels

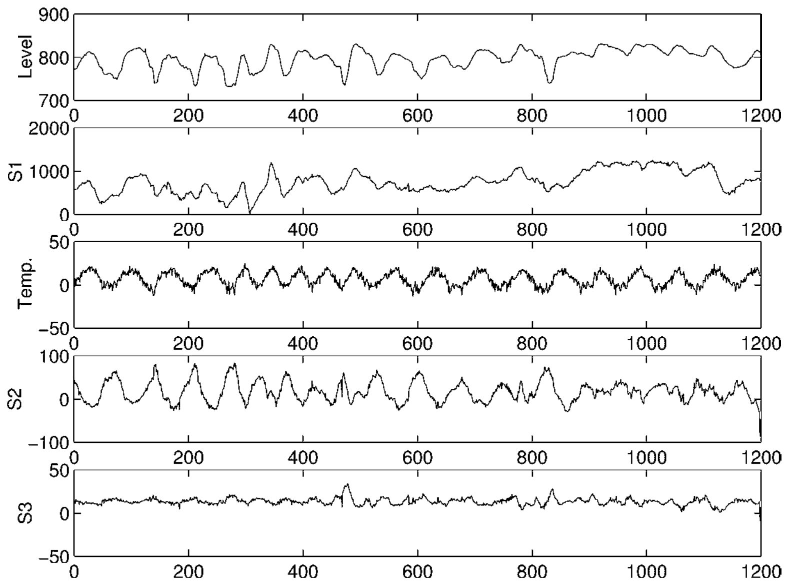


Figure 5. Lake level, estimated source 1, temperature, estimated source 2 and source 3 assimilated with the irreversible component

The results represent only a preliminary analysis of the dam under study. More experiments and data analysis by different methods are necessary for a complete investigation of the dam behaviour.

5. Conclusions

The paper presented some methods and algorithms for independent component analysis based on second-order statistics and high-order statistics, to perform the blind separation of stationary or cyclostationary sources. The SOBI (Second Order Blind Identification) algorithm is described in detail and it is applied in an application having as subject displacements monitoring of an instrumented dam.

References

[1] Belouchrani, A., K. Abed Meraim, J.F. Cardoso and E. Moulines (1997), A blind source separation technique based on second order statistics, *IEEE Trans. on Signal Processing*, **45**, pp. 434-444.

[2] Cardoso, J. F. (1993), and A. Souloumiac, Blind beamforming for non Gaussian signals, *IEE Proceedings - F*,**140**, pp. 362-370.

[3] Cardoso, J. F. (1998), Blind signal separation: statistical principles, *Proceedings of the IEEE*, **9**, pp. 2009-2025.

[4] Comon, P. (1994), Independent Component analysis - a new concept ?, *Signal Processing*, **36**, pp. 287-314.

[5] Golub, G. H. and C.F.V. Loan (1989), *Matrix Computation*, The John Hopkins University Press.

[6] Hyvärinen, A. and E. Oja (1997), A fast fixed-point algorithm for independent component analysis, *Neural Computation*,**9**, pp. 1483-1492.

[7] Hyvärinen, A., J. Karhunen and E. Oja (2001), *Independent Component Analysis*, John Wiley & Sons, Inc., New York.

[8] Ispas, D., C. Scumpu, D. Hulea and Th. Popescu (2000), "Dams and their foundations monitoring by statistic methods", *Hidrotehnica*, Special issue edited by The Romanian Committee on Large Dams, **45**, pp. 37 - 44.

[9] Mazenot, P. (1971), *Methode generale d'interpretation des mesures de surveillance des barrages en exploitation a Electricite de France*, Division Technique Generale.

[10] Popescu, Th. (2002), Dams Displacements Monitoring Using Second Order Blind Identification Algorithm, *Proc. IEEE International Symposium on Intelligent Control (ISIC)*, Vancouver, British Columbia, Canada, 27-30 October.

[11] Souloumiac, A. and J.F. Cardoso (1991), Comparaison de methodes de separation de sources, *Proc. GRETSI*, Juan les Pines.

[12] Souloumiac, A. and J.F. Cardoso (1994), Givens angles for simultaneous diagonalization, *SIAM J. Matrix Anal. Appl.*.

[13] Wax, M. and T. Kailath (1983), Determining the number of signals by information theoretic criteria", *Workshop on spectral estimation II*, Florida, pp. 192-196.

[14] Yin, Y. and P. Krishnaiah (1987), Methods for detection of the number of signals, *IEEE Trans. on ASSP*,**35**, pp. 1533-1538.

FUZZY CONTROLLERS WITH DYNAMICS, A SYSTEMATIC DESIGN APPROACH

Stefan Preitl and Radu-Emil Precup

"Politehnica" University of Timisoara, Dept. of Automation
Bd. V. Parvan 2, RO-1900 Timisoara, Romania
Phone: +40-256-4032-24, -29, -26, Fax: +40-256-403214
E-mail: spreitl@aut.utt.ro, rprecup@aut.utt.ro

Abstract The paper presents aspects concerning the systematic design of fuzzy controllers (of Mamdani type and Takagi-Sugeno type) with dynamics. There are considered PI and PID fuzzy controllers resulting in fuzzy control systems which are type-II and type-III fuzzy systems according to Koczy (1996) and Sugeno (1999). The fuzzy controllers are applicable to a wide range of applications.

Keywords: PI controllers, fuzzy controllers, dynamics, design, digital simulation

1. Introduction

The "classical" engineering approach to the reality is essentially a qualitative and quantitative one, based on a more or less "accurate" mathematical modeling. In this context the elaboration of the control strategy and of the controller requires an "as accurate as possible" quantitative modeling of controlled plant (CP). Some advanced control strategies require even the permanent reassessment of the models and of the parameters values characterizing these (parametric) models. By many aspects the fuzzy control is more pragmatic by the capability to use a linguistic characterization of the quality of CP behavior and to adapt it as function of the concrete conditions of CP operation.

The basic fuzzy controllers (FCs) with dynamics have a specific nonlinear behavior, accompanied by anticipative, derivative, integral and – more general – predictive effects and adaptation possibilities to the concrete operating conditions. The "coloring" of the linguistic characterization of CP evolution – based on experience – will be done by means of parameters which enable the modification of FC features.

In some applications the development of fuzzy control systems (FCSs) is often done by heuristic means that can be sometimes accompanied by failures. A systematic design approach can be advantageous; some design procedures developed by the authors and presented in this paper compensate the lack of general design methods applicable to certain categories of systems.

The authors thank the colleagues from the "Gh. Asachi" Technical University of Iasi for the possibility to meet and share opinions concerning controller development techniques, and for the opportunity offered by the Symposium on Automatic Control and Computer Science where the authors have presented research results; part of these results are included in this paper.

2. Of fuzzy controllers: continuous time analysis

The shape of the non-linearity (Driankov, *et al.*, 1993) of a FC can be modeled in a large variety of forms by an adequate choice of the variable parameters taking part to the FC informational modules. The FCs can obtain dynamic features by additional dynamic processing of some of system variables in terms of differentiation and / or integration. The effects of these components can be reflected either in permanent regimes – by the disturbance rejection or just the alleviation of the control error – or in dynamic ones, by improving the phase margin (in generalized sense), reducing the overshoot, the settling time, and / or relaxing the stability conditions.

The derivative (D) and integral (I) components can be implemented in conventional digital version; these components can create a quasi-continuous (Q-C) equivalent of the analogue D and I components, respectively. Two methods for the accomplishment of Q-C D and I components are presented as follows.

Firstly, for the D component, the usual computation relation is given by the relation (1):

$$d_k = \frac{1}{T_s} \cdot (e_k - e_{k-1}), \quad k \in N^*, \tag{1}$$

with T_s – the sampling period. In the case of a rapid variation of the input variable $e(t)$ which could be harmful on the implementation of the D component, then either e_k can be pre-filtered in terms of a first order delay (PT1) law, or the D component is obtained as function of the actual sample e_k and of an "old sample" e_{k-m}.

Secondly, for the I-component, a version of computation relation is given by the relation (2):

$$\sigma_k = \sum_{i=0}^{k} e_i = e_k + \sum_{i=0}^{k-1} e_i, \quad \text{or} \quad \sigma_k = x_k + e_k \quad \text{with} \quad x_k = \sum_{i=0}^{k-1} e_i. \tag{2}$$

Such characterizations will also permit a relative Q-C equivalence of the digital case; by using the first order Pàde approximation, these two components can be expressed as:

$$d(s) \approx \frac{s}{1+s\cdot T_s/2}\cdot e(s)\,, \qquad \sigma(s) \approx \frac{1+s\cdot T_s/2}{s\cdot T_s}\cdot e(s)\,. \tag{3}$$

The equations (3) ensure a continuous pseudo-transfer function for the FC with dynamics; it justifies the analysis of FCs in the linear case (Siler and Ying, 1989).

By employing the widely accepted experience in the design of PI controllers and the very good control features offered by these controllers (zero steady-state, enhancement of control system (CS) dynamics – alleviation of the settling time and / or of the overshoot – by the pole-zero cancellation technique), the knowledge on linear PI controllers can be incorporated in the properties of strictly speaking FCs (without dynamics).

The PI fuzzy controllers (PI-FCs) are very useful because starting from the features of a basic linear PI controller can systematically develop them. But, the arbitrary introduction of dynamic components in the FC structure creates a lot of difficulties mainly concerning the interpretation of introducing the dynamics in CS behavior in different regimes, and the increase of the number of the degrees of freedom in controller design and implementation. The analysis of the behavior of some FCSs has been performed in (Precup and Preitl, 1995).

There will be obtained two versions of PI-FCs, the position type and the velocity type. The position type PI-FC can be further accomplished in two versions obtaining the integral component on either the output or the input of the FC, in structures of fuzzy controllers of Mamdani (Mamdani, 1974) or of Takagi-Sugeno type (Takagi and Sugeno, 1985).

The Mamdani version of position type PI-FC – presented here – is characterized by the presence of the integral component on FC output, with the basic relation:

$$u(t) = k_i \cdot \int_0^t [k_{F1}\cdot e(\tau) + k_d \cdot k_{F2}\cdot \dot{e}(\tau)]\cdot d\tau\,. \tag{4}$$

The relation (4) characterizes a typical dependence for a PI controller. By expressing (4) in its operational form, the Q-C equivalent of the PI-FC is obtained:

$$u(s) \approx k_i \cdot \frac{1+s\cdot T_s/2}{s\cdot T_s}\cdot\left(k_{F1} + k_{F2}\cdot k_d \cdot \frac{s}{1+s\cdot T_s/2}\right)\cdot e(s)\,. \tag{5}$$

Therefore, the expression of the pseudo-transfer function can be expressed:

$$H_c(s) \approx \frac{k_c}{s} \cdot (1 + s \cdot T_i) \text{ with } k_c = \frac{k_i \cdot k_{F1}}{T_s}, \quad T_i = \frac{k_{F2}}{k_{F1}} \cdot k_d + \frac{T_s}{2}. \quad (6)$$

The position type PI-FC is characterized by the following discrete time equation obtained by differentiating (4) and using (1):

$$u_k = u_{k-1} + T_s \left(k_i \cdot k_{F1} \cdot e_k + k_d \cdot k_{F2} \cdot \frac{e_k - e_{k-1}}{T_s} \right). \quad (7)$$

Hence, the discrete time equation of an incremental PI-FC obtains immediately the form (8):

$$\Delta u_k = (k_i \cdot k_{F1} \cdot T_s + k_d \cdot k_{F2}) \cdot e_k - k_d \cdot k_{F2} \cdot e_{k-1}, \quad (8)$$

where $\Delta u_k = u_k - u_{k-1}$ stands for the increment of control signal.

Using the presented approach there can be also developed many versions of PD fuzzy controllers (PD-FCs) and of PID fuzzy controllers (PID-FCs) (see, for example, Tang and Mulholland (1987) Kawaji, *et al.* (1991), Galichet and Foulloy (1995), Moon (1995), Mann, *et al.* (1999).

3. Details regarding a design method for Mamdani PI fuzzy controllers

The standard version of the Mamdani type PI-FC with integration of output / control signal, Fig.1, is based on the numerical differentiation of the control error e_k under the form of the increment of control error, $\Delta e_k = e_k - e_{k-1}$, and on the numerical integration of the increment of control signal Δu_k. The FCSs with Mamdani type PI-FCs are type-II fuzzy systems (Koczy, 1996; Sugeno, 1999).

The design of this controller starts with expressing the discrete time equation of the Q-C digital PI controller (PI-C) in its incremental (velocity type) version:

$$\Delta u_k = K_P \cdot \Delta e_k + K_I \cdot e_k = K_P \cdot (\Delta e_k + \alpha \cdot e_k), \quad (9)$$

where the parameters $\{K_P, K_I, \alpha\}$ are functions of $\{k_C, T_i\}$:

$$H_C(s) = \frac{k_C}{s \cdot T_i} \cdot (1 + s \cdot T_i), \quad K_P = k_C \cdot \left(1 - \frac{T_s}{2 \cdot T_i}\right), \quad K_I = \frac{k_C \cdot T_s}{T_i}, \quad \alpha = \frac{K_I}{K_P} = \frac{2 \cdot T_s}{2 \cdot T_i - T_s}. \quad (10)$$

On the basis of (10) and of the representation of Δu_k in the phase plane $<\Delta e, e_k>$, Fig.2, the pseudo-fuzzy features of the Q-C digital PI-C are worthwhile:

- there exists a “zero control signal line” $\Delta u_k = 0$, having the equation (11):

$$\Delta e_k + \alpha \cdot e_k = 0; \quad (11)$$

- this line divides the phase plane in two half-planes, with $\Delta u_k > 0$ and $\Delta u_k < 0$;
- the distance from any point of the phase plane to the "zero control signal line" corresponds to the absolute value of the increment of control signal $|\Delta u_k|$; it is influenced by the properties of the strictly speaking FC.

The fuzzification can be solved as follows: for the input linguistic variables (LVs) e_k and Δe_k there are chosen 5 (or more, but an odd number) linguistic terms (LTs) with regularly distributed triangular type membership functions (m.f.s) having an overlap of 1, and for the output LV Δu_k there are chosen 7 LTs with regularly distributed singleton type m.f.s, Fig.3, corresponding to the specific strictly positive parameters of this PI-FC, $\{B_e, B_{\Delta e}, B_{\Delta u}\}$. These parameters are in connection with the shapes of the m.f.s of the LTs corresponding to the input and output LVs. The complete rule base can be expressed as a decision table in the form of Table 1. The inference and defuzzification methods represent the designer's option (Driankov, *et al.*, 1993).

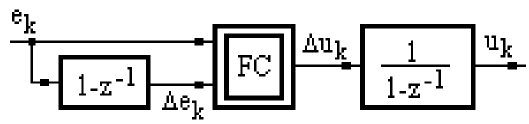


Figure 1. Structure of PI-FC with integration on controller output

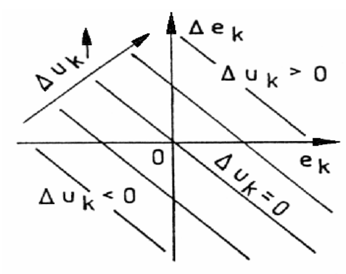


Figure 2. Phase plane representation of quasi-continuous digital PI controller

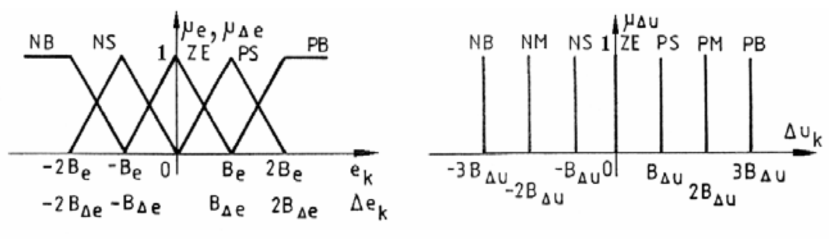


Figure 3. Shapes of membership functions of Mamdani PI fuzzy controller with integration on controller output

Table 1. Decision table of Mamdani PI-FC with integration on controller output

$\Delta e_k \setminus e_k$	NB	NS	ZE	PS	PB
PB	ZE	PS	PM	PB	PB
PS	NS	ZE	PS	PM	PB
ZE	NM	NS	ZE	PS	PM
NS	NB	NM	NS	ZE	PS
NB	NB	NB	NM	NS	ZE

The main steps of the design method are:

- the dependence (11) is valid for the "zero control signal line", resulting in:

$$\alpha = \frac{\Delta e_k}{e_k} = \frac{B_{\Delta e}}{B_e}; \tag{12}$$

- a further the condition in the form of (13) is fulfilled along the "constant control signal line", $\Delta u_k = B_{\Delta u}$:

$$B_{\Delta u} = \Delta u_k = K_P \cdot (\Delta e_k + \alpha \cdot e_k) = K_P \cdot B_{\Delta e}; \tag{13}$$

- the condition (13) can be transformed into:

$$B_{\Delta u} = K_P \cdot \alpha \cdot B_e = K_I \cdot B_e; \tag{14}$$

based on designer's experience; one of the parameters, for example B_e, is chosen, and the other two parameters, $B_{\Delta e}$ and $B_{\Delta u}$, result from (13) and (14).

It must be highlighted that by applying this method for tuning the FC parameters, $\{B_e, B_{\Delta e}, B_{\Delta u}\}$, the parameters of the basic linear PI-C (10), k_C and T_i, are taken into consideration in the design relations (13) and (14). Such controllers have been applied in several papers including (Precup and Preitl, 2001).

The obtained control signal in its incremental form Δu_k can be further used in the CS: directly, if the actuator contains the integral component (I), or by computing the actual value of control signal according to (15):

$$u_k = u_{k-1} + \Delta u_k. \tag{15}$$

4. Details regarding a design method for Takagi-Sugeno PI fuzzy controllers

The structure of a Takagi-Sugeno PI fuzzy controller is similar to that presented in Fig.1, but the FCSs with these FCs are type-III fuzzy systems (Koczy, 1996; Sugeno, 1999). The specific feature of Takagi-Sugeno FCs with dynamics is in the fact that the consequent of the rule base can contain expressions of conventional controllers resulting in a blend of conventional controllers due to the interpolative property of the fuzzy control rules (Babuska and Verbruggen, 1996).

The FCS comprising a Takagi-Sugeno PI-FC presented in this Section will ensure desired behaviours of the FCSs in dynamic regimes with respect to the step modifications of the reference input (w) and of four types of disturbance inputs (v). This is ensured for the beginning by the separate design of two continuous time linear PI controllers of type (10). For the design of the Takagi-Sugeno PI-FC it is necessary to discretize these two continuous linear PI controllers. The use of Tustin's method results in two incremental Q-C digital PI-Cs:

$$\Delta u_k = \Delta u_k^w = K_P^w \cdot \Delta e_k + K_I^w \cdot e_k \ ,$$

$$\Delta u_k = \Delta u_k^v = K_P^v \cdot \Delta e_k + K_I^v \cdot e_k \ , \tag{16}$$

where the parameters of these two incremental digital PI controllers, $\{K_P^w, K_I^w\}$ and $\{K_P^v, K_I^v\}$, are computed in terms of (17):

$$K_P^w = k_C^w \cdot \left(1 - \frac{T_s}{2 \cdot T_i^w}\right),\ K_I^w = \frac{k_C^w \cdot T_s}{T_i^w}\ ,\ K_P^v = k_C^v \cdot \left(1 - \frac{T_s}{2 \cdot T_i^v}\right),\ K_I^v = \frac{k_C^v \cdot T_s}{T_i^v}\ . \tag{17}$$

The structure of the proposed Takagi-Sugeno PI-FC is presented in Fig.4, and it consists of: the strictly speaking PI-FC, the additional fuzzy block FB1 for computing the current regime r_k, the fuzzy block FB2 for computing the current status s_k, and the linear blocks with dynamics.

The blocks {PI-FC, FB1, FB2} are Takagi-Sugeno fuzzy systems, and the inference and defuzzification methods can be selected according to the designer's option. The fuzzification is done by the m.f.s from Fig.5 ($\Delta w_k = w_k - w_{k-1}$ – increment of reference input) outlining the parameters of the Takagi-Sugeno PI-FC to be determined by the design method: $\{B_e, B_{\Delta e}, B_{\Delta w}, B_s, B_w, B_v\}$.

The fuzzy block FB1 has the role of observing the dynamic regime by computing the variable r_k. The linguistic terms "WR" and "VR" correspond to the dynamic regimes caused by the modification of w (wr) and v (vr), respectively. The fuzzy block FB2 that operates in parallel with PI-FC, computes the variable s_k characterizing the current status of the fuzzy control system. The linguistic term "ZE" corresponds to an accepted steady-state regime with almost zero e_k and Δe_k, and the linguistic term "P" corresponds to the situations when either e_k is non-zero or e_k is zero but it has the tendency to modify. The rule bases by Precup and Preitl (2002), expressed as decision tables, assist the inference engines of FB1 and FB2.

The inference engine of the strictly speaking PI-FC employs the rule base gathered in the decision table from Table 2. Such a decision table ensures quasi-PI behaviour of the PI-FC. An additional parameter α was introduced, $\alpha \in (0, 1]$, for the sake of FCS performance enhancement.

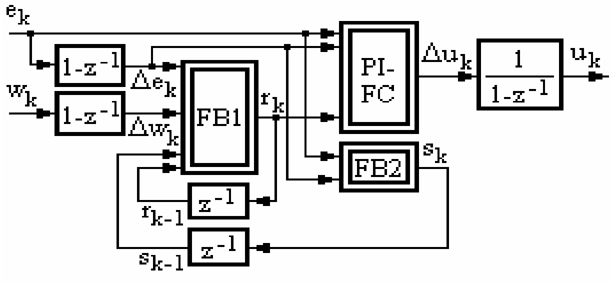


Figure 4. Structure of Takagi-Sugeno PI-FC

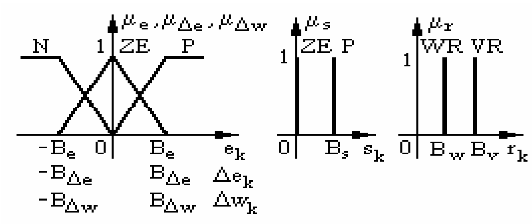


Figure 5. Accepted input membership functions

Table 2. Decision table of Takagi-Sugeno PI-FC

		r_k					
		WR			VR		
		e_k			e_k		
		N	ZE	P	N	ZE	P
	P	Δu_k^w	Δu_k^w	$\alpha\ \Delta u_k^w$	Δu_k^v	Δu_k^v	$\alpha\ \Delta u_k^v$
Δe_k	ZE	Δu_k^w	Δu_k^w	Δu_k^w	Δu_k^v	Δu_k^v	Δu_k^v
	N	$\alpha\ \Delta u_k^w$	Δu_k^w	Δu_k^w	$\alpha\ \Delta u_k^v$	Δu_k^v	Δu_k^v

In comparison with the previous Section, the fuzzy controller design becomes more complex due to the increased number of fuzzy blocks, and it consists mainly in the following steps:

- choose the values of the parameters B_w and B_v; since these values have to be different in order to create a clear difference between the two regimes, *wr* and *vr*; this is achieved by choosing $B_w = 1$ and $B_v = 2$;
- choose the values of the parameters $B_{\Delta w}$ and B_s; these values must be sufficiently small to clearly point out the constant values of w_k, and of e_k and Δe_k, respectively; for a unit step modification of w and a 2% settling time is accepted the recommended values for these two parameters are $B_{\Delta w} = 0.02$ and $B_s = 0.02$;
- design two continuous PI-Cs (with respect to the reference input and with respect to the disturbance input) and compute the value of the

parameter B_e there is applied the modal equivalences principle (Galichet and Foulloy, 1995) resulting in:

$$B_{\Delta e} = 2T_s \cdot B_e (2T_i^m - T_s) , \qquad (18)$$

where T_i^m represents the medium value of the integral time constant, and B_e is chosen in accordance with the experience of the control system designer. The relation (18) will ensure the approximate equivalence between the Takagi-Sugeno PI-FC and the separately designed two linear PI controllers (Precup and Preitl, 2002).

5. Advanced structures of fuzzy controllers

By starting with the PI fuzzy controllers presented in Sections 3 and 4, there will be presented here design aspects regarding the controllers with fuzzy adaptation of conventional controller parameters and predictive fuzzy controllers.

5.1. Controllers with fuzzy adaptation of conventional PI controller parameters

This class of controllers belongs to the class of self-tuning nonlinear controllers having some features. Firstly, the (basic) conventional controller used in the CS is developed on the basis of a classical design method, for a settled steady-state operating point. Secondly, depending on the modification of CP operating conditions, the parameters of the conventional controller are on-line adapted by the fuzzy adaptation block (F-AB) based on a specific fuzzy adaptation strategy (see also the results of De Silva (1991) and Zhao, *et al.* (1993)).

The structure of a CS comprising a Q-C digital PI-C (in incremental version (9)) with fuzzy adaptation of the parameters $\{k_C, T_i\}$ or $\{K_P, K_I\}$ as function of e_k and Δe_k is presented in Fig.6-a. This Q-C digital PI-C has at least two control features:

- firstly, the integral term, $K_I \cdot e_k$, mainly affects the overshoot; therefore, the value of K_I is adjusted as function of the control error;
- secondly, the proportional term, $K_P \cdot \Delta e_k$, affects both the first settling time / the rise time and the overshoot; the increase of K_P results in the increase of the overshoot and in the decrease of the first settling time.

In accordance with these features, the F-AB can be defined and designed by taking into account the following aspects:

- the F-AB1 module adjusts the coefficient K_P, and the F-AB2 module adjusts the coefficient K_I; both modules of the F-AB admit as inputs the LVs e_k and Δe_k; each input LV has 3 or 5 LTs with triangular type m.f.s having initial distribution with an overlap of 1 (Fig.6-b); the specific parameters are $\{B_e^P, B_{\Delta e}^P\}$ and $\{B_e^I, B_{\Delta e}^I\}$, respectively;

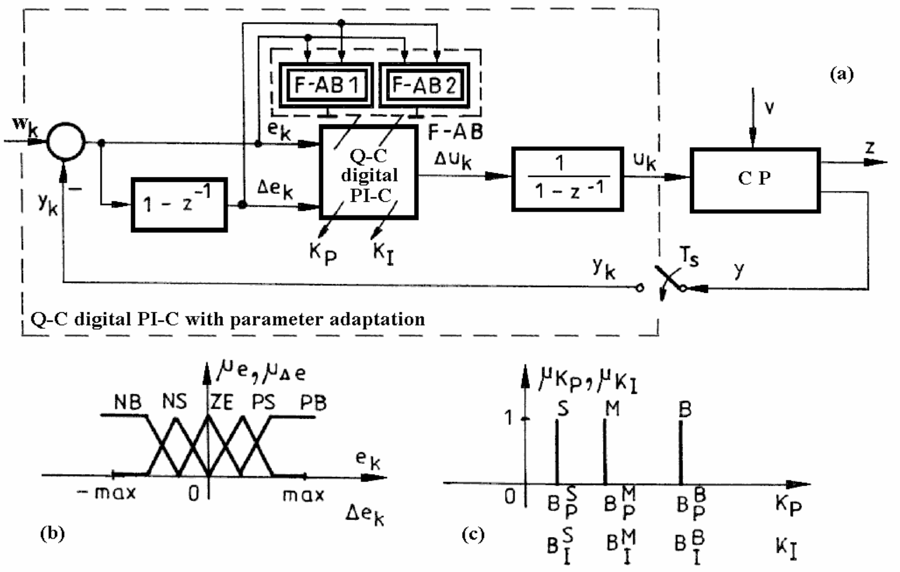


Figure 6. CS with fuzzy adaptation of the parameters of a Q-C digital PI-C

- the LTs corresponding to the output LVs, K_P and K_I, respectively, have regularly or non-regularly initially distributed singleton type m.f.s (Fig.6-c); the specific tuning parameters are $\{B_P^S, B_P^M, B_P^B\}$ and $\{B_I^S, B_I^M, B_I^B\}$, respectively;
- the rule base has to be defined by taking into account the particular features of the application involved.

5.2 PID fuzzy predictive controllers

The PID fuzzy predictive controllers (defined here in the sense given by Tzafestas (1985)) can be developed by starting with the incremental version of the Q-C digital PID controller (PID-C):

$$\Delta u_k = K_I\, e_k + K_P\, \Delta e_k + K_D\, \Delta^2 e_k,$$
$$\Delta^2 e_k = \Delta e_k - \Delta e_{k-1} = e_k - 2e_{k-1} + e_{k-2}, \tag{19}$$

where $\Delta^2 e_k$ represents the second increment of control error, and $\{K_P, K_I, K_D\}$ are the parameters of the Q-C digital PID-C. These parameters can be derived from the parameters of the conventional continuous PID-C by a discretization method.

Two versions of PID fuzzy predictive controllers have been developed by the authors, the PID fuzzy predictive controller with first order prediction, and the PID fuzzy controller with second order prediction, based on the prediction of the control error. The approach is quite different from that usually known under the name of fuzzy predictive control, which mainly deals with the use of fuzzy goals and fuzzy constraints in predictive control (Skrjanc, *et al.*, 1996).

The PID fuzzy predictive controller with first order prediction is designed by starting with the prediction of the control error defined by:

$$e_k = 2 \cdot e_{k-1} - e_{k-2}, \tag{20}$$

and on the substitution of it into (19). The result is in the discrete time equation (21):

$$\Delta u_k = K_{PI} \cdot (\Delta e_{k-1} + \beta \cdot e_{k-1}), \tag{21}$$

with the parameters K_{PI} and β computed by Precup and Preitl (1994).

This version of PID fuzzy controller is characterised by the fact that (21) is similar to the discrete time equation of a Q-C digital PI-C, (9). Therefore, an FC approximately equivalent to the controller characterised by (21) can be designed; this FC has two input LVs (e_{k-1} and Δe_{k-1}) and one output LVs (Δu_k), and the rest of elements are as in Sections 3 and 4.

The PID fuzzy predictive controller with first order prediction is designed by starting with the prediction of the control error in terms of:

$$e_k = 2.5 \cdot e_{k-1} - 2 \cdot e_{k-2} + 0.5 \cdot e_{k-3}. \tag{21}$$

Then, the substitution of e_k from (21) into (19) yields:

$$\Delta u_k = 0.5 \cdot K_{PID} \cdot \Delta^2 e_{k-1} + K_{PI} \cdot (\Delta e_{k-1} + \beta \cdot e_{k-1}), \tag{22}$$

with K_{PID} as in (Precup and Preitl, 1994).

The equation (22) can be transposed in a fuzzy manner (for the sake of CS performance enhancement) by two ways resulting in two versions of FCs:

- the first version: Δu_k is obtained by the addition of the crisp term $\Delta u_k^c = 0.5 \cdot K_{PID} \cdot \Delta^2 e_{k-1}$ to the increment of control signal Δu_k^f given by the PID fuzzy predictive controller with first order prediction presented before;
- the second version: the term Δu_k^c can be provided by another FC resulting in a parallel connection of two FCs.

Both PID fuzzy predictive controllers can be implemented in either Mamdani versions or Takagi-Sugeno ones, and their design is performed by using the results from Section 3 and Section 4.

6. Case study; implementation aspects

The case study corresponds to the class of CPs with the transfer functions:

$$H_P(s)=k_P/[s(1+sT_\Sigma)] \quad \text{(a)}, \quad H_P(s)=k_P/[s(1+sT_1)(1+sT_\Sigma)] \quad \text{(b)},(23)$$

with T_Σ – small time constant or time constant corresponding to the sum of parasitic time constants, $T_\Sigma < T_2 < T_1$, which characterize well enough many control applications with electrical drives playing the role of CPs.

The goal of the case study is to design a Takagi-Sugeno PI-FC based on two methods for optimal tuning of controller parameters meant for controlling the low order benchmarks (23) with integral character, by taking into consideration the conventional CS structure (Fig.6, with a certain controller in the framed part).

The classical design approach is the ESO method by Preitl and Precup (1999) providing a PI or a PID controller that can ensure very good CS performance. In both cases, the open-loop and closed-loop transfer functions with respect to the reference input (w) have unified forms with the design parameter β chosen by the designer as a compromise between desired all control system performance.

An example of digital simulation results of the designed CSs with respect to w ensured by the Takagi-Sugeno PI-FC in comparison with the PI-C is illustrated in Fig.7 when controlling the benchmark (23), with $k_P = 1$ and T_Σ =1 sec.

Regarding the adaptive control structure in Fig.6 employing Mamdani PI-FC there appear problems at the implementation because the parameters of the C are modified.

For ensuring a bump-less transfer from a digital PI-C to another one with different parameter sets, there are previously computed the "past values" which are necessary to the digital PI-C having the new set of parameters, Fig.8, with (Preitl and Precup, 2001):

- for the old digital PI-C:

$$u_k = u_{k-1} + (K_P + K_I)\cdot e_k - K_I \cdot e_{k-1}\,, \tag{24}$$

- for the new digital PI-C:

$$u_k = u_{k-1} + (K_P^* + K_I^*)\cdot e_k - K_I^* \cdot e_{k-1}\,, \tag{25}$$

where $\{K_P, K_I\}$ represent the parameters of the old digital PI-C, $\{K_P^*, K_I^*\}$

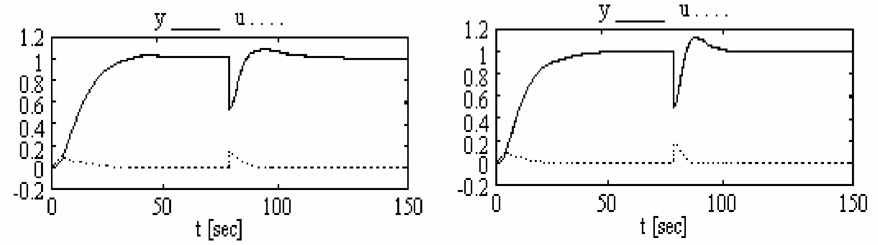


Figure 7. Digital simulation results for CSs with PI-C and Takagi-Sugeno PI-FC

Figure 8. Transfer from the old digital PI-C to the new one

are the parameters of the new digital PI-C and $\{e_{k-1}^*\}$ stand for the new initial conditions ("past values").

This is a general implementation problem which appears also in variable structure Mamdani FCs. As a matter of principle, in the case of Takagi-Sugeno FCs this problem is reported to be guaranteed by the FC operation principle itself. However, in the case of Mamdani FCs the relations (24) and (25) can be seen as linear equivalents of the FCs over the past and actual sampling interval (with appropriately computed parameters), and this is a version to ensure a bump-less transfer from one FC to another.

7. Conclusions

The paper presents some aspects regarding attractive systematic design approaches to the development of FCs with dynamics offering design and implementation recommendations. For some applications the FCS performance indices are approximately guaranteed by the design methods.

The presented transparent methods, focussed on PI-FCs, enable the design of other FCs with dynamics including the PD and the PID ones. All presented PI fuzzy controllers can be extended without any difficulties to PID fuzzy controllers due to the fact that the conventional PID-C can be expressed – in some conditions – as a series connection of two conventional controllers, a PD one (implemented in fuzzy manner like the PI-FC but without the integration of controller output) and a PI one; such a solution reduces strongly the dimension of the rule base in comparison with the situation of PID fuzzy controllers with three input LVs.

The case study can correspond to the speed control of a separately excited DC drive, and validates the design approaches and controller structures.

References

Babuska, R. and H.B. Verbruggen (1996). An Overview on Fuzzy Modeling for Control. *Control Engineering Practice*, **4**, 1593-1606.

De Silva, C.W. (1991). Simulation Studies of an Analytical Fuzzy Tuner for a PID Servo. *Proceedings of ACC'91*, 2100-2105.

Driankov, D., H. Hellendoorn and M. Reinfrank (1993). *An Introduction to Fuzzy Control.* Springer-Verlag, Berlin, Heidelberg, New York.

Galichet, S. and L. Foulloy (1995). Fuzzy Controllers: Synthesis and Equivalences. *IEEE Trans. on Fuzzy Systems*, **3**, 140-148.

Kawaji, S., T. Maeda and N. Matsunaga (1991). Design of Fuzzy Control System Based on PD Control Scheme. *Proceedings of 4th IFSA World Congress*, Brussels, 77-80.

Mamdani, E.H. (1974). Application of Fuzzy Algorithms for Control of a Simple Dynamic Plant. *Proceedings of IEE*, **121**, 1585-1588.

Mann, G.K.I., B.-G. Hu and R.G. Gosine (1999). Analysis of Direct Action Fuzzy PID Controller Structure. *IEEE Trans. on SMC – part B*, **29**, 371-388.

Moon, B.S. (1995). Equivalence between Fuzzy Logic Controllers and PI Controllers for Single Input Systems. *Fuzzy Sets and Systems*, **69**, 105-113.

Koczy, L.T. (1996). Fuzzy If-Then Rule Models and Their Transformation into One Another. *IEEE Trans. on Systems, Man, and Cybernetics – part A*, **26**, 621-637.

Precup, R.-E. and S. Preitl (1993). Fuzzy Control of an Electro-hydraulic Servo-system under Nonlinearities Constraints. *Proceedings of 1st EUFIT'93 European Congress*, Aachen, **3**, 1524-1530.

Precup, R.-E. and S. Preitl (1994). On a Fuzzy Digital PID Predictive Controller. *Proceedings of 2nd IEEE Mediterranean Symposium on "New Directions in Control and Automation"*, Chania, Crete, 669-673.

Precup, R.-E. and S. Preitl (1995). On the Transfer Properties of a SISO Fuzzy Controller with Adaptable Static Characteristic. *Proceedings of 5th Symposium on "Automatic Control and Computer Science – SACCS'95"*, Iasi, **1**, 132-137.

Precup, R.-E. and S. Preitl (2001). On the Development of Fuzzy Controllers for Electro-hydraulic Systems. *Proceedings of 7th Symposium on "Automatic Control and Computer Science – SACCS'93"*, Iasi, CD-ROM, 4 pp.

Precup, R.-E. and S. Preitl (2002). Development Method for a Takagi-Sugeno PI Fuzzy Controller. *Preprints of 15th IFAC World Congress*, Barcelona, CD-ROM, paper s390, 6 pp.

Preitl, S. and R.-E. Precup (1999). An Extension of Tuning Relations after Symmetrical Optimum Method for PI and PID Controllers. *Automatica*, **35**, 1731-1736.

Preitl, S. and R.-E. Precup (2001). *Introduction to Control Engineering.* Editura Politehnica Publishers, Timisoara (in Romanian).

Siler, W. and H. Ying (1989). Fuzzy Control Theory – the Linear Case. *Fuzzy Sets and Systems*, **33**, 275-290.

Skrjanc, I., K. Kavsek-Biasizzo, D. Matko and O. Hecker (1996). Fuzzy Predictive Control Based on Relational Matrix Model. *Computers in Chemical Engineering*, **20**, S931-S936.

Sugeno, M. (1999). On Stability of Fuzzy Systems Expressed by Fuzzy Rules with Singleton Consequents. *IEEE Trans. on Fuzzy Systems*, **7**, 201-224.

Takagi, T. and M. Sugeno (1985). Fuzzy Identification of Systems and Its Application to Modeling and Control. *IEEE Trans. on S MC*, **15**, 116-132.

Tang, K.L. and R.J. Mulholland (1987). Comparing Fuzzy Logic with Classical Controller Designs. *IEEE Trans. on SMC*, **17**, 1085-1087.

Tzafestas, S.G. (1985). *Applied Digital Control.* North-Holland, Amsterdam.

Zhao, Z.-Y., M. Tomizuka and S. Isaka (1993). Fuzzy Gain Scheduling of PID Controllers. *IEEE Trans. on SMC*, **23**, 1392-1398.

DISCRETE TIME LINEAR PERIODIC HAMILTONIAN SYSTEMS AND APPLICATIONS

Vladimir Răsvan
Department of Automatic Control,
University of Craiova,
A. I. Cuza Str. No. 13, RO 200585 Craiova, ROMANIA,
e-mail : vrasvan@automation.ucv.ro

Abstract Linear canonical (Hamiltonian) systems are familiar to the engineering community both from Rational Mechanics and Control. In Rational Mechanics an "evergreen" problem is that of the λ-zones of stability in connection with parametric resonance while in Control these systems belong to the field of Linear Quadratic Theory, being strongly connected to Matrix Riccati Equation. In both cases some robustness problems are met but they deal with different classes of systems: totally stable in the first and hyperbolic in the second case.
The present survey gives an account of these topics especially of their discrete-time counterpart.

Keywords: Hamiltonian systems, periodic coefficients, robustness

1. Introduction. Motivation and basic problems

Recently more attention is paid again to the theory of linear canonical/Hamiltonian systems, with a special reference to the discrete time case. There are several reasons for such an attention and we shall mention here but a few. We start with a topic that is familiar to control engineers - the linear quadratic control theory. More precisely, let us consider the differential controlled system

$$\dot{x} = A(t)x + B(t)u(t)\ ,\ x(t_0) = x_0, \tag{1}$$

and an associated integral index

$$J_{x_0,u}(t_0, t_1) = \int_{t_0}^{t_1} \mathcal{F}(t, u(t), x(t))dt + \mathcal{G}(x(t_1)), \tag{2}$$

where the quadratic forms $\mathcal{F}$ and $\mathcal{G}$ are defined as

$$\begin{array}{rcl} \mathcal{F}(t,u,x) & = & u^*K(t)u + u^*L(t)^*x + x^*L(t)u + x^*M(t)x \\ \mathcal{G}(x) & = & x^*Gx, \end{array} \qquad (3)$$

the matrices $K(t)$, $M(t)$, G being Hermitian. The integral in (2) is viewed as defined along the pairs $(u\ ,\ x)$ satisfying (1). Two basic problems are stated for the system defined by (1), (2) and (3)

i) *the Liapunov function problem* : find a quadratic form $\mathcal{V}(t,x) = x^*H(t)x$ such that there exists some $\delta > 0$ in order that

$$\begin{array}{rl} x^*\dot{H}(t)x & + \quad (A(t)x + B(t)u)^*H(t)x + x^*H(t)(A(t)x + B(t)u) \\ & \geq \quad -\mathcal{F}(t,u,x) + \delta(|x|^2 + |u|^2)\ ,\ \forall x \in \mathbb{C}^n,\ u \in \mathbb{C}^m; \end{array} \quad (4)$$

ii) *the optimization problem* of minimizing (2) along all admissible pairs satisfying (1).

These two problems, see, e.g. (Yakubovich, 1986;1991) are in fact connected with some very actual problems in control theory : *optimal stabilization* (problem ii) with $G = 0$ and $t_1 \to \infty$), *stability radius and absolute stability, forced oscillations a.o.*.

The above mentioned topics are connected with matrix Riccati differential equation and it is a well established fact that the Riccati differential equation is associated to a *linear canonical (Hamiltonian in the complex coefficient case) system*. The discrete-time counterparts of these topics are also well known to the researchers, especially the Riccati equation occurring from the dynamic programming approach to the optimization problem (Halanay, 1962; Halanay, 1963; Tou, 1963; Halanay and Răsvan, 2000).

As pointed out by Yakubovich (1991) there are several other problems leading to Hamiltonian systems : non-oscillatory/oscillatory behavior in differential equations, some self-adjoint boundary value problems, total stability of linear Hamiltonian systems, parametric resonance.

Especially oscillation and boundary value problems are now studied in the discrete-time framework due to the efforts of Erbe and Yan (1992a; 1992b; 1993; 1995), Bohner(1996), Bohner and Došlý(1997), Bohner, Došlý and Kratz (to appear); see also the books (Ahlbrandt and Peterson, 1996; Kratz, 1995).

Parametric resonance represents an interesting applied topic that has to be viewed in the context of the extensions of the results concerning Hill equation and of some classical results due to Žukovskii(1891/1893) and Liapunov(1899a; 1899b).

These topics send all to the theory of λ-zones of stability for linear Hamiltonian systems which has been considered in the monumental paper of M.G. Krein(1955). Other references are the survey of

Krein and Yakubovich(1963) and the monograph by Yakubovich and Staržinskii(1972).

The extension of the theory of λ-zones to the discrete-time case is an urgent task and it is in progress (Halanay and Răsvan, 1999; Răsvan, 2000; Răsvan, 2002).

The two main topics mentioned here - the one connected to the linear quadratic theory and the other one concerned with λ-zones - require properties of the associated Hamiltonian systems that are quite different in their nature. For this reason we shall present in this survey a model problem for each case.

2. Forced oscilations in systems with sector restricted nonlinearities

The study of forced oscillations in discrete-time affine systems is motivated by such applications as digital signal processing by nonlinear signal processors (Wade, 1994). Here also the "almost linear behavior" i.e. existence of a unique bounded on the whole real axis solution that is exponentially stable and of the same type as the forcing term is of interest. We would like to insist on almost periodic signals since they correspond to modulated signals; in the discrete-time case almost periodic sequences (discrete signals) are obtained in a natural way by sampling periodic signals when the sampling period and the period of the continuous time signal are in an irrational ratio (Halanay and Wexler, 1968).

We shall consider here the system

$$x_{k+1} = A_k x_k - b_k \phi_k (c_k^* x_k) + f_k \tag{5}$$

under the following basic assumptions: i) the matrix sequences $\{A_k\}_k$, the vector sequences $\{b_k\}_k$, $\{c_k\}_k$ and the sequence $\{\phi_k(\cdot)\}_k$ are N-periodic sequences; ii) $\phi_k(\sigma)$ are continuous with respect to σ and satisfy

$$0 \leq \frac{\phi_k(\sigma_1) - \phi_k(\sigma_2)}{\sigma_1 - \sigma_2} \leq \overline{\Phi} \tag{6}$$

for any $\sigma_1 \neq \sigma_2$ and $k = \overline{0, N-1}$; iii) f_k has bounded components for all integers k, possibly periodic or almost periodic. Also in the periodic case the period of f_k may equal N, the period of system's coefficients, but this is not compulsory.

In order to state the main result on discrete-time systems, we need introduction of the following linear discrete-time Hamiltonian system:

$$\begin{aligned} x_{k+1} &= (A_k - \tfrac{1}{2\overline{\Phi}} b_k c_k^*) x_k - \tfrac{1}{\overline{\Phi}} b_k b_k^* p_{k+1} \\ p_k &= -\tfrac{1}{4\overline{\Phi}} c_k c_k^* x_k + (A_k - \tfrac{1}{2\overline{\Phi}} b_k c_k^*)^* p_{k+1}. \end{aligned} \tag{7}$$

We may now state:

Theorem 1 *Consider system (5) under the basic assumptions i) - iii) and assume additionally the following: iv) the multipliers of A_k are inside the unit disk $\mathbb{D}_1$ of the complex plane i.e. A_k defines an exponentially stable evolution; v) the triple (A_k, b_k, c_k) and $\overline{\Phi}$ are such that*

$$det(A_k - \frac{1}{2\overline{\Phi}} b_k c_k^*) \neq 0,\ 0 \leq k \leq N-1. \tag{8}$$

vi) the Hamiltonian system (7) is exponentially dichotomic and strongly disconjugate (non-oscillatory).

Then there exists a bounded sequence satisfying (5) for all $k \in \mathbb{Z}$, which is periodic if f_k is periodic and almost periodic if f_k is almost periodic. Moreover this solution of (5) is exponentially stable.

An explanation of the terms is here necessary : a linear periodic Hamiltonian system with the general form

$$\begin{aligned} y_{k+1} - y_k &= B_k y_k + D_k z_{k+1} \\ z_{k+1} - z_k &= -A_k y_k - B_k^* z_{k+1}\ ,\ A_k = A_k^*,\ D_k = D_k^* \end{aligned} \tag{9}$$

is called *exponentially dichotomic* if its multipliers are not located on the unit circle. Let $y,\ z$ be m-dimensional and consider the m linearly independent solutions $x_k^1,\ \dots\ ,x_k^m$ corresponding to the multipliers which are located inside the unit disk (here x is the $2m$-column vector having y and z as component vectors). If the $m \times m$ matrix $(y_k^1\ \dots\ ,y_k^m)$ has rank m for $k = 0, 1,\ \dots\ , N-1$, the Hamiltonian is called *strongly disconjugate.*

Let us remark that if the Hamiltonian system (7) is such (i.e. exponentially dichotomic and strongly disconjugate) then it can be shown (Halanay and Ionescu, 1994) that the associated discrete-time Riccati matrix equation

$$\begin{aligned} H_k - A_k^* H_{k+1} A_k - \\ -(\frac{1}{\overline{\overline{\Phi}}} - b_k^* H_{k+1} b_k)^{-1} (\frac{1}{2} c_k - A_k^* H_{k+1} b_k)(\frac{1}{2} c_k - A_k^* H_{k+1} b_k)^* = 0 \end{aligned} \tag{10}$$

has a N-periodic global solution such that

$$\frac{1}{\overline{\overline{\Phi}}} - b_k^* H_{k+1} b_k > 0 \tag{11}$$

and this periodic solution is stabilizable in the following sense: if the controlled system

$$x_{k+1} = A_k x_k + b_k \mu_k \tag{12}$$

is considered, by choosing the control (input) sequence as follows

$$\mu_k = -(\frac{1}{\Phi} - b_k^* H_{k+1} b_k)^{-1}(\frac{1}{2}c_k - A_k^* H_{k+1} b_k)^* x_k \tag{13}$$

the "closed loop" linear system

$$x_{k+1} = [A_k - b_k(\frac{1}{\Phi} - b_k^* H_{k+1} b_k)^{-1}(\frac{1}{2}c_k - A_k^* H_{k+1} b_k)^*]x_k \tag{14}$$

is exponentially stable.

Let us remark also that exponential dichotomy and strong disconjugacy are *robust* with respect to system's coefficients perturbations (Yakubovich, 1990; 1991). This property is important for computational purposes: the matrix Riccati equation (9) may be replaced by the discrete-time Riccati inequality

$$\begin{gathered} H_k - A_k^* H_{k+1} A_k - \\ -(\frac{1}{\Phi} - b_k^* H_{k+1} b_k)^{-1}(\frac{1}{2}c_k - A_k^* H_{k+1} b_k)(\frac{1}{2}c_k - A_k^* H_{k+1} b_k)^* \geq \delta I \end{gathered} \tag{15}$$

and this inequality by a Linear Matrix inequality (LMI)

$$\begin{pmatrix} H_k - A_k^* H_{k+1} A_k & \frac{1}{2}c_k - A_k^* H_{k+1} b_k \\ (\frac{1}{2}c_k - A_k^* H_{k+1} b_k)^* & \frac{1}{\phi} - b_k^* H_{k+1} b_k \end{pmatrix} \geq \delta I_{n+1} \tag{16}$$

together with the condition $H_N = H_0$. In fact this is a Dynamic Linear Matrix Inequality but since we assumed tht H_k is N-periodic, a simple dimension augmentation reduces (16) to a $N(n+1) \times N(n+1)$ LMI that may be solved using the existing software.

3. Total stability and λ-zones. The second order system

Total stability means boundedness of all solutions on $\mathbb{R}$ ($\mathbb{Z}$). For linear canonical systems total stability analysis goes back to Žukovskii(1891/1893) and Liapunov(1899a; 1899b) who considered the simplest case of a scalar equation

$$y'' + \lambda^2 p(t) y = 0, \tag{17}$$

where $p(t)$ is T-periodic and λ is real. Obviously this reduces to the simplest canonical system.

We call λ_0 a λ-*point of stability* of (17) if for $\lambda = \lambda_0$ all solutions of (17) are bounded on $\mathbb{R}$. If moreover all solutions of any equation of (17) type but with $p(t)$ replaced by $p_1(t)$ sufficiently close to $p(t)$ (in some sense) are also bounded for $\lambda = \lambda_0$ then λ_0 is called a λ-*point of strong (robust) stability*.

Remark that we might take $p_1(t) = \lambda p(t)$ with $\lambda \neq \lambda_0$. In this case it was established by Liapunov himself(Liapunov, 1899a; 1899b) that the set of the λ-points of strong stability of (17) is open and if it is nonempty it decomposes into a system of disjoint open intervals called λ-*zones of strong stability*.

Equation (17) belongs to the more general class of linear periodic Hamiltonian systems described by

$$\dot{x} = \lambda J H(t) x \tag{18}$$

with $H(t)$ a T-periodic Hermitian $2m \times 2m$ matrix and

$$J = \begin{pmatrix} 0 & I_m \\ -I_m & 0 \end{pmatrix}. \tag{19}$$

For this system the results of Liapunov have been generalized by M. G. Krein (1955), Gelfand and Lidskii (1955), V. A. Yakubovich and many other; the final part of this long line of research was the book of Yakubovich and Staržinskii(1972). As pointed out by Krein and Yakubovich(1963) this research is motivated by various problems in contemporary physics and engineering (e.g. dynamic stability of structures, parametric resonance both in Mechanical and Electrical Engineering, Quantum-Mechanical treatment of the motion of the electron in a periodic field - see the book of Eastham(1973) - and other).

There exist discrete counterparts of the results concerning total stability and λ-zones; the research is in progress (Halanay and Răsvan, 1999; Răsvan, 2000; Răsvan, 2002).

Let us remark that the development of this research which follows closely the line of M.G. Krein is *top-down* i.e. from most general framework to its applications. Here we shall present one of the simplest cases aiming to a better understanding of various generalizations. Consider the real scalar version of (9) but with a parameter λ

$$\begin{aligned} y_{k+1} - y_k &= \lambda(b_k y_k + d_k z_{k+1}) \\ z_{k+1} - z_k &= -\lambda(a_k y_k + b_k z_{k+1}) \end{aligned} \tag{20}$$

with a_k, b_k, d_k being real and N-periodic. If we denote $x = col(y\ ,\ z)$ then (20) may be written as the recurrence $x_{k+1} = C_k(\lambda) x_k$ with $C_k(\lambda)$ defined by

$$C_k(\lambda) = \begin{pmatrix} 1 & -\lambda d_k \\ 0 & 1 + \lambda b_k \end{pmatrix}^{-1} \begin{pmatrix} 1 + \lambda b_k & 0 \\ -\lambda ak & 1 \end{pmatrix} = \tag{21}$$

$$= \frac{1}{1 + \lambda b_k} \begin{pmatrix} (1 + \lambda b_k)^2 - \lambda^2 d_k a_k & \lambda d_k \\ -\lambda ak & 1 \end{pmatrix}.$$

Obviously this is a matrix with rational items, having a real pole at $\lambda = -1/b_k$. At the same time $\det C_k(\lambda) \equiv 1$ hence it is *an unimodular matrix*. As known, for periodic systems the structure and the stability properties are given by *system's multipliers - the eigenvalues of the monodromy matrix* $U_N(\lambda) = C_{N-1}(\lambda)\ldots C_1(\lambda)C_0(\lambda)$. As a product of rational unimodular matrices $U_N(\lambda)$ is also *rational and unimodular (unlike the continuous-time case when it is an entire matrix function).* It follows that the characteristic equation of $U_N(\lambda)$ in this case is

$$\varrho^2 - 2A(\lambda)\varrho + 1 = 0, \tag{22}$$

where $2A(\lambda) = tr(U_N(\lambda))$ - the trace of the unimodular monodromy matrix of (20); the function$A(\lambda)$ is called *characteristic function* of the canonical system. Its properties are essential for defining and computing the λ - zones. In the continuous time case $A(\lambda)$ is an entire function while in *the case of* (20) *it is a rational function* with its poles the real numbers $-1/b_k, k = \overline{0, N-1}$ (these poles may not be distinct). In the following we shall see, once more, that not all properties of $A(\lambda)$ in the continuous time case are valid *mutatis mutandis* in the discrete time case.

In the following we shall assume that (20) *is of positive type in the sense of Krein* (1955) i.e. $H_k \geq 0, \forall k, \sum_0^{N-1} H_k > 0$. This positivity assumption allows obtaining some basic properties of the characteristic function $A(\lambda)$ and of system's multipliers - the roots of (22)

i) all zeros of the rational function $A(\lambda) - \alpha$, where $|\alpha| \leq 1$, are real and for $|\alpha| \neq 1$ are simple i.e. $A^{'}(\lambda) \neq 0$ for those λ such that $|A(\lambda)| < 1$; the roots of $A(\lambda) \pm 1$ are at most double ;

ii) the non-real multipliers with $|\varrho| = 1$ (but $\varrho \neq \pm 1$) are of definite type in the sense of Krein (1955) i.e. $(Ju, u) \neq 0$ where (,) denotes the usual Euclidean scalar product, u is the eigenvector associated to the multiplier and J, defined by (19), is taken for $m = 1$.

Since the two non-real multipliers are conjugate, their eigenvectors are such: if $\varrho(\lambda)$ has u as eigenvector, then $\overline{\varrho(\lambda)}$ has $\bar{u}$ as eigenvector. We deduce that one multiplier is K-positive i.e. with $\jmath(Ju, u) > 0$ while the other one is K-negative i.e. with $\jmath(J\bar{u}, \bar{u}) < 0$. (The term of K-positive was coined by Ekeland(1991); the original terms of Krein were 1st kind for K-positive and 2nd kind for K-negative).

There are also other properties of $A(\lambda)$ and of $F(\lambda) = A^2(\lambda) - 1$ which depend strongly on the structure of these functions. Generally speaking, the elements of $U_N(\lambda)$ - the monodromy matrix - are rational functions with $-b_k^{-1}$ as real poles. Using the analysis of $A(\lambda)$ and $F(\lambda)$ around $\lambda = 0$ and the properties mentioned above we obtain

$$
\begin{aligned}
A(\lambda) &= \frac{(1-\lambda/\alpha_1)(1-\lambda/\alpha_2)\dots(1-\lambda/\alpha_K)}{(1+\lambda b_1)^{\mu_1}(1+\lambda b_2)^{\mu_2}\dots(1+\lambda b_r)^{\mu_r}} \\
F(\lambda) &= A^2(\lambda)-1 = a_2\lambda^2\frac{(1-\lambda/\lambda_1)^{\nu_1}(1-\lambda/\lambda_2)^{\nu_2}\dots(1-\lambda/\lambda_q)^{\nu_q}}{(1+\lambda b_1)^{\mu_1}(1+\lambda b_2)^{\mu_2}\dots(1+\lambda b_r)^{\mu_r}},
\end{aligned} \tag{23}
$$

where $a_2 < 0$, $\nu_i \leq 2$ and $\sum_1^r \mu_i = N$. We shall have

$$
\frac{d^2}{d\lambda^2}(\ln A(\lambda)) = \frac{A^{''}(\lambda)A(\lambda)-(A^{'}(\lambda))^2}{(A(\lambda))^2} =
$$

$$
= -\sum_1^K \frac{1}{(\lambda-\alpha_i)^2} + \sum_1^r \frac{\mu_j b_j^2}{(1+\lambda b_j)^2}.
$$

If λ_* is a critical point of $A(\lambda)$ i.e. $A^{'}(\lambda_*) = 0$ then

$$
\frac{d^2}{d\lambda^2}(\ln A(\lambda)) \mid_{\lambda=\lambda_*} = \frac{A^{''}(\lambda_*)A(\lambda_*)}{(A(\lambda_*))^2}.
$$

From now on we have to consider two different cases.

i) $\mu_j = 0,\ \forall j$ i.e. $A(\lambda)$ and $F(\lambda)$ are entire functions of polynomial type; this was also the case for continuous time Hamiltonian systems analyzed by Krein (1955), Yakubovich (e.g. Yakubovich and Staržinskii, 1972) and others. For any critical point we shall have $A^{''}(\lambda_*)A(\lambda_*) < 0$; the critical points are extrema - maxima or minima - and $|A(\lambda_*)| \geq 1$; remark that in any case $\lambda_* = 0$ is a maximum of $A(\lambda)$, $A(0) = 1$.

Let the non-zero roots of $F(\lambda)$ with their multiplicities (at most 2 each) be ordered as follows

$$
\dots < \lambda_{-4} \leq \lambda_{-3} < \lambda_{-2} \leq \lambda_{-1} < 0 < \lambda_1 \leq \lambda_2 < \lambda_3 \leq \lambda_4 < \dots \tag{24}
$$

the sequence being obviously finite. The graphic of $A(\lambda)$ is as in fig. 1 and its properties may be summarized as follows

Theorem 2 *Let $b_k = 0$, $\forall k$; all zeros of the polynomial $A^2(\lambda) - 1$ are real and among them at least one is positive and one negative. These zeros may be indexed as in (24) and we have $F(\lambda)$ as in fig.1. On any interval $(\lambda_{2k}, \lambda_{2k+1})$, $k \geq 0$ or $(\lambda_{2k-1}, \lambda_{2k})$, $k \leq 0$ the function $(-1)^k A(\lambda)$ is increasing and $-1 < A(\lambda) < 1$. On any interval $(\lambda_{2k-1}, \lambda_{2k})$, $k \geq 1$ or $(\lambda_{2k-2}, \lambda_{2k-1})$, $k \leq 0$ two cases are possible : either $(-1)^k A(\lambda) > 1$ on that interval and the interval contains a maximum of $(-1)^k A(\lambda)$ and only one, or $\lambda_{2k-1} = \lambda_{2k} (\lambda_{2k-2} = \lambda_{2k-1})$ and then the maximum is $(-1)^k A(\lambda_{2k-1}) = 1$.*

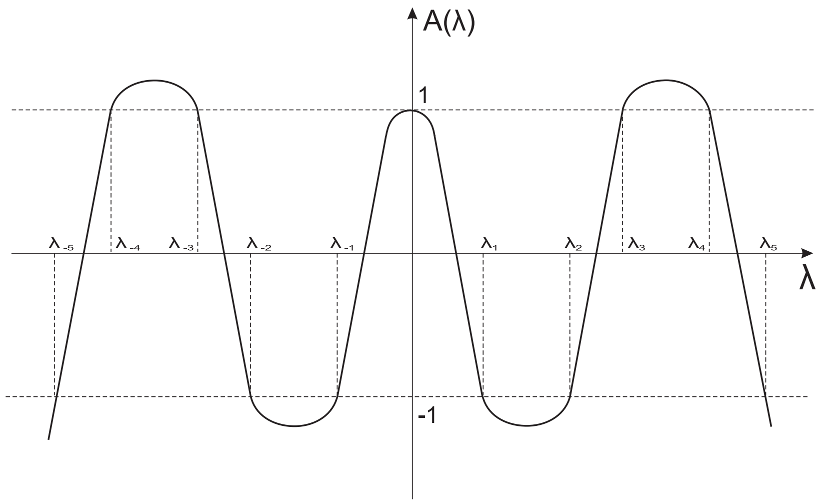


Figure 1. The graphic of an entire $A(\lambda)$

If we use the formulae for the multipliers - roots of (22) - namely

$$\varrho_{1,2}(\lambda) = A(\lambda) \pm \sqrt{A^2(\lambda) - 1}$$

it is obvious that the intervals $(\lambda_{2k}, \lambda_{2k+1}),\ k \geq 0$ or $(\lambda_{2k-1}, \lambda_{2k}),\ k \leq 0$ correspond to stability zones ($-1 < A(\lambda) < 1$ imply $\varrho_{1,2}(\lambda) = \alpha \pm \jmath\sqrt{1-\alpha},\ |\varrho_j(\lambda)| = 1$) while the other ones correspond to instability zones; the "degenerate" intervals described by $\lambda_{2k-1} = \lambda_{2k}(\lambda_{2k-2} = \lambda_{2k-1})$ are included in the stability zones.

$\boxed{\text{ii) } \mu_j \geq 1}$ i.e. assume that $A(\lambda)$ and $F(\lambda)$ are rational functions with real poles. This situation is specific for discrete-time systems, showing that not all properties of continuous-time systems migrate *mutatis-mutandis* to discrete-time ones. Nevertheless some properties of $A(\lambda)$ and $F(\lambda)$ from the previous case are indeed valid now also: the absence of critical points of $A(\lambda)$ within $(-1, 1)$ and the multiplicity of the zeros of $F(\lambda)$ which is at most 2. But the poles of $A(\lambda)$ induce the fact that the behavior of $A(\lambda)$ outside $(-1, 1)$ might be more complicated than previously. The fact that $A(\lambda)$ is monotonic on $(-1, 1)$ implies *the alternance of stability and instability zones in this case also.* The central zone - around the maximum $A(0) = 1$ - keeps its form and it is a stability zone. Also *any pole belongs to an instability zone* but an instability zone may contain more than one pole; instability zones without poles are also possible as well as more than one critical point in the stability zone. A possible representation of $A(\lambda)$ is given in fig. 2. We may state

Theorem 3. *Assume some $b_k \neq 0$. All zeros of $A^2(\lambda) - 1$ are real, among them being at least one positive and one negative. The non-zero* *roots of $F(\lambda)$ - which are at most of multiplicity 2 - may be ordered as in (24), their sequence being finite, and the representations of (23) are valid. On any interval $(\lambda_{2k}, \lambda_{2k+1}),\ k \geq 0$ or $(\lambda_{2k-1}, \lambda_{2k}),\ k \leq 0$ $A(\lambda)$ is monotone and $-1 < A(\lambda) < 1$. On any interval $(\lambda_{2k-1}, \lambda_{2k}),\ k \geq 1$ or $(\lambda_{2k-2}, \lambda_{2k-1}),\ k \leq 0$ one of the following situations may occur : a) the interval does not contain any pole of $A(\lambda)$ and $|A(\lambda)| > 1$ but with finite values; if the interval is just a point ($\lambda_{2k} = \lambda_{2k+1},\ k \geq 0$ or $\lambda_{2k} = \lambda_{2k-1},\ k \leq 0$) this point corresponds to an extremum equal to ± 1; b) the interval contains at least a pole of $A(\lambda)$ and such an interval may contain extrema of $A(\lambda)$ (but not compulsory).*

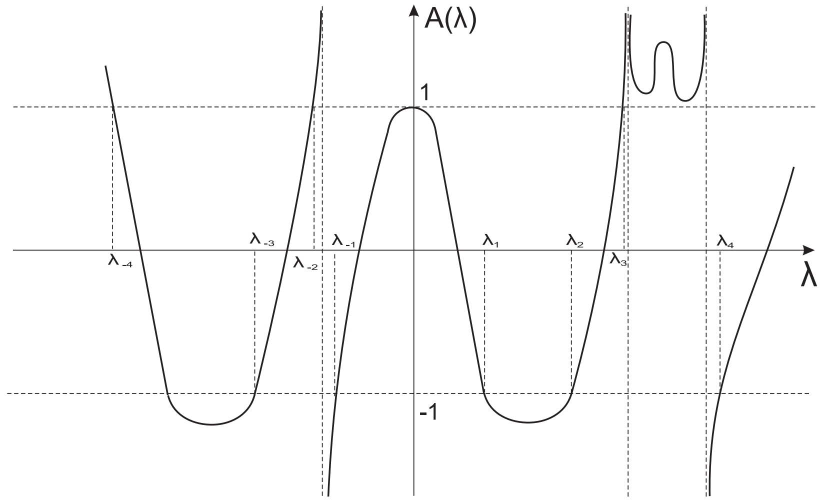


Figure 2. The graphic of $A(\lambda)$ having real poles

4. Multipliers' "Traffic rules"

The term "multiplier motion" was introduced by M. G. Krein and reflects a property called by him *strong stability* and which turns out to represent for the contemporary control engineer a special (and very interesting) case of robust stability (for a class of linear systems).

For a better explanation we shall start from the well known problem of robust stability (in the sense of Liapunov) of some linear systems with constant coefficients. The necessary and sufficient condition of stability is given by the location of the roots of system's characteristic equation in $\mathbb{C}^-$ (for continuous-time systems) or inside the unit disk $\mathbb{D}_1 \subset \mathbb{C}$ (for discrete-time systems). A well known fact - the continuous dependence of the roots of a polynomial on its coefficients - is the basis of another well known fact : sufficiently small perturbations of the coefficients do not modify the half-plane to which the roots belong (or, in the other case, the roots do not leave the interior or the exterior of $\mathbb{D}_1$) ; for this reason the exponential stability of linear systems with constant coefficients is robust.

The same type of analysis is valid for linear systems with periodic coefficients (both continuous and discrete-time) since their solutions also have a structure which is defined by the roots of a certain characteristic equation - the characteristic equation of the monodromy matrix - the multipliers of the system. In both cases - continuous and discrete-time - the multipliers have to be located inside the unit disk $\mathbb{D}_1 \subset \mathbb{C}$ in order to have exponential stability.

Let us turn now to the Hamiltonian systems with periodic coefficients. Here the multipliers have a certain symmetry: they occur in pairs - one inside $\mathbb{D}_1$, another outside $\mathbb{D}_1$. Consequently exponential stability is not possible but only boundedness on $\mathbb{R}$ ($\mathbb{Z}$) otherwise the system has an exponential dichotomy (half of its linearly independent solutions tend exponentially to 0 for
$t \to \infty$ while half tend exponentially to 0 for $t \to -\infty$). This boundedness of solutions on $\mathbb{R}$ ($\mathbb{Z}$) is called stability and, due to linearity, it corresponds to (non-asymptotic) stability. Therefore all multipliers have to be located on the unit circle, being either *simple* or of *simple type* (with simple elementary divisors).

Generally speaking such a location is not robust with respect to perturbations of the coefficients: the roots of the equation have to jump either inside $\mathbb{D}_1$ or outside $\mathbb{D}_1$. But the symmetry which is specific to Hamiltonian systems introduces some corrections to these simple facts. Indeed, if the multipliers are all simple the stability is robust since they cannot leave the circle under small perturbations because this would break the above mentioned symmetry.

If in some point of the unit circle $\partial\mathbb{D}_1 = \{z \in \mathbb{C} | z = e^{\jmath\theta}\}$ there is a multiplier of multiplicity larger than 1 it could leave the circle without affecting the symmetry since the multiple multiplier would split. But this representation is far of being complete: speaking about multiplier "motion" we viewed them just as points in $\mathbb{C}$ - a purely geometric vision. In fact it is necessary that these points (complex numbers) were multipliers of a neighboring Hamiltonian system with respect to the basic one. But such a perturbation is not always possible; in fact we require existence of sufficiently small perturbations which preserve the Hamiltonian character of the system and this requirement depends essentially on the kind of the multipliers that coincide at some point of $\partial\mathbb{D}_1$.

If these multipliers are of different kinds - such perturbed Hamiltonians *do* exist. But if they are of the same kind i.e. the non-simple multiplier is definite (K-positive or K-negative, as well as its conjugate which has the opposite kind) then such a perturbed Hamiltonian - which would imply the leaving of the unit circle $\partial\mathbb{D}_1$ - cannot exist.

We deduce that if one would imagine a continuous perturbation of the Hamiltonian leading to an "encounter" of multipliers of the same kind on $\partial\mathbb{D}_1$, they could not "jump" from the circle; on the contrary, if the "encounter" is of multipliers of different kinds resulting in a non-simple multiplier of mixed type, a certain perturbation (whose existence is now ensured) would imply leaving of the circle and destabilizing the Hamiltonian system.

We may conceive some kind of kinematic representation: since any $2m$ - dimensional strongly stable Hamiltonian system has exactly m multipliers of one kind and m of the opposite, they might be numbered in such a way that a multiplier with given number ("label") would "move" continuously on $\partial\mathbb{D}_1$ when the Hamiltonian would be perturbed by some admissible perturbation. In their motion on the circle (fig. 3a) these multipliers obey the so-called "multipliers' traffic rules" formulated by Krein (1955) :

$\kappa 1$) if two multipliers of the same kind "cross" on $\partial\mathbb{D}_1$ they cannot "jump" from the circle since, due to the spectral symmetry (Liapunov - Poincaré theorem) they should reach one the inside and the other the outside of $\mathbb{D}_1$; but since they are of the same kind they could reach only one side. This is a direct consequence of Theorem 1.2 of (Krein, 1955) which gives also a sufficient condition of strong stability : *all multipliers should be located on the unit circle and be of definite kind* (the necessity of this condition has been proved later by Gelfand and Lidskii (1955).

$\kappa 2$) if two multipliers of different kinds "cross" on $\partial\mathbb{D}_1$ then spectral symmetry does not restrict any longer the "leave" of the unit circle.

5. The multiplier traffic in the scalar case

We consider again the second order system of Section 3. This system has 2 multipliers and if λ belongs to a stability zone i.e. $|A(\lambda)| < 1$ then one of them is K-positive and the other one K-negative. Let $\lambda > 0$ and $\lambda \in (\lambda_{2k}, \lambda_{2k+1})$. For the K-*positive multiplier* we shall have (Yakubovich and Staržinskii, 1972; Răsvan, 2002)

$$\begin{aligned} \jmath\varphi_1'(\lambda) &= -\frac{1}{(Ju^1(\lambda), u^1(\lambda))} \sum_0^{N-1} \Big[a_k|y_k^1(\lambda)|^2 + 2b_k \Re(\overline{y_k^1(\lambda)} z_{k+1}^1(\lambda)) \\ &+ d_k|z_{k+1}^1(\lambda)|^2 \Big] \end{aligned} \tag{25}$$

and therefore $\varphi_1^{'}(\lambda) < 0$; here $\varphi_1(\lambda)$ is the phase of the multiplier $\varrho_1(\lambda) = \exp(\jmath\varphi_1(\lambda))$. It follows that the multiplier moves clockwise on the lower semi-circle of $\partial\mathbb{D}_1$, from $+1$ to -1 if $A(\lambda)$ decreases from $+1$ to -1 and on the upper semi-circle from -1 to $+1$ if $A(\lambda)$ increases from -1 to $+1$. The K-negative multiplier moves counter-clockwise on the complementary semi-circle.

Let now λ belong to an instability zone i.e. $|A(\lambda)| > 1$. The two multipliers are real, given by the formulae of Section 3 ; moreover

$$\frac{d\varrho_1}{dA} = 1 + \frac{A}{\sqrt{A^2-1}} > 0, \quad \frac{d\varrho_2}{dA} = 1 - \frac{A}{\sqrt{A^2-1}} < 0. \qquad (26)$$

These equations show that the multipliers move on the real axis outside or inside the unit disk, keeping the well known symmetry with respect to the unit circle. We have now to consider separately the two cases analyzed previously at Section 3:

i) if $A(\lambda)$ is a polynomial then in any instability zone $(\lambda_{2k-1}, \lambda_{2k})$ it has a unique extremum. When λ covers this instability zone the multipliers leave the encounter point $(A(\lambda_{2k-1}), 0)$ from $\partial\mathbb{D}_1$ to move outside and inside the unit disk on the real axis up to some extremal positions and further to return to the same point of the circle since $(A(\lambda_{2k-1}), 0) = (A(\lambda_{2k}), 0)$. Then the motion will continue on the circle since a new stability zone follows. Since a stability zone with increasing $A(\lambda)$ is followed by an instability zone and next by another stability zone with decreasing $A(\lambda)$ a.s.o. it is obtained the image of a continuous motion of the two multipliers which move each on its own semi-circle, meet in $(\pm 1, 0)$, leave the circle remaining on the real axis inside and outside the unit disk, return to the same point and follow the motion on the other semi-circle preserving the sense of motion (fig.3b). If λ_{2k} is an extremum i.e. $A^{'}(\lambda_{2k}) = 0$ the multipliers continue the motion on the circle since λ remains in a stability zone;

ii) if $A(\lambda)$ is rational then an instability zone may contain poles or may be *pole-free*. Assume first that the instability zone is such : in this case the multiplier traffic is exactly as previously.

Let now be a single pole of even multiplicity in the instability zone. This case is much alike to the pole-free case but the "extremum" is infinite. The multiplier of first kind may reach the origin while that of second kind may reach $\pm\infty$ returning to the splitting point and continuing on the corresponding semi-circles.

If the single pole within the instability zone is of odd multiplicity then the multiplier of first kind crosses the origin, the one of second kind "jumps from one infinity to the other" and the two multipliers

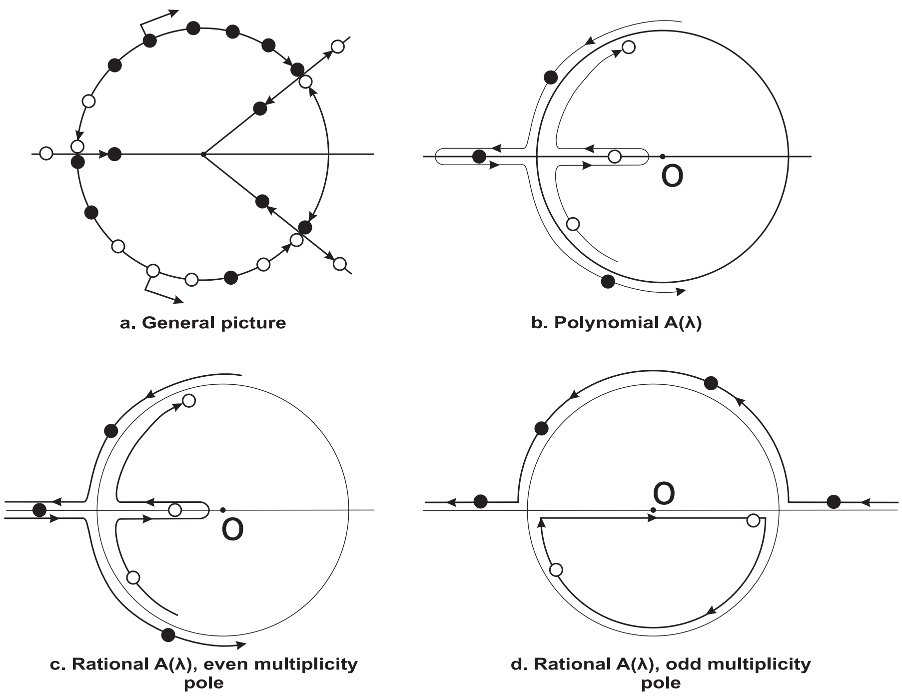


Figure 3. Multiplier "traffic" (scalar case)

return to the point of the circle which is the diametrically opposed on the real axis to the splitting one.

If within the instability zone there are several poles of $A(\lambda)$ the multipliers may oscillate on the real axis in the sense that they tend to $\pm\infty$ (and 0), return approaching the circle, move away again etc. It is quite clear the essential role of the summarized multiplicity of the poles within the instability zone : if it is even the return to the circle passes through the splitting point (fig.3c) and if it is odd - through the diametrically opposed on the real axis (fig.3d).

6. The scalar equation

The discretized version of (17) is obtained by taking the symmetric difference which preserves the Hamiltonian character

$$y_{k+1} - 2y_k + y_{k-1} + \lambda^2 p_k y_k = 0. \tag{27}$$

We may introduce

$$y_{k+1} - y_k = \lambda z_{k+1}$$

to obtain the system

$$
\begin{aligned}
y_{k+1} - y_k &= \lambda z_{k+1} \\
z_{k+1} - z_k &= -\lambda p_k y_k
\end{aligned}
\tag{28}
$$

which is alike (20) but with $b_k = 0$; in this case $A(\lambda)$ is polynomial and we may refer to fig.1 and to considerations made at Section 5, Case i). Moreover, as pointed out in (Răsvan, 2002), the end points of the central stability zone being the first(largest) negative and the first (smallest) positive characteristic numbers of the skew -periodic boundary value problem defined by (28) and the boundary conditions $y_0 = -y_N,\ z_0 = -z_N$ the estimates for the width of the central stability zone of Krein type given in (Răsvan, 2002) are valid. Among them we would like to mention the discrete version of the well known Liapunov criterion formulated for (17)(Liapunov, 1899a).

Proposition 1 *(Răsvan, 2002) All solutions of (27) are bounded provided* $p_k \geq 0, \quad \sum_0^{N-1} p_k > 0$ *and* $\lambda^2 < 4/N(\sum_0^{N-1} p_k)$.

In this way all assertions of Liapunov's paper (Liapunov, 1899a) have been extended to the discrete-time case using the general framework developed in (Krein, 1955). Worth mentioning that even in this case the Liapunov criterion is only a sufficient estimate of the stability zone while not very conservative. The exact width of the central stability zone is given by the inequality (Halanay and Răsvan, 1999).

$$
\lambda^2 < \pi/N(\sum_0^{N-1} p_k).
$$

As pointed out by Krein (1955), the results of Liapunov for the central stability zone of (17) have been extended to the case when $p(t)$ has values of both signs (Liapunov, 1899b) but the cited reference contained no proofs. The proofs are to be found following the line of (Krein, 1955)(see Section 9 of this reference or (Yakubovich and Staržinskii, 1972)); the discrete version can be obtained in an analogous way following the hints contained in the cited references and using the results of (Halanay and Răsvan, 1999).

7. Concluding remarks

We have surveyed throughout the paper some basic results on discrete-time periodic Hamiltonian systems with particular reference to robustness. The field is filled up with open problems. We do not have a complete discrete analogue of the theorem of Yakubovich on linear quadratic theory (Yakubovich, 1986) and many of the Krein type results for the theory of λ-zones and its applications are still under research for the discrete-time case e.g. discrete-time parametric resonance.

References

[1] Ahlbrandt, D.C. and A.C. Peterson (1996). *Discrete Hamiltonian Systems: Difference Equations, Continued Fractions and Riccati Equations*, Kluwer Academic Publishers, Dordrecht-Boston.

[2] Bohner, M.(1996). Linear Hamiltonian difference systems: disconjugacy and Jacobi conditions. *J. Math. Anal. Appl.* **199**, 804-826.

[3] Bohner, M. and O. Došlý(1997). Disconjugacy and transformations for symplectic systems. *Rocky Mountain J. Math.* **27**, 707-743.

[4] Bohner, M., O. Došlý and W. Kratz (to appear). An oscillation theorem for discrete eigenvalue problems. *Rocky Mountain J. Math.*

[5] Eastham, M.S.(1973). *The spectral theory of periodic differential equations.* Scottish Acad. Press, Edinburgh.

[6] Ekeland, I.(1991). *Convexity Methods in Hamiltonian Mechanics.* Springer Verlag, Berlin-Heidelberg-New York.

[7] Erbe, L. and P. Yan (1992a). Disconjugacy for linear Hamiltonian difference systems. *J. Math. Anal. Appl.* **167**, 355-367.

[8] Erbe, L. and P. Yan (1992b). Qualitative properties of Hamiltonian difference systems. *J. Math. Anal. Appl.* **171**, 334-345.

[9] Erbe, L. and P. Yan (1993). Oscillation criteria for Hamiltonian matrix difference systems. *Proc. Amer. Math. Soc.* **119**, 525-533.

[10] Erbe, L. and P. Yan (1995). On the discrete Riccati equation and its applications to discrete Hamiltonian systems. *Rocky Mountain J. Math.* **25**, 167-178.

[11] Gelfand, I.M. and V.B. Lidskii (1955). On the structure of the regions of stability of linear canonical systems of differential equations with periodic coefficients (in Russian). *Uspehi Mat. Nauk* **10**, 3-40 (English version *AMS Translations* **8(2)**, 143-181, 1958).

[12] Halanay, A.(1962). Un problème d'optimisation pour les systèmes aux différences finies. *C.R. Acad.Sci. Paris* **254**, 2512-2513.

[13] Halanay, A.(1963). An optimization problem for discrete-time systems (in Romanian). In: *Probleme de Automatizare* **V**, 103-109. Editura Academiei, Bucharest.

[14] Halanay, A. and V. Ionescu (1994). *Time-Varying Discrete Linear Systems.* Birkhäuser Verlag, Basel-Boston-Berlin.

[15] Halanay, A. and Vl. Răsvan (1999). Stability and boundary value problems for discrete-time linear Hamiltonian systems. *Dynam. Systems Appl.* **8**, 439-459 (Special Issue on "Discrete and Continuous Hamiltonian Systems" edited by R. P. Agarwal and M. Bohner).

[16] Halanay, A. and Vl. Răsvan (2000). Oscillations in Systems with Periodic Coefficients and Sector-restricted Nonlinearities. In *Operator Theory; Advances and Applications* **117** (V. Adamian and H. Langer. (Eds)), 141-154. Birkhäuser Verlag, Basel.

[17] Halanay, A. and D.Wexler (1968). *Q*ualitative Theory of pulse systems (in Romanian). Editura Academiei, Bucharest (Russian Edition by Nauka, Moscow, 1971).

[18] Kratz, W. (1995). *Quadratic Functionals in Variational Analysis and Control Theory*, Mathematical Topics vol. **6**, Akademie Verlag, Berlin.

[19] Krein, M.G.(1955). Foundations of the theory of λ-zones of stability of a canonical system of linear differential equations with periodic coeffi- cients (in Russian). In *In Memoriam A.A.Andronov*, 413-498. USSR Acad. Publ.House, Moscow (English version *AMS Translations* **120(2)**, 1-70, 1983).

[20] Krein, M.G. and V.A. Yakubovich (1963). Hamiltonian systems of linear differential equations with periodic coefficients (in Russian). In *Proceedings of Int'l Conference on Nonlin.Oscillations* **1**, 277-305. Ukrainian SSR Acad.Publ.House, Kiev (English version *AMS Translations* **120(2)**, 139-168, 1983).

[21] Liapunov, A.M.(1899a). Sur une équation différentielle linéaire du second ordre. *C.R. Acad. Sci. Paris* **128**, 910-913.

[22] Liapunov, A.M.(1899b): Sur une équation transcendente et les équations différentielles linéaires du second ordre à coefficients périodiques. *C.R. Acad. Sci. Paris* **128**, 1085-1088.

[23] Răsvan, Vl.(2000). Stability zones for discrete time Hamiltonian systems. *Archivum mathematicum* **tomus 36**, 563-573 (CDDE2000 issue).

[24] Răsvan, Vl.(2002). Krein-type Results for λ-Zones of Stability in the Discrete-time Case for 2-nd Order Hamiltonian Systems. *Folia FSN Universitatis Masarykianae Brunensis, Mathematica* **10**, 1-12 (CDDE2002 issue)

[25] Tou, J.T.(1963). *Optimum design of digital control systems.* Academic Press, N.Y.

[26] Žukovskii, N.E.(1891/1893). Conditions for the finiteness of integrals of the equation $d^2y/dx^2 + py = 0$ (in Russian). *Matem. Sbornik* **16**, 582-591.

[27] Yakubovich, V.A.(1986). Linear quadratic optimization problem and frequency domain theorem for periodic systems I. *Sib.Mat.Ž*, **27**, 181-200 (in Russian).

[28] Yakubovich, V.A.(1990). Linear quadratic optimization problem and frequency domain theorem for periodic systems II. *Sib.Mat.Ž*, **31**, 176-191 (in Russian).

[29] Yakubovich, V.A.(1991). Nonoscillation of Linear Periodic Hamiltonian Equations and Related Topics. *St.Petersburg Math. J.*, **3**, 1165-1188.

[30] Yakubovich, V.A. and V.M. Staržinskii (1972). *Linear differential equations with periodic coefficients* (in Russian). Nauka Publ. House, Moscow (English version by J. Wiley, 1975).

[31] Wade, G.(1994) *S*ignal coding and processing. Cambridge Univ.Press, Cambridge.

STABILITY OF NEUTRAL TIME DELAY SYSTEMS: A SURVEY OF SOME RESULTS

S. A. Rodrìguez, J.-M. Dion, L. Dugard
Laboratoire d'Automatique de Grenoble (INPG CNRS UJF) ENSIEG,
BP 46, 38402, St. Martin d'Hères, FRANCE.
Email: Luc.Dugard@inp.fr

Abstract In this work, some of the recent developments concerning stability and robust stability analysis of neutral systems with uncertain parameters and uncertain delays are presented. Then, solutions and specific stability properties of neutral systems are discussed. The aim of this chapter, without being completely exhaustive, is to present important tools used to derive stability and robust stability properties for neutral systems.

Keywords: time-delay system, neutral system, stability, robust stability

1. Introduction

A great variety of systems can be modeled by time-delay systems (Kolmanovskii 1996), i.e. the "future" states depend not only on the "present" states, but also on the "delayed" states

$$\dot{x}(t) = f(t, x_t).$$

Indeed, the delay naturally occurs in the dynamical behavior of systems in many fields: mechanics, physics, etc. Even if the systems themselves do not have internal delays, closed loop systems may involve delay phenomena, because of actuators, sensors and computation time.

A neutral system is a general class of "time-delay systems" characterized by the fact that the behavior of the system depends both on the "delayed state" and on its derivative

$$\dot{x}(t) = f(t, x_t, \dot{x}_t).$$

Some examples of such neutral systems are given in (Brayton 1976), (Niculescu and Brogliato 1995), (Logemann and Townley 1996), (Mounier *et al.,* 1997), (Bellen *et al.,* 1999), (Hu and Davinson 2000). Neutral systems represent a very general class which includes as particular cases, ordinary time-delay systems. In this chapter, we will focus on linear neutral systems.

After the publication of two seminal papers (Hale and Cruz 1970), (Cruz and Hale 1970) which gave strong mathematical basis as well as some stability results, many studies have been dealt with time-delay systems of neutral type. In the last years, interesting works have been concerned with neutral systems see e.g. , (Logemann and Townley 1996), (Dugard and Verriest 1997), (Bellen *et al.,* 1999), (Tchangani *et al.,*1999), (Niculescu 2001), (Hale and Verduyn Lunel 2002), (Loiseau *et al.,* 2002), (Fridman and Shaked 2002).

Several works have been focused on the stability analysis of neutral systems either in the time domain approach, see for example (Infante and Castelan 1978), (Hale and Verduyn Lunel 1993), (Verriest and Niculescu 1997), or in the frequency domain approach, see for example (Chen 1995), (Verriest and Niculescu 1997). In these studies, the attention was mainly focused in giving conditions for delay independent stability (i.e. stability for any value of the time delay), which are rather conservative when the delays are unknown. It is then of interest to consider delay-dependent stability analysis, see (Chen 1995), (Ivanescu *et al.,* 2003), (Rodriguez *et al.,* 2001), (Rodriguez *et al.,* 2002).

In practice, model parameters are not precisely known, leading to the study of the robustness of the stability w.r.t. parameter uncertainties, see (Kharitonov 1998). Some physical systems can be represented by uncertain neutral models: for example, lossless transmission line models (Kolmanovskii and Myshkis 1999). It is then of interest to consider here uncertain neutral systems models.

The content of this chapter is as follows: in section 2, introductory example and the initial value problem are presented for neutral systems. Section 3 is devoted to stability analysis; some model transformations that allow to perform stability analysis are detailed; finally some delay-independent and delay-dependent stability results are given. Some concluding remarks end the chapter.

Notation. By I_m we denote the identity matrix of dimension m. Let vector $x \in \mathbb{R}^n$, then $\|x\|$ denotes the Euclidean norm of the vector. Let $\mathcal{C}\left([a,b],\mathbb{R}^n\right) \equiv \mathcal{C}\left[a,b\right]$ the space of continuous functions from $[a,b]$ (in general we take $[a,b]=[-r,0]$) to $\mathbb{R}^n$, with norm

$$\|\varphi\|_{\mathcal{C}} = \sup_{a \leq \theta \leq b} \|\varphi(\theta)\|.$$

$\mathcal{L}_2\left[-r,0\right]$ is the space of Lebesgue square integrable functions defined on $[-r,0]$ and $\mathcal{W}_p^1\left[-r,0\right]$ is the Sobolev space of absolutely continuous vector functions with p-th power integrable derivative. "$\cdot$" on the variable denotes the right-hand derivative at t, " $'$ " on the variable denote the upper right-hand derivative at t respectively. The functions x_t and $\dot{x}_t$

denote the restriction of $x(t)$ and $\dot{x}(t)$ to interval $[t-r,t]$, so that x_t is an element of $\mathcal{C}$ defined by $x_t(\theta) = x(t+\theta)$, $\theta \in [-r,0]$, and $\dot{x}_t$is an element of $\mathcal{W}^{1,2}$ or $\dot{x}(t)$ is differentiable defined by $\dot{x}_t(\theta) = \dot{x}(t+\theta)$ for $\theta \in [-r,0]$.

2. Neutral systems

In this section we introduce time-delay and neutral systems, then we present in a detailed way one example, we emphasize the functional initial condition value problem, and we introduce the mathematical description proposed in (Hale and Verduyn Lunel 1993).

A time delay system is a system of retarded type in which the derivative of the "state" at the present time is specified as a function of the past values of the "state" in some interval, for example:

$$\dot{x}(t) = Ax(t) + Bx(t-r)\ ,\ t > 0\ ,\ r > 0.$$

A neutral time delay system is a system in which the derivative of the "state" at the present time is specified not only as a function of the past values of the "state" in some interval but also of the derivative of the "state", for example:

$$\dot{x}(t) = Ax(t) + Bx(t-r) + C\,\dot{x}(t-r)\ ,\ t > 0\ ,\ r > 0.$$

Some times, open loop systems $\dot{x}(t) = Ax(t) + Bu(t)$ do not have internal delays but in closed loop, they may have: $\dot{x}(t) = Ax(t) + Bx(t-r)$, $t > 0$, $r > 0$, because of actuators, sensors and computer time. However in practice, many time-delay systems are obtained from models with partial differential equations of parabolic type, while time-delay systems of neutral type (or neutral systems) are obtained from models with partial differential equations of hyperbolic type. In this section, this fact is stressed with one example.

2.1. Lossless transmission line

Consider the following example, similar to the one given in (Brayton 1976), (Kolmanovskii and Myshkis 1999) which is a lossless transmission line, at the end ($x=0$) of which there is an external source of constant voltage E, while the other end ($x=l$) is grounded by means of a tunnel diode (such diodes are widely used in high-frequency amplifiers of electronic oscillators for example). The current $i(\cdot,\cdot)$ and voltage $v(\cdot,\cdot)$ are functions of t and x, and they satisfy the system of telegraph equations

$$L\frac{\partial i}{\partial t} + \frac{\partial v}{\partial x} = 0,\ C\frac{\partial v}{\partial t} + \frac{\partial i}{\partial x} = 0\ ,\ t > 0\ ,\ 0 \leq x \leq l, \qquad (1)$$

with boundary conditions

$$E = R_0 i(t,0) + v(t,0) \ , \ i(t,l) = C_l \frac{\partial v(t,l)}{\partial t} + i_d(v(t,l)) \, , \qquad (2)$$

where l is the length of the line, L, C are the inductance and capacity of the conductor per length, R is the resistance at the input, C_l is the capacity at the output, $i_d(v)$ is the current-voltage characteristic of the diode (a nonlinear polynomial function).

The equilibrium points (v_0, i_0) satisfy

$$E = v_0 + R_0 i_0 \ , \ i_0 = i_d(v_0) \, .$$

Suppose that the working point given by the first order approximation is

$$i(v_d) \approx i(v_0) + m(v_d - v_0) \ , \ m = \text{ constant } > 0.$$

The general solution of (1.1) is given by the d'Alembert's formula

$$\begin{aligned} v(t,x) - v_0 &= \varphi(t - bx) + \psi(t + bx) \, , \\ i(t,x) - i_0 &= Z^{-1}\left[\varphi(t - bx) - \psi(t + bx)\right] , \\ b \ &: = \sqrt{LC} \ , \ Z := \sqrt{LC^{-1}} , \end{aligned}$$

with the first boundary condition (1.2) leading to

$$\varphi(t) = (R_0 - Z)(R_0 + Z)^{-1} \psi(t) \, .$$

The second condition (1.2) gives the following linear neutral system

$$\psi'(\xi) - C\psi'(\xi - r) = A\psi(\xi) + B\psi(\xi - r) \, , \qquad (3)$$

where $A = \frac{m - \sqrt{\frac{C}{L}}}{C_\ell}$, $B = -\frac{\left(\sqrt{\frac{L}{C}} - R_0\right)\left(m + \sqrt{\frac{C}{L}}\right)}{\left(\sqrt{\frac{L}{C}} + R_0\right) C_l}$, $C = \frac{\left(\sqrt{\frac{L}{C}} - R_0\right)}{\left(\sqrt{\frac{L}{C}} + R_0\right)}$, $r = 2\sqrt{LC}l$. This shows a transformation from the partial differential equation (1.1) of hyperbolic type to the time-delay system of neutral type (1.3).

Now an "impulse response" for the lossless transmission line is given in the figure 1.1. The initial condition is $\psi(\xi) \equiv 0$ for $\xi < 0$ and $\psi(0) = 1$. It shows that the dynamics for linear neutral systems are very different from time-delay systems or linear time invariant systems.

Now some classical results about the initial value problem for neutral system are given.

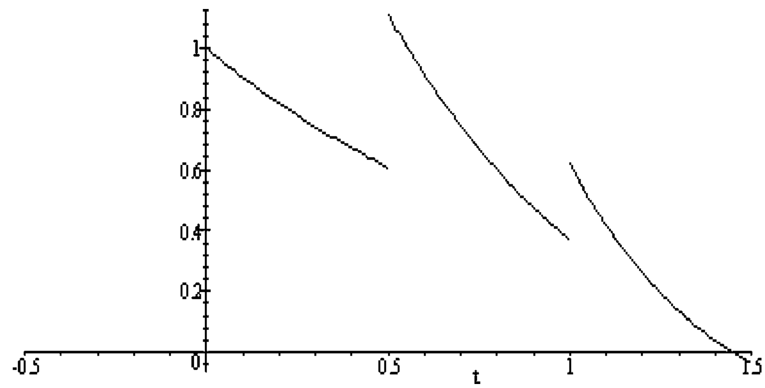


Figure 1. An impulse response for the lossless transmission line

2.2. The initial value problem and solutions

As equation (1.3) can be written in the form $[\psi(\xi) - C\psi(\xi - r)]' = A\psi(\xi) + B\psi(\xi - r)$, we present neutral systems in the Hale's form (Hale and Cruz 1970), (Kolmanovskii and Myshkis 1999) and the solutions in the simple space of continuous functions $\mathcal{C}[-r, 0]$. So there are other spaces in which the initial value problem may be considered, for example in the Sobolev space of absolutely continuous functions $\mathcal{W}_p^1[-r, 0]$ in which $\psi(\xi)$ is differentiable almost everywhere (see (Kolmanovskii and Nosov 1986) for $\mathcal{W}_2^1$ and (Henry 1974) for $\mathcal{W}_\infty^1$ and $\mathcal{W}_p^1$), or in the product space $\mathcal{M}_p = \mathbb{R}^n \times \mathcal{L}^p[-r, 0]$, in which only the difference $[\psi(\xi) - C\psi(\xi - r)]$ is differentiable almost everywhere.

Let Ω be an open set in $\mathbb{R}\times\mathcal{C}$. Consider the following neutral system, written in the form proposed by (Hale and Cruz 1970), (Hale and Verduyn Lunel 1993), :

$$\frac{d}{dt}[D(t)x_t] = F(t, x_t),\ t \geq t_0, \tag{4}$$

with the difference operator

$$D(t)\varphi := [\varphi(0) - G(t, \varphi)], \tag{5}$$

and initial condition

$$x_{t_0} \equiv \phi(\theta),\ \phi \in \mathcal{C},\ -r \leq \theta \leq 0, \tag{6}$$

where $F, G : \Omega \to \mathbb{R}^n$ are supposed to be continuous, $|F(t, \varphi)|, |G(t, \varphi)|$ are bounded uniformly in t for φ in compact sets of $\mathcal{C}$, with $F(t, 0) \equiv 0$,

$G(t,0) \equiv 0$ so that $x = 0$ is a solution of the System (1.4)-(1.6), and in order to have a well posed neutral system suppose $G(t,\varphi)$ non-atomic at zero (Hale and Cruz 1970). In particular, if $G(t,\varphi)$ depends only upon values of $\varphi(\theta)$ for $-r \leq \theta \leq -\varepsilon < 0$, then $G(t,\varphi)$ is non-atomic at zero.

Remark 1. In this approach, the state is $x_t(\theta) = x(t+\theta) \in \mathcal{C}$, $-r \leq \theta \leq 0$, and not $x(t) \in \mathbb{R}$. This is natural because the system is infinite dimensional. □

Under the previous conditions, there is a unique solution x_t, to the system (1.4)-(1.6), that needs not be differentiable since only the difference operator $D(\cdot)$ is differentiable on (t_0, ∞) (and right hand differentiable at t_0).

An important class of neutral systems concerns generalized linear autonomous neutral systems:

$$\frac{d}{dt}[Dx_t] = Lx_t\ ,\ t \geq 0, \tag{7}$$

$$D\varphi := \varphi(0) - \int_{-r}^{0} d\mu(\theta)\,\varphi(\theta)\,,$$

where D and L are bounded linear maps from $\mathcal{C}$ into $\mathbb{R}^n$. The initial data is given by (1.6) and μ is continuous at zero. For example, the following neutral system is a particular case of (1.7)

$$\frac{d}{dt}[x(t) - Cx(t-r)] = Ax(t) + Bx(t-r)\ ,\ t \geq 0\ ,\ r > 0. \tag{8}$$

where A, B, C are $n \times n$ constant matrices.

In practice, model parameters are not precisely known, leading to the study of the robustness of the stability w.r.t. parameter uncertainties, see (Kharitonov 1998). It is then of interest to consider the following uncertain neutral system:

$$\frac{d}{dt}[x(t) - Cx(t-r_1) - \Delta_C] = Ax(t) + Bx(t-r_2) + \Delta_A + \Delta_B, t \geq 0, \tag{9}$$

where $r_1 > 0$, $r_2 > 0$, and the uncertainty terms Δ_A, Δ_B, Δ_C will be precised later.

Now we give some stability results for neutral systems.

3. Stability of neutral systems

In this section we introduce the different stability notions and illustrate with a simple example the notion of delay-dependent stability. We present the different model transformations which are used for the stability analysis and then focus on both cases, delay-dependent stability and delay-independent stability.

3.1. Stability

Let us recall some basic facts about stability.

Definition 1. The trivial solution $x(t) = 0$ of (1.4)-(1.6) is said to be stable if for any $\varepsilon > 0$, $t_0 \in \mathbb{R}$, there is $\delta = \delta(\varepsilon, t_0)$ such that $\|\varphi\|_{\mathcal{C}} \leq \delta(\varepsilon, t_0)$ implies $\|x(t; t_0, \phi)\| \leq \varepsilon$. The zero solution of (1.4)-(1.6) is said to be asymptotically stable if it is stable and $x(t; t_0, \phi) \to 0$ as $t \to \infty$. The solution $x(t) = 0$ of (1.4)-(1.6) is said to be exponentially stable if there exist $\alpha > 0$, $\beta \geq 1$ such that every solution $x(t; t_0, \phi)$ of (1.4)-(1.6) with initial condition ϕ satisfies the inequality

$$\|x(t; t_0, \phi)\| \leq \beta \|\phi\|_{\mathcal{C}} e^{-\alpha t}, \forall\, t > t_0. \square$$

The system (1.7) is exponentially stable if its characteristic function (Hale and Verduyn Lunel 1993) given by

$$h(s) = \det\left(s\left(I - \int_{-r}^{0} e^{st} d[\mu(t)]\right) - \int_{-r}^{0} e^{st} d[\eta(t)]\right), \ s \in \mathbb{C}, \quad (10)$$

has no zero with nonnegative real parts. For example, the system (1.8) is exponentially stable if its characteristic function

$$p\left(s, e^{-rs}\right) = \det\left(s\left(I - Ce^{-rs}\right) - A - Be^{-rs}\right) \quad (11)$$

satisfies

$$\sup\left\{Re(s) : \det\left(s\left(I - Ce^{-rs}\right) - A - Be^{-rs}\right) = 0\right\} < 0.$$

Remark 2. [Hale and Verduyn Lunel 2002] For some linear neutral systems, asymptotic stability does not imply exponential stability as it happens for time-delay systems (Kharitonov 1998). For example (Hale and Verduyn Lunel 2002)

$$\frac{d}{dt}[x(t) - x(t-1)] = -ax(t), \ t \geq 0, \quad (12)$$

$$x_0 = \phi \in \mathcal{C},$$

with $a > 0$ implies $x \to 0$ as $t \to \infty$, but system (1.12) is not exponentially stable. $\square$

Now consider the following equation

$$D(t)\,x_t = D(t_0)\,\phi + H(t) - H(t_0)\ ,\ t \geq t_0, \tag{13}$$

$$x_{t_0} = \phi \tag{14}$$

$H \in \mathcal{C}([t_0,\infty),\mathbb{R}^n)$, $\phi \in \mathcal{C}([-r,0),\mathbb{R}^n)$ and a quasilinear difference operator $D(t)$ satisfying all the conditions given in section 1.2.3.

Definition 2. [Cruz and Hale 1970] Suppose that $\mathcal{H} \subset \mathcal{C}([t_0,\infty),\mathbb{R}^n)$. The difference operator is uniformly stable with respect to $\mathcal{H}$ if there are constants K, Λ such that for any $\phi \in \mathcal{C}$, $t_0 \in \mathbb{R}$, and H in $\mathcal{H}$, the solution $x(t_0;t,\phi,H)$ of (1.13)-(1.14) satisfies

$$\|x(t_0;t,\phi,H)\| \leq K\,\|\phi\|_{\mathcal{C}} + \Lambda \sup_{t_0\leq\tau\leq t} \|H(\tau) - H(t_0)\|\ ,\ t \geq t_0.\square \tag{15}$$

When the difference operator is independent of t, condition (1.15) implies that the roots of the equation

$$\det\left(I - \int_{-r}^{0} [d\mu(\theta)]\,\rho^{\theta}\right) = 0$$

have modulii $\leq (1-\varepsilon)$, $\varepsilon > 0$.

Then the difference operator $D\varphi := \varphi(0) - C\varphi(r)$ for (1.8) is stable if and only if C is a stable Schur-Cohn matrix (eigenvalues inside the unit circle).

Let $V : [t_0,\infty) \times \mathcal{C} \in\!\rightarrow \mathbb{R}^+$ be a continuous functional; the upper right-hand derivative of V along the solution of the system (1.4)-(1.6) is defined by

$$\dot{V}(t,\phi) = \lim_{h\rightarrow 0} \sup \frac{1}{h}\left[V(t+h, x_{t+h}(t,\phi)) - V(t,\phi)\right].$$

Recall the following Lyapunov-Krasovskii functional approach:

Theorem 1 *[Cruz and Hale 1970] Consider the Neutral System (1.4)-(1.6). Assume that $D(t)$ is uniformly stable with respect to $\mathcal{C}$ and that there exist non decreasing continuous functions $v_i : \mathbb{R}^+ \rightarrow \mathbb{R}^+$, $i = 1,2,3$ such that $v_i(0) = 0$ and $v_i(s) > 0$, for all $s > 0$ and $i = 1,2,3$. Then, the zero solution of (1.4)-(1.6) is asymptotically stable if there exists a continuous functional $V : [t_0,\infty) \times \mathcal{C} \rightarrow \mathbb{R}^+$ such that:*

$$v_1(\|D(t)\varphi\|) \leq V(t,\varphi) \leq v_2(\|\varphi\|_{\mathcal{C}})$$

$$\dot{V}(t,x_t) \leq -v_3(\|D(t)x_t\|)\ ,\forall\, t \geq t_0.$$

Now we show with a simple example how stability of time-delay systems depends on the systems parameters and may depend or not on the delays. Consider the following simple scalar time-delay system

$$\dot{x}\,(t) = -ax\,(t) - bx\,(t-r)\ ,\ t \geq 0\ ,\ r > 0\ ,\ (a,b) \in \mathbb{R} \times \mathbb{R}.$$

Under appropriate initial conditions, the characteristic equation associated to this system is

$$s + a + be^{-st} = 0,$$

it is a transcendental equation having an infinite number of solutions.

The use of the D-decomposition method (Kolmanovskii and Nosov 1986) gives a parametrization of the space (a, b) in several regions, each region being characterized by the same number of roots with positive real parts (see also (Niculescu *et al.*, 1997), (El'sgots and Norkin 1973)). Furthermore, each region is bounded by a "hypersurface" (here a first order one), in which at least one root of the characteristic equation lies on the imaginary axis for the corresponding parameters a, b, r. Simple computation proves that the corresponding regions are as shown in Figure 2. In S_∞ the stability is ensured independently of the size of the

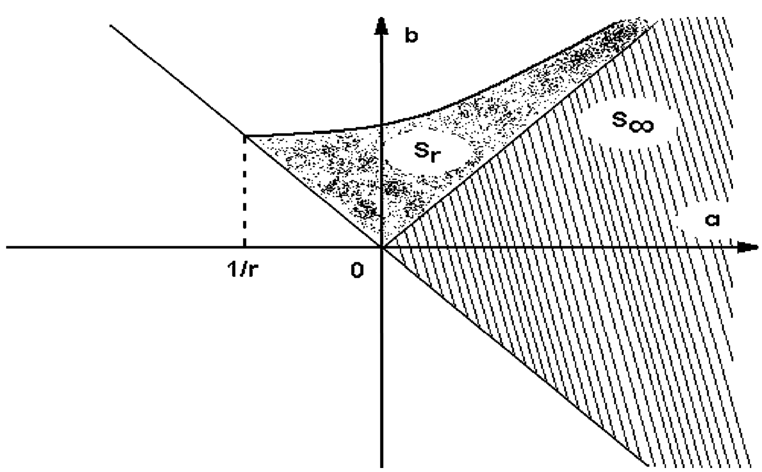


Figure 2. Stability regions, S_r and S_∞

delay and in $S_r \cup S_\infty$, the stability is ensured for the delay less than or equal to r. Then, in a general framework, given a criteria for stability of system (1.4), it can be classified into: delay-independent stability or delay-dependent stability, according to their dependence upon the size of delays (Dugard and Verriest 1997).

3.2. Transformations

In this part, we present some transformations in the frequency and in the time domain approaches that are used by many authors, see (Rekasius 1980), (Gu and Niculescu 1999) in order to prove stability or to deduce some important properties.

3.2.1. Transformation in the frequency domain

First, consider the time delay system obtained from (1.8) when $C = 0$,

$$\dot{x}(t) = Ax(t) + Bx(t-r), \; t \geq 0. \tag{16}$$

The stability properties can be studied by the analysis of the transfer matrix

$$H_{xu}(s) = (sI - A)^{-1} B, \tag{17}$$

associated to the simple linear time invariant system:

$$\dot{x}(t) = Ax(t) + Bu(t), \; t \geq 0. \tag{18}$$

Another transformation is to transform the characteristic function (1.10) by taking the Möbius transformation (Mathews and Howell 1996)

$$s = \frac{1-w}{1+w}, \; w \in \mathbb{C}. \tag{19}$$

It maps the left and right half planes of s to the inner and outer regions of the unit circle of w, and the imaginary axis of s to the unit circle $|w| = 1$.

Finally a two variable model is presented, in the variable s and in "the delay term" $z = e^{-rs}$. Then we can consider the characteristic polynomial in (1.11), not in the variable s, but in two independent variables (s, z), (Hu and Hu 1996)

$$p(s,z) = \det((sI - A) - (B + Cs)z), \; |z| \leq 1, \; (s, z \in \mathbb{C}). \tag{20}$$

3.2.2. Transformation in the time domain

Stability of time delay systems has been studied using model transformation under the assumption that the "state" is differentiable. In this case, some authors introduce transformation to prove stability, for example transform a time delay system of the form $\dot{x}(t) = \sum_{i=0}^{m} B_i x(t - h_i)$, $h_i \geq 0$, into a neutral time delay system written in the Hale's form, by integration over one delay interval (Kolmanovskii and Richard 1998)

$$\frac{d}{dt}\left[x(t) + \sum_{i=0}^{m} B_i \int_{t-h_i}^{t} x(\tau)\, d\tau\right] = \sum_{i=0}^{m} B_i x(t). \tag{21}$$

Note that this transformation is also applicable to the following neutral system (1.22) with $u = 0$, $r_i > 0$, (see (Lien *et al.,* 2000)); however here we propose the "inverse transformation", i.e. transform a neutral system into a time delay system. Consider the following neutral system

$$\frac{d}{dt}\left[x(t) - \sum_{i=1}^{m} C_i x(t - r_i)\right] = A_0 x(t) + \sum_{i=1}^{m} A_i x(t - r_i) + u(t - r), \quad r_i > 0, \; t \geq t_0, \tag{22}$$

with $r := \max\{r_i, i = 1, 2, ...m\}$. Now, if we choose the feedback as

$$u(t - r) = -\sum_{i=1}^{m} [A_0 C_i + A_i] \sum_{i=1}^{m} C_i x(t - r_i - r), \tag{23}$$

we obtain the following time delay system

$$\dot{y}(t) = A_0 y(t) + \sum_{i=1}^{m} [A_0 C_i + A_i] y(t - r), \tag{24}$$

where $y(t) := x(t) - \sum_{i=1}^{m} C_i x(t - r_i)$.

Another possible transformation is obtained when the neutral system

$$\dot{x}(t) = f(t, x_t, \dot{x}_t),$$

can be rewritten in the Hale's form (1.4) (Kolmanovskii and Myshkis 1999).

Now we consider the Leibnitz'rule in the state x_t

$$x(t) - x(t - r) = \int_{-r}^{0} \dot{x}(t + \theta)\, d\theta, \; t \geq 0, \; \theta \in [-r, 0], \tag{25}$$

then, for example, the time delay system (1.16) with pointwise delays is rewritten as a distributed system

$$\dot{x}(t) = (A + B) x(t) - B \int_{-r}^{0} [Ax(t + \theta) + Bx(t + \theta - r)]\, d\theta. \tag{26}$$

However, the transfer function associated to (1.26) is the product

$$H(s) = \left(sI - A - e^{-sr}B\right)\left(I - \frac{1 - e^{-sr}}{s} B\right),$$

i.e. the original transfer function associated to (1.16), $(sI - A - e^{-sr}B)$, and the additional dynamics,

$$\left(I - \frac{1 - e^{-sr}}{s} B\right), \tag{27}$$

introduced by the above transformation (Gu and Niculescu 1999). This idea can be extended to neutral systems by the supposition that $x(t)$ is differentiable i.e.

$$\dot{x}(t) = Ax(t) + Bx(t-r) + C\,\dot{x}(t-r)\ ,\ t \geq 0\ ,\ r > 0, \tag{28}$$

and (1.25) holds. Note that in this approach, the initial condition should be differentiable (Ivanescu *et al.,* 2003). This technical problem can be avoided if the Leibnitz'rule in the difference operator $D(t)\,x_t$, is used (Rodriguez *et al.,* 2002), (Carter and van Brunt 2000)

$$D(t)(x_t - x_{t-r}) = \int_{-r}^{0} d_\theta [D(t)\,x_{t+\theta}] = \int_{-r}^{0} \frac{d}{dt}[D(t)\,x_{t+\theta}]\,d\theta, \tag{29}$$

and the new system is given by

$$\begin{aligned} \frac{d}{dt}Dx_t &= (A+B)\,x(t) - BCx(t-r) + BCx(t-2r) \\ &\quad -B\int_{-r}^{0}[Ax(t+\theta) + BAx(t+\theta-r)]\,d\theta. \end{aligned}$$

This is a more general class of problem since the solution x may have discontinuities, as long as the difference $[x(t) - Cx(t-r)]$ is differentiable (Ivanescu *et al.,* 2003). In fact, this transformations holds in many interesting spaces (Rodriguez *et al.,* 2003) where the initial value problem is well posed: $\mathcal{W}_p^1([-r,0],\mathbb{R}^n)$ (Henry 1974), $\mathcal{C}([-r,0],\mathbb{R}^n)$ (Hale and Verduyn Lunel 1993), $\mathbb{R}^n \times \mathcal{L}^p([-r,0],\mathbb{R}^n)$ (Burns *et al.,* 1983). So, after applying this transformation to the neutral system (1.8), the transformed system has the same additional dynamics (1.27), as for the time-delay case (Ivanescu *et al.,* 2003).

In the next section, all the mentioned transformations in the frequency and time domain are applied to derive delay-independent and delay dependent stability results to neutral systems.

3.3. Delay-Independent Stability

In the last section, a transformation from neutral system (1.22)-(1.23) into a time delay system (1.24) was proposed in a time domain framework; here, its stability is studied.

Theorem 2 *Consider the system (1.24). Suppose* $0 < r_j \leq r$ *are real numbers such that the ratios* r_i/r_j *are rational if* $m > 1$ *and all roots of the equation*

$$\det\left[I - \sum_{i=1}^{m} C_i \rho^{-r_i}\right] = 0, \tag{30}$$

have modulii less than 1. Then, if the time delay system (1.24) is asymptotically stable (equivalently exponential stable (Kharitonov 1998)), then the closed loop system (1.22)-(1.23) is uniformly stable.

Proof. If (1.24) is asymptotically stable, then $\|y(t;t_0,\phi)\| \leq \beta \|\phi\|_{\mathcal{C}} e^{-\alpha t}$, i.e.

$$\left\| x(t) - \sum_{i=1}^{m} C_i x(t - r_i) \right\| \leq \beta \|\phi\|_{\mathcal{C}} e^{-\alpha t} \ , \ \alpha > 0 \ , \ \beta \geq 1.$$

Now since (1.30) holds, the difference operator

$$D\varphi := \varphi(0) - \sum_{i=1}^{m} C_i \varphi(-r_i)$$

is uniformly stable, with respect to $\mathcal{C}([t_0,\infty), \mathbb{R}^n)$ (see (Cruz and Hale 1970)), and then $x(t) \to 0$ uniformly implies that the closed loop neutral system (1.22)-(1.23) is uniformly stable. ■

In theorem 2, it is supposed that the time-delay system (1.24) is stable then the small gain theorem criteria can be invoked to prove stability (Dugard and Verriest 1997), (Datko 1985).

Now, in the frequency domain approach, (Hu and Hu 1996) uses the Möbius transformation (1.19) and the characteristic polynomial (1.20) to prove the following lemma (see also (Siljak 1975)):

Lemma 1 *Consider the two variable polynomial (1.20). If the conditions that $p(s,0) \neq 0$ for s such that $Re(s) \geq 0$ and $p(s,z) \neq 0$ for (s,z) such that $Re(s) = 0$ and $|z| \leq 1$ hold, then $p(s,z) \neq 0$ for (s,z) such that $Re(s) \geq 0$ and $|z| \leq 1$.*

This lemma can be considered as the extension to the neutral case of proposition 10 in (Niculescu *et al.*, 1997) to time delay systems (Hale *et al.*, 1985).

Also in (Hu and Hu 1996) some easy to check delay-independent stability results are proposed, one of them is the next theorem:

Theorem 3 [*Hu and Hu 1996*] *The system (1.28) is asymptotically stable if the following two conditions hold*

$$\mu(A) + \|B\| + \frac{\|CA\| + \|CB\|}{1 - \|C\|} < 0 \ ; \ \|C\| < 1.$$

Here $\mu(A) = \lim_{\epsilon \to 0} \frac{1}{\epsilon}(\|I + \epsilon A\| - 1)$, *and* $\|\cdot\|$ *the matrix norm.*

In the time domain or Lyapunov-Krasovskii approach (Lien *et al.*, 2000), one has the following delay-independent stability result:

Theorem 4 [*Lien et al., 2000*] *The system (1.22) with $u = 0$ is globally uniformly asymptotically stable provided that $\sum_{i=1}^{m} \|C_i\| < 1$ and that there exist some matrices $R_i > 0$, $i = 1, 2, ...m$, such that*

$$\begin{pmatrix} A_0 + A_0^\top + \sum_{i=1}^{m} R_i & A_1 - A_0^\top C_1 & \cdots & A_m - A_0^\top C_m \\ A_1^\top - C_1^\top A_0 & -C_1^\top A_1- & \cdots & -C_1^\top A_m \\ & -A_1^\top C_1 - R_1 & & \\ A_2^\top - C_2^\top A_0 & -A_2^\top C_1 & \cdots & -C_2^\top A_m \\ \vdots & \vdots & \ddots & \vdots \\ A_m^\top - C_m^\top A_0 & -A_m^\top C_1 & \cdots & -C_m^\top A_m- \\ & & & -A_m^\top C_m - R_m \end{pmatrix} < 0.$$

3.4. Delay-Dependent Stability

In the time domain or Lyapunov-Krasovskii approach, the work presented in (Lien *et al.,* 2000) uses the transformation (1.21) to neutral system (1.22) with $u = 0$ to obtain the following delay-independent stability result:

Theorem 5 [*Lien et al., 2000*] *The system (1.22) with $u = 0$ is globally uniformly asymptotically stable provided that $\sum_{i=1}^{m} (\|C_i\| + r_i \|A_i\|) < 1$ and for some matrices $Q > 0$ and $R_i > 0$, $i = 1, 2, ...m$, there exists a solution $P > 0$ for the following Riccati equation*

$$\overline{A}^\top P + P\overline{A} + \\ + \sum_{i=1}^{m} \left[(r_i + m_i) R_i + r_i \overline{A}^\top P A_i R_i^{-1} \overline{A}_i^\top P \overline{A} + \overline{A}^\top P C_i R_i^{-1} C_i^\top P \overline{A} \right] = -Q,$$

where $\overline{A} = A_0 + \sum_{i=1}^{m} A_i$, $m_i = 0$ if $C_i = 0$, $m_i = 1$ if $C_i \neq 0, i = 1, 2, ...m$.

Now one presents the delay dependent robust stability problem. (Hu and Davinson 2000) equation (1.9) is considered with $r_1 = r_2$ in the frequency framework, where uncertainties $\Delta_J := E\delta_J F_J$, $J \in \{A, B, C\}$ and $\delta_A \in \mathbb{R}^{m \times p_A}$, $\delta_B \in \mathbb{R}^{m \times p_B}$, $\delta_C \in \mathbb{R}^{m \times p_C}$ denote the perturbation matrices, and $E \in \mathbb{R}^{n \times m}$, $F_A \in \mathbb{R}^{p_A \times m}$, $F_B \in \mathbb{R}^{p_B \times m}$, $F_C \in \mathbb{R}^{p_C \times m}$ are known scaling matrices. Let $\delta := [\delta_A,\ \delta_B,\ \delta_C]$.

Definition 3. The real structured stability radius of (1.9) is

$$r_{\mathbb{R}}(A, B, C; E, F_A, F_B, F_C) := \\ = \inf \left\{ \overline{\sigma}(\delta) : \delta \in \mathbb{R}^{n \times (p_A + p_B + p_C)} \textit{ and (1.9) is unstable.} \right\}$$

Here $\sigma_i(M)$, $i = 1, 2, ..., \min\{n, (p_A + p_B + p_C)\}$ denotes the singular values of $M \in \mathbb{C}^{n \times (p_A + p_B + p_C)}$, nonincreasingly ordered and $\sigma_1(M) := \overline{\sigma}(M)$. □

Theorem 6 [*Hu and Davinson 2000*] *Let the nominal system (1.9) be stable, then the real structured stability radius of (1.9) is given by*

$$r_{\mathbb{R}} = \left\{ \sup_{s\in\mathbb{C}_g} \inf_{\gamma\in(0,1]} \sigma_2 \left(\begin{bmatrix} ReQ & -\gamma ImQ \\ \gamma^{-1} ImQ & ReQ \end{bmatrix} \right) \right\}^{-1},$$

where

$$Q(s) := \begin{pmatrix} F_A \\ F_B \\ F_C \end{pmatrix} \left(sI - A - Be^{r_1 s} - Ce^{r_1 s}\right)^{-1} E \in \mathbb{C}^{n\times(p_A+p_B+p_C)}.$$

The proof of this theorem uses the main result of (Qiu *et al.*, 1995)

Note that in theorem 6, the nominal system (1.9) is assumed to be stable; it is a common practice in the study of robust stability. The next result uses the same assumption.

In (Kharitonov *et al.*, 2002), a constructive time domain approach was proposed to build a Lypunov-Krasovskii funtional to neutral system (1.28), in the space $\mathcal{W}_2^1$ for the initial condition ϕ.

Lemma 2 [*Kharitonov et al., 2002*] *Let the system (1.28) be stable. Given positive definite $n \times n$ matrices W_1, W_2, W_3, one can define the functional*

$$w(\varphi) := \phi^\top(0)W_1\phi(0) + \int_{-h}^{0} \phi^\top(\theta)W_2\phi(\theta)d\theta + \phi^\top(-h)W_3\phi(-h). \quad (31)$$

Then the Lyapunov-Krasovskii functional $v(x_t)$ defined by

$$v(x_t) = x^\top(t)[U(0) - C^\top U(r) - U^\top(r)C + C^\top U(0)C]x(t) +$$
$$+2x^\top(t)\int_{-r}^{0} [U^\top(r+\theta) - C^\top U^\top(\theta)]\,[Bx(t+\theta) + C\dot{x}(t+\theta)]\,d\theta +$$
$$+\int_{-r}^{0}\int_{-r}^{0} [Bx(t+\theta_1) + C\dot{x}(t+\theta_1)]^\top U(\theta_1-\theta_2)[Bx(t+\theta_2) +$$
$$+C\dot{x}(t+\theta_2)]d\theta_1 d\theta_2 + \int_{-r}^{0} x^\top(t+\theta)[(r+\theta)W_2 + W_3]x(t+\theta)d\theta.$$

is solution of

$$\frac{d}{dt}v(x_t)|_{(1.28)} = -w(x_t), \quad (32)$$

along the solutions of the system (1.28). Here

$$U(\xi) = \int_{0}^{\infty} K^\top(\tau)[W_1 + rW_2 + W_3]K(\tau+\xi)d\tau \quad (33)$$

is the Lyapunov matrix valued function for equation (1.28) and K *is the fundamental matrix associated to equation (1.28) (Bellman and Cooke 1963). Since the neutral system (1.28) is stable, such a functional always exists and satisfies the following bounds*

$$\omega_1 \left(\|\varphi(0)\|\right) \leq w(\varphi) \leq \omega_2 \left(\|\varphi\|_{\mathcal{W}}\right). \tag{34}$$

In (Rodriguez *et al.*, 2003 bis) the delay-dependent robust stability is given with the help of this Lypunov-Krasovskii funtional.

Theorem 7 [*Rodriguez et al., 2003 bis*] *Let the system (1.28) be exponentially stable, then the perturbed system (1.9) (with* $r_1 \equiv r_2$, $\Delta_C \equiv 0$*) remains stable for all uncertainties satisfying*

$$\Delta_A^{\mathsf{T}} R_A \Delta_A \leq \rho_A I_n \, , \tag{35}$$

$$\Delta_B^{\mathsf{T}} R_B \Delta_B \leq \rho_B I_n,$$

if the matrix C *is Schur-Cohn stable and if there exist positive matrices* W_1, W_2, W_3, *and a positive scalar* μ *such that*

$$\begin{aligned} &i)\ W_1 > \mu U(0) \left(R_A^{-1} + R_B^{-1}\right) U(0) + \tfrac{2}{\mu} \left(\rho_A I + \rho_A r I\right); \\ &ii)\ W_2 > \mu B^{\mathsf{T}} U(\theta + r) \left(R_A^{-1} + R_B^{-1}\right) U^{\mathsf{T}}(\theta + r) B + \\ &+\mu C^{\mathsf{T}} U'(\theta + r) \left(R_A^{-1} + R_B^{-1}\right) [U'(\theta + r)]^{\mathsf{T}} C \ , \ \forall \theta \in (-r, 0)\, ; \\ &iii)\ W_3 > \mu C^{\mathsf{T}} U(0) \left(R_A^{-1} + R_B^{-1}\right) U(0) C + \tfrac{2}{\mu} \left(\rho_B I + \rho_B r I\right). \end{aligned} \tag{36}$$

Finally consider the neutral system (1.9) with initial condition (1.6) and bounds satisfying the following assumptions by

$$\begin{gathered} \Delta_A \left(t, \varphi\left(0\right)\right) := E_A \delta_A \left(t, \varphi\left(0\right)\right), \\ \delta_A^T \left(t, \varphi\left(0\right)\right) \delta_A \left(t, \varphi\left(0\right)\right) \leq \varphi^T \left(0\right) W_A^T W_A \varphi\left(0\right), \\ \Delta_B \left(t, \varphi\left(-r_2\right)\right) := E_B \delta_B \left(t, \varphi\left(-r_2\right)\right), \\ \delta_B^T \left(t, \varphi\left(-r_2\right)\right) \delta_B \left(t, \varphi\left(-r_2\right)\right) \leq \varphi^T \left(-r_2\right) W_B^T W_B \varphi\left(-r_2\right), \\ \Delta_C \left(t, \varphi\left(-r_1\right)\right) := E_C \delta_C \left(t\right) \varphi\left(-r_1\right), \\ W_C \pm \delta_C \left(t\right) \geq 0, \\ \forall \left(t, \varphi\right) \in \mathbb{R}^+ \times \mathcal{C}, \end{gathered} \tag{37}$$

and suppose also that $\Delta_C(t, \varphi)$ depends only upon values of $\varphi(\theta)$ for $-r \leq \theta \leq -\varepsilon < 0$, see (Hale and Cruz 1970).

The matrices E_A, E_B and $E_C \equiv I_n$ are known, and the matrices W_A, W_B and W_C are given weighting matrices. The unknown mappings δ_A, δ_B satisfy the conditions

$$\delta_A(t,0) \equiv 0 \ , \ \delta_B(t,0) \equiv 0, \tag{38}$$

so that $x = 0$ is a solution of the neutral differential equation (1.9) with initial condition (1.6). Then we have the following result (close to (Rodriguez *et al.*, 2002)):

Theorem 8 [*Rodriguez et al., 2003*] *The Neutral System (1.9), (1.6), (1.38) is robustly delay-dependent asymptotically stable for any* $r_2 \leq r_2^*$ *if*

1 $A_1 := A + B$ *is a Hurwitz stable matrix;*

2 *The difference operator* $D(t)\varphi := [\varphi(0) - C\varphi(-r_1) + \Delta_C(t,\varphi)]$ *is linear in* φ*, continuous and uniformly stable with respect to* $\mathcal{C}$ *and* $\Delta_C(t,\varphi)$ *is nonatomic at zero; vspace-0.25cm*

3 *there exist a real positive number* r_2^* *and positive definite matrices* P*,* $S_i > 0$*,* $i = \overline{1,7}$ *such that the following LMIs hold:*

$$\Gamma := \begin{pmatrix} Q(r_2^*) & \Omega_{12} & 0 & \Omega_{14} & \overline{S}(r_2^*) \\ \Omega_{12}^\top & \Omega_{22} & 0 & 0 & 0 \\ 0 & 0 & \Omega_{33} & 0 & 0 \\ \Omega_{14}^\top & 0 & 0 & \Omega_{44} & 0 \\ \overline{S}^\top(r_2^*) & 0 & 0 & 0 & R \end{pmatrix} < 0, \tag{39}$$

$$S := W_A^\top W_A + \sum_{i=1}^{2} S_i + r_2^* \sum_{i=3}^{5} S_i > 0, \tag{40}$$

where

$$Q(r_2^*) := \Omega_{11} + \overline{S}(r_2^*) R^{-1} \overline{S}^\top(r_2^*), \tag{41}$$

$$\Omega_{11} := PA_1 + A_1^\top P + 2S - \overline{S}(r_2^*) R^{-1} \overline{S}^\top(r_2^*) \tag{42}$$

$$\overline{S}(r_2^*) := \begin{pmatrix} \overline{S}_1 & \overline{S}_2(r_2^*) \end{pmatrix}, \tag{43}$$

$$\overline{S}_1 := \begin{pmatrix} PA_1 & \sqrt{2}PB & PE_A & PE_B \end{pmatrix}, \tag{44}$$

$$\overline{S}_2(r_2^*) := \sqrt{r_2^*} \begin{pmatrix} PBE_A & PBE_B & PBA & PB^2 \end{pmatrix}, \tag{45}$$

$$R^{-1} := \begin{pmatrix} -I_{6n} & 0 & 0 \\ 0 & -S_5^{-1} & 0 \\ 0 & 0 & -S_6^{-1} \end{pmatrix}, \tag{46}$$

$$\Omega_{12} := \left(PA + W_A^\top W_A + \sum_{i=1}^{2} S_i + r_2^* \sum_{i=3}^{5} S_i \right) C, \tag{47}$$

$$\Omega_{22} := 2W_C^\top W_C - S_1 + S_7 + C^\top SC + 3W_C^\top SW_C \tag{48}$$

$$\Omega_{33} := W_B^\top W_B - S_2 + r_2^* S_6, \tag{49}$$

$$\Omega_{44} := W_C^\top W_C - S_7, \tag{50}$$

$$\Omega_{14} := PBC. \tag{51}$$

Notice that condition 1 is necessary and directly follows from the satisfaction of condition 3. If $\Delta_C \equiv 0$ then $E_C \equiv 0$, then (1.39) is a LMI equivalent to the one given in (Rodriguez *et al.*, 2002); however if $\Delta_C \neq 0$, the result given in (Rodriguez *et al.*, 2002) cannot be checked by using LMI Tools (Boyd et *et al.*, 1994). Condition 2 consists in checking the uniformly stability of the difference operator $D(t)\varphi := [\varphi(0) - C\varphi(-r_1) + \Delta_C(t,\varphi)]$ with respect to $\mathcal{C}([\sigma,\infty), \mathbb{R}^n)$, see Section 1.3.

Remark 3. Theorem (8) does not suppose that the nominal system is stable as theorems (6) and (2). □

4. Concluding remarks

This work covers a part of the large number of contributions in modeling, delay-dependent stability and delay-independent stability of linear neutral systems. Two approaches are presented: the frequency domain and the time domain. Important practical tools (Lyapunov matrix, LMIs, measure matrix, Riccati equation, ...) are given in order to conclude stability and robust stability results via some model transformations.

An effort has been made to present a difficult subject in a comprehensive manner with an example, while giving a fairly wide bibliography on the subject.

References

[1] Bellen A., N. Guglielmi and A.E. Ruehli (1999) Methods for Linear Systems of Circuits Delays Differential Equations of Neutral Type. IEEE Trans. Circuits and Sys. 46, 1:212–216.

[2] Bellman R. and K.L. Cooke (1963) Differential Difference Equations. Academic Press, New York.

[3] Boyd S., L.El Ghaoui, E. Feron and V. Balakrishnan (1994) Linear Matrix Inequalities in System and Control Theory. Society for Industrial and Applied Mathematics.

[4] Brayton R. (1976) Nonlinear Oscillations in a Distributed Network, Quart. Appl. Math., 24: 289-301.

[5] Burns J.A. Herdman T.L and Stech H.W (1983) Linear Functional Differential Equations As Semigroups On Products Spaces. SIAM. J. Math. Anal., 14:98–116.

[6] Carter M. and B. Van Brunt (2000). The Lebesgue-Stieljes Integral A Practical Introduction. Springer.

[7] Chen J. (1995) On Computing the Maximal Delay Intervals for Stability of Linear Delay Systems. IEEE Trans. Automat. Contr., 40:1087–1093.

[8] Cruz M.A. and J.K. Hale (1970) Stability of functional differential equations of neutral type. J. Differential Eqns. 7:334–355.

[9] Datko R.A. (1985) Remarks Concerning the Asymptotic and Stabilization of Linear Delay Differential Equations. J. Math. Anal. Appl., 111: 571–584.

[10] Dugard L. and E. Verriest, (Editors) (1997) Stability and Control of Time-Delay Systems. Springer Verlag.

[11] El'sgots and Norkin (1973) Introduction to Theory and Applications of Differential Equations with deviating arguments. Academic Press.

[12] Fridman and Shaked (2002) LMI Approach To H Infinity Filtering Of Linear Neutral Systems With Delay. 15th Triennial World Congress of the IFAC.

[13] Infante E.F. and W.B. Castelan (1978) A Liapunov Functional for a Matrix Difference-Differential Equation. Jounal Differential Equations. 29:439–451.

[14] Hale J.K. and M.A. Cruz (1970) Existence, Uniqueness and Continuous Dependence for Hereditary Systems. Ann. Mat. Pura Appl. 85, 4:63–82.

[15] Hale J.K., E.F. Infante and F.S.P. Tsen (1985) Stability of Linear Delay Equations. J. Math. Anal. Appl., 105:533–555.

[16] Hale J.K. and S.M. Verduyn Lunel (1993) Introduction to Functional Differential Equations. Springer-Verlag.

[17] Hale J.K. and S.M. Verduyn Lunel (2000)

[18] Hale J.K. and S.M. Verduyn Lunel (2002) Strong Stabilization of Neutral Functional Differential Equations. IMA Jounal of Mathematical Control and Information:5–23.

[19] Henry D. (1974) Linear Autonomous Neutral Functional Differential Equations. J. Diff. Equ., 15:106–128.

[20] Hu G. and E. Davinson (2000) Real Stability Radii for Linear Neutral Systems. 2nd IFAC Workshop on Linear Time-Delay Systems. Ancona Italy:117–122.

[21] Hu Guang-Di and Guang-Da Hu (1996) Some Simple Criteria for Stability of Neutral Delay-Differential Systems. App. Math. and Comp. 80:257–271.

[22] Ivanescu D., S. Niculescu, L. Dugard, J.M. Dion and E.I. Verriest (2003) On Delay Dependent Stability for Linear Neutral Systems. Automatica. 39:255–261.

[23] Gu K. and Niculescu S. (19989) Additional Dynamics in Transformed Time-Delay Systems. Proc. CDC, Phoenix, Arizona,USA:4673–4677.

[24] Kharitonov V.L. (1998) Robust Stability Analysis of Time-Delay Systems: a Survey. Proc. IFAC Syst. Struct.

[25] Kharitonov V.L., Rodriguez S.A., J.M. Dion and L. Dugard. (2002) Lyapunov-Krasovskii Functionals for Neutral Systems: A Constructive Approach. Congreso

Latinoamericano de Control Automatico CLCA la 2002 Guadalajara México: 3-6 Dic 2002.

[26] Kolmanovskii V.B. (1996) The stability of Hereditary Systems of neutral type. J. Appl. Maths. Mechs. 60, 2:205–216.

[27] Kolmanovskii V.B. and A.D. Myshkis (1999) Introduction to the Theory and Applications of Functional Differential Equations. Kluwer Academic Publishers.

[28] Kolmanovskii V.B. and V.R. Nosov (1986) Stability of Functional Differential Equations. Kluwer Academic Publishers.

[29] Kolmanovskii V.B. and J.P. Richard (1998) Stability of Systems With Pure, Discrete Multi-Delays. IFAC Conference System Structure and Control Nantes, France:13–18.

[30] Lien C.-H., K.W. Yu and J.-G. Hsieh (2000) Stability Conditions for a Class of Neutral Systems With Multiple Time Delays. J. Math. Anal. and Appl. 245:20–27.

[31] Logemann H. and S. Townley (1996) The effect of Small Delays in the Feedback Loop on the Stability of Neutral Systems. Syst. Contr. Lett. 27:267–274.

[32] Loiseau J.J., M. Cardelli and X. Dusser (2002). Neutral-Type-Time-Delay Systems That Are Not Formally Stable Are Not BIBO Stabilizable. IMA Jounal of Mathematical Control and Information:217–227.

[33] Mathews J.H. and R.W. Howell (1996) Complex Analysis for MAthematics and Engineering. Wm C. Brown Publishers.

[34] Mounier H., P. Rouchon and J. Rudolph, (1997) Some Examples of Linear Systems with Delays. RAIRO-APII-JESA (Jounal Europeen des Systemes Automatises). 31,6:9111–925.

[35] Niculescu S.I. (2001) On Robust Stability of Neutral Systems, Special issue On Time-Delay Systems. Kybernetica 37:253–263.

[36] Niculescu S.I., E.I.Verriest, L. Dugard and J.-M. Dion (1997) Delay-Independent Stability of LNS: A Riccati Equation Approach, in Stability and Control of Time-Delay Systems. L. Dugard and E.I. Verriest, Eds. Springer 22:92–100.

[37] Niculescu S.I. and B. Brogliato (1995) On Force Measurement Time-Delay Systems, Proc. IFAC Syst. Struc. Contr. Nantes France:266–271.

[38] Rekasius Z.V. (1980) A Stability Test for Systems with Delays. Proc. Joint Automatic Control Conf. TP9-A

[39] Richard J.P. (1998) Some Trends and Tools for the Study of Time Delay Systems. Proc. IEEE-IMACs Conference Comp. Eng. in Sys. App., Nabeul, Tunisia.

[40] Rodriguez S.A., J.M. Dion, L. Dugard and D. Ivănescu (2001) On delay-dependent robust stability of neutral systems. 3rd IFAC Workshop on Time Delay Systems, Santa Fe USA:101–106.

[41] Rodriguez S.A., J.M. Dion and L. Dugard (2002) Robust Stability Analysis of Neutral Systems Under Model Transformation. 41st IEEE Conference on Decision and Control, Las Vegas, Nevada, USA:1850–1855.

[42] Rodriguez S.A., J.M. Dion and L. Dugard (2003) Robust Delay Dependent Stability Analysis of Neutral Systems. CNRS-NSF Workshop , 22-24 Enero. Submited to Springer.

[43] Rodriguez S.A., V. Kharitonov, J.M. Dion and L. Dugard (2003) Robust Stability of Neutral Systems: A Lyapunov-Krasovskii Constructive Approach. Internal report.

[44] Qiu L., B. Bernhardsson, A. Rantzer, E.J. Davison, P.M. Young and J.C. Doyle (1995). A Formula for Computation of the Real Stability Radius. Automatica:879–890.

[45] Siljak D.D. Criteria for Two-Variables Polynomials (1975). IEEE Trans. Circuit and Systems, CAS-22, 3:185–189.

[46] Tchangani (1999)

[47] Verriest E.I. and S.I. Niculescu (1997) Delay-Independent Stability of LNS: A Riccati Equation Approach, in Stability and Control of Time-Delay Systems. (L. Dugard and E.I. Verriest, Eds.), Springer-Verlag L. 228:92–100.

SLICOT-BASED ADVANCED AUTOMATIC CONTROL COMPUTATIONS

Vasile Sima
National Institute for Research
Development in Informatics
71316 Bucharest 1, Romania,
E-mail: vsima@iciadmin.ici.ro

Abstract Recent advances in the SLICOT Library for computer-aided control systems analysis and design (CACSD) computations are addressed. Functional and performance capabilities of the SLICOT algorithms and software are presented. The main emphasis is put on the reliability, accuracy, and efficiency of the computational tools. For many basic control problems, SLICOT calculations are performed several times faster than with the currently used MATLAB toolboxes, at comparable accuracy, and with increased reliability.

Keywords: Computer-aided control system design; linear multivariable system; Lyapunov equation; numerical algorithm; numerical linear algebra; Riccati equation; ningular value decomposition; ntate-space representation.

1. Introduction

Many relevant practical control problems have high dimensionality and often involve ill-conditioned subproblems, causing failures of the theoretically well-developed, traditional methods. Moreover, for large-scale problems, it is extremely important to achieve the best computational efficiency, by exploiting any special structure, and by using the potential of modern high-performance computer architectures. Therefore, developing reliable and efficient algorithms for control systems analysis and design is an objective of primary importance. Recent advances in systems and control theory, numerical linear algebra, and scientific computations should be taken into account in order to achieve this objective.

Improved routines (in terms of numerical stability, efficiency, and functionality) implementing basic algorithms for *computer-aided control system design*—CACSD—have been recently developed and incorporated in the Fortran 77 **S**ubroutine **L**ibrary **in** **Co**ntrol **T**heory—SLICOT[1]— [4], within the

European "Numerics in Control" network. Initially developed by the Working Group on Software (WGS), partly with external contributors, the first two SLICOT releases, [16, 17], based on the NAG mathematical library, have been commercial versions, distributed by the Numerical Algorithms Group, Ltd. (NAG). Starting with Release 3 (1997) on, SLICOT became copyrighted freeware and can be downloaded from the Internet addresses `ftp://wgs.esat.kuleuven.ac.be/`, directory `pub/WGS/SLICOT/`, or `http://www.win.tue.nl /niconet/`. On line `.html` documentation files are available.

SLICOT library is currently built on a nucleus of basic numerical linear algebra subroutines from the state-of-the-art software packages LAPACK (Linear Algebra Package) [1] and the three levels of BLAS (Basic Linear Algebra Subprograms) [6, 7, 13], and partly on their extensions for parallel computers, ScaLAPACK and PBLAS. Besides the inherited capabilities to exploit some modern high-performance computer architectures, new algorithmic developments for the control systems analysis and design have been included. The conversion of the library to a public-domain software package offered the opportunity to improve the modularity, functionality, reliability, as well as the performance of the codes, e.g., by using calls to BLAS Level 3 and LAPACK block algorithms whenever possible, and/or by exploiting any special problem structure. Moreover, an impressive number of new routines have been added since 1998, and interfaces to MATLAB [15] and Scilab [9] are provided for many typical CACSD calculations. These features turned SLICOT into a powerful and convenient computational engine for control-related applications, covering system identification, analysis, model reduction, system transformations, and control synthesis. Often, SLICOT computations are several times faster than MATLAB computations, at comparable or better accuracy.

This paper presents an overview of the current version of the SLICOT Library. Its functional capabilities are summarized, and typical performance results are given.

2. Reliability and accuracy

Current research in numerical algorithms focuses on the exploitation of any structural information of the underlying computational problem. Structure-preserving algorithms ensure that structural properties of a problem are preserved during finite precision computations. Therefore, the computed result can be thought as the exact solution of the original problem with perturbed input data. Besides increasing the reliability of the returned results, this often improves their accuracy, as shown below.

Table 1. Results computed using MATLAB `eig` and SLICOT routines

Ex.	MATLAB	SLICOT
(1)	1.999999999999996e+1	2.000000000000000e+01
	−2.642122482758924e-16+1.258258465000505e-7i	0.000000000000000e+00
	−2.642122482758924e-16−1.258258465000505e-7i	0.000000000000000e+00
	−2.000000000000000e+1	−2.000000000000000e+01
(2)	−1.414213562373094e+1	1.414213562373095e+01
	−1.694239404281922e-16+7.465277149295360e-8i	0.000000000000000e+00
	−1.694239404281922e-16−7.465277149295360e-8i	0.000000000000000e+00
	1.414213562373095e+1	−1.414213562373095e+01
(3)	−1.282269210672000e+1	−1.282269210672000e+01
	1.282269210672000e+1	1.282269210672000e+01
	−2.220446049250313e-16+7.837182712157440e-0i	7.837182712157435e+0i
	−2.220446049250313e-16−7.837182712157440e-0i	−7.837182712157435e+0i

Example 1. Consider the real matrices

$$H = \begin{pmatrix} A & F \\ Q & -A^T \end{pmatrix}, \qquad J = \begin{pmatrix} 0 & I_n \\ -I_n & 0 \end{pmatrix},$$

where A, F, $Q \in \mathbb{R}^{n\times n}$ and F and Q are symmetric matrices, $F = F^T$, $Q = Q^T$. Matrix H satisfies $JH = (JH)^T$; such matrices are called *Hamiltonian.* But $J^T H^T J = -H$, since $J^T = -J = J^{-1}$. Hence, $-H$ has the same eigenvalues as H^T and H, and so, if $\lambda \in \lambda(H)$, the set of eigenvalues of H, then $-\lambda \in \lambda(H)$, with the same (algebraic) multiplicity. However, if the eigenvalues are computed numerically with a standard eigensolver (for instance, `eig` from MATLAB, or DGEES from LAPACK), this pairing property can no longer be guaranteed, as shown for the three very simple numerical examples below:

$$A = \begin{pmatrix} 10 & 10 \\ 0 & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 0 & 10 \\ 10 & 10 \end{pmatrix}, \quad Q = F, \tag{1}$$

$$A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 10 & 10 \\ 10 & 10 \end{pmatrix}, \quad Q = \begin{pmatrix} 0 & 10 \\ 10 & 0 \end{pmatrix}, \tag{2}$$

$$A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}, \quad F = \begin{pmatrix} 10 & 0 \\ 0 & 10 \end{pmatrix}, \quad Q = \begin{pmatrix} 0 & 10 \\ 10 & 10 \end{pmatrix}. \tag{3}$$

Table 1 presents the results computed in double precision arithmetic on an IBM-PC computer (machine accuracy about $2.22 \cdot 10^{-16}$) using `eig` and SLICOT routines `MB04ZD` and `MB03SD`, based on a structure-preserving algorithm [25].

While the results produced by `eig` are relatively close to the exact eigenvalues for each example (1)–(3), the complex conjugate pair is qualitatively wrong, since both real parts have the same sign. Of course, in such very simple cases, the real parts could be set to zero, but for larger problems there could be no clue on how the computed eigenvalues should be paired to have meaningful results. On the other hand, SLICOT results are always qualitatively good, since they follow the needed pattern. This is important, e.g., for Riccati equation solvers (see, e.g., [19]).

Another example refers to Grammian matrices and Hankel singular values. These values are input-output invariants of stable linear time-invariant (LTI) systems and play a fundamental role in finding balanced realizations and in model reduction. They are defined as the square roots of the eigenvalues of a product P_cP_o, where P_c and P_o are the non-negative definite *controllability* and *observability Grammians*, respectively. For instance, for a stable state-space realization (A, B, C), $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, $C \in \mathbb{R}^{p\times n}$, of a discrete-time LTI system, the Grammians are given by the solutions of the discrete-time Lyapunov equations, also called *Stein equations*

$$AP_cA^T - P_c = -BB^T, \qquad A^TP_oA - P_o = -C^TC. \tag{4}$$

As the eigenvalues of P_cP_o are all real non-negative, so are the Hankel singular values. By solving the equations (4) without exploiting the symmetry and semi-definiteness of the solutions, round-off errors can cause the computed Grammians to become nonsymmetric and/or nondefinite. This can then result in negative or even complex Hankel singular values—a complete nonsense regarding the system-theoretic properties.

Example 2. For the following discrete-time system matrices

$$A = \begin{pmatrix} 0.01954 & -0.00773 & -0.0083 \\ -0.02273 & -0.01274 & -0.0161 \\ -0.03394 & 0.01878 & 0.05661 \end{pmatrix}, \quad B = \begin{pmatrix} -0.23571 \\ 0.81074 \\ 2.41953 \end{pmatrix},$$
$$C = \begin{pmatrix} 0.29033 & 0.33462 & -2.40456 \end{pmatrix},$$

`eig(Pc*Po)` gives (with 5 significant digits) 3.1865e+1, -1.2801e-15, 4.2270e-5. Hence, the computed Hankel "singular values" are the positive square roots of these eigenvalues, i.e., 5.6449e+0, 3.5778e-8i, 6.5015e-3. On the other hand, the SLICOT function `AB13AD` returned the following values 5.6449e+0, 6.5015e-3, 2.7081e-8, by exploiting the fact that the Hankel singular values can be equivalently computed as the singular values of the product R_cR_o, where the upper triangular matrices R_c and R_o are the Cholesky factors of the Grammians, satisfying $P_c = R_cR_c^T$ and $P_o = R_o^TR_o$. Here, R_c and R_o are obtained by solving the equations (4) directly for these factors using the algorithm in [10].

An essential feature of reliable algorithms is to allow an assessment of the accuracy of computed results. The SLICOT codes provide some upper bounds, in terms of condition numbers, on the errors of the computed solutions for basic computational problems in control theory, including Lyapunov and algebraic Riccati equations (AREs).

3. Functional capabilities

The SLICOT Library currently covers the following chapters: Analysis Routines, Benchmark and Test Problems, Data Analysis, Filtering, Identification, Mathematical Routines, Synthesis Routines, Transformation Routines, and Utility Routines. A chapter-by-chapter SLICOT Library contents (with sections and subsections) is given in [24]. A summary follows.

System analysis routines perform various tasks, such as: finding reduced forms revealing structural properties; invariant zeros of a system; various system norms; model reduction [26]. Analysis of generalized state-space systems is also covered.

Benchmark and test problems routines generate benchmark examples for LTI dynamical systems, as well as for (generalized) Lyapunov and algebraic Riccati equations.

Data analysis routines include convolution or deconvolution of two signals and various transforms of real or complex signals.

Filtering routines cover several time-varying or time-invariant filters, including the conventional Kalman filter and fast recursive least-squares filter.

Identification routines deal with both LTI state-space systems and Wiener systems, and include fast solvers (see, e.g., [21, 20]).

Mathematical routines implement special algorithms for computations encountered in many applications (not only in CACSD), which are not available in LAPACK or similar packages. Examples are (structured) matrix factorizations, including those for (block) Toeplitz matrices, evaluation of matrix exponentials, solving (structured) nonlinear least-squares problems, and calculations related to (matrix) polynomials.

System synthesis routines cover the computational problems in control systems design: pole assignment; solution of continuous-time or discrete-time algebraic Riccati equations, using the Schur vector method [12], the method of deflating subspaces [23], or–for continuous-time equations–the matrix sign function method [5]; solution of continuous-time or discrete-time Lyapunov equations [3, 2]; factored solution of stable non-negative definite continuous-time or discrete-time Lyapunov equations [10]; solution of generalized Lyapunov equations [18], using extensions of Bartels/Stewart and Hammarling's methods; solution of the continuous-time or discrete-time Sylvester equations $AX + XB = C$, $AXB + X = C$, using the Hessenberg-Schur method [8]; solution (for R and L) of the generalized Sylvester equation [11] $AR - LB = C$, $DR - LE = F$; minimum norm feedback matrix for "deadbeat control"; spectral factorization of transfer-function matrices; H_∞ (sub)optimal, and H_2 optimal state controller, as well as positive feedback controller for both continuous-time and discrete-time systems.

Transformation routines mainly cover conversions between various system representations: conversions between state-space representations, such as balancing a system matrix corresponding to a triplet (A, B, C), calculation of the controller or observer Hessenberg form, or minimal block Hessenberg realization for a state-space representation; transformations from a given representation to another representation, e.g., state-space to polynomial, or rational matrix representation (or conversely), polynomial representation to frequency response, etc. Generalized state-space transformations for descriptor systems are also included.

The functional flexibility of the basic computational tools is illustrated with an example: the solver for stable non-negative definite Lyapunov equations based on [10]. Using $\mathrm{op}(M)$ to denote either the matrix M or its transpose, M^T, this solver computes the factored solution, $X = \mathrm{op}(U)^T \mathrm{op}(U)$, with U upper triangular, of either stable continuous-time Lyapunov equation

$$\mathrm{op}(A)^T X + X\,\mathrm{op}(A) = -\sigma^2 \mathrm{op}(B)^T \mathrm{op}(B), \tag{5}$$

or convergent discrete-time Lyapunov equation

$$\mathrm{op}(A)^T X\,\mathrm{op}(A) - X = -\sigma^2 \mathrm{op}(B)^T \mathrm{op}(B), \tag{6}$$

where the scalar σ is a scaling factor, set less than or (usally) equal to one, in order to prevent solution overflowing. Note that both forms of $\mathrm{op}(A)$ in (5) or (6) could be needed in an application, for instance, in Example 2, for solving the equations (4). The solver ability to deal with the $\mathrm{op}(\cdot)$ operator is advantageous in this context since only one reduction to the real Schur form is needed to solve both equations.

Similar capabilities are implemented in other solvers. The codes for solving algebraic Riccati equations have options for various scaling strategies, and for sorting the eigenvalues (also anti-stabilizing Riccati solutions can be obtained); linear quadratic optimization problems with coupling terms can be optionally solved. Special cases of matrices in real Schur form, and/or Hessenberg form can efficiently be dealt with by some solvers. Estimates of the reciprocal condition numbers for Riccati and Lyapunov equations can be computed.

The current version of the SLICOT Library includes

- 379 documented routines, user-callable or programmer-callable, plus 36 partially documented routines for low level computations;
- 193 example programs, with associated files with data and results;
- 41 MATLAB/Scilab MEX-files;
- 196 MATLAB/Scilab M-files.

The compressed files slicot.tar.gz (for Unix platforms), and slicotPC.zip (for Windows platforms) each contain over 2250 files.

4. MATLAB-Gateways

The essential functionality and performance of SLICOT routines are made accessible from the high-level software environments MATLAB and Scilab by a large collection of MEX- and M-function gateways developed during the SLICOT implementation process. These MEX- and M-functions are used as interfaces between the powerful Fortran routines and these popular user-friendly design environments. There are multiple, conflicting requirements the CACSD software should satisfy, such as, functionality – simplicity, flexibility – easy-of-use, complexity – performance, etc. The basic strategy used to ensure the needed trade-off between these objectives was to develop a reduced number of MEX-function interfaces, each covering an extended functionality, and several M-function interfaces, each solving a specific control systems analysis or design problem. Each MEX-function calls several related SLICOT routines, while each M-function calls the appropriate MEX-function to perform its specific task. For instance, the MEX-function `aresol` calls SLICOT routines `SB02MD`, `SB02MT`, `SB02ND`, and `SB02OD`, for solving either continuous-time or discrete-time algebraic Riccati equations (CARE/DARE) using standard or generalized Schur vector methods applied on suitably built Hamiltonian or symplectic matrices or matrix pencils, while the M-functions `slcaregs`, `slcares`, `sldaregs`, `sldares`, and `sldaregsv` call `aresol` to solve a specific CARE or DARE equation using a certain method. Allocatable Fortran 90 arrays are employed in the MEX-files, to reduce the storage requirements. The MEX interfaces are necessarily quite complex, to cope with the extended functionality and flexibility, and so, they are primarily intended for expert use and further software developments. On the other hand, the M interfaces are much simpler, hiding any computational details, and therefore they are destined to all categories of users. A minimal input data has to be specified, and default option values are automatically set. For convenience, demonstration packages for PC-Windows platforms are made available on the SLICOT ftp or Web sites.

A partial list of the currently available SLICOT M-functions is given below. The functions related to benchmarks (5 functions), data analysis (7 functions), structured matrix factorizations (7 functions), model and controller reduction (20 functions), test functions, etc., are omitted due to lake of space.

Canonical forms and system transformations

slconf	Controllability staircase form of a system (A, B, C).
slobsf	Observability staircase form of a system (A, B, C).
slminr	Minimal realization of a system (A, B, C).
slsbal	Balance the system matrix for a state-space system (A, B, C).
slsdec	Additive spectral decomposition of a system (A, B, C) with respect to a given stability domain.
slsorsf	Transform the state matrix of a state space system to a specified eigenvalue-ordered real Schur form.
slsrsf	Transform the state matrix A to a real Schur form.

Riccati equations

slcaregs	Solve CARE with generalized Schur method on an extended pencil.
slcares	Solve CARE with Schur method.
slcaresc	Solve CARE with refined Schur method and estimate condition.
sldaregs	Solve DARE with generalized Schur method on an extended pencil.
sldares	Solve DARE with Schur method.
sldaresc	Solve DARE with refined Schur method and estimate condition.
sldaregsv	Solve DARE with generalized Schur method on a symplectic pencil.
slgcare	Solve a descriptor CARE with generalized Schur method.
slgdare	Solve a descriptor DARE with generalized Schur method.
carecond	Estimate reciprocal condition number of a CARE.
darecond	Estimate reciprocal condition number of a DARE.

Sylvester and Lyapunov-like equations

slsylv	Solve continuous-time Sylvester equations.
sldsyl	Solve discrete-time Sylvester equations.
sllyap	Solve continuous-time Lyapunov equations.
slstei	Solve Stein equations.
slstly	Solve stable continuous-time Lyapunov equations.
slstst	Solve stable Stein equations.
slgesg	Solve generalized linear matrix equation pairs.
slgely	Solve generalized continuous-time Lyapunov equations.
slgest	Solve generalized Stein equations.
slgsly	Solve stable generalized continuous-time Lyapunov equations.
slgsst	Solve stable generalized Stein equations.
lyapcond	Estimate reciprocal condition number of a Lyapunov equation.
steicond	Estimate reciprocal condition number of a Stein equation.

Discrete-time LTI and Wiener multivariable state-space systems identification

slmoen4	Find the system matrices and the Kalman gain of a discrete-time system, using combined MOESP and N4SID subspace identification techniques.

slmoesm	Idem, using combined MOESP and simulation techniques.
slmoesp	Idem, using the MOESP technique.
sln4sid	Idem, using the N4SID technique.
findR	Preprocess the input-output data using Cholesky, QR, or fast QR factorization [14], and estimate the system order.
findAC	Find the system matrices A and C of a system, given the system order n and the relevant part of the R factor of the QR factorization.
findBDK	Find the system matrices B and D and the Kalman gain of a system, given n, the matrices A and C, and the relevant part of the R factor.
findABCD	Find the system matrices and the Kalman gain, given n and the relevant part of the R factor.
findx0BD	Estimate the initial state and/or the matrices B and D of a system, given the system matrices A, C, and a set of input/output data.
inistate	Estimate the initial state of a system, given the (estimated) system matrices, and a set of input/output data.
NNout	Compute the output of a set of neural networks used to model the nonlinear part of a Wiener system.
dsim	Compute the output response of a linear discrete-time system (much faster than the MATLAB function lsim).
o2s	Transform a linear discrete-time system given in the output normal form to a state-space representation.
s2o	Transform a state-space representation of a linear discrete-time system into the output normal form.

Miscellaneous functions

slinorm	Compute the L-infinity system norm.
slH2norm	Compute the H2/L2 norm of a system.
slHknorm	Compute the Hankel-norm of a stable projection of a system.
slstabr	Compute the complex stability radius.
Hameig	Compute the eigenvalues of a Hamiltonian matrix.
pass	Perform (partial) pole assignment.
persch	Compute the periodic Hessenberg or periodic Schur decomposition of a matrix product.

5. Performance results

This section presents typical performance results for the SLICOT Sylvester solver, in comparison with equivalent computations performed by the MATLAB function `sylv`. The calculations have been done on an IBM PC computer at 500 MHz, with 128 Mb memory, using Compaq Visual Fortran 6.5, non-optimized BLAS, and MATLAB 6.5 (R13). Other performance results for specific analysis and synthesis tasks appeared, e.g., in [4, 22, 24]. The results show that, at comparable accuracy, SLICOT gateways are several times faster than MATLAB computations.

Figure 1 shows the execution times in seconds for SLICOT function `slsylv` and MATLAB function `sylv`, for solving ten randomly generated Sylvester equations, $AX + XB = C$, with known solutions, $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{m\times m}$, with $n = 30 : 30 : 300$, and $m = n$. In the right hand side, the equations have the matrices A and B in real Schur form. Clearly, `slsylv` is impressibly much faster than `sylv` in this case, since it could exploit the problem structure.

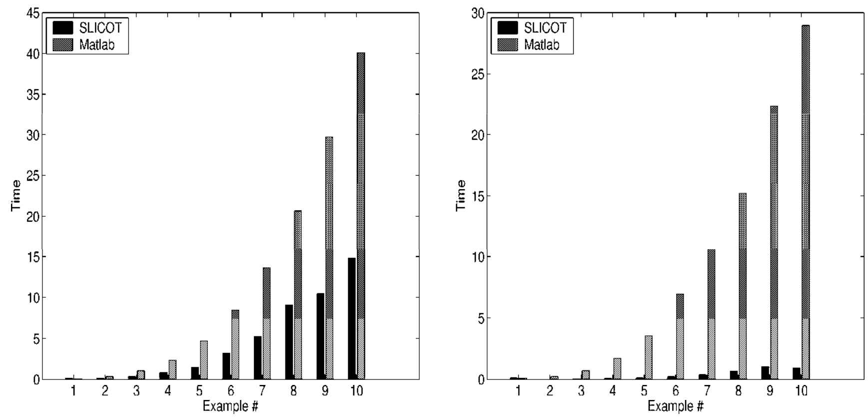


Figure 1. Timing comparison: SLICOT `slsylv` versus MATLAB `sylv` for random Sylvester equations with $n = 30 : 30 : 300$, and $m = n$. Left: general matrices; Right: A and B are in real Schur form.

Figure 2 shows the relative residuals of the solutions of the same Sylvester equations. These residuals are almost always better for `slsylv` than for `sylv`.

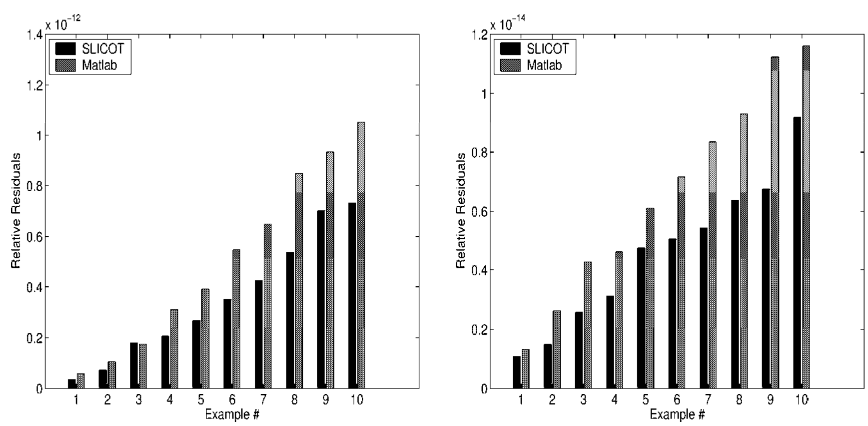


Figure 2. Relative residuals of solutions of the same Sylvester equations as in Fig. 1.

Figure 3 shows the relative error of the solutions of the same Sylvester equations.

Other SLICOT calculations reveal similar performances, or even larger speed-up factors, when fast algorithms are used.

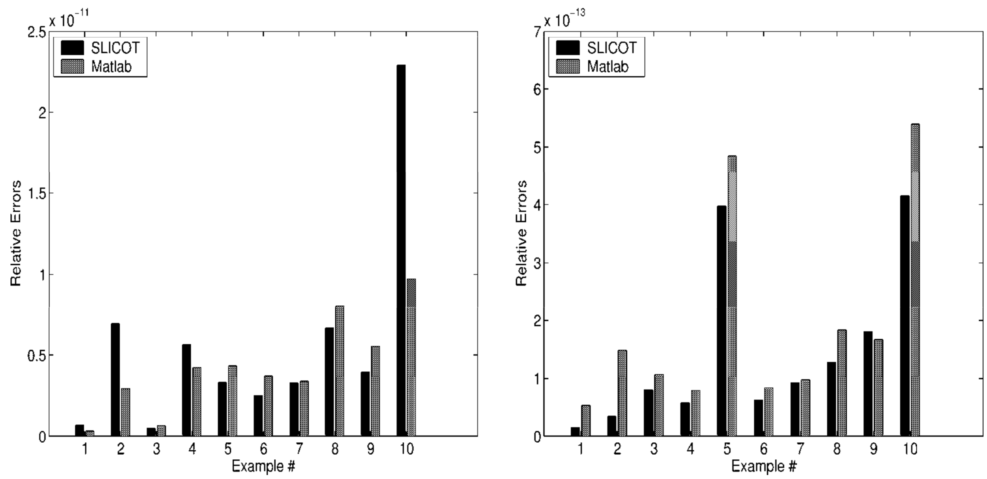


Figure 3. Relative error of solutions of the same Sylvester equations as in Fig. 1.

6. Conclusions

Algorithmic improvements in systems and control computations have been incorporated into the freely-available version of the Subroutine Library in Control Theory, SLICOT. This library enables to exploit the potential of modern high-performance computer architectures. Performance comparisons of some SLICOT components and equivalent MATLAB functions show that SLICOT computations are several times faster than MATLAB computations, at comparable accuracy.

Future work could include the development of a parallel version of the SLICOT Library for massively parallel architectures, based on the ScaLAPACK version of the LAPACK package. This is partly done for model reduction routines.

Acknowledgements

The associated work was partially supported by the European Community BRITE-EURAM III *Thematic Networks Programme NICONET* (project BRRT –CT97-5040).

Notes

1. The author is currently the "SLICOT librarian".

References

[1] Anderson, E., Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney and D. Sorensen (1999). *LAPACK Users' Guide. Third Edition.* SIAM, Philadelphia.

[2] Barraud, A.Y. (1977). A numerical algorithm to solve $A^T XA - X = Q$. *IEEE Trans. Automat. Control*, **AC-22**(5), 883–885.

[3] Bartels, R.H. and G.W. Stewart (1972). Algorithm 432: Solution of the matrix equation $AX + XB = C$. *Comm. ACM*, **15**(9), 820–826.

[4] Benner, P., V. Mehrmann, V. Sima, S. Van Huffel and A. Varga (1999). SLICOT — A subroutine library in systems and control theory. In: *Applied and Computational Control, Signals, and Circuits* (B.N. Datta, Ed.), **1**, 499–539. Birkhäuser, Boston.

[5] Byers, R. (1987). Solving the algebraic Riccati equation with the matrix sign function. *Lin. Alg. Appl.*, **85**(1), 267–279.

[6] Dongarra, J.J., J. Du Croz, I.S. Duff and S. Hammarling (1990). Algorithm 679: A set of Level 3 Basic Linear Algebra Subprograms. *ACM Trans. Math. Softw.*, **16**, 1–17, 18–28.

[7] Dongarra, J.J., J. Du Croz, S. Hammarling and R.J. Hanson (1988). Algorithm 656: An extended set of Fortran Basic Linear Algebra Subprograms. *ACM Trans. Math. Softw.*, **14**, 1–17, 18–32.

[8] Golub, G.H., S. Nash and C.F. Van Loan (1979). A Hessenberg-Schur method for the problem $AX + XB = C$. *IEEE Trans. Automat. Control*, **AC–24**(6), 909–913.

[9] Gomez, C., Ed. (1999). *Engineering and Scientific Computing with Scilab*. Birkhäuser, Boston.

[10] Hammarling, S.J. (1982). Numerical solution of the stable, non-negative definite Lyapunov equation. *IMA J. Numer. Anal.*, **2**, 303–323.

[11] Kågström, B. and P. Poromaa (1996). LAPACK-style algorithms and software for solving the generalized Sylvester equation and estimating the separation between regular matrix pairs. *ACM Trans. Math. Softw.*, **22**(1), 78–103.

[12] Laub, A.J. (1979). A Schur method for solving algebraic Riccati equations. *IEEE Trans. Automat. Control*, **AC–24**(6), 913–921.

[13] Lawson, C.L., R.J. Hanson, D.R. Kincaid and F.T. Krogh (1979). Basic Linear Algebra Subprograms for Fortran usage. *ACM Trans. Math. Softw.*, **5**(3), 308–323.

[14] Mastronardi, N., D. Kressner, V. Sima, P. Van Dooren and S. Van Huffel (2001). A fast algorithm for subspace state-space system identification via exploitation of the displacement structure. *J. Comput. Appl. Math.*, **132**(1), 71–81.

[15] The MathWorks, Inc. (1999). *Using MATLAB. Version 5*. 24 Prime Park Way, Natick.

[16] NAG—The Numerical Algorithms Group Ltd. (1991). *NAG SLICOT Library Manual, Release 1*. Wilkinson House, Jordan Hill Road, Oxford OX2 8DR, U.K.

[17] NAG—The Numerical Algorithms Group Ltd. (1993). *NAG SLICOT Library Manual, Release 2*. Wilkinson House, Jordan Hill Road, Oxford OX2 8DR, U.K.

[18] Penzl, T. (1998). Numerical solution of generalized Lyapunov equations. *Advances in Comp. Math.*, **8**, 33–48.

[19] Sima, V. (1996). *Algorithms for Linear-Quadratic Optimization*. Vol. **200** of *Pure and Applied Mathematics: A Series of Monographs and Textbooks*. Marcel Dekker, Inc., New York.

[20] Sima, V. (2003). Fast numerical algorithms for Wiener systems identification. In: *Proceedings of Analysis and Optimization of Differential Systems, September 10–14, 2002, "Ovidius" University, Constanţa, România* (D. Tiba V. Barbu, I. Lasiecka and C. Varsan, Eds.). Kluwer Academic Publishers, Boston/Dordrecht/London.

[21] Sima, V., D.M. Sima and S. Van Huffel (2002). SLICOT system identification software and applications. In: *Proceedings of the 2002 IEEE International Conference on Control Applications and IEEE International Symposium on Computer Aided Control System Design, CCA/CACSD 2002, September 18–20, 2002, Glasgow, Scotland, U.K.* (P.R. Kalata, Ed.), 45–50. Omnipress.

[22] Sima, V. and S. Van Huffel (1999). High-performance algorithms and software for systems and control computations. In: *Proceedings of the 1999 IEEE International Conference on Control Applications and IEEE International Symposium on Computer Aided Control System Design, August 22–27, 1999, Hapuna-Beach Prince Hotel, Kohala Coast-Island of Hawai'i, Hawai'i, USA* (O. Gonzales, Ed.), 85–90. Omnipress, Old Dominion University, Norfolk.

[23] Van Dooren, P. (1981). A generalized eigenvalue approach for solving Riccati equations. *SIAM J. Sci. Stat. Comput.*, **2**(2), 121–135.

[24] Van Huffel, S. and V. Sima (2002). SLICOT and control systems numerical software packages. In: *Proceedings of the 2002 IEEE International Conference on Control Applications and IEEE International Symposium on Computer Aided Control System Design, CCA/CACSD 2002, September 18–20, 2002, Glasgow, Scotland, U.K.* (P.R. Kalata, Ed.), 39–44. Omnipress.

[25] Van Loan, C.F. (1984). A symplectic method for approximating all the eigenvalues of a Hamiltonian matrix. *Lin. Alg. Appl.*, **61**, 233–251.

[26] Varga, A. (2001). Model reduction software in the SLICOT library. In: *Applied and Computational Control, Signals, and Circuits* (B.N. Datta, Ed.), **2**, 239–282. Kluwer Academic Publishers, Boston.

ON THE CONNECTION BETWEEN RICCATI INEQUALITIES AND EQUATIONS IN H^∞ CONTROL PROBLEMS

Adrian Stoica

University "Politehnica" of Bucharest, Faculty of Aerospace Engineering
Str. Splaiul Independentei No. 313, RO-77226 Bucharest, Romania
e-mail: amstoica@fx.ro

Abstract The paper presents some connections between the solvability conditions expressed in terms of linear matrix inequalities and the ones using Riccati equations. It is shown that the methodology based on the Bounded Real Lemma, mainly used in the singular H^∞ control theory, can be successfully employed in nonsingular problems, providing solvability conditions in terms of the stabilizing solutions to algebraic Riccati equations.

Keywords: H^∞ control, linear matrix inequalities, Riccati equations, Bounded Real Lemma, stabilizing solutions.

1. Introduction

The H^∞ theory captured a major interest in control engineering and in applied mathematics over the last two and half decades. This interest is determined by the fact that the H^∞ optimization provides elegant solutions for many practical problems. Robust control, tracking, filtering, fault detection and identification are only some of the most known areas of applications. From an historical perspective, there were several directions in which the study of this problem has been orientated. Among them one can mention the methods based on operator theory and interpolation, mostly used in the early period (Adamjan *et al.*, 1978), (Ball and Helton, 1983), reduction to Nehari problems (Francis, 1986), polynomial approaches (Kwakernaak, 1986). A major contribution is brought by the state-space results derived in (Glover and Doyle, 1988), (Doyle *et al.*, 1989) providing necessary and sufficient solvability conditions for the so-called *nonsingular* H^∞ *control problem*. These conditions are expressed in terms of the stabilizing solutions of some game-theoretic *algebraic Riccati equations* (ARE). Explicit formulae and parameterization of the set of all solutions have been also obtained. State-space solutions for the nonsingular H^∞ control problem are also derived by a different technique using the generalized Popov-Yakubovich theory in (Ionescu *et al.*, 1999), (Ionescu and Stoica, 1999). The *singular* H^∞ *control problem* arises from the nonsingular case by removing some assumptions. The approaches develop in (Gahinet and Apkarian, 1994), (Boyd *et al.*, 1994), (Iwasaki and Skelton, 1994) use a different methodology, based on the *Bounded Real Lemma* in inequality form. In this case the solvability conditions are expressed in terms of the feasibility of some specific *linear matrix inequalities* (LMI). In contrast with the nonsingular H^∞ control problem where the solvability conditions and the solution are derived starting from some particular problems fully exploiting the nonsingularity assumptions, in the singular case the developments are easily performed directly for the general case. Although efficient numerical algorithms to determine the solutions of the corresponding LMI have been developed, there are applications when ill conditioned numerical computations and instabilities occur (see *e.g.* (Gahinet and Apkarian, 1994)). In such cases it is preferable to replace, if possible, the conditions based on Riccati inequalities by Riccati equations for which several alternative algorithms to compute their stabilizing solutions are available.

The aim of this paper is to present some connections between the solvability conditions expressed in terms of LMI and the ones using ARE. It is shown that the general methodology based on the Bounded Real Lemma, mainly used in the singular H^∞ control theory, can be successfully employed in nonsingular problems, providing solvability conditions in terms of the stabilizing solutions to ARE.

2. Problem formulation and preliminary results

Consider the *generalized two-input, two-output system* T with the state space equations:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + B_1u_1(t) + B_2u_2(t) \\ y_1(t) &= C_1x(t) + D_{11}u_1(t) + D_{12}u_2(t) \\ y_2(t) &= C_2x(t) + D_{21}u_1(t), \end{aligned} \tag{1}$$

where $x(t) \in \mathbf{R}^n$ is the state vector, $u_1(t) \in \mathbf{R}^{m_1}$ is the external input vector,

$u_2(t) \in \mathbf{R}^{m_2}$ denotes the control variable, $y_1(t) \in \mathbf{R}^{p_1}$ includes the regulated output variables and $y_2(t) \in \mathbf{R}^{p_2}$ is the measured outputs vector. Then, the H^∞ control problem consists in finding a controller K such that the resulting system obtained by (1) taking $u_2 = Ky_2$ is internally stable and $\left\|T_{u_1 y_1}\right\|_\infty < \gamma$, where $\gamma > 0$ is a given *level of attenuation*. $T_{y_1 u_1}$ denotes the transfer function from u_1 to y_1 of the resulting system and the H^∞ norm of a stable system with the transfer function $G(s)$ is defined as $\left\|G(s)\right\|_\infty = \sup_{-\infty<\omega<\infty} \lambda_{\max}{}^{1/2}\left(G^T(-j\omega)G(j\omega)\right)$, $\lambda_{\max}(\cdot)$ denoting the maximal eigenvalue of $(\cdot)$. In the following, two preliminary results are stated. The first one is also known as the *Bounded Real Lemma* and an early proof of it is given in (Anderson and Vongpanitlerd, 1973).

Lemma 1 (Bounded Real Lemma) *Let* $H(s) := (A,B,C,D)$ *be a stable system with* $A \in \mathbf{R}^{n\times n}, B \in \mathbf{R}^{n\times m}, C \in \mathbf{R}^{p\times n}$ *and* $D \in \mathbf{R}^{p\times m}$. *Then the following assertions are equivalent:*

i) $\left\|H(s)\right\|_\infty < \gamma$;

ii) There exists $X > 0$ *such that:*

$$\begin{bmatrix} A^{\mathrm{T}}X + XA & XB & C^{\mathrm{T}} \\ B^{\mathrm{T}}X & -\gamma I_m & D^{\mathrm{T}} \\ C & D & -\gamma I_m \end{bmatrix} < 0;$$

iii) $\gamma^2 I_m - D^{\mathrm{T}}D > 0$ *and the Riccati equation*

$$A^{\mathrm{T}}X + XA + \left(XB + C^{\mathrm{T}}D\right)\left(\gamma^2 I_m - D^{\mathrm{T}}D\right)^{-1}\left(B^{\mathrm{T}}X + D^{\mathrm{T}}C\right) + C^{\mathrm{T}}C = 0$$

has a stabilizing solution.

Based on the above result, the following theorem providing necessary and sufficient solvability conditions to the H^∞ control problem is proved in (Gahinet and Apkarian, 1994):

Theorem 1 *Consider the generalized system (1) and a scalar* $\gamma > 0$. *Then the following assertions are equivalent:*

i) It exists an n_k*-order controller such that the resulting system* $T_{u_1 y_1}$ *is stable and* $\left\|T_{u_1 y_1}\right\|_\infty < \gamma$;

ii) There exist the symmetric matrices $R, S \in \mathbf{R}^{n\times n}, R > 0, S > 0$, *such that:*

$$\begin{bmatrix} \mathcal{N}_{11} & 0 \\ \mathcal{N}_{12} & 0 \\ 0 & I_{m_1} \end{bmatrix}^T \begin{bmatrix} AR + RA^T & RC_1^T & B_1 \\ C_1 R & -\gamma I_{p_1} & D_{11} \\ B_1^T & D_{11}^T & -\gamma I_{m_1} \end{bmatrix} \begin{bmatrix} \mathcal{N}_{11} & 0 \\ \mathcal{N}_{12} & 0 \\ 0 & I_{m_1} \end{bmatrix} < 0 \tag{2}$$

$$\begin{bmatrix} \mathcal{N}_{21} & 0 \\ \mathcal{N}_{22} & 0 \\ 0 & I_{p_1} \end{bmatrix}^T \begin{bmatrix} A^T S + SA & SB_1 & C_1^T \\ B_1^T S & -\gamma I_{m_1} & D_{11}^T \\ C_1 & D_{11} & -\gamma I_{p_1} \end{bmatrix} \begin{bmatrix} \mathcal{N}_{21} & 0 \\ \mathcal{N}_{22} & 0 \\ 0 & I_{p_1} \end{bmatrix} < 0 \tag{3}$$

$$\begin{bmatrix} R & I_n \\ I_n & S \end{bmatrix} \geq 0, \quad \text{rank}\begin{bmatrix} R & I_n \\ I_n & S \end{bmatrix} \leq n + n_k, \tag{4}$$

where $\begin{bmatrix} \mathcal{N}_{11} \\ \mathcal{N}_{12} \end{bmatrix}$ *and* $\begin{bmatrix} \mathcal{N}_{21} \\ \mathcal{N}_{22} \end{bmatrix}$ *denotes bases of the null spaces of the matrices* $\begin{bmatrix} B_2^T & D_{12}^T \end{bmatrix}$ *and* $\begin{bmatrix} C_2 & D_{21} \end{bmatrix}$*, respectively.*

A crucial role for the further developments is played by the next proposition which proof can be found for example in (Gahinet, 1992):

Proposition 1 *If the pair* (C, A) *where* $A \in \mathbf{R}^{n\times n}, C \in \mathbf{R}^{p\times n}$ *has no imaginary unobservable modes then the following assertions are equivalent:*
i) The Riccati type inequality

$$AR + RA^T + RC^T CR + Q < 0 \tag{5}$$

with $Q = B_1 B_1^T - B_2 B_2^T \in \mathbf{R}^{n\times n}$ *symmetric has a symmetric matrix solution* $\hat{R} > 0$*;*
ii) The Riccati equation

$$A^T X + XA + XQX + C^T C = 0 \tag{6}$$

has a stabilizing solution $\tilde{X}$.
If the conditions i) or ii) are accomplished then $0 \leq \tilde{X} < \hat{R}^{-1}$.

3. The nonsingular H^∞ control problem

In this section it is shown that the solvability conditions in the nonsingular H^∞ case expressed in terms of Riccati equations can be directly recovered via Proposition 1 by the general result stated in Theorem 1.

Theorem 2 *Assume that the following conditions hold:*
A1) $\text{rank}\, D_{12}^T D_{12} = m_2$ *and* $\text{rank}\, D_{21} D_{21}^T = p_2$*;*

A2) The system $\left(A,B_2,C_1,D_{12}\right)$ *has no invariant zero on the imaginary axis, that is:*

$$\operatorname{rank}\begin{bmatrix} A-j\omega I_n & B_2 \\ C_1 & D_{12} \end{bmatrix} = n+m_2,\ \forall\omega\in\mathbf{R}\ .$$

A3) The system $\left(A,B_1,C_2,D_{21}\right)$ *has no invariant zero on the imaginary axis, that is:*

$$\operatorname{rank}\begin{bmatrix} A-j\omega I_n & B_1 \\ C_2 & D_{21} \end{bmatrix} = n+p_2,\ \forall\omega\in\mathbf{R}.$$

Then the H^∞ *control problem has an* n*-order solution if and only if:*

$$\gamma^2 I_{p_1} > \hat{D}_{11}\hat{D}_{11}{}^T \quad \textit{and} \quad \gamma^2 I_{m_1} > \tilde{D}_{11}{}^T\tilde{D}_{11} \tag{7}$$

and the game-theoretic Riccati equations:

$$\begin{aligned}&\left[\hat{A}+\hat{B}_1\hat{D}_{11}{}^T\left(\gamma^2 I_{p_1}-\hat{D}_{11}\hat{D}_{11}{}^T\right)^{-1}\hat{C}_1\right]^T X+\left[\hat{A}+\hat{B}_1\hat{D}_{11}{}^T\left(\gamma^2 I_{p_1}-\hat{D}_{11}\hat{D}_{11}{}^T\right)^{-1}\hat{C}_1\right]X\\&+X\left[\hat{B}_1\left(\gamma^2 I_{p_1}-\hat{D}_{11}\hat{D}_{11}{}^T\right)^{-1}\hat{B}_1^T-\hat{B}_2\hat{B}_2{}^T\right]X+\gamma^2\hat{C}_1^T\left(\gamma^2 I_{p_1}-\hat{D}_{11}\hat{D}_{11}{}^T\right)^{-1}\hat{C}_1=0\end{aligned} \tag{8}$$

$$\begin{aligned}&\left[\tilde{A}+\tilde{B}_1\tilde{D}_{11}{}^T\left(\gamma^2 I_{m_1}-\tilde{D}_{11}\tilde{D}_{11}{}^T\right)^{-1}\tilde{C}_1\right]Y+\left[\tilde{A}+\tilde{B}_1\tilde{D}_{11}{}^T\left(\gamma^2 I_{m_1}-\tilde{D}_{11}\tilde{D}_{11}{}^T\right)^{-1}\tilde{C}_1\right]Y\\&+Y\left[\tilde{C}_1^T\left(\gamma^2 I_{m_1}-\tilde{D}_{11}\tilde{D}_{11}{}^T\right)^{-1}\tilde{C}_1-\tilde{C}_2^T\tilde{C}_2\right]Y+\gamma^2\tilde{B}_1\left(\gamma^2 I_{m_1}-\tilde{D}_{11}{}^T\tilde{D}_{11}\right)^{-1}\tilde{B}_1^T=0\end{aligned} \tag{9}$$

have the stabilizing solutions $X\geq 0$ *and* $Y\geq 0$ *respectively, satisfying the condition:*

$$\rho\left(XY\right)<\gamma^2, \tag{10}$$

where $\rho\left(\cdot\right)$ *denotes the spectral radius of* $\left(\cdot\right)$ *and:*

$$\begin{aligned}&\hat{A}:=A-B_2D_{12}{}^+C_1,\ \hat{B}_1:=B_1-B_2D_{12}{}^+D_{11},\ \hat{B}_2:=B_2D_{12}{}^+\\&\hat{C}_1:=\left(I_{p_1}-D_{12}D_{12}{}^+\right)C_1,\ \hat{D}_{11}:=\left(I_{p_1}-D_{12}D_{12}{}^+\right)D_{11},\\&\tilde{A}:=A-B_1D_{21}{}^+C_2,\ \tilde{B}_1:=B_1\left(I_{m_1}-D_{21}{}^+D_{21}\right),\\&\tilde{C}_1:=C_1-D_{11}D_{21}{}^+C_2,\ \tilde{C}_2:=D_{21}{}^+C_2,\ \tilde{D}_{11}:=D_{11}\left(I_{m_1}-D_{21}{}^+D_{21}\right).\end{aligned} \tag{11}$$

Proof. The proof is based on Proposition 1 and on Theorem 1. Thus, Theorem 1 shows that the general H^{∞} problem has a solution if and only if conditions (2)-(4) hold. In the particular case of the nonsingular H^{∞} problem studied in this section, a basis of the null space of the matrix $\begin{bmatrix} B_2{}^T & D_{12}{}^T \end{bmatrix}$ is given by:

$$\begin{bmatrix} \mathcal{N}_{11} \\ \mathcal{N}_{12} \end{bmatrix} = \begin{bmatrix} I_n & 0 \\ -D_{12}{}^{+T} B_2{}^T & U_{12} \end{bmatrix}, \tag{12}$$

where U_{12} denotes a basis of the null space of $D_{12}{}^T$, that is $D_{12}{}^T U_{12} = 0$. Then the inequality (2) yields:

$$\begin{bmatrix} \hat{A}R + R\hat{A}^T - \gamma \hat{B}_2 \hat{B}_2{}^T & \left(RC_1{}^T + \gamma \hat{B}_2\right) U_{12} & \hat{B}_1 \\ U_{12}{}^T \left(C_1 R + \gamma B_2{}^T\right) & -\gamma U_{12}{}^T U_{12} & U_{12}{}^T D_{11} \\ \hat{B}_1{}^T & D_{11}{}^T U_{12} & -\gamma I_{m_1} \end{bmatrix} < 0 \tag{13}$$

with $R > 0$. Further two properties of the matrix U_{12} are further presented. According with the Singular Value Decomposition Theorem, it exists a unitary matrix $U \in \mathbf{R}^{p_1 \times p_1}$ such that

$$D_{12} = U \begin{bmatrix} \Sigma \\ 0_{(p_1 - m_2) \times m_2} \end{bmatrix}, \tag{14}$$

where, by virtue of assumption A1, Σ is nonsingular. Thus it results that:

$$D_{12}{}^{+} = \begin{bmatrix} \Sigma^{-1} & 0_{m_2 \times (p_1 - m_2)} \end{bmatrix} U^T \tag{15}$$

and

$$U_{12} = U \begin{bmatrix} 0 \\ I_{p_1 - m_2} \end{bmatrix}. \tag{16}$$

Taking into account that U is unitary, by the above equation it follows that

$$U_{12}{}^T U_{12} = I_{p_1 - m_2} \tag{17}$$

and

$$U_{12} U_{12}{}^T = I - D_{12} D_{12}{}^{+}. \tag{18}$$

On the other hand, by (13) one deduces that

$$\begin{bmatrix} -\gamma U_{12}{}^T U_{12} & U_{12}{}^T D_{11} \\ D_{11}{}^T U_{12} & -\gamma I_{m_1} \end{bmatrix} < 0.$$

Writing the Schur complement of the element (1,1) and taking into account (17) and (18), from the above inequality one obtains:

$$-\gamma I_{m_1} + \gamma^{-1} D_{11}{}^T \left(I_{p_1} - D_{12} D_{12}{}^{+} \right) D_{11} < 0.$$

Based on the general properties of the pseudo-inverse one deduces that

$$I_{p_1} - D_{12} D_{12}{}^{+} = \left(I_{p_1} - D_{12} D_{12}{}^{+} \right)^T \left(I_{p_1} - D_{12} D_{12}{}^{+} \right), \qquad (19)$$

which shows that the above inequality is equivalent with the first inequality (7). On the other hand, notice that $\hat{B}_2 U_{12} = 0$. Using (17) and (18), the Schur complement of the element (2,2) of the matrix in the left side of (13) becomes:

$$\begin{bmatrix} \begin{matrix} \hat{A}R + RA^T - \gamma \hat{B}_2 \hat{B}_2{}^T \\ +\gamma^{-1} R C_1{}^T \left(I_{p_1} - D_{12} D_{12}{}^{+} \right) C_1 R \end{matrix} & \vdots & \hat{B}_1 + \gamma^{-1} R C_1{}^T \left(I_{p_1} - D_{12} D_{12}{}^{+} \right) D_{11} \\ \cdots\cdots\cdots & \vdots & \cdots\cdots\cdots \\ \hat{B}_1{}^T + \gamma^{-1} D_{11}{}^T \left(I_{p_1} - D_{12} D_{12}{}^{+} \right) C_1 R & \vdots & -\gamma I_{m_1} + \gamma^{-1} D_{11}{}^T \left(I_{p_1} - D_{12} D_{12}{}^{+} \right) D_{11} \end{bmatrix} < 0.$$

Taking into account (19) and the notations (11) one directly obtains that the above inequality is equivalent with:

$$\begin{bmatrix} \hat{A}R + R\hat{A}^T - \gamma \hat{B}_2 \hat{B}_2{}^T + \gamma^{-1} R \hat{C}_1{}^T \hat{C}_1 R & \hat{B}_1 + \gamma^{-1} R \hat{C}_1{}^T \hat{D}_{11} \\ \hat{B}_1{}^T + \gamma^{-1} \hat{D}_{11}{}^T \hat{C}_1 R & -\gamma I_{m_1} + \gamma^{-1} \hat{D}_{11}{}^T \hat{D}_{11} \end{bmatrix} < 0. \qquad (20)$$

The Schur complement of the element (2,2) of the matrix in the left side of the above inequality has the property:

$$\left[\hat{A} + \hat{B}_1 \hat{D}_{11}{}^T \left(\gamma^2 I_{p_1} - \hat{D}_{11} \hat{D}_{11}{}^T \right)^{-1} \hat{C}_1 \right] R + R \left[\hat{A} + \hat{B}_1 \hat{D}_{11}{}^T \left(\gamma^2 I_{p_1} - \hat{D}_{11} \hat{D}_{11}{}^T \right)^{-1} \hat{C}_1 \right]^T$$
$$+\gamma R \hat{C}_1{}^T \left(\gamma^2 I_{p_1} - \hat{D}_{11} \hat{D}_{11}{}^T \right)^{-1} \hat{C}_1 R + \gamma \left[\hat{B}_1 \left(\gamma^2 I_{p_1} - \hat{D}_{11} \hat{D}_{11}{}^T \right)^{-1} \hat{B}_1{}^T - \hat{B}_2 \hat{B}_2{}^T \right] < 0.$$

Denoting $\hat{R} := \gamma^{-1} R$, from the above inequality one obtains that:

$$\left[\hat{A}+\hat{B}_1\hat{D}_{11}{}^T\left(\gamma^2 I_{p_1}-\hat{D}_{11}\hat{D}_{11}{}^T\right)^{-1}\hat{C}_1\right]\hat{R}+\hat{R}\left[\hat{A}+\hat{B}_1\hat{D}_{11}{}^T\left(\gamma^2 I_{p_1}-\hat{D}_{11}\hat{D}_{11}{}^T\right)^{-1}\hat{C}_1\right]^T$$
$$+\gamma^2\hat{R}\hat{C}_1^T\left(\gamma^2 I_{p_1}-\hat{D}_{11}\hat{D}_{11}{}^T\right)^{-1}\hat{C}_1 R+\hat{B}_1\left(\gamma^2 I_{p_1}-\hat{D}_{11}\hat{D}_{11}{}^T\right)^{-1}\hat{B}_1^T-\hat{B}_2\hat{B}_2^T<0. \quad (21)$$

In the following one shows that the pair:

$$\left(\left(\gamma^2 I_{p_1}-\hat{D}_{11}\hat{D}_{11}{}^T\right)^{-\frac{1}{2}}\hat{C}_1,\hat{A}+\hat{B}_1\hat{D}_{11}{}^T\left(\gamma^2 I_{p_1}-\hat{D}_{11}\hat{D}_{11}{}^T\right)^{-1}\hat{C}_1\right) \quad (22)$$

has no unobservable modes on the imaginary axis. Indeed, assuming that it exists $\omega\in\mathbf{R}$ such that

$$\operatorname{rank}\begin{bmatrix} j\omega I_n-\hat{A}-\hat{B}_1\hat{D}_{11}{}^{\mathrm{T}}\left(\gamma^2 I_{p_1}-\hat{D}_{11}\hat{D}_{11}{}^{\mathrm{T}}\right)^{-1}\hat{C}_1 \\ \left(\gamma^2 I_{p_1}-\hat{D}_{11}\hat{D}_{11}{}^{\mathrm{T}}\right)^{-\frac{1}{2}}\hat{C}_1 \end{bmatrix}<n,$$

by the above condition it results that:

$$\operatorname{rank}\begin{bmatrix} I_n & \hat{B}_1\hat{D}_{11}{}^{\mathrm{T}}\left(\gamma^2 I_{p_1}-\hat{D}_{11}\hat{D}_{11}{}^{\mathrm{T}}\right)^{-\frac{1}{2}} \\ 0 & \left(\gamma^2 I_{p_1}-\hat{D}_{11}\hat{D}_{11}{}^{\mathrm{T}}\right)^{\frac{1}{2}} \end{bmatrix}\begin{bmatrix} j\omega I_n-\hat{A}-\hat{B}_1\hat{D}_{11}{}^{\mathrm{T}}\left(\gamma^2 I_{p_1}-\hat{D}_{11}\hat{D}_{11}{}^{\mathrm{T}}\right)^{-1}\hat{C}_1 \\ \left(\gamma^2 I_{p_1}-\hat{D}_{11}\hat{D}_{11}{}^{\mathrm{T}}\right)^{-\frac{1}{2}}\hat{C}_1 \end{bmatrix}$$
$$=\operatorname{rank}\begin{bmatrix} j\omega I_n-\hat{A} \\ \hat{C}_1 \end{bmatrix}<n.$$

Using the definitions of $\hat{A}$ and $\hat{C}_1$ it results that:

$$rank\begin{bmatrix} j\omega I_n - A + B_2 D_{12}{}^{+} C_1 \\ \left(I_{p_1} - D_{12}D_{12}{}^{+}\right)C_1 \end{bmatrix}<n$$

or equivalently, there exists $v\in\mathbf{R}^n$, $v\neq 0$ such that

$$\begin{bmatrix} j\omega I_n - A + B_2 D_{12}{}^{+} C_1 \\ \left(I_{p_1} - D_{12}D_{12}{}^{+}\right)C_1 \end{bmatrix}v=0.$$

On the other hand, from the above equation it follows that:

$$\begin{bmatrix} -j\omega I_n + A - B_2 D_{12}{}^{+} C_1 & B_2 \\ C_1 - D_{12} D_{12}{}^{+} C_1 & D_{12} \end{bmatrix} \begin{bmatrix} v \\ 0 \end{bmatrix} = 0,$$

that is,

$$\begin{bmatrix} -j\omega I_n + A & B_2 \\ C_1 & D_{12} \end{bmatrix} \begin{bmatrix} I_n & 0 \\ -D_{12}{}^{+} C_1 & I_{m_2} \end{bmatrix} = 0. \tag{23}$$

Since the vector:

$$w := \begin{bmatrix} I_n & 0 \\ -D_{12}{}^{+} C_1 & I_{m_2} \end{bmatrix} \begin{bmatrix} v \\ 0 \end{bmatrix} = \begin{bmatrix} v \\ -D_{12}{}^{+} C_1 v \end{bmatrix} \neq 0,$$

by (23) it results that

$$\operatorname{rank} \begin{bmatrix} -j\omega I_n + A & B_2 \\ C_1 & D_{12} \end{bmatrix} < n + m_2,$$

contradicting thus the assumption A2. Therefore the pair (22) has no imaginary unobservable modes and then, applying Proposition 1 for the inequality (21) it follows that the Riccati equation (8) has a stabilizing solution $X \geq 0$.

The proof of the fact that the Riccati equation (9) has a stabilizing solution is similar and it is based on Proposition 1 together with the assumption A3. On the other hand, by Proposition 1 it also results that $0 \leq X < \hat{R}^{-1} = \gamma R^{-1}$ and $0 \leq Y < \hat{S}^{-1} = \gamma S^{-1}$. The first condition (4) in Theorem 2 is equivalent with $RS - I_n \geq 0$, from which (10) directly follows. The second condition (4) is automatically fulfilled for the case $n_k = n$. The sufficiency part of the theorem simply results by the same steps of the necessity part, in reversed order. ■

4. A class of H^{∞} controllers

The result presented in this section is inspired by the paper (Sampei *et al.*, 1990). It gives a strictly proper solution to the H^{∞} control problem and it may be applied both in the singular and in the nonsingular case.

Theorem 3 *It exists an* n*-order strictly proper controller* H^{∞} *controller* $K := (A_k, B_k, C_k, 0)$ *if and only if the two following conditions are accomplished*

i) $\gamma^2 I_{m_1} - D_{11}{}^{T} D_{11} > 0$*, and*

ii) There exist the matrices $F \in \mathbf{R}^{m_2 \times n}$ *and* $K \in \mathbf{R}^{n \times p_2}$ *such that:*
a) Either the Riccati equations:

$$\mathfrak{R}_F(X) = 0 \tag{24}$$

and

$$\mathfrak{S}_K(Y) = 0 \ , \tag{25}$$

where

$$\begin{aligned}\mathfrak{R}_F(X) &:= (A + B_2F)^T X + X(A + B_2F) + \left[XB_1 + (C_1 + D_{12}F)^T D_{11}\right] \\ &\times \left(\gamma^2 I_{m_1} - D_{11}{}^T D_{11}\right)^{-1} \left[XB_1 + (C_1 + D_{12}F)^T D_{11}\right]^T + (C_1 + D_{12}F)^T (C_1 + D_{12}F)\end{aligned} \tag{26}$$

and

$$\begin{aligned}\mathfrak{S}_K(Y) &:= (A + KC_2)Y + Y(A + KC_2)^T + \left[YC_1{}^T + (B_1 + KD_{21})D_{11}{}^T\right] \\ &\times \left(\gamma^2 I_{p_1} - D_{11}D_{11}{}^T\right)^{-1} \left[YC_1{}^T + (B_1 + KD_{21})D_{11}{}^T\right]^T + (B_1 + KD_{21})(B_1 + KD_{21})^T\end{aligned} \tag{27}$$

have the stabilizing solutions $\tilde{X} \geq 0$ *and* $\tilde{Y} \geq 0$ *such that*

$$\rho\left(\tilde{X}\tilde{Y}\right) < \gamma^2 \tag{28}$$

or, equivalently
b) The Riccati inequalities

$$\mathfrak{R}_F(X) < 0 \tag{29}$$

and

$$\mathfrak{S}_K(Y) < 0 \tag{30}$$

have the solutions $\hat{X} > 0$ *and* $\hat{Y} > 0$ *satisfying the condition*

$$\rho\left(\hat{X}\hat{Y}\right) < \gamma^2 \, . \tag{31}$$

Moreover, if the conditions i) and ii) are satisfied then an H^∞ *controller has the following realization in terms of the solutions to the Riccati inequalities (29) and (30)*

$$\begin{aligned}A_k &= A + B_2F + \left(\gamma^2 I_n - \hat{Y}\hat{X}\right)^{-1}\left(\gamma^2 KC_2 - \hat{Y}M\right) \\ B_k &= -\gamma^2\left(\gamma^2 I_n - \hat{Y}\hat{X}\right)^{-1} K \\ C_k &= F,\end{aligned} \tag{32}$$

where

$$\begin{aligned} M &:= F\left[B_2^T \hat{X} + D_{12}^T\left(C_1 + D_{12}F\right)\right] \\ &+\left[\left(\gamma^2 I_n - \hat{Y}\hat{X}\right)^{-1}\hat{Y}\left(B_k D_{21} - B_1\right) + F D_{12}^T D_{11}\right] \\ &\times\left(\gamma^2 I_{m_1} - D_{11}^T D_{11}\right)^{-1}\left[B_1^T \hat{X} + D_{11}^T\left(C_1 + D_{12}F\right)\right] - \mathfrak{R}_F\left(\hat{X}\right). \end{aligned} \tag{33}$$

Proof. Necessity. The condition i) immediately follows taking into account that $\left\|T_{y_1 u_1}\right\|_\infty < \gamma$ implies

$$\lambda_{\max}\left[C_R\left(-j\omega I - A_R\right)^{-1} B_R + D_R\right]^T\left[C_R\left(j\omega I - A_R\right)^{-1} B_R + D_R\right] < \gamma^2,$$

$\forall\ \omega \in \mathbf{R} \cup \{\pm\infty\}$, from which making $\omega \to \infty$, it results that $D_R^T D_R < \gamma^2 I_{m_1}$. Since $D_R = D_{11}$, i) directly follows. Further one proves ii). Assume that it exists a strictly proper controller ($D_k = 0$) solving the H^∞ control problem for the generalized system (1). Then according with the Bounded Real Lemma it follows that it exists $X_R > 0$ such that:

$$\Pi < 0\ , \tag{34}$$

where

$$\begin{aligned} \Pi &:= A_R^T X_R + X_R A_R + \left(X_R B_R + C_R^T D_R\right)\left(\gamma^2 I_{m_1} - D_R^T D_R\right)^{-1} \\ &\times\left(B_R^T X_R + D_R^T C_R\right) + C_R^T C_R \end{aligned} \tag{35}$$

and

$$A_R = \begin{bmatrix} A & B_2 C_k \\ B_k C_2 & A_k \end{bmatrix},\ B_R = \begin{bmatrix} B_1 \\ B_k D_{21} \end{bmatrix},\ C_R = \begin{bmatrix} C_1 & D_{12} C_k \end{bmatrix},\ D_R = D_{11}. \tag{36}$$

Consider the partitions:

$$X_R = \begin{bmatrix} S & N \\ N^{\mathrm{T}} & \tilde{S} \end{bmatrix};\ S, \tilde{S} \in \mathbf{R}^{n\times n} \text{ and } X_R^{-1} = \begin{bmatrix} R & M \\ M^{\mathrm{T}} & \tilde{R} \end{bmatrix};\ R, \tilde{R} \in \mathbf{R}^{n\times n}. \tag{37}$$

Without reducing the generality of the reasoning one can consider a solution $X_R > 0$ of (34) such that the matrix M introduced above is nonsingular. Indeed, if M is singular one can always consider a small enough perturbation such that M becomes nonsingular, X_R^{-1} remains positive definite and condition (34) holds. Therefore, in the following it is assumed that M is invertible. Define:

$$T := \begin{bmatrix} R & M \\ I_n & 0 \end{bmatrix}. \tag{38}$$

Since M invertible, T is invertible, too. Then the inequality (34) is equivalent with:

$$\tilde{\Pi} := T\Pi T^T < 0. \tag{39}$$

Performing the partition

$$\tilde{\Pi} := \begin{bmatrix} \Pi_{11} & \Pi_{12} \\ \Pi_{12}^T & \Pi_{22} \end{bmatrix} < 0, \ \Pi_{11} \in \mathbf{R}^{n\times n}, \tag{40}$$

direct algebraic computations using (35), (36), (38), and the equations obtained from the equality $X_R X_R^{-1} = I$, give:

$$\begin{aligned} \Pi_{11} &= R(A + B_2 F)^T + (A + B_2 F)R \\ &\quad + \left[B_1 + R(C_1 + D_{12}F)^T D_{11} \right] \left(\gamma^2 I_{m_1} - D_{11}^T D_{11} \right)^{-1} \\ &\quad \times \left[B_1^T + D_{11}^T (C_1 + D_{12}F) R \right] + R(C_1 + D_{12}F)^T (C_1 + D_{12}F) R \end{aligned} \tag{41}$$

and

$$\begin{aligned} \Pi_{22} &= S(A + KC) + (A + KC)^T S + \left[S(B_1 + KD_{21}) + C_1^T D_{11} \right] \\ &\quad \times \left(\gamma^2 I_{m_1} - D_{11}^T D_{11} \right)^{-1} \left[(B_1 + KD_{21})^T S + D_{11}^T C_1 \right] + C_1^T C_1, \end{aligned} \tag{42}$$

where the following notations have been introduced:

$$F := C_k M^T R^{-1}, \qquad K := S^{-1} N B_k. \tag{43}$$

Since $\tilde{\Pi} < 0$, it follows that $\Pi_{11} < 0$ and $\Pi_{22} < 0$. Further one can write Π_{11} given by (41) in the equivalent form:

$$\begin{aligned} &\left[A + B_2 F + B_1 \left(\gamma^2 I_{m_1} - D_{11}^T D_{11} \right)^{-1} D_{11}^T (C_1 + D_{12}F) \right] R \\ &+ R \left[A + B_2 F + B_1 \left(\gamma^2 I_{m_1} - D_{11}^T D_{11} \right)^{-1} D_{11}^T (C_1 + D_{12}F) \right]^T \\ &+ R(C_1 + D_{12}F)^T \left[I_{p_1} + D_{11} \left(\gamma^2 I_{m_1} - D_{11}^T D_{11} \right)^{-1} D_{11}^T \right] (C_1 + D_{12}F) R \\ &+ B_1 \left(\gamma^2 I_{m_1} - D_{11}^T D_{11} \right)^{-1} B_1^T < 0. \end{aligned} \tag{44}$$

Taking into account that $\gamma^2 I_{m_1} - D_{11}{}^T D_{11} > 0$ and $R > 0$, one deduces from (44) that the matrix

$$A + B_2 F + B_1 \left(\gamma^2 I_{m_1} - D_{11}{}^T D_{11}\right)^{-1} D_{11}{}^T \left(C_1 + D_{12} F\right)$$

is Hurwitz. Then applying Proposition 1 it results that the Riccati equation

$$\begin{aligned}
&\left[A + B_2 F + B_1 \left(\gamma^2 I_{m_1} - D_{11}{}^T D_{11}\right)^{-1} D_{11}{}^T \left(C_1 + D_{12} F\right)\right]^T X \\
&+X\left[A + B_2 F + B_1 \left(\gamma^2 I_{m_1} - D_{11}{}^T D_{11}\right)^{-1} D_{11}{}^T \left(C_1 + D_{12} F\right)\right] \\
&+XB_1\left(\gamma^2 I_{m_1} - D_{11}{}^T D_{11}\right)^{-1} B_1{}^T X + \left(C_1 + D_{12} F\right)^T \\
&\quad\times\left[I_{p_1} + D_{11}\left(\gamma^2 I_{m_1} - D_{11}{}^T D_{11}\right)^{-1} D_{11}{}^T\right]\left(C_1 + D_{12} F\right) = 0
\end{aligned}$$

has a stabilizing solution $0 \le \tilde{X} < R^{-1}$. Since the above equation coincides with (24) it follows that this equation has a stabilizing solution $\tilde{X} \ge 0$. Further Π_{22} given by (42) can be written in the equivalent form:

$$\begin{aligned}
\Pi_{22} = &\left[A + KC + \left(B_1 + KD_{21}\right) D_{11}{}^T \left(\gamma^2 I_{p_1} - D_{11} D_{11}{}^T\right)^{-1} C_1\right]^T S \\
&+ S\left[A + KC + \left(B_1 + KD_{21}\right) D_{11}{}^T \left(\gamma^2 I_{p_1} - D_{11} D_{11}{}^T\right)^{-1} C_1\right] \\
&+ \gamma^{-2} S\left(B_1 + KD_{21}\right) D_{11}{}^T \left(\gamma^2 I_{p_1} - D_{11} D_{11}{}^T\right)^{-1} D_{11}\left(B_1 + KD_{21}\right)^T S \\
&+ \gamma^{-2} S\left(B_1 + KD_{21}\right)\left(B_1 + KD_{21}\right)^T S + \gamma^2 C_1^T \left(\gamma^2 I_{p_1} - D_{11} D_{11}{}^T\right)^{-1} C_1 < 0
\end{aligned} \tag{45}$$

with $S > 0$. Applying again Proposition 1 it results that the Riccati equation:

$$\begin{aligned}
&\left[A + KC + \left(B_1 + KD_{21}\right) D_{11}{}^T \left(\gamma^2 I_{p_1} - D_{11} D_{11}{}^T\right)^{-1} C_1\right]^T Y \\
&+ Y\left[A + KC + \left(B_1 + KD_{21}\right) D_{11}{}^T \left(\gamma^2 I_{p_1} - D_{11} D_{11}{}^T\right)^{-1} C_1\right] \\
&+ \gamma^2 Y C_1^T \left(\gamma^2 I_{p_1} - D_{11} D_{11}{}^T\right)^{-1} C_1 Y \\
&+ \gamma^{-2}\left(B_1 + KD_{21}\right)\left[I_{m_1} + D_{11}{}^T \left(\gamma^2 I_{p_1} - D_{11} D_{11}{}^T\right)^{-1} D_{11}\right]\left(B_1 + KD_{21}\right)^T = 0
\end{aligned}$$

has a stabilizing solution $0 \le \overline{Y} < S^{-1}$. The above Riccati equation can be rewritten as

$$\left(A+KC_2\right)\overline{Y}+\overline{Y}\left(A+KC_2\right)^T+\gamma^{-2}\left[\gamma^2\overline{Y}C_1^{\ T}+\left(B_1+KD_{21}\right)D_{11}^{\ T}\right]$$
$$\times\left(\gamma^2 I_{p_1}-D_{11}D_{11}^{\ T}\right)^{-1}\left[\gamma^2 C_1\overline{Y}+D_{11}\left(B_1^{\ T}+D_{21}^{\ T}K^T\right)\right]$$
$$+\gamma^{-2}\left(B_1+KD_{21}\right)\left(B_1+KD_{21}\right)^T=0.$$

Multiplying the above equality by γ^2 and denoting $\tilde{Y}:=\gamma^2\overline{Y}$, one directly obtains (25) and thus one concludes that it has the stabilizing solution $\tilde{Y}\ge 0$. On the other hand, since $X_R>0$, it results that $TX_RT^T>0$. Taking into account again (37), and the equations obtained from the equality $X_R X_R^{\ -1}=I_n$ one obtains:

$$\begin{bmatrix} R & I_n \\ I_n & S \end{bmatrix}<0, \tag{46}$$

from which it results $R>S^{-1}$ or equivalently, $R^{-1}S^{-1}<I_n$. Since, as shown earlier $0\le\tilde{X}<R^{-1}$ and $0\le\tilde{Y}=:\gamma^{-2}\overline{Y}<S^{-1}$, the condition (31) in the statement directly follows. The inequalities (29), (30) and (31) follows by pre and post multiplication of (44) and (45) by R^{-1} and S^{-1}, respectively.
Sufficiency. Consider the nonsingular transform

$$\hat{T}=\begin{bmatrix} I_n & I_n \\ -I_n & 0 \end{bmatrix} \tag{47}$$

and define

$$X_R=\begin{bmatrix} \gamma^2\hat{Y}^{-1} & -\left(\gamma^2\hat{Y}^{-1}-\hat{X}\right) \\ -\left(\gamma^2\hat{Y}^{-1}-\hat{X}\right) & \gamma^2\hat{Y}^{-1}-\hat{X} \end{bmatrix}, \tag{48}$$

where $\hat{X}>0$ and $\hat{Y}>0$ satisfy (29), (30) and (31). Then it follows that

$$\hat{T}X_R\hat{T}^T=\begin{bmatrix} \hat{X} & -\hat{X} \\ -\hat{X} & \gamma^2\hat{Y}^{-1} \end{bmatrix}.$$

By a Schur complement argument, it results using (31) that $\hat{T}X_R\hat{T}^T>0$ and therefore $X_R>0$. Consider now the controller with the state-space realization defined in (32) and (33). Computing $\hat{\Pi}:=\hat{T}\Pi\hat{T}^T$ with $\hat{T}$ defined by (47) and with Π having the expression (35), direct algebraic computations give:

$$\hat{\Pi} = \begin{bmatrix} \mathfrak{R}_F\left(\hat{X}\right) & 0 \\ 0 & \gamma^2 \hat{Y}^{-1} \mathfrak{S}_F\left(\hat{Y}\right) \hat{Y}^{-1} \end{bmatrix},$$

Taking into account (29), (30) it follows that $\hat{\Pi} < 0$ and thus $\Pi < 0$. Since $X_R > 0$, according with the Bounded Real Lemma it results that the strictly proper controller $\left(A_k, B_k, C_k\right)$ is a solution to the H^∞ control problem. ■

5. Conclusions

The theoretical developments in Section 2 show that the well-known solvability conditions for the nonsingular H^∞ control problem can be easily obtained in a general form by Theorem 1 and Proposition 1. For $F = -B_2^T X$ and $K = -YC_2^T$, (24) and (25) coincide with the Riccati equations derived in (Doyle *et al.*, 1989).

References

Adamjan, V.M., D.Z. Arov and M.G. Krein (1978). Infinite block Hankel matrices and related extension problems. *AMS Transl.*, Vol. 111, 133-156.

Anderson, B.D.O. and S. Vongpanitlerd (1973). *Network Analysis*. Prentice Hall, Englewood Cliffs.

Ball, J.A. and J.W. Helton (1983). A Beurling-Lax theorem for the Lie group $U\left(m,n\right)$ which contains most classical interpolation theory. *J. Op. Theory*, Vol. 9, 107-142.

Boyd, S., L. El-Ghaoui, E. Feron and V. Balakrishnan (1994). Linear matrix inequalities in systems and control theory. *Studies in Applied Mathematics, SIAM*, Philadelphia, PA, Vol. 15.

Doyle, J.C., K. Glover, P. Khargonekar and P. Francis (1989). State-space solutions to standard H_2 and H_∞ control problem. *IEEE-Trans-Autom. Control*, Vol. 34, 831-848.

Francis, B.A. (1986). *A course in* H_∞ *control theory*. Springer Verlag.

Gahinet, P. (1992). A New Representation of H_∞ Suboptimal Controllers. *Rapport de Recherche, No. 1641*, INRIA.

Gahinet, P. and P. Apkarian (1994). A linear matrix inequality approach to H_∞ control. *International Journal Robust and Nonlinear Control*, No. 4, 421-448.

Glover, K. and J.C. Doyle (1988). State-space formulae for all stabilizing controllers that satisfy an H_∞-norm bound and relations to risk sensitivity. *Systems & Control Letters*, Vol. 11, 167-172.

Ionescu, V., C. Oara and M. Weiss (1999). *Generalised Riccati Theory and Robust Control: A Popov Function Approach*, John Wiley, 1999.

Ionescu, V. and A. Stoica (1999). *Robust Stabilisation and H^∞ Problems*. Kluwer Academic Publishers, 1999.

Iwasaki, T. and R.E. Skelton (1994). All controllers for the general H_∞ control problem: LMI existence conditions and state pace formulas. *Automatica*, Vol. 30, 1307-1317.

Kwakernaak, H. (1986). A polynomial approach to minimax frequency domain optimization of multivariable feedback systems, *Int. J. Contr.*, Vol. 44, 117-156.

Sampei, M., T. Mita and M. Nakamichi (1990). An algebraic approach to H_∞ output feedback control problems. *Systems & Control Letters*, *Vol. 14*, 13-24.

NEW COMPUTATIONAL APPROACH FOR THE DESIGN OF FAULT DETECTION AND ISOLATION FILTERS

A. Varga
German Aerospace Center,
DLR - Oberpfaffenhofen
Institute of Robotics and Mechatronics
D-82234 Wessling, Germany.
E-mail: Andras.Varga@dlr.de

Abstract We propose a numerically reliable computational approach for the design of residual generators for fault detection and isolation filters. The new approach is based on computing solutions of least dynamical orders of linear equations with rational matrix coefficients in combination with special rational factorizations. The main computational ingredients are the orthogonal reduction of the associated system matrix pencil to a certain Kronecker-like staircase form, the solution of a minimal dynamic cover design problem, and the computation of stable and proper rational factorizations with diagonal denominators. For all these computations we discuss numerically reliable algorithms relying on matrix pencil and descriptor system techniques. The proposed residual generator design approach is completely general, is applicable to both continuous- and discrete-time systems, and can easily handle even unstable and/or improper systems.

Keywords: Fault detection, fault isolation, linear rational equations, rational factorizations, numerical methods.

1. Introduction

In the model based fault diagnosis, the fault detection task is achieved by detecting discrepancies between the outputs of the monitored plant and the predictions obtained with a mathematical model. These discrepancies - called also *residuals* - are indications of faults and are produced by special devices called *residual generators*. From a system theoretic point of view, the residual generators are physically realizable systems having as inputs the measured outputs and the control inputs of the monitored system, and as outputs the generated residuals. The residual generators are usually implemented as parts of control algorithms or as independent monitoring procedures.

Several algorithms underlying the design of residual generators require the manipulation of rational matrices. For low dimensional systems, this is possible to some extent by using symbolic manipulation software as provided by tools like Maple or Mathematica. However for large order systems, symbolic computation is not anymore applicable because of tremendous manipulation efforts, and therefore the use of numerical algorithms is the only possible option. The need to address the numerical issues encountered in designing fault detection and isolation filters has been already recognized by (3). By using recently developed numerically reliable descriptor system algorithms, many of computational problems in the fault detection field can be addressed for high dimensional systems.

We propose a new computational approach for the design of residual generators for fault detection and isolation filters based on solving linear equations with rational matrix coefficients. Additionally, rational factorization techniques are employed to ensure the properness and stability of the resulting residual generators. For the solution of these computational problems we propose numerically reliable algorithms relying on descriptor system techniques. The main computational ingredient in solving linear equations with rational matrices is the orthogonal reduction of the associated system matrix pencil to a certain Kronecker-like staircase form. Using this form a solution can be easily constructed, without the need to explicitly invert any rational or polynomial matrix. To determine stable and proper solutions of least dynamical orders, minimal dynamic cover design techniques in combination with coprime factorization procedures are employed. The proposed computational approach to design residual generators for fault detection and isolation filters is completely general, is applicable to both continuous- and discrete-time systems, and can easily handle even unstable and/or improper systems. The design procedure of residual generators can be easily implemented using the robust numerical tools available in the DESCRIPTOR SYSTEMS Toolbox developed by the author[1].

2. Design of fault detection and isolation filters

Consider the linear time-invariant system described by the input-output relations

$$\mathbf{y}(\lambda) = G_p(\lambda)\mathbf{u}(\lambda) + G_f(\lambda)\mathbf{f}(\lambda) + G_d(\lambda)\mathbf{d}(\lambda), \quad (1)$$

where $\mathbf{y}(\lambda)$, $\mathbf{u}(\lambda)$, $\mathbf{f}(\lambda)$, and $\mathbf{d}(\lambda)$ are Laplace- or Z-transformed vectors of the p-dimensional system output vector $y(t)$, m-dimensional plant input vector $u(t)$, q-dimensional fault signal vector $f(t)$, and r-dimensional disturbance vector $d(t)$, respectively, and where $G_p(\lambda)$, $G_f(\lambda)$ and $G_d(\lambda)$ are the *transfer-function matrices* (TFMs) from the plant inputs to outputs, fault signals to outputs, and disturbances to outputs, respectively. According to the system type, $\lambda = s$ in the case of a continuous-time system or $\lambda = z$ in the case of a discrete-time system.

The *fault detection and isolation* (FDI) problem can be formulated as follows: determine a linear residual generator (or detector) of least dynamical order having the general form

$$\mathbf{r}(\lambda) = R(\lambda) \begin{bmatrix} \mathbf{y}(\lambda) \\ \mathbf{u}(\lambda) \end{bmatrix} \tag{2}$$

such that: (i) $r(t) = 0$ when $f(t) = 0$; and (ii) $r_i(t) \neq 0$ when $f_i(t) \neq 0$, for $i = 1, \ldots, q$. Besides the requirement that the TFM of the detector $R(\lambda)$ has least possible McMillan degree, it is also necessary, for physical realizability, that $R(\lambda)$ is a proper and stable TFM.

One possibility to determine a least order $R(\lambda)$ is to solve the following minimal design problem (18): choose a suitable diagonal $M(\lambda)$ (i.e., stable, proper and invertible) and find a least McMillan degree solution $R(\lambda)$ of the linear equation with rational matrices

$$R(\lambda) \begin{bmatrix} G_f(\lambda) & G_d(\lambda) & G_p(\lambda) \\ O & O & I_m \end{bmatrix} = \begin{bmatrix} M(\lambda) & O & O \end{bmatrix} \tag{3}$$

which is stable and proper. This equation arises by imposing for a detector of the general form (2) the condition that $r(\lambda) = M(\lambda)f(\lambda)$ for all $d(\lambda)$ and $u(\lambda)$.

To solve the above equation for properly chosen $M(\lambda)$, the minimum degree algorithm of (23) can be considered as basis for a possible numerical approach using polynomial techniques. Alternatively, provided $[\, G_f(\lambda) \;\; G_d(\lambda) \,]$ is left invertible, a numerically reliable inversion based procedure has been proposed by (18). Here we compute first (using a state space based approach) a least order left-inverse $G^+(\lambda)$ of

$$G(\lambda) = \begin{bmatrix} G_f(\lambda) & G_d(\lambda) & G_p(\lambda) \\ O & O & I_m \end{bmatrix}$$

and then determine $R(\lambda)$ as the numerator of a stable and proper fractional representation $G_1^+(\lambda) = M(\lambda)^{-1}R(\lambda)$, where $G_1^+(\lambda)$ represents the first q rows of $G^+(\lambda)$ and $M(\lambda)$ is diagonal. Note that determining a least order $R(\lambda)$ is part of the computation of the left-inverse $G^+(\lambda)$, and can be explicitly addressed.

In this paper we propose an alternative approach based on solving the rational equation (3) to obtain a least order stable and proper solution $R(\lambda)$ by choosing an appropriate $M(\lambda)$. This computation can be performed in several steps involving manipulation of rational matrices. For each step we propose matrix pencil based reliable numerical algorithms which allow to determine the solution by computing exclusively with real matrices of state space models. By using the proposed approach, the fault detection and isolation filter design problem can be solved in the most general setting. Thus our approach represents a completely general solution to the FDI problem, being a numerically reliable computational alternative to various inversion based methods (4, 5, 10, 13, 8, 3, 9).

3. Solving rational equations

For the design of residual generators in the most general setting, we have to solve a rational equation of the form (3), where we have the additional freedom of choosing a diagonal $M(\lambda)$ such that the resulting $R(\lambda)$ is proper and stable. Since the solution is in general non-unique, we would like to compute a solution which has the least McMillan degree.

In order to solve this problem, we consider, for convenience, the more general dual problem to solve a linear rational system of the form

$$G(\lambda)X(\lambda) = F(\lambda)M(\lambda), \tag{4}$$

where $G(\lambda)$ and $F(\lambda)$ are given $p \times m$ and $p \times q$ rational TFMs, respectively, and we need to choose an invertible diagonal $M(\lambda)$ such that the resulting solution $X(\lambda)$ is proper, stable and has the least possible McMillan degree. It is a well known fact that the system (4) has a solution provided the rank condition

$$\text{rank}\, G(\lambda) = \text{rank}\,[\, G(\lambda)\ F(\lambda)\,] \tag{5}$$

is fulfilled. We assume throughout the paper that this condition holds.

For a given $M(\lambda)$ the general solution of (4) can be expressed as

$$\widehat{X}(\lambda) = X_0(\lambda) + X_N(\lambda)Y(\lambda),$$

where $X_0(\lambda)$ is a particular solution of (4) and $X_N(\lambda)$ is a rational basis matrix for the right nullspace of $G(\lambda)$. Thus a straightforward procedure to solve (4) would be to compute first $X_0(\lambda)$ and $X_N(\lambda)$ for $M(\lambda) = I$, then to determine a suitable $Y(\lambda)$ to obtain a solution $\widehat{X}(\lambda)$ of least McMillan degree, and finally to choose an appropriate $M(\lambda)$ ensuring the stability and properness of $X(\lambda) = \widehat{X}(\lambda)M(\lambda)$.

The main difficulty using this approach is the computation of $Y(\lambda)$ in the case when $X_0(\lambda)$ is not proper. In this case the corresponding $Y(\lambda)$ can be improper as well and for this computation there is no known computational procedure. In contrast, if $X_0(\lambda)$ and $X_N(\lambda)$ are proper rational matrices, then the resulting proper $Y(\lambda)$ can be determined by employing the approach proposed by (12) based on minimal cover design techniques. The following conceptual procedure is merely a refining of the above steps in order to guarantee the applicability of the approach of (12). For this, we determine $M(\lambda)$ in a factored form $M(\lambda) = M_f(\lambda)M_s(\lambda)$, where $M_f(\lambda)$ is a proper and stable factor chosen to ensure the properness of $X_0(\lambda)$ and $M_s(\lambda)$ is a proper and stable factor chosen to ensure the stability of the solution $X(\lambda)$. In what follows we formalize the main steps of the solution procedure and subsequently we discuss suitable computational methods based on pencil manipulation techniques to perform these steps.

1. Compute a particular solution $X_0(\lambda)$ satisfying $G(\lambda)X_0(\lambda) = F(\lambda)$.
2. Compute a proper rational basis $X_N(\lambda)$ of the right nullspace of $G(\lambda)$.
3. Compute a diagonal $M_f(\lambda)$ having the least McMillan degree such that $\widehat{X}_0(\lambda) := X_0(\lambda)M_f(\lambda)$ is proper.
4. Determine a proper $Y(\lambda)$, such that the solution
$$\widehat{X}(\lambda) = \widehat{X}_0(\lambda) + X_N(\lambda)Y(\lambda)$$
has the least possible McMillan degree.
5. Determine a diagonal $M_s(\lambda)$ having the least McMillan degree such that $X(\lambda) := \widehat{X}(\lambda)M_s(\lambda)$ is stable.

In what follows we discuss numerically reliable state space computational algorithms for each step of the above procedure.

3.1. Computation of $X_0(\lambda)$

Let assume that the compound TFM $[\,G(\lambda)\ F(\lambda)\,]$ has a minimal descriptor realization of order n of the form

$$\begin{aligned} E\lambda x(t) &= Ax(t) + B_G u(t) + B_F \nu(t) \\ \xi(t) &= Cx(t) + D_G u(t) + D_F \nu(t) \end{aligned} \tag{6}$$

satisfying

$$[\,G(\lambda)\ F(\lambda)\,] = C(\lambda E - A)^{-1}[\,B_G\ B_F\,] + [\,D_G\ D_F\,]. \tag{7}$$

According to the system type, λ also represents here either the differential operator $\lambda x(t) = \dot{x}(t)$ in the case of a continuous-time system or the advance operator $\lambda x(t) = x(t+1)$ in the case of a discrete-time system. Note that for most of practical applications $[\,G(\lambda)\ F(\lambda)\,]$ is a proper TFM, thus we can always choose a realization such that $E = I$. However, for the sake of generality, we only assume that the pencil $A - \lambda E$ is regular, without assuming E is nonsingular. In this way, we will also cover the most general case of solving rational linear systems.

Let $S_G(\lambda)$ and $S_F(\lambda)$ be the system matrix pencils associated to the realizations of $G(\lambda)$ and $F(\lambda)$

$$S_G(\lambda) = \left[\begin{array}{c|c} A-\lambda E & B_G \\ \hline C & D_G \end{array}\right], \quad S_F(\lambda) = \left[\begin{array}{c|c} A-\lambda E & B_F \\ \hline C & D_F \end{array}\right].$$

Using the straightforward relations

$$\begin{bmatrix} A-\lambda E & B_G \\ O & G(\lambda) \end{bmatrix} = \begin{bmatrix} I_n & O \\ -C(A-\lambda E)^{-1} & I_p \end{bmatrix} S_G(\lambda)$$

$$\begin{bmatrix} A-\lambda E & B_F \\ O & F(\lambda) \end{bmatrix} = \begin{bmatrix} I_n & O \\ -C(A-\lambda E)^{-1} & I_p \end{bmatrix} S_F(\lambda)$$

it is easy to see that $X(\lambda)$ is a solution of $G(\lambda)X(\lambda) = F(\lambda)$ if and only if

$$Y(\lambda) = \begin{bmatrix} Y_{11}(\lambda) & Y_{12}(\lambda) \\ Y_{21}(\lambda) & X(\lambda) \end{bmatrix}$$

satisfies

$$S_G(\lambda)Y(\lambda) = S_F(\lambda). \tag{8}$$

The existence of the solution of (8) is guaranteed by (5), which is equivalent to

$$\operatorname{rank} S_G(\lambda) = \operatorname{rank}\,[\,S_G(\lambda)\ S_F(\lambda)\,]. \tag{9}$$

It follows that instead of solving the rational equation $G(\lambda)X(\lambda) = F(\lambda)$, we can solve the polynomial equation (8) and take

$$X(\lambda) = [\ O \quad I_m\ ]\, Y(\lambda) \begin{bmatrix} O \\ I_q \end{bmatrix}.$$

In fact, since we are interested in the second block column $Y_2(\lambda)$ of $Y(\lambda)$, we need only to solve

$$\begin{bmatrix} A-\lambda E & B_G \\ C & D_G \end{bmatrix} Y_2(\lambda) = \begin{bmatrix} B_F \\ D_F \end{bmatrix} \tag{10}$$

and compute $X(\lambda)$ as

$$X(\lambda) = [\ O \quad I_m\ ]\, Y_2(\lambda).$$

The condition (9) for the existence of a solution becomes

$$\operatorname{rank} \begin{bmatrix} A-\lambda E & B_G \\ C & D_G \end{bmatrix} = \operatorname{rank} \begin{bmatrix} A-\lambda E & B_G & B_F \\ C & D_G & D_F \end{bmatrix}. \tag{11}$$

To solve (10), we isolate a full rank part of $S_G(\lambda)$ by reducing it to a particular Kronecker-like form. Let Q and Z be orthogonal matrices to reduce $S_G(\lambda)$ to the Kronecker-like form

$$\overline{S}_G(\lambda) := QS_G(\lambda)Z = \left[\begin{array}{cccc} B_r & A_r - \lambda E_r & A_{r,reg} - \lambda E_{r,reg} & * \\ 0 & 0 & A_{reg} - \lambda E_{reg} & * \\ 0 & 0 & 0 & A_l - \lambda E_l \end{array} \right], \tag{12}$$

where $A_{reg} - \lambda E_{reg}$ is a regular subpencil, the pair $(A_r - \lambda E_r, B_r)$ is controllable with E_r nonsingular and the subpencil $A_l - \lambda E_l$ has full column rank. The above reduction can be computed by employing numerically stable algorithms as those proposed in (15, 1).

If $\overline{Y}_2(\lambda)$ is a solution of the reduced equation

$$\overline{S}_G(\lambda)\overline{Y}_2(\lambda) = Q \left[\begin{array}{c} B_F \\ D_F \end{array} \right]. \tag{13}$$

then $Y_2(\lambda) = Z\overline{Y}_2(\lambda)$ and thus

$$X(\lambda) = [\ O \quad I_m\ ]\, Z\overline{Y}_2(\lambda)$$

is a solution of the equation $G(\lambda)X(\lambda) = F(\lambda)$. Partition

$$Q \left[\begin{array}{c} -B_F \\ -D_F \end{array} \right] = \left[\begin{array}{c} \overline{B}_1 \\ \overline{B}_2 \\ \overline{B}_3 \end{array} \right]$$

in accordance with the row structure of $\overline{S}_G(\lambda)$. Since $A_l - \lambda E_l$ has full column rank, it follows from (11) that $\overline{B}_3 = 0$. Thus, we can choose $\overline{Y}_2(\lambda)$ of the form

$$\overline{Y}_2(\lambda) = \left[\begin{array}{c} \overline{Y}_{12}(\lambda) \\ \overline{Y}_{22}(\lambda) \\ \overline{Y}_{32}(\lambda) \\ O \end{array} \right],$$

where the partitioning of $\overline{Y}_2(\lambda)$ corresponds to the column partitioning of $\overline{S}_G(\lambda)$. Choosing $\overline{Y}_{12}(\lambda) = 0$, we obtain

$$\left[\begin{array}{c} \overline{Y}_{22}(\lambda) \\ \overline{Y}_{32}(\lambda) \end{array} \right] = \left[\begin{array}{cc} \lambda E_r - A_r & \lambda E_{r,reg} - A_{r,reg} \\ O & \lambda E_{reg} - A_{reg} \end{array} \right]^{-1} \left[\begin{array}{c} \overline{B}_1 \\ \overline{B}_2 \end{array} \right].$$

Let partition $[\,O\ I_m\,]Z$ in accordance with the column structure of $S_G(\lambda)$ as

$$[\,O\ I_m\,]Z = [\,D_r\ C_r\ C_{reg}\ C_l\,] \tag{14}$$

and denote

$$\overline{A} - \lambda\overline{E} = \left[\begin{array}{cc} A_r - \lambda E_r & A_{r,reg} - \lambda E_{r,reg} \\ O & A_{reg} - \lambda E_{reg} \end{array}\right], \quad \overline{B} = \left[\begin{array}{c} \overline{B}_1 \\ \overline{B}_2 \end{array}\right],$$

$$\overline{C} = [C_r \; C_{reg}]$$

Then a particular solution $X_0(\lambda)$ of the equation $G(\lambda)X(\lambda) = F(\lambda)$ can be expressed in form of a descriptor realization

$$X_0(\lambda) = \overline{C}(\lambda\overline{E} - \overline{A})^{-1}\overline{B}.$$

To compute $X_0(\lambda)$ we employed exclusively orthogonal similarity transformations. Therefore, this step is numerically stable, because we can easily show that the computed system matrices in the presence of roundoff errors are exact for an original problem with slightly perturbed data.

Some properties of $X_0(\lambda)$ can be easily deduced from the computed Kronecker -like form. The poles of $X_0(\lambda)$ are among the generalized eigenvalues of the pair $(\overline{A}, \overline{E})$ and are partly fixed, and partly freely assignable. The fixed poles represent the controllable eigenvalues of the pair $(\overline{B}_2, A_{reg} - \lambda E_{reg})$. The generalized eigenvalues of the pair (A_r, E_r) are called the "spurious" poles, and they originate from the column singularity of $G(\lambda)$. These poles are in fact freely assignable by appropriate choice of a (non-orthogonal) right transformation matrix (18).

If $G(\lambda)$ and $F(\lambda)$ have no common zeros then the pair $(\overline{B}_2, A_{reg} - \lambda E_{reg})$ is controllable. This condition is always fulfilled in the case of solving a system (4) originating from FDI problems, where $F(\lambda) = [I \;\; 0]^T$ is a constant full column rank matrix. In this case, the solution $X_0(\lambda)$ will be proper if $[\,G_f(\lambda) \;\; G_d(\lambda)\,]$ (see (1)) has no infinite zeros (i.e., all infinite eigenvalues of the matrix pair (A_{reg}, E_{reg}) are simple). Moreover, a stable and proper solution will exist provided $[\,G_f(\lambda) \; G_d(\lambda)\,]$ is additionally minimum-phase.

More generally, the solution $X(\lambda)$ of $G(\lambda)X(\lambda) = F(\lambda)$ will have no pole in γ (finite or infinite) if $c_\gamma(G) = c_\gamma([G \;\; F])$, where $c_\gamma(G)$ is the *content* of $G(\lambda)$ in γ as defined by (21). Roughly, this is equivalent to say that the pole and zero structures of $G(\lambda)$ and $[G(\lambda) \;\; F(\lambda)]$ at γ coincide. For practical computations, this implies that some or all of common zeros of $G(\lambda)$ and $[G(\lambda) \;\; F(\lambda)]$ will cancel. This cancellation can be done either explicitly by removing the uncontrollable eigenvalues of the pair $(\overline{B}_2, A_{reg} - \lambda E_{reg})$ using orthogonal staircase algorithms (14), or implicitly at the next steps, during determining $\widehat{X}_0(\lambda)$ and $\widehat{X}(\lambda)$.

Remark 1. In this moment, we can easily determine a stable and proper solution of (4) by choosing an invertible diagonal $M(\lambda)$ such that $X(\lambda) := X_0(\lambda)M(\lambda)$ is stable and proper. The computation of $M(\lambda)$ can be done using methods discussed in (18). Note however that the resulting solution is usually not of least McMillan degree. Therefore, the next steps of the proposed procedure address exclusively the least order aspect.

3.2. Computation of $X_N(\lambda)$

Using the same reduction of $S_G(\lambda)$ to $\overline{S}_G(\lambda)$ as in (12), a nullspace basis $X_N(\lambda)$ of $G(\lambda)$ can be computed from a nullspace basis $\overline{Y}_N(\lambda)$ of $\overline{S}_G(\lambda)$ as

$$X_N(\lambda) = [\,O\ I_m\,]Z\overline{Y}_N(\lambda)$$

We can determine $\overline{Y}_N(\lambda)$ in the form

$$\overline{Y}_N(\lambda) = \left[\begin{array}{c} I \\ (\lambda E_r - A_r)^{-1}B_r \\ O \\ O \end{array}\right].$$

With C_r and D_r defined in (14), we obtain a descriptor realization of $X_N(\lambda)$ as

$$X_N(\lambda) = C_r(\lambda E_r - A_r)^{-1}B_r + D_r.$$

Note that $X_N(\lambda)$ is a proper TFM which has least McMillan degree (19). Moreover, the poles of $X_N(\lambda)$ are freely assignable by appropriately choosing the transformation matrices Q and Z to reduce the system pencil $S_G(\lambda)$.

Remark 2. We can express $X_N(\lambda)$ to have the same state, descriptor and output matrices as $X_0(\lambda)$. If we denote $\overline{B}_r = \left[\begin{array}{c} B_r \\ O \end{array}\right]$, then $X_N(\lambda)$ can be also expressed as

$$X_N(\lambda) = \overline{C}(\lambda\overline{E} - \overline{A})^{-1}\overline{B}_r + D_r \qquad (15)$$

This representation is evidently not minimal, since all generalized eigenvalues of the pair (A_{reg}, E_{reg}) are uncontrollable.

3.3. Selecting $M_f(\lambda)$

In principle, this computation can be done simply by solving q independent proper *right coprime factorization* (RCF) problems for the single-input systems corresponding to each of q columns of $X_0(\lambda)$. Assuming $X_{0,i}(\lambda)$ is the i-th column of $X_0(\lambda)$, we can compute the proper RCF

$$X_{0,i}(\lambda) = \frac{\widehat{X}_{0,i}(\lambda)}{m_{f,i}(\lambda)},$$

where $m_{f,i}(\lambda)$ and $\widehat{X}_{0,i}$ are both proper and mutually coprime. The resulting scalar transfer-function $m_{f,i}(\lambda)$ and rational vector $\widehat{X}_{0,i}(\lambda)$ are the i-th diagonal element of $M_f(\lambda)$ and the i-th column of $\widehat{X}_0(\lambda)$, respectively.

The transfer functions $m_{f,i}(\lambda)$ can be chosen, for example, in the form

$$m_{f,i}(\lambda) = \frac{1}{(\lambda+\alpha)^{n_{\infty,i}}},$$

where $n_{\infty,i}$ is the number of infinity zeros of $X_{0,i}(\lambda)$ and α is an arbitrary value, representing a desired stability degree for the solution. It is possible to determine $n_{\infty,i}$ efficiently from the resulting Kronecker-like form of the system pencil $S_G(\lambda)$. We can assume that $A_{reg} - \lambda E_{reg}$ and $\overline{B}_2$ are partitioned conformably and have the structure

$$A_{reg} - \lambda E_{reg} = \left[\begin{array}{cc} A_f - \lambda E_f & A_{f,\infty} - \lambda E_{f,\infty} \\ O & A_\infty - \lambda E_\infty \end{array}\right], \quad \overline{B}_2 = \left[\begin{array}{c} \overline{B}_f \\ \overline{B}_\infty \end{array}\right],$$

where $A_f - \lambda E_f$ and $A_\infty - \lambda E_\infty$ contain the finite and infinite invariant zeros of $S_G(\lambda)$. If we denote $\overline{b}_{\infty,i}$ the i-th column of $\overline{B}_\infty$, then $n_{\infty,i}+1$ is just the order of the controllable part of the pair $(A_\infty - \lambda E_\infty, \overline{b}_{\infty,i})$. To compute $n_{\infty,i}$ we can apply the generalized controllability staircase algorithm of (14) to the pair $(E_\infty - \lambda A_\infty, \overline{b}_{\infty,i})$ (note that A_∞ and E_∞ are interchanged).

After having determined a minimal state-space realization $(A_{M_f} - \lambda I, B_{M_f}, C_{M_f}, 0)$ for $M_f(\lambda)$, it is necessary to compute a proper descriptor representation of $\widehat{X}_0(\lambda) = X_0(\lambda)M_f(\lambda)$. This can be done in two steps: first, remove all uncontrollable infinite eigenvalues from the state-space realization of $X_0(\lambda)M_f(\lambda)$ applying the algorithm of (14) and then remove the non-dynamic part applying standard techniques (22). Both steps can be performed efficiently by exploiting the inherited structure of the system matrices of $X_0(\lambda)$ from the Kronecker-like structure of $S_G(\lambda)$. Note that in the first step we also eliminate the uncontrollable infinite eigenvalues originating from the common infinite poles and zeros of $G(\lambda)$ and $F(\lambda)$. We omit further details here because of lack of space.

3.4. Computation of a least order $\widehat{X}(\lambda)$

We assume that $\widehat{X}_0(\lambda)$ and $X_N(\lambda)$ are proper TFMs and possess state-space representations sharing the same descriptor, state and output matrices

$$\left[\begin{array}{cc} \widehat{X}_0(\lambda) & X_N(\lambda) \end{array}\right] = \left[\begin{array}{c|cc} \widehat{A} - \lambda \widehat{E} & \widehat{B} & \widehat{B}_r \\ \hline \widehat{C} & \widehat{D} & \widehat{D}_r \end{array}\right] \qquad (16)$$

with $\widehat{E}$ non-singular. This can be easily achieved by performing all relevant transformations employed to eliminate the non-proper part of $X_0(\lambda)M_f(\lambda)$ also on the non-minimal realization (15) of $X_N(\lambda)$.

It was shown by (12) that computing a least order solution $\widehat{X}(\lambda) = \widehat{X}_0(\lambda) + X_N(\lambda)Y(\lambda)$ by choosing an appropriate proper $Y(\lambda)$ is equivalent to determine a feedback matrix $\widehat{F}_r$ and a feedforward matrix $\widehat{L}_r$ to cancel the maximum number of unobservable and uncontrollable poles of

$$\widehat{X}(\lambda) = \left[\begin{array}{c|c} \widehat{A} + \widehat{B}_r\widehat{F}_r - \lambda\widehat{E} & \widehat{B} + \widehat{B}_r\widehat{L}_r \\ \hline \widehat{C} + \widehat{D}_r\widehat{F}_r & \widehat{D} + \widehat{D}_r\widehat{L}_r \end{array}\right].$$

It can be shown that if we start with a minimal realization of $[\,G(\lambda)\;\; F(\lambda)\,]$, then we can not produce any unobservable poles in $\widehat{X}(\lambda)$ via state-feedback. Therefore, we only need to determine the matrices $\widehat{F}_r$ and $\widehat{L}_r$ to cancel the maximum number of uncontrollable poles.

(12) has shown that this problem can be solved as a minimal order dynamic cover design problem. Consider the set

$$\mathcal{J} = \{\mathcal{V} : \mathrm{Im}\,\overline{B} + \overline{A}\mathcal{V} \subset \mathrm{Im}\,\overline{B}_r + \mathcal{V}\},$$

where $\overline{A} := \widehat{E}^{-1}\widehat{A}$, $\overline{B} := \widehat{E}^{-1}\widehat{B}$, and $\overline{B}_r := \widehat{E}^{-1}\widehat{B}_r$. Let $\mathcal{J}^*$ denote the set of subspaces in $\mathcal{J}$ of least dimension. If $\mathcal{V} \in \mathcal{J}^*$, then a pair $(\widehat{F}_r, \widehat{L}_r)$ can be determined such that

$$(\overline{A} + \overline{B}_r\widehat{F}_r)\mathcal{V} + \mathrm{Im}\,(\overline{B} + \overline{B}_r\widehat{L}_r) \subset \mathcal{V}.$$

Thus, determining a minimal dimension $\mathcal{V}$ is equivalent to a minimal order cover design problem, and a conceptual approach to solve it has been indicated by (12). The outcome of his method is, besides $\mathcal{V}$, the pair $(\widehat{F}_r, \widehat{L}_r)$ which achieves a maximal order reduction by forcing pole-zero cancellations. This approach has been turned into a numerically reliable procedure by (20). In this procedure $\widehat{F}_r$ and $\widehat{L}_r$ are determined from a special controllability staircase form of the pair $(\widehat{A} - \lambda\widehat{E}, [\,\widehat{B}_r\;\; \widehat{B}\,])$ obtained by using a numerically reliable method relying on both orthogonal and non-orthogonal similarity transformations. An additional feature of this procedure is that all uncontrollable eigenvalues of the pair $(\widehat{A} - \lambda\widehat{E}, \widehat{B})$, arising from common poles or zeros of $G(\lambda)$ and $F(\lambda)$ are also eliminated.

3.5. Selecting $M_s(\lambda)$

The computation of $M_s(\lambda)$ can be done simply by solving q stable RCF problems for the single-input systems corresponding to each column $\widehat{X}_i(\lambda)$ of $\widehat{X}(\lambda)$

$$\widehat{X}_i(\lambda) = \frac{X_i(\lambda)}{m_{s,i}(\lambda)}.$$

One distinctive feature of these single-input factorization problems is that each $\widehat{X}_i(\lambda)$, has generally an uncontrollable descriptor realization. This aspect is handled automatically when employing the Algorithm GRCF-P of (16). Since each of resulting $m_{s,i}(\lambda)$ has least McMillan degree, the resulting diagonal matrix $M_s(\lambda)$ has least McMillan degree as well.

4. Examples

Example 1. Consider the following simple continuous-time example taken from (6):

$$G_p(s) = \begin{bmatrix} \frac{1}{s+1} \\ \frac{1}{(s+1)^2} \end{bmatrix}, \; G_f(s) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \; G_d = 0.$$

A minimal order left-inverse of $G_f(s)$ is $Q(s) = [\; 0 \;\; 1 \;]$, which is proper and stable. According to (4), a residual generator can be determined in the observer-like form

$$\mathbf{r}(s) = Q(s)(\mathbf{y}(s) - G_p(s)\mathbf{u}(s)).$$

This leads to a second order stable and proper detector

$$R(s) = Q(s)[\, I \;\; - G_p(s)\,] = \begin{bmatrix} 0 & 1 & -\frac{1}{(s+1)^2} \end{bmatrix}$$

which is however not of least possible order.

We apply now the proposed approach to compute a least order detector. For this simple model we will explicitly manipulate rational matrices instead of state space matrices. The TFMs defining the equation (4) are given by

$$G(s) = \begin{bmatrix} 0 & 1 & 0 \\ \frac{1}{s+1} & \frac{1}{(s+1)^2} & 1 \end{bmatrix}, \quad F = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

A particular solution $X_0(s)$ of the equation $G(s)X(s) = F$ and a rational nullspace basis $X_N(s)$ of $G(s)$ are

$$X_0(s) = \begin{bmatrix} 0 \\ 1 \\ -\frac{1}{(s+1)^2} \end{bmatrix}, \quad X_N(s) = \begin{bmatrix} -1 \\ 0 \\ \frac{1}{s+1} \end{bmatrix}.$$

Note that $X_0^T(s)$ is the second order detector determined previously. If we choose $Y(s) = \frac{1}{s+1}$, then we obtain a first order stable and proper detector

$$R(s) = (X_0(s) + X_N(s)Y(s))^T = \begin{bmatrix} -\frac{1}{s+1} & 1 & 0 \end{bmatrix}$$

having the least possible McMillan degree.

Example 2. This example is the descriptor system described in (11) corresponding to a linearized three-links planar manipulator model. This model has state vector dimension $n = 11$, command input vector dimension $m = 3$, fault vector dimension $q = 2$, no disturbance input, and output vector dimension $p = 4$. This model is not minimal and a minimal realization has order 5 and is *proper*.

The method proposed by (11) is essentially equivalent to design two independent FDI filters. By considering the fault input 1 as fault and fault input 2 as disturbance, a 4-th order FDI filter $R_1(s)$ has been designed by (11). Similarly, by considering fault input 2 as fault and fault input 1 as disturbance, he obtained a 4-th order FDI filter $R_2(s)$. In this way, a FDI filter of order 8 has been determined by stacking the two designed filters

$$R(s) = \begin{bmatrix} R_1(s) \\ R_2(s) \end{bmatrix}.$$

By using the new approach proposed in this paper we can determine a residual generator which has a least order equal to 2. In the rational system (4) to be solved $G(s)$ is 5×7 and $F(s)$ is 5×2, thus $R(s) = X^T(s)$ will be a 2×7 matrix. A particular solution $X_0(s)$ has been determined having a state space realization of order 10, with the pair $(\overline{A}, \overline{E})$ having 3 finite and 7 infinite generalized eigenvalues. The nullspace basis $X_N(s)$ is 7×2 and has dynamical order 3. With $M_f(s)$ of the form $M_f(s) = \frac{1}{s+1} I_2$, we can eliminate all infinite poles of $X_0(s)$ and the resulting proper solution $\widehat{X}_0(s) = X_0(s) M_f(s)$ has order 5. After performing the minimum cover design, we get a stable solution $X(s) = \widehat{X}(s)$ of order 2 with both eigenvalues stable and equal to -1. For reference purposes we give the resulting 2-nd order FDI filter

$$R(s) = \left[\begin{array}{ccc} -\dfrac{0.01042s + 0.04455}{s+1} & \dfrac{0.03462}{s+1} & -\dfrac{0.03899s + 1.936}{s+1} \\ \dfrac{1}{s+1} & \dfrac{s}{s+1} & 0 \end{array} \right.$$
$$\cdots \left. \begin{array}{cccc} -\dfrac{0.02753s + 1.377}{s+1} & 0 & \dfrac{0.03899}{s+1} & \dfrac{0.02753}{s+1} \\ 0 & -\dfrac{1}{s+1} & 0 & 0 \end{array} \right].$$

5. Conclusions

We proposed numerically reliable approaches to solve several basic computational problems encountered in the design of FDI filters, namely: (1) the solution of linear rational equations; (2) the computation of rational nullspace bases of rational matrices; (3) the reduction of the dynamical orders of the solutions by employing minimal dynamic cover design techniques; and (4) the computation of stable and proper rational factorizations with diagonal denominators. Each of these computations can be performed using numerically stable or numerically reliable algorithms. Using such algorithms, the FDI problem can be solved in the most general setting. Our approach provides, for the first time, a satisfactory numerical solution to this problem. Note that least order residual generator design algorithms have been already proposed to solve the simpler fault detection problem (i.e., without isolation) by (7) using a polynomial basis approach, and by (19) using state space computational techniques.

For the implementation of the proposed residual generator design approach, all necessary basic numerical software is available in the DESCRIPTOR SYSTEMS Toolbox for MATLAB (17), as for example, the computation of Kronecker -like staircase forms, computation of standard and special controllability forms (required in minimum cover design), computation of poles and zeros of descriptor systems, determination of minimal realizations, stable coprime factorization, etc. The basic computational tools in this toolbox are several functionally rich *MEX*-functions, representing MATLAB interfaces to powerful and numerically robust Fortran subroutines partly available in the control and systems library SLICOT (2).

Notes

1. `http://www.robotic.dlr.de/control/num/desctool.html`

References

[1] Beelen, T. and P. Van Dooren (1988). An improved algorithm for the computation of Kronecker's canonical form of a singular pencil. *Lin. Alg.* & *Appl.* **105**, 9–65.

[2] Benner, P., V. Mehrmann, V. Sima, S. Van Huffel and A. Varga (1999). SLICOT – a subroutine library in systems and control theory. In: *Applied and Computational Control, Signals and Circuits* (B. N. Datta, Ed.), Vol. 1., pp. 499–539. Birkhäuser.

[3] Chen, J. and R. J. Patton (1999). *Robust Model-Based Fault Diagnosis for Dynamic Systems,* Kluwer Academic Publishers, London.

[4] Ding, X. and P. M. Frank (1990). Fault detection via factorization. *Systems* & *Control Lett.* **14**, 431–436.

[5] Frank, P. M. and X. Ding (1994). Frequency domain approach to optimally robust residual generation and evaluation for model-based fault diagnosis. *Automatica* **30**, 789–804.

[6] Frisk, E. (2000). Order of residual generators – bounds and algorithms. *Prepr. IFAC Symp. SAFEPROCESS'2000, Budapest, Hungary,* pp. 599–604.

[7] Frisk, E. and M. Nyberg (2001). A minimal polynomial basis solution to residual generation for fault diagnosis in linear systems. *Automatica* **37**, 1417–1424.

[8] Gertler, J. (1998). *Fault Detection and Diagnosis in Engineering Systems,* Marcel Dekker, New York.

[9] Gertler, J. (2000). Designing dynamic consistency relation for fault detection and isolation. *Int. J. Control* **73**, 720–732.

[10] Gertler, J. J. and R. Monajemy (1995). Generating directional residuals with dynamic parity relations. *Automatica* **31**, 627–635.

[11] Hou, M. (2000). Fault detection and isolation for descriptor systems. In: *Issues of Fault Dyagnosis for Dynamic Systems* (R. J. Patton, P. M. Frank and R. N. Clark, Eds.), Springer Verlag, London. pp. 115–144.

[12] Morse, A. S. (1976). Minimal solutions to transfer matrix equations. *IEEE Trans. Autom. Control* **21**, 131–133.

[13] Peng, Y., A. Youssouf, Ph. Arte and M. Kinnaert (1997). A complete procedure for residual generation and evaluation with application to a heat exchanger. *IEEE Trans. Control Systems Technology* **5**, 542–555.

[14] Varga, A. (1990). Computation of irreducible generalized state-space realizations. *Kybernetika* **26**, 89–106.

[15] Varga, A. (1996). Computation of Kronecker-like forms of a system pencil: Applications, algorithms and software. *Proc. CACSD'96 Symposium, Dearborn, MI*, pp. 77–82.

[16] Varga, A. (1998). Computation of coprime factorizations of rational matrices. *Lin. Alg. & Appl.* **271**, 83–115.

[17] Varga, A. (2000). A descriptor systems toolbox for MATLAB. *Proc. CACSD'2000 Symposium, Anchorage, Alaska*.

[18] Varga, A. (2002). Computational issues in fault-detection filter design. *Proc. of CDC'2002, Las Vegas, Nevada*.

[19] Varga, A. (2003*a*). On computing least order fault detectors using rational nullspace bases. *Prepr. of IFAC Symp. SAFEPROCESS'2003, Washington D.C.*

[20] Varga, A. (2003*b*). Reliable algorithms for computing minimal dynamic covers. *(submitted to CDC'2003)*.

[21] Verghese, G. and T. Kailath (1981). Rational matrix structure. *IEEE Trans. Autom. Control* **26**, 434–439.

[22] Verghese, G., B. Lévy and T. Kailath (1981). A generalized state-space for singular systems. *IEEE Trans. Autom. Control* **26**, 811–831.

[23] Wang, S.-H. and E. J. Davison (1973). A minimization algorithm for the design of linear multivariable systems. *IEEE Trans. Autom. Control* **18**, 220–225.

SETTING UP THE REFERENCE INPUT IN SLIDING MOTION CONTROL AND ITS CLOSED-LOOP TRACKING PERFORMANCE

Mihail Voicu and Octavian Pastravanu

Department of Automatic Control and Industrial Informatics
Technical University "Gh. Asachi" of Iaşi
Blvd. Mangeron 53A, 700050 Iaşi, Romania, Phone / Fax: 40-232-230751
e-mail: mvoicu@ac.tuiasi.ro, opastrav@.ac.tuiasi.ro

Abstract On the background of the design of sliding motion controller based on flow – invariance method, the paper presents a solution for setting up the reference input. It uses a state transformation which contains, as a component of the new state vector, the error between the reference input $v(t)$ and the plant output $y(t)$. It is shown that the feedback sliding mode controller ensures the closed-loop tracking performance $y(t) = v(t) = V\sin\Omega t$ under certain conditions in the amplitude – frequency domain (V, Ω).

Keywords: reference input, feedback control, sliding motion control, flow – invariance method, tracking performance

1. Introduction

In this paper, based on two necessary and sufficient conditions, obtained via flow – invariance method, which allow a unified design of the sliding domain reaching and sliding motion, one approaches the problem of setting up the reference input.

Some preliminaries on sliding motion control and the necessary and sufficient conditions for sliding motion and for sliding domain reaching are presented in section 2. In section 3 one presents a solution for setting up the reference input which takes the natural advantage of the feedback control explicitly and directly dependent on the error between the reference input and the plant output. Some concluding remarks are formulated in section 3.

2. Reaching and sliding motion control

Consider the linear time – invariant system:

$$\begin{cases} \dot{\boldsymbol{x}} = \boldsymbol{Ax} + \boldsymbol{b}u + \boldsymbol{Dz}, t \in \mathbf{R}_+, \boldsymbol{x} \in \mathbf{R}^n, u \in \mathbf{R}, \boldsymbol{z} \in \mathbf{R}^q, & (1) \\ y = \boldsymbol{hx}, \ y \in \mathbf{R}, & (2) \end{cases}$$

where $\boldsymbol{x}$ is the state (completely and directly available for measurement), u and y are the scalar control and output respectively, and $\boldsymbol{z}$ is a disturbing input; $\boldsymbol{A} =: (a_{ij})$, $\boldsymbol{b} =: (b_j)$, $\boldsymbol{D} =: (d_{ij})$ and $\boldsymbol{h} =: (h_j)$ are constant matrices of appropriate dimensions.

Let us associate with system (1), (2) the following switching hyperplane:

$$S =: \left\{ \boldsymbol{x} \in \mathbf{R}^n; \ s = \boldsymbol{c}^T \boldsymbol{x} = 0 \right\}, \tag{3}$$

where $\boldsymbol{c} =: \begin{bmatrix} c_1 & \dots & c_n \end{bmatrix}^T$, with $c_n = 1$, is a constant vector and $(\,)^T$ symbolizes the transposition.

Defining the state transformation from $\boldsymbol{x} =: \begin{bmatrix} x_1 & x_{n-1} & x_n \end{bmatrix}^T$ to $\tilde{\boldsymbol{x}} = \begin{bmatrix} x_1 & \dots & x_{n-1} & s \end{bmatrix}^T$, with $s = \boldsymbol{c}^T \boldsymbol{x}$, by:

$$\tilde{\boldsymbol{x}} = \boldsymbol{Px}, \ \boldsymbol{P} =: \left[\begin{array}{c|c} \boldsymbol{I}_{n-1} & 0 \\ \hline \boldsymbol{c}_{(n)}^T & 1 \end{array} \right], \tag{4}$$

where $\boldsymbol{I}_{n-1}$ is the unit matrix of order n–1, and $\boldsymbol{c}_{(n)}$ denotes the vector obtained by deleting the n-th element of $\boldsymbol{c}$, system (1), (2) becomes:

$$\begin{cases} \dot{\boldsymbol{x}}_{(n)} = \boldsymbol{Ex}_{(n)} + \boldsymbol{a}_{(n)}^n \boldsymbol{s} + \boldsymbol{b}_{(n)} u + \boldsymbol{D}_{(n)} z, & (5) \\ \dot{\boldsymbol{s}} = \boldsymbol{c}^T \left[\boldsymbol{A}_{(n)} - \boldsymbol{a}^n \boldsymbol{c}_{(n)}^T \right] \boldsymbol{x}_{(n)} + \boldsymbol{c}^T \boldsymbol{a}^n \boldsymbol{s} + \boldsymbol{c}^T \boldsymbol{b} u + \boldsymbol{c}^T \boldsymbol{Dz}, & (6) \end{cases}$$

$$y = \boldsymbol{hP}^{-1} \tilde{\boldsymbol{x}}. \tag{7}$$

In (5) – (7) $\boldsymbol{E} =: (a_{ij} - a_{in} c_j)$ is an $(n-1) \times (n-1)$ matrix, $\boldsymbol{A}_{(n)}$ and $\boldsymbol{D}_{(n)}$ denote the matrices obtained by deleting the n – th row and column of $\boldsymbol{A}$ and $\boldsymbol{D}$ respectively, $\boldsymbol{a}^n$ is the n-th column of $\boldsymbol{A}$; $\boldsymbol{x}_{(n)}$, $\boldsymbol{a}_{(n)}^n$ and $b_{(n)}$ denote the vectors obtained by deleting the n-the component of $\boldsymbol{x}$, $\boldsymbol{a}^n$ and $\boldsymbol{b}$, respectively.

The achievement of sliding motion of $\mathbf{x}$ on S towards an equilibrium point in $S \subset \mathbf{R}^n$ consists in the synthesis of control $u(t, \ \tilde{\boldsymbol{x}})$, discontinuous on

S, such that the following three requirements are fulfilled:

1. For every initial pair $(t_0, \boldsymbol{x}_0) \in \mathbf{R}_+ \times (\mathbf{R}^n \setminus S)$, with $\boldsymbol{x}(t_0) = \boldsymbol{x}_0$, system (1) evolves towards *S*, i.e. its state reaches hyperplane *S*, in a *reaching point*, after a finite time interval $[t_0, \tau]$, $\tau > t_0$. This is the *reaching condition*.
2. Since the reaching instant τ, corresponding to the reaching point, the state of system (1), (2) remains to evolve on *S*. This evolution is called the *ideal sliding motion* and *S* is the *ideal sliding domain* for system (1), (2). This is the *ideal sliding motion condition*.
3. The *ideal sliding motion* (on *S*) must be asymptotically stable towards an equilibrium point (usually or conventionally $\boldsymbol{x} = \tilde{\boldsymbol{x}} = 0$) belonging to *S*. (Subsidiary and implicitly the disturbance rejection is (eventually) desirable.) This is the *stability condition* of the ideal sliding motion.

Chronologically (i.e. in the dynamics of system (1), (2)), conditions 1° and 2° must be successively fulfilled, while 2° and 3° must be simultaneously satisfied. This means that for each pair $(t_0, \boldsymbol{x}_0) \in \mathbf{R}_+ \times (\mathbf{R}^n \setminus S)$ the whole evolution of system (1), (2) covers two concatenated time intervals: first, a finite one, $[t_0, \tau]$, according to condition 1°, followed by the second one, (τ, t_f), finite or not, according to conditions 2° and 3°.

This essential and, as a matter of fact, natural concatenation of the reaching process followed by the ideal sliding motion has been approached in a unified manner by Voicu and Moroşan (1989; 1991; 1997), Moroşan and Voicu (1994), and Moroşan (1994) using the flow – invariance method, see (Voicu, 1984a; 1984b; 1987; Voicu and Pastravanu, 2003). According to (Voicu and Moroşan, 1997), the starting point of this approach was formalized by Definition 1 (pertaining to the *ideal sliding domain*) and correspondingly characterized by Theorem 4 (based on the flow – invariance method), and, via the flow structure of the state space induced by sliding motion control (unfolded by Theorem 1), by Theorem 5 which characterizes the reaching process as a natural flowing precursor of the ideal sliding motion.

In view of Definition 1 and Theorems 4 and 5 (Voicu and Moroşan, 1997) it follows that for the control design only equation (6) rewritten in the following equivalent form is to be used:

$$\dot{s} = \sum_{i=1}^{n-1} \boldsymbol{c}^T (\boldsymbol{a}^i - c_i \boldsymbol{a}^n) x_i + \boldsymbol{c}^T \boldsymbol{a}^n s + \boldsymbol{c}^T \boldsymbol{b} u + \boldsymbol{c}^T \boldsymbol{D} \boldsymbol{z}, \tag{8}$$

where $\boldsymbol{a}^i$ is the *i*-th column of **A**. Taking into account the possibility that $\boldsymbol{c}^T(\boldsymbol{a}^i - c_i \boldsymbol{a}^n) = 0$ for some *i*, let us define the index sets:

$$I = \{1,\ 2,\ \ldots,\ n-1\}, \quad J = \{i \in I;\ \ c^T (\boldsymbol{a}^i - c_i \boldsymbol{a}^n) \neq 0\}. \tag{9}$$

For $\boldsymbol{c}^T\boldsymbol{b} \neq 0$, one can synthesize the following control algorithm:

$$u = -u_r(s) - u_s(\boldsymbol{x}_{(n)}) - u_z(\boldsymbol{x}_{(n)}), \tag{10}$$

where

$$u_r(s) = \rho s + \delta_r, \quad \delta_r = \begin{cases} \delta_0, & s < 0, \\ 0, & s = 0, \\ -\delta_0, & s > 0, \end{cases} \tag{11}$$

has to control the reaching process;

$$u_s(\boldsymbol{x}_{(n)}) = \sum_{i \in J} \Psi_i x_i, \; \Psi_i = \begin{cases} \beta_i, & x_i s < 0, \\ \alpha_i, & x_i s > 0, \end{cases} \tag{12}$$

is the *sliding motion control*;

$$u_z(\boldsymbol{x}_{(n)}) = \begin{cases} \beta_0, & s < 0, \\ \alpha_0, & s > 0, \end{cases} \tag{13}$$

ensures the *disturbance rejection*.

Clearly, ρ, δ_0, α_i, β_i $(i \in J)$ and α_0, β_0 are the design parameters of the *variable structure controller* which are to be adjusted in order to fulfill the requirements $1^{\circ} - 3^{\circ}$.

Using inequalities (15) and (16) of Theorem 4 (Voicu and Moroşan, 1997), for equation (8) the following result can be formulated.

Theorem 1 *Hyperplane S is the ideal sliding domain for system (1) – (4), (10) – (13), with* $\boldsymbol{c}^T\boldsymbol{b} \neq 0$, *if and only if*

$$\begin{cases} \alpha_i \boldsymbol{c}^T\boldsymbol{b} \geq \boldsymbol{c}^T(\boldsymbol{a}^i - \boldsymbol{c}_i\boldsymbol{a}^n) \\ \beta_i \boldsymbol{c}^T\boldsymbol{b} \leq \boldsymbol{c}^T(\boldsymbol{a}^i - \boldsymbol{c}_i\boldsymbol{a}^n), \quad i \in J, \end{cases} \tag{14}$$

$$\begin{aligned} &\alpha_0 \boldsymbol{c}^T\boldsymbol{b} \geq \sup_t \boldsymbol{c}^T\boldsymbol{D}\boldsymbol{z}(t) \\ &\beta_0 \boldsymbol{c}^T\boldsymbol{b} \leq \inf_t \boldsymbol{c}^T\boldsymbol{D}\boldsymbol{z}(t). \end{aligned} \tag{15}$$

In order to characterize the reaching process as a precursor of the sliding motion, let us introduce the following reaching function:

$$r(t) = \begin{cases} \left[(|s_0| + \delta)e^{-\lambda(t-t_0)} - \delta\right] \; \mathrm{sgn}(s), & t \in [t_0, \; \tau], \\ 0 \;, & t \in (\tau, \; t_f), \end{cases} \tag{16}$$

where $s_0 =: s(\boldsymbol{x}_0) = s(\boldsymbol{x}(t_0))$, and $\delta > 0$, $\lambda > 0$, $\tau > t_0$ are pre-assignable parameters. It is easy to verify that the reaching function (16) satisfies the statements 1° and 2° of Theorem 5 (Voicu and Moroşan, 1997), with $\tau = t_0 + \lambda^{-1} \ln(1 + |s_0|/\delta) > t_0$ (τ being the reaching instant, i.e. $r(\tau) = 0$). Using the statement 3 of Theorem 5 (loc cit) with (11) – 13), under (14), (15), one can formulate the following result.

Theorem 2 *For each pair* $(t_0, \mathbf{x}_0) \in \mathbf{R}_+ \times (\mathbf{R}^n \setminus S)$ *the state of system (1) – (4), (10) – (13), with* $\boldsymbol{c}^T\boldsymbol{b} \neq 0$*, reaches ideal sliding domain S if and only if*

$$\begin{aligned} \rho \boldsymbol{c}^T \boldsymbol{b} &\geq \boldsymbol{c}^T \boldsymbol{a}^n + \lambda \\ \delta_0 \boldsymbol{c}^T \boldsymbol{b} &\leq \boldsymbol{c}^T \boldsymbol{a}^n . \end{aligned} \tag{17}$$

Equations (10) – (13) and inequalities (14), (15) and (17) define the (variable) structure and the adjustable parameters of the controller according to requirements 1° and 2°.

Using the equivalent control method (Utkin, 1978), i.e. by solving the equation:

$$\dot{s} = 0 \tag{18}$$

(according to (6)) with respect to u and replacing the result, with $s = 0$, into (5), one obtains the *ideal sliding equations*:

$$\begin{cases} \dot{\boldsymbol{x}}_{(n)} = \boldsymbol{A}^1 \boldsymbol{x}_{(n)} + \boldsymbol{D}^1 \boldsymbol{z}, & (19) \\ s = 0 & (20) \end{cases}$$

where matrices $\mathbf{A}^1$ and $\mathbf{D}^1$ can be appropriately calculated.

If

$$\text{rank}\, \boldsymbol{b} = \text{rank}\, [\boldsymbol{b}, \ \boldsymbol{D}], \tag{21}$$

then (Utkin, 1978)

$$\boldsymbol{D}^1 = 0 \tag{22}$$

and the disturbance rejection is ensured.

The stability of system (19) depends on $\boldsymbol{A}^1$ only and may be improved (according to requirement 3°) by an additional state/output feedback in (10). Under these circumstances the state (and output) trajectories of system (1) – (4), (10) – (13) evolve with pre-assignable velocity (for the state only) from each initial state $\boldsymbol{x}_0 \in \mathbf{R}^n \setminus S$ towards ideal sliding domain S and then towards the final equilibrium point $\boldsymbol{x} = \tilde{\boldsymbol{x}} = 0$ (and $y = 0$ respectively).

3. Setting up the reference input

The problem of setting up the reference input in the control (variable) structure with a sliding motion controller may be solved in the following two manners.

The first solution, presented and analyzed by Utkin (1978), consists in the introduction of an additional term in equation (10), depending on the scalar reference input v. This means that (10) has to be replaced by:

$$u = u_v + u_c, \tag{23}$$

where

$$u_v = \alpha v \tag{24}$$

is the reference input term with α= constant, and

$$u_c = -u_r(s) - u_s(\boldsymbol{x}_{(n)}) - u_z(\boldsymbol{x}_{(n)}) \tag{25}$$

is the reaching, sliding motion and rejection control. Because output y should track reference input v, coefficient α is to be chosen, e.g. pertaining to the steady state of the whole system, such that:

$$y_{ss} = v_{ss} \tag{26}$$

(subscript $_{ss}$ denotes the steady state), i.e.

$$\alpha = -(\boldsymbol{h}\boldsymbol{A}^{-1}\boldsymbol{b})^{-1},\ \det \boldsymbol{A} \neq 0\ . \tag{27}$$

Unlike this solution, the second one, which is addressed in the sequel, takes the natural (and classical) advantage of the feedback control explicitly and directly dependent on the error between reference input v and plant output y, i.e. (for instance):

$$x_1 =: v - y\ . \tag{28}$$

In order to expose this solution, let us consider the plant:

$$\begin{cases} \dot{\overline{\boldsymbol{x}}} = \overline{\boldsymbol{A}}\overline{\boldsymbol{x}} + \overline{\boldsymbol{b}}u + \overline{\boldsymbol{D}}\boldsymbol{z}, t \in \mathbf{R}_+, \overline{\boldsymbol{x}} \in \mathbf{R}^n, u \in \mathbf{R}, \boldsymbol{z} \in \mathbf{R}^q & (29) \\ y = \overline{\boldsymbol{h}}\overline{\boldsymbol{x}}\ , & (30) \end{cases}$$

where $\overline{\boldsymbol{x}}$ is the state (completely and directly available for measurement), and $\overline{\boldsymbol{A}}$, $\overline{\boldsymbol{b}}$, $\overline{\boldsymbol{D}}$ and $\overline{\boldsymbol{h}}$ are time – invariant matrices of appropriate dimensions.

The trivial case $\overline{\boldsymbol{h}} =: \left[\overline{h}_1 \quad \ldots \quad \overline{h}_n\right] = 0$ being excluded, for the sake of simplicity, let us suppose that:

$$\overline{h}_1 \neq 0 \tag{31}$$

(as it was subsidiary assumed by (28)). In this case, error (28), with (30), may be taken into consideration, for instance, by the transformation:

$$\boldsymbol{x} = \boldsymbol{i}^1 v - \boldsymbol{Q}\overline{\boldsymbol{x}} \tag{32}$$

where

$$\boldsymbol{Q} =: \left[\begin{array}{c|c} \bar{h}_1 & \bar{\boldsymbol{h}}_{(1)} \\ \hline 0 & \boldsymbol{I}_{n-1} \end{array}\right], \tag{33}$$

$\boldsymbol{i}^1$ is the first column of $\boldsymbol{I}_n$ and $\bar{\boldsymbol{h}}_{(1)}$ is obtained by deleting the first element of row matrix $\bar{\boldsymbol{h}}$.

Using the transformation (32) with (33) for the plant described by (29) and (30), one obtains:

$$\begin{cases} \dot{\boldsymbol{x}} = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}u + \boldsymbol{D}\boldsymbol{z} + \boldsymbol{F}\boldsymbol{w}, & (34) \\ y = \boldsymbol{h}\boldsymbol{x} + v, & (35) \end{cases}$$

where

$$\boldsymbol{A} =: \boldsymbol{Q}\bar{\boldsymbol{A}}\boldsymbol{Q}^{-1}, \ \boldsymbol{b} =: -\boldsymbol{Q}\bar{\boldsymbol{b}}, \ \ \boldsymbol{D} = -\boldsymbol{Q}\bar{\boldsymbol{D}}, \\ \boldsymbol{F} =: \left[-\boldsymbol{a}^1, \ \ \boldsymbol{i}^1\right], \ \boldsymbol{h} =: \left[-1 \ \ 0 \ \ \ldots \ \ 0\right] \tag{36}$$

$$\boldsymbol{w} =: \left[v \ \ \dot{v}\right]^T \in \mathbf{R}^2 . \tag{37}$$

In comparison with equation (1), equation (36) contains the disturbing term $\boldsymbol{Fw}$ and, in control equation (10) the term u_z has to be replaced with u_{zw} accordingly. Under these circumstances Theorem 1 must be adequately completed as follows.

Theorem 3 *Hyperplane S is the ideal sliding domain for system (34), (35), (3), (4), (10) – (13), with* $\boldsymbol{c}^T\boldsymbol{b} \neq 0$*, if and only if (14), (15) are met and*

$$\begin{cases} \alpha_0 \boldsymbol{c}^T \boldsymbol{b} \geq \sup_t \boldsymbol{c}^T \boldsymbol{F}\boldsymbol{w}(t) \\ \beta_0 \boldsymbol{c}^T \boldsymbol{b} \leq \inf_t \boldsymbol{c}^T \boldsymbol{F}\boldsymbol{w}(t) \end{cases}. \tag{38}$$

At the same time, for the reaching function (16), Theorem 2 is to be applied to system (34), (35), (3), (4), (10) – (13).

According to (19) and (20), the ideal sliding equations are now:

$$\begin{cases} \dot{\boldsymbol{x}}_{(n)} = \boldsymbol{A}^1 \boldsymbol{x}_{(n)} + \boldsymbol{D}^1 \boldsymbol{z} + \boldsymbol{F}^1 \boldsymbol{w}, & (39) \\ s = 0, & (40) \end{cases}$$

where matrices $\boldsymbol{A}^1$, $\boldsymbol{D}^1$ and $\boldsymbol{F}^1$can be appropriately calculated.

In addition to (21), (22), one has to point out that if

$$\operatorname{rank} \boldsymbol{b} = \operatorname{rank}\,[\boldsymbol{b}, \boldsymbol{F}] \tag{41}$$

then

$$\boldsymbol{F}^1 = 0 \tag{42}$$

and the rejection of disturbance $\boldsymbol{w}$ is also ensured.

Clearly, related to the steady state of the whole system, one may assert that if $\boldsymbol{A}^1$ is Hurwitzian, then $\boldsymbol{x} = \tilde{\boldsymbol{x}} = 0$, together with (32), (33), as a final result of the control (10), yield the steady state (26) as a natural consequence of the feedback control explicitly and directly dependent on the error between v and y. Moreover, for any $\boldsymbol{w}$ (see (37)) satisfying conditions (38), the system (34), (35), (3), (4), (10) – (13), after a settling time depending on x_0, λ and δ (from (16)) and the greatest real part of the eigenvalues of Hurwitzian matrix $\boldsymbol{A}^1$, achieves the closed-loop tracking performance:

$$y(t) = v(t)\,. \tag{43}$$

4. Closed-loop tracking performance

The employment of (15) and (38) pertaining to the choice of parameters α_0 and β_0 demands some knowledge about disturbance z and reference input v (see (37)). In order to illustrate the tracking performance of the closed-loop system and, correspondingly, to determine the tracking conditions in the amplitude – frequency domain, let us consider that the system is not disturbed by $\boldsymbol{z}$, i.e. $\boldsymbol{z} = 0$, and the reference input has the following form:

$$v(t) = V\sin \Omega t. \tag{44}$$

Using (44) in (38), with (36), (37), and by assuming that $\boldsymbol{c}^T\boldsymbol{b} > 0$ and $\alpha_0 = -\beta_0 > 0$, one successively obtains:

$$\boldsymbol{c}^T\boldsymbol{b}\alpha_0 = -\boldsymbol{c}^T\boldsymbol{b}\beta_0 \geq \sup_t V(-\boldsymbol{c}^T\boldsymbol{a}^1 \sin\Omega t + c_1\Omega\cos\Omega t),$$

$$\boldsymbol{c}^T\boldsymbol{b}\alpha_0 = -\boldsymbol{c}^T\boldsymbol{b}\beta_0 \geq V\left(\left(\boldsymbol{c}^T\boldsymbol{a}^1\right)^2 + \left(c_1\Omega\right)^2\right)^{1/2}. \tag{45}$$

Tacking into account that usually

$$0 < \alpha_0 = -\beta_0 \leq \alpha_{\max}, \tag{46}$$

where $\alpha_{\max}$ depends on some technological limits (saturation, limited control power) characterizing the plant (30), from (45) and (46) one obtains for the closed-loop tracking performance (43) the following admissible amplitude – frequency domain:

$$V \le \boldsymbol{c}^T \boldsymbol{b} \alpha_0 \left(\left(\boldsymbol{c}^T \boldsymbol{a}^1 \right)^2 + \left(c_1 \Omega \right)^2 \right)^{-1/2} \le \boldsymbol{c}^T \boldsymbol{b} \alpha_{\max} \left(\left(\boldsymbol{c}^T \boldsymbol{a}^1 \right)^2 + \left(c_1 \Omega \right)^2 \right)^{-1/2}, \quad (47)$$

which is ensured by the feedback control explicitly and directly dependent on the error (28) between reference input v and plant output y.

Finally, let us remark that suitable transformations such as (32), which include the input – output error (28) and some other errors pertaining to the velocity, acceleration etc, may be taken into consideration in order to improve the tracking behavior of the feedback control systems with sliding motion controller.

5. Concluding remarks

In this paper, on the background of two necessary and sufficient conditions, obtained via flow – invariance method, which allow a unified design of the (variable structure) state feedback controller ensuring the sliding – domain reaching and the sliding motion, one gives a solution for setting-up the reference input for the whole system containing a variable structure controller. Unlike the usual solution (which consists in an additional term dependent on the reference input, in the plant control), the solution presented in the paper uses a state transformation (e.g. (32), (33)) which contains, as a component of the new state vector, the error between the reference input v and the plant output y. Under these circumstances the feedback controller based on this new state vector, under certain conditions and after an inherent settling time, ensures the closed-loop tracking performance $y(t) = v(t)$.

References

Moroşan, B.I. (1994). *Digital Variable – Structure Systems Studied by Flow – Invariance Method*. Doctoral thesis (in Romanian), Technical University of Iaşi.

Moroşan, B.I., Voicu, M. (1994). General sliding mode systems analysis and design via flow – invariance method. *Studies in Informatics and Control*, **vol. 3**, 4, pp. 347 – 366.

Utkin, V.I. (1978). *Sliding Modes and Their Application in Variable Structure Systems*. MIR, Moscow.

Voicu, M. (1984a). Free response characterization via flow – invariance. *9-th World Congress of IFAC*, Budapest. Preprints, **vol. 5**, pp. 12 – 17.

Voicu, M. (1984b). Componentwise asymptotic stability of linear constant dynamical systems. *IEEE Trans. on AC*, **AC – 29**, 10, pp. 937 – 939.

Voicu, M. (1987). On the application of the flow – invariance method in control theory and design. *10 – th World Congress of IFAC*, Munich. Preprints, **vol. 8**, pp. 364 – 369.

Voicu, M., Moroşan, B.I. (1989). State space flow structure induced by sliding motion control. *Bul. Inst. Politehnic din Iaşi, III*, **t. XXXV (XXXIX)**, 3 – 4, pp. 25 – 29.

Voicu, M., Moroşan, B.I. (1991). Variable – structure controller for feedback positioning of d.c. electric motor. *Rev Roum. Sci. Techn., ser. Electrot. Energet.*, **t. 36**, 2, pp. 247 – 264.

Voicu, M., Moroşan, B.I. (1997). Variable – structure control characterization and design via flow – invariance method. *Int. J. Automation Austria*, **Jg. 5**, 1, pp. 12 - 24.

Voicu M., Pastravanu O. (2003). Flow-invariance method in control – a survey of some results. In Voicu M. (Ed.), *Advances in Automatic Control*, Kluwer, pp. 393 - 434 (in this volume).

FLOW-INVARIANCE METHOD IN CONTROL - A SURVEY OF SOME RESULTS

Mihail Voicu and Octavian Păstrăvanu

Department of Automatic Control and Industrial Informatics

Technical University of Iasi

Blvd. D. Mangeron 53A, RO-700050 Iasi, Romania

E-mail: {mvoicu,opastrav}@ac.tuiasi.ro

Abstract This paper represents a survey of the main contributions of the two authors to the application of the flow-invariance method in control system analysis and design. By considering time-dependent state hyper-intervals, a componentwise approach has been developed for the evaluation of the behaviour of continuous-time dynamical systems, in general, and of their stability (regarded as a free response) in particular. Thus, a special type of asymptotic stability, called componentwise (exponential) asymptotic stability was introduced, which allows individual monitoring of each state variable, unlike the standard concept of asymptotic stability offering a global characterisation in terms of vector norms. Consequently, new and refined instruments were created for the analysis of linear and nonlinear dynamical systems, with constant or interval type coefficients. These instruments are equally useful in synthesis for dealing with componentwise stabilizability and detectability. The same framework constructed on the flow-invariance background is also able to accommodate the design of the sliding motion control. The paper does not contain proofs for the enounced results in order to keep the size of the text under some moderate limits, but these proofs can be found in the cited references.

Keywords: control systems, time dependent flow-invariance, componentwise asymptotic (exponential) stability, componentwise absolute stability, robustness, interval matrix systems, nonlinear uncertain systems, componentwise detectability and stabilizability, sliding motion control

1. Introduction

Generally speaking, the object of flow-invariance theory is the “stream”, in the state space X of a dynamical system Σ (described by a differential equation), of state trajectories $\boldsymbol{x}(t)$, $t \in [t_0, t_f], t_f > t_0$, starting in the *initial states* $x(t_0) = x_0 \in X$ and arriving in the corresponding *current states* $x(t) \in X$. More precisely, the evolution operator $\Phi_t : X \to X$, which generates the *flow* of Σ, is the *application* that transforms the (set of) initial states into (the set) of current states and, under certain assumptions and details, Σ itself can be represented by $\mathbf{x}(t) = \Phi_t(\mathbf{x}_0)$. As expected, the notion of flow and the entire conceptual vision around it, based on the investigations of Φ_t and by means of Φ_t, offer more refined instruments for the qualitative analysis of differential equations that may unfold more subtle structural properties and characterizations of dynamical systems. In this respect the classical attributive (joint) concept of *invariance* opened a special field of research that studies the existence and characterizations of some (nonempty) *flow-invariant* subsets $X_i \subseteq X$ (not necessarily time-invariant) having the property that for each $\mathbf{x}_0 \in X_i$ system Σ fulfils $\mathbf{x}(t) \in X_i$ for each $t \in [t_0, t_f]$. It is worth to be noticed here that such behaviour, interesting by itself because it certainly occasions a deeper insight into the structure of dynamical systems, characterizes large classes of *real* dynamical systems encountered in engineering (electric circuits and networks, neuronal networks, control systems), biology, ecology, pharmacokinetics etc. In these cases the system evolution must satisfy certain state and control constraints concerning and conditioning their normal functioning and their very physical existence respectively. In this context the theoretical research can ensure a meaningful non-conventional knowledge of significant classes of real systems and, at the same time, accurately enlighten subtle aspects of theoretical and practical interest for various applications.

The *method* of flow-invariance emerged, in the theory of differential equations, from the pioneering research developed by Nagumo [1] and Hukuhara [2] at the middle of the last century. Further significant contributions have been brought by many well-known mathematicians, among which Brezis [3], Crandall [4], Martin [5]. Two remarkable monographs on this field are due to Pavel [6], and Motreanu and Pavel [7]. Voicu, in [8] – [20], by using the flow-invariance method for state hyper-intervals, defines and characterizes the componentwise asymptotic stability (CWAS) and the componentwise exponential asymptotic stability (CWEAS) and analyses their subsidiary aspects and connections with some other properties pertaining to the sphere of stability concept. Recent results extend the concepts of CW(E)AS for interval matrix systems and for a certain class of nonlinear systems, including the analysis of robustness and of the preservation of CW(E)AS under certain disturbances, (Păstrăvanu and Voicu, [21] – [29]), for time-delay systems and 1D and 2D linear discrete-time systems (Hmamed, [30], [31]), for non-symmetrical hyper-intervals (Hmamed and Benzaouia, [32]), and for time-discrete neural networks (Chu, [33]). Further results on the synthesis problems of the CWEAS state feedback controller and of the CWEAS state observer are reported by Voicu [34] – [40], Voicu and Bulea [41], and by Voicu [42]. A unified approach, due to Păstrăvanu and Voicu [43], efficiently solves these two problems by using a convex optimisation procedure with a cost function expressed by a matrix infinity norm equivalently involving CWEAS. According to Voicu and Moroşan [44] – [50], Moroşan [51] – [56], and Moroşan and Voicu [57], [58], another field of successful application of the concept of CWEAS is that of variable structure (sliding motion) control. Recently, the flow-invariance principles have been adapted by Păstrăvanu and Voicu in order to address refined objectives in the analysis [21] - [23] and synthesis [43] of linear systems with discrete-time dynamics.

The current paper is a survey that collects (partially according to the leading idea of [59]) the principal results generated by the application of the flow-invariance method (on time-dependent state hyper-intervals) for the componentwise evaluation of the evolution and of the stability of dynamical systems (Sections 2 – 4, and 5). The remainder of the paper deals with the characterization of componentwise detectability and stabilizability (Section 6) and the synthesis of the sliding motion control based on the flow-invariance approach (Section 7). Motivated, on the one hand, by the natural necessity of a cursive and compact presentation of the results (definitions and the afferent theorems, pertinently commented) and, on the other hand, by some constraints viewing the maximum length of the text, the paper does not contain any proof of the enounced theorems. For these proofs the cited references are recommended. Finally (Section 8), some concluding remarks are formulated and further research directions are identified.

2. Constrained Evolution of Dynamical Control Systems

2.1. General Definition and Characterizations, [15]

Consider the nonlinear continuous-time dynamical system:

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}, \mathbf{u}), t \in \mathbb{R}_+, \mathbf{x} \in \mathbb{R}^n, \mathbf{u} \in \mathbb{R}^m, \tag{2.1}$$

where $\mathbf{x}$ is the state, $\mathbf{u}$ is the control, and $\mathbf{f}$ is a vector function.

According to the usual attributes of a dynamical control system, assume that $\mathbf{u}$ belongs to the following set of admissible controls:

$$\mathcal{U} =: \left\{ \mathbf{u} \in \overline{C}^0; \quad \mathbf{u}(t) \in U(t) \subset \mathbb{R}^m, \quad t \in T \right\}, \tag{2.2}$$

where $\overline{C}^0$ is the set of the piecewise continuous vector functions (i.e. left discontinuous in a number of points) on the time interval:

$$T =: [t_0, t_1) \subseteq \mathbb{R}_+, \quad t_1 > t_0, \tag{2.3}$$

and $U(t)$ is a compact (time-dependent) subset. In compliance with these control constraints, one has to assume that $\mathbf{x}(t)$ also belongs to a certain compact (time-dependent) subset $X(t) \subset \mathbb{R}^n$. Under these circumstances the problem concerning the constrained evolution of control system (2.1) on $T \times \mathcal{U} \times X$ is consistent and can be dealt with the flow-invariance method.

Definition 2.1 The evolution of control system (2.1) is called $T\mathcal{U}X$–*constrained evolution* if for each $\mathbf{x}(t_0) =: \mathbf{x}_0 \in X(t_0)$ and for each $\mathbf{u} \in \mathcal{U}$ the response of control system (2.1) satisfies:

$$\mathbf{x}(t) \in X(t) \quad \forall t \in T. \ \Box \tag{2.4}$$

According to [60] the $T\mathcal{U}X$–constrained evolution of control system (2.1), under the condition that the Cauchy solution is unique (i.e. $\mathbf{f}$ is continuous and locally Lipschtzian in $\mathbf{x}$), is equivalent with the flow-invariance of $X(t)$ for each $\mathbf{u} \in \mathcal{U}$ on T. This means that the general result on flow-invariance given in [1] (apud [6]) may be used in order to characterize the $T\mathcal{U}X$–constrained evolution. For the simplicity of writing of the next result, let us denote by $d(\mathbf{v}; V)$ the distance from $\mathbf{v} \in \mathbb{R}^n$ to the set $V \subset \mathbb{R}^n$.

Theorem 2.1 *The control system (2.1), with* $\mathbf{f}$ *continuous and locally Lipschitzian in* $\mathbf{x}$*, has a* $T\mathcal{U}X$*–constrained evolution if and only if:*

$$\liminf_{h \downarrow 0} h^{-1} d\left(\mathbf{v} + h\mathbf{f}(t, \mathbf{v}, \mathbf{u}(t)); X(t+h)\right) = 0 \quad \forall (t, \mathbf{u}, \mathbf{v}) \in T \times \mathcal{U} \times X. \tag{2.5}$$

This general result may be used both for analysis and design of control system (2.1) for previously given $\mathbf{f}$ and specifiable T, U and X. In this respect many possibilities to investigate the dynamics of control systems may be considered. As in other cases of general results that have in view the applications too, the *tangential condition* (2.5) must be appropriately converted into more transparent and easier to handle formulas. A possibility, involving a very simple calculation of the distance d, consists in the specialization of $X(t)$ in form of a *hyper-interval* in $\mathbb{R}^n$ which leads to an evaluation of system constrained evolution by its state components.

2.2. Componentwise Constrained Evolution, [15]

For a concise formulation of the next results, let us introduce some notations. Let $\mathbf{v} =: (v_i)$, $\mathbf{w} =: (w_i) \in \mathbb{R}^k$; $|\mathbf{v}| =: (|v_i|)$; $\mathbf{v} \leq \mathbf{w}$ ($\mathbf{v} <$

$\mathbf{w}$) or $\mathbf{v} \geq \mathbf{w}$ ($\mathbf{v} > \mathbf{w}$) mean the componentwise inequalities: $v_i \leq w_i$ ($v_i < w_i$) or $v_i \geq w_i$ ($v_i > w_i$), respectively. $V \subset \mathbb{R}^n$ is a compact set and $\mathbf{g}: V \to \mathbb{R}^n$, with $\mathbf{g} =: (g_i)$, is continuous, and $\mathbf{z} =: (z_i) \in V$ is fixed; $\mathcal{C}_v^z$ denotes the operator which "*catches*" $\mathbf{g}(\mathbf{v})$ at $\mathbf{z} \in V$ in a diagonal manner, i.e. $\mathcal{C}_v^z \{\mathbf{g}(\mathbf{v})\} =: [g_1(z_1, v_2, ..., v_n), ..., g_i(v_1, ..., z_i, ..., v_n), ..., g_n(v_1, ..., v_{n-1}, z_n)]^*$ ($[\,]^*$ means transposition); further, $\mathrm{ext}_V \mathcal{C}_v^z \{\mathbf{g}(\mathbf{v})\}$ denotes the vector with the components $\mathrm{ext}_V g_i(v_1, ..., z_i, ..., v_n)$, where ext may be min or max. Let $\underline{\mathbf{a}} : T \to \mathbb{R}^n$, $\overline{\mathbf{a}} : T \to \mathbb{R}^n$, $\underline{\mathbf{a}}(t) \leq \overline{\mathbf{a}}(t)$, be differentiable, and let $\underline{\mathbf{b}} : T \to \mathbb{R}^m$, $\overline{\mathbf{b}} : T \to \mathbb{R}^m$, $\underline{\mathbf{b}}(t) \leq \overline{\mathbf{b}}(t)$, be continuous. Usually both $\mathbf{x}$ and $\mathbf{u}$ are subject to certain prescribed constraints which in the most cases have in view their components. Thus, with the two pairs of functions the following time-dependent hyper-intervals may be respectively defined:

$$X(t) =: \{\mathbf{v} \in \mathbb{R}^n;\ \ \underline{\mathbf{a}}(t) \leq \mathbf{v} \leq \overline{\mathbf{a}}(t)\}\,, \quad t \in T, \tag{2.6}$$

$$U(t) =: \left\{\mathbf{w} \in \mathbb{R}^m;\ \ \underline{\mathbf{b}}(t) \leq \mathbf{w} \leq \overline{\mathbf{b}}(t)\right\}, \quad t \in T. \tag{2.7}$$

They are associated with the dynamical control system (2.1) in order to confer to its constrained evolution (due to physical, constructive and/or technological reasons) a more pragmatic meaning, namely by expressing it *componentwise*. Such an approach allows a more subtle and detailed evaluation of the dynamical behaviour instead of the global evaluation by means of the vector norm. This may be necessary especially when the state and control components are physically different and/or of different importance for the normal process evolution. At the same time, $X(t)$ (see (2.6)) allows an explicit conversion of the tangential condition (2.5), as it will be shown in the sequel. On the other hand, by this conversion the specific form of $U(t)$ (for instance (2.7)) does not play an essential role.

Theorem 2.2 *The control system (2.1), with* $\mathbf{f}$ *continuous and locally Lipschitzian in* $\mathbf{x}$*, has a* $T\,\mathcal{U}\,X$*–constrained evolution for* $X(t)$ *given by (2.6) and a compact* $U(t)$ *if and only if:*

$$\min_{T \times \mathcal{U} \times X} [\mathcal{C}_{\mathbf{v}}^{\underline{\mathbf{a}}} \{\mathbf{f}(t, \mathbf{v}, \mathbf{u}(t))\} - \dot{\underline{\mathbf{a}}}(t)] \geq 0, \tag{2.8}$$

$$\max_{T \times \mathcal{U} \times X} \left[\mathcal{C}_{\mathbf{v}}^{\overline{\mathbf{a}}} \{\mathbf{f}(t, \mathbf{v}, \mathbf{u}(t))\} - \dot{\overline{\mathbf{a}}}(t)\right] \leq 0. \tag{2.9}$$

The inequality form of (2.8) and (2.9) means that certain classes of systems associated with the compacts $X(t)$ (see (2.6)) and $U(t)$ have to be considered. As a matter of fact, for the given $X(t)$ (with given $\underline{\mathbf{a}}(t)$ and $\overline{\mathbf{a}}(t)$ in (2.6)) and $U(t)$ (e.g. (2.7) with given $\underline{\mathbf{b}}(t)$ and $\overline{\mathbf{b}}(t)$) inequalities (2.8) and (2.9) may provide, on the one hand, classes of solutions $\mathbf{f}(t, \mathbf{x}, \mathbf{u})$ and, on the other hand, open the approach of the *componentwise constrained evolution* (CCE) (and its robustness) as a practicable form of $T\,\mathcal{U}\,X$–constrained evolution.

The hyper-intervals (2.6) and (2.7) may be respectively replaced by:

$$X(t) =: \{\mathbf{v} \in \mathbb{R}^n; \underline{\mathbf{a}}(t) \le \mathbf{p}(\mathbf{v}) \le \overline{\mathbf{a}}(t)\},$$

$$U(t) =: \left\{\mathbf{w} \in \mathbb{R}^m; \underline{\mathbf{b}}(t) \le \mathbf{q}(\mathbf{w}) \le \overline{\mathbf{b}}(t)\right\}, t \in T,$$

where $\mathbf{p} : \mathbb{R}^n \to \mathbb{R}^n$ and $\mathbf{q} : \mathbb{R}^m \to \mathbb{R}^m$ are invertible and derivable and respectively invertible and continuous. Thus, the transformations $\tilde{\mathbf{x}} = \mathbf{p}(\mathbf{x})$, $\tilde{\mathbf{u}} = \mathbf{q}(\mathbf{u})$ translate the discussion in terms of the equivalent hyper-intervals $\tilde{X}(t)$ and $\tilde{U}(t)$ and of equivalent CCE for $\tilde{\mathbf{x}}$, $\tilde{\mathbf{u}}$.

2.3. Linear Time-Invariant Control Systems, [15]

To obtain more applicative formulas, consider the system:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \quad t \in \mathbb{R}_+, \; \mathbf{x} \in \mathbb{R}^n, \; \mathbf{u} \in \mathbb{R}^m, \tag{2.10}$$

where $\mathbf{A}$ and $\mathbf{B}$ are constant matrices of appropriate dimensions. $T = \mathbb{R}_+$ and for (2.6) and (2.7) the following conditions are respectively assumed:

$$-\underline{\mathbf{a}}(t) = \overline{\mathbf{a}}(t) =: \mathbf{a} > 0, \quad t \in \mathbb{R}_+, \tag{2.11}$$

$$-\underline{\mathbf{b}}(t) = \overline{\mathbf{b}}(t) =: \mathbf{b} > 0, \quad t \in \mathbb{R}_+. \tag{2.12}$$

Clearly, the control system (2.10) has in $\mathcal{U} \times X$ a symmetry point, namely $\mathbf{u} = 0$, $\mathbf{x} = 0$, and the *constant* CCE (i.e. $X(t)$ and $U(t)$ are time-invariant (constant); see (2.6), (2.7)) is simply expressed by:

$$|\mathbf{x}(t)| \le \mathbf{a}, \quad |\mathbf{u}(t)| \le \mathbf{b}, \quad t \in \mathbb{R}_+, \tag{2.13}$$

for each $\mathbf{x}(t_0)$ and each $\mathbf{u} \in \overline{C}^0$, both compatible with (2.13). Now conditions (2.8) and (2.9) are expressing a certain relation between $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{a}$, $\mathbf{b}$. To formulate it in a simple manner, supplementary notations are needed. $\mathbf{M} =: (m_{ij})$ being a real $(q \times r)$ matrix, then $|\mathbf{M}| =: (|m_{ij}|)$, and $\overline{\mathbf{M}} =: (\overline{m}_{ij})$ is the matrix with $\overline{m}_{ii} = m_{ii}$ and $\overline{m}_{ij} = |m_{ij}|$, $i \ne j$.

Theorem 2.3 *The linear control system (2.10) has a constant CCE if and only if:*

$$\overline{\mathbf{A}}\mathbf{a} + |\mathbf{B}|\,\mathbf{b} \le 0. \tag{2.14}$$

From (2.14) it is obvious that the constant CCE necessarily implies that all the diagonal elements of $\overline{\mathbf{A}}$ (and of $\mathbf{A}$ too) be strictly negative. This condition leads to the investigation of the case $\mathbf{u} = 0$, i.e. that one of the CCE of the free response of both control systems (2.1) and (2.10).

3. Stability Analysis Via CCE

The $T\,\mathcal{U}\,X$–constrained evolution may be used, under certain assumptions on $X(t)$ and $U(t)$, in order to derive some special type of stability results. For instance, the constant CCE, according to (2.13), defines the *componentwise bounded input – bounded state stability* and the inequality (2.14) is a necessary and sufficient condition for this kind of stability. Similar results may be obtained for the internal stability.

3.1. Componentwise Asymptotic Stability, [15]

Consider the free nonlinear dynamical system:

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}), \quad t \in \mathbb{R}_+, \ \mathbf{x} \in \mathbb{R}^n, \tag{3.1}$$

with the assumption $\mathbf{f}(t, 0) = 0$, $t \in \mathbb{R}_+$. Since $\mathbf{x} = 0$ is an equilibrium state, one may associate with system (3.1) the symmetric hyper-interval:

$$X^{\mathbf{h}}(t) =: \{\mathbf{v} \in \mathbb{R}^n;\ |\mathbf{v}| \leq \mathbf{h}(t)\}, \quad t \in \mathbb{R}_+, \tag{3.2}$$

where $\mathbf{h} : \mathbb{R}_+ \to \mathbb{R}^n$ is differentiable and $\mathbf{h}(t) > 0$, $t \in \mathbb{R}_+$. Concerning the asymptotic behaviour of $\mathbf{x}(t)$ the following condition plays an essential role:

$$\lim_{t \to \infty} \mathbf{h}(t) = 0. \tag{3.3}$$

Definition 3.1 The equilibrium state $\mathbf{x} = 0$ of system (3.1) is called *componentwise asymptotically stable* with respect to (w.r.t.) $\mathbf{h}(t)$ (CWAS$_\mathbf{h}$) if for each $t_0 \geq 0$ and for each $\mathbf{x}_0$, with $|\mathbf{x}_0| \leq \mathbf{h}(t_0)$, the system (3.1) satisfies:

$$|\mathbf{x}(t)| \leq \mathbf{h}(t), \quad t \geq t_0. \ \square \tag{3.4}$$

Notice that, under condition (3.3), the inequality (3.4) holds only if the equilibrium state $\mathbf{x} = 0$ is asymptotically stable, i.e. CWAS$_\mathbf{h}$ of $\mathbf{x} = 0$ implies its asymptotic stability. Inequality (3.4) does not evidence only certain bounds for $\mathbf{x}(t)$. CWAS$_\mathbf{h}$ actually belongs to the sphere of stability, as it is obvious from its definition. The main advantage of CWAS$_\mathbf{h}$ consists just in the componentwise evaluation, which is more subtle and detailed than the global evaluation (by means of a norm) involved by the asymptotic stability. An extension of Definition 3.1 referring to CWAS$_\mathbf{h}$ with $X^{\mathbf{h}}(t)$ of polyhedral type is considered and discussed in [61].

In general the asymptotic stability of $\mathbf{x} = 0$ does not imply CWAS$_\mathbf{h}$. And this is because in Definition 3.1 the condition "for each $\mathbf{x}_0$, with $|\mathbf{x}_0| \leq \mathbf{h}(t_0)$" (motivated also by the flow-invariance approach to be afterwards performed), corresponds in the standard definition of the

(Liapunov) stability, formulated in (ε, δ)–terms, to the special case $\varepsilon = \delta$. Remind that the stability (necessary to be fulfilled for the asymptotic stability) is defined, [14], as follows: "equilibrium state $\mathbf{x} = 0$ is called *stable* if for each $\varepsilon > 0$ there exists $\delta > 0$ such that $\|\mathbf{x}_0\| < \delta$ implies $\|\mathbf{x}(t)\| < \varepsilon, t \geq t_0$" ($\| \; \|$ denotes any norm in $\mathbb{R}^n$). Clearly, it follows that actually $\delta \leq \varepsilon$ and by setting $\delta = \varepsilon$ a special type of stability definition is considered, not necessarily satisfied by all the systems which fulfil the original definition. More details are given in [62].

The characterization of $\text{CWAS}_{\mathbf{h}}$ can be easily derived from Theorem 2.2 for $\mathbf{u}(t) = 0$ and $X^{\mathbf{h}}(t)$ given by (3.2).

Theorem 3.1 *The equilibrium state $\mathbf{x} = 0$ of system (3.1) is $CWAS_{\mathbf{h}}$ if and only if:*

$$\max_{t \geq 0, \, |\mathbf{v}| \leq \mathbf{h}(t)} \left[\mathcal{C}_{\mathbf{v}}^{\pm \mathbf{h}(t)} \left\{ \pm \mathbf{f}(t, \mathbf{v}) \right\} - \dot{\mathbf{h}}(t) \right] \leq 0. \tag{3.5}$$

Inequality (3.5) is a sufficient condition for the asymptotic stability of $\mathbf{x} = 0$ and the set $X^{\mathbf{h}}_{\text{AS}} =: \left\{ \mathbf{v} \in \mathbb{R}^n; \; |\mathbf{v}| \leq \max_{\mathbb{R}_+} \mathbf{h}(t) \right\}$ is one of its asymptotic stability regions. The advantage of $\text{CWAS}_{\mathbf{h}}$ previously evoked is reinforced now by (3.5), which is a necessary and sufficient condition for it.

Related to $X^{\mathbf{h}}_{\text{AS}}$ one can state now the question of some kind of *global* $\text{CWAS}_{\mathbf{h}}$. A possibility of its consistent definition is the following.

Definition 3.2 Replace $\mathbf{h}(t)$ in Definition 3.2 with $\rho \mathbf{h}(t)$, $\rho \geq 1$. The equilibrium state $\mathbf{x} = 0$ of system (3.1) is *globally* $\text{CWAS}_{\mathbf{h}}$ (or simply the system (3.1) is $\text{CWAS}_{\mathbf{h}}$) if $\mathbf{x} = 0$ is $\text{CWAS}_{\mathbf{h}}$ for all $\rho \geq 1$. □

Theorem 3.2 *The system (3.1) is $CWAS_{\mathbf{h}}$ if and only if:*

$$\max_{t \geq 0, \; |\mathbf{v}| \leq \mathbf{h}(t), \; \rho \geq 1} \left[\frac{1}{\rho} \mathcal{C}_{\mathbf{v}}^{\pm \mathbf{h}(t)} \left\{ \pm \mathbf{f}(t, \rho \mathbf{v}) \right\} - \dot{\mathbf{h}}(t) \right] \leq 0. \tag{3.6}$$

This is also a sufficient condition for the asymptotic stability in the large.

In order to instrument Theorem 3.2 and to open a way towards some more significant and concrete evaluations of type (3.4), [10], [11], let us consider now the free evolution of linear system (2.10), described by:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x}, \quad t \in \mathbb{R}_+, \; \mathbf{x} \in \mathbb{R}^n. \tag{3.7}$$

Theorem 3.3 *The linear system (3.7) is $CWAS_{\mathbf{h}}$ if and only if:*

$$\max_{\mathbb{R}_+} \left[\overline{\mathbf{A}} \mathbf{h}(t) - \dot{\mathbf{h}}(t) \right] \leq 0. \tag{3.8}$$

The solvability of $\overline{\mathbf{A}} \mathbf{h}(t) - \dot{\mathbf{h}}(t) \leq 0$ w.r.t. $\mathbf{h}(t)$ leads to the following.

Theorem 3.4 *The linear system (3.7) is CWAS*$_{\mathbf{h}}$ *if and only if:*

$$\mathbf{h}(t) \geq e^{\overline{\mathbf{A}}(t-t_0)}\mathbf{h}(t_0), \text{ for each } t_0 \geq 0 \text{ and for each } t \geq t_0. \qquad (3.9)$$

Theorem 3.5 *A necessary and sufficient condition for the existence of* $\mathbf{h}(t)$ *such that the linear system (3.7) be CWAS*$_{\mathbf{h}}$ *is that* $\overline{\mathbf{A}}$ *be a Hurwitzian matrix.*

Let H be the Abelian semigroup of the solutions of (3.8) or (3.9) with $\overline{\mathbf{A}}$ Hurwitzian. Linear system (3.7) is CWAS$_{\mathbf{h}}$ for each $\mathbf{h} \in H$, and for each $\mathbf{h}_1, \mathbf{h}_2 \in H$, CWAS$_{\mathbf{h}_1}$ is equivalent to CWAS$_{\mathbf{h}_2}$. In this context and taking into account the right hand term in (3.9), one has to investigate if H may contain (or be endowed with) simpler exponential functions.

3.2. Componentwise exponential asymptotic stability, [10], [11], [15]

Consider

$$\mathbf{h}(t) = \mathbf{d}e^{-\beta(t-t_0)}, \quad t \geq t_0 \geq 0. \qquad (3.10)$$

where $\mathbf{d} =: [d_1\, d_2 ... d_n]^* > 0$ and $\beta > 0$ (scalar).

Definition 3.3 The equilibrium state $\mathbf{x}=0$ of system (3.1) is called *componentwise exponential asymptotically stable* (CWEAS) if there exist $\mathbf{d} > 0$ and $\beta > 0$ such that for each $t_0 \geq 0$ and for each $|\mathbf{x}_0| \leq \mathbf{d}$ the system (3.1) satisfies:

$$|\mathbf{x}(t)| \leq \mathbf{d}e^{-\beta(t-t_0)}, \quad t \geq t_0. \ \square \qquad (3.11)$$

Definition 3.4 Replace $\mathbf{d}$ in Definition 3.3 with $\rho\mathbf{d}$, $\rho \geq 1$. The equilibrium state $\mathbf{x}=0$ of system (3.1) is called *globally* CWEAS (or simply the system (3.1) is CWEAS) if $\mathbf{x}=0$ is CWEAS for all $\rho \geq 1$. $\square$

Theorems 3.1 and 3.2, with (3.10), lead to the following results.

Theorem 3.6 *The equilibrium state* $\mathbf{x} = 0$ *of system (3.1) is CWEAS if and only if:*

$$\max_{t\geq 0,\, |\mathbf{v}|\leq \mathbf{d}} \left[e^{\beta t}\mathcal{C}_{\mathbf{v}}^{\pm\mathbf{d}} \left\{ \pm\mathbf{f}(t, \mathbf{v}e^{-\beta t}) \right\} \right] \leq -\beta\mathbf{d}. \qquad (3.12)$$

Theorem 3.7 *The system (3.1) is CWEAS if and only if:*

$$\max_{t\geq 0,\, |\mathbf{v}|\leq \mathbf{d},\, \rho\geq 1} \left[\frac{1}{\rho} e^{\beta t}\mathcal{C}_{\mathbf{v}}^{\pm\mathbf{d}} \left\{ \pm\mathbf{f}(t, \rho\mathbf{v}e^{-\beta t}) \right\} \right] \leq -\beta\mathbf{d}. \qquad (3.13)$$

In order to derive some results for the linear system (3.7), pertinent notations are needed, namely:

$$\mathbf{A}_{\mathbf{d}} =: \operatorname{diag}\{1/d_1,\, 1/d_2, ..., 1/d_n\}\mathbf{A}\operatorname{diag}\{d_1, d_2, ..., d_n\};$$

$$G_i(\mathbf{A_d}) =: \left\{ s\mathbb{C};\ |s - a_{ii}| \leq \frac{1}{d_i} \sum_{j=1,\ldots,n, j\neq i} |a_{ij}|\, d_j \right\}, \quad i = 1, ..., n,$$

are the **d**-Gershgorin's discs associated with **A** ($\mathbb{C}$ is the set of complex numbers); $\overline{A}_k$, $k = 1, ..., n$, are the leading principal minors of matrix $\overline{\mathbf{A}}$; for any matrix $\mathbf{M}$ the inequality $\mathbf{M} > 0$ means $m_{ij} > 0$ for all i and j.

Theorem 3.8, [10], [11] *The next statements are equivalent:*
1° *Linear system (3.7) is CWEAS.*
2° $\overline{\mathbf{A}}\mathbf{d} \leq -\beta\mathbf{d}$.
3° $0 < \beta \leq \min_i \left(-a_{ii} - \frac{1}{d_i} \sum_{j=1,\ldots,n,\, j\neq i} |a_{ij}| d_j \right)$.
4° $\overline{\mathbf{A}}\mathbf{d} < 0$.
5° $-\overline{\mathbf{A}}$ *is an M-matrix.*
6° $\overline{\mathbf{A}}$ *is Hurwitzian.*
7° $\cup_{i=1,\ldots,n} G_i(\mathbf{A_d}) \subset \{s\mathbb{C};\ \Re e\, s < 0\}$.
8° $(-1)^k \overline{A}_k > 0$, $k = 1, ..., n$.
9° $\det \mathbf{A} \neq 0$, $\left(-\overline{\mathbf{A}}\right)^{-1} \geq 0$.

For linear systems, $\text{CWAS}_\mathbf{h}$ is equivalent to CWEAS and this is a special type of asymptotic stability. CWEAS depends on the vector basis in $\mathbb{R}^n$ (see Theorem 3.8-4°) and there exist vector bases in which a given asymptotically stable linear system is CWEAS. It is a row property of **A** in the sense of a certain dominance of the first diagonal elements along the corresponding rows, necessarily implying that the first diagonal elements are strictly negative (see Theorem 3.8-3°). Using definitions that specialize Definition 2.1 for nonsymmetrical hyper-intervals, detailed results for CWEAS of continuous-time delay linear systems and for 1D and 2D linear discrete-time systems have been reported in [30], [31] and [32].

3.3. Componentwise absolute stability, [15]

The inequality form of condition (3.12) and Theorem 3.8 suggest a special approach for the following class of nonlinear systems:

$$\dot{\mathbf{x}} = \mathbf{F}(t, \mathbf{x})\mathbf{x}, \quad t \in \mathbb{R}_+,\ \mathbf{x} \in \mathbb{R}^n, \tag{3.14}$$

which are encountered in engineering (transistor circuits, electrical nets), economics, biology, ecology, pharmacokinetics etc. For system (3.14) the *asymptotic connective stability* via Liapunov direct method and based on a direct elementwise boundedness of $\mathbf{F}(t, \mathbf{x})$ has been studied in [63] and [64]. Unlike this study the CWEAS based approach allows to consider a class of $(n \times n)$ continuous matrices $\mathbf{F}(t, \mathbf{x})$ that are bounded in the following sense: for a given real constant $(n \times n)$ matrix $\mathbf{A}$ there exist $\mathbf{d} > 0$, $(\mathbf{d} \in \mathbb{R}^n)$ and $\beta > 0$ (scalar) such that the next elementwise inequality holds:

$$\mathcal{C}_{\mathbf{v}}^{\pm\mathbf{d}} \left\{ \overline{F}(t, \rho \mathbf{v} e^{-\beta t}) \right\} \leq \overline{\mathbf{A}}, \quad t \geq 0, \ |\mathbf{v}| \leq \mathbf{d}, \ \rho \geq 1, \tag{3.15}$$

where $\mathcal{C}_v^{\pm\mathbf{d}}$ is to apply to each column of $\overline{\mathbf{F}}(t, \mathbf{x})$. Clearly, there exists a nonempty class $\mathcal{F}_{\overline{\mathbf{A}}}$ of continuous matrices $\mathbf{F}(t, \mathbf{x})$, which satisfy (3.15). Accordingly, the linear control system (3.7) is called the *linear elementwise $\mathcal{C}$-majorant* (LE$\mathcal{C}$M) of systems (3.14) with $\mathbf{F} \in \mathcal{F}_{\overline{\mathbf{A}}}$.

Definition 3.5 The nonlinear system (3.14) is called *componentwise absolutely stable* (CAS) if it is CWEAS for all $\mathbf{F} \in \mathcal{F}_{\overline{\mathbf{A}}}$. □

Theorem 3.9 *The nonlinear system (3.14), with $\mathbf{F} \in \mathcal{F}_{\overline{\mathbf{A}}}$, is CAS if and only if its LE$\mathcal{C}$M (3.7) is CWEAS ($\overline{\mathbf{A}}$ is Hurwitzian).*

3.4. Robustness analysis of CWEAS, [22], [23]

Let us explore the preservation of the componentwise asymptotic stability for general classes of perturbations affecting the dynamics of the linear system (3.7). To this goal, let us get a deeper insight into the mutual relationship between the compatibility of inequality system:

$$\overline{\mathbf{A}}\mathbf{d} \leq -\beta \mathbf{d}, \quad \mathbf{d} \in \mathbb{R}^n, \ \mathbf{d} > 0, \ \beta > 0, \tag{3.16}$$

and the spectrum location of matrix $\overline{\mathbf{A}}$ which are referred to in Theorem 3.8, item (2°) and (6°), respectively.

Theorem 3.10, [21] *Given an arbitrary square matrix $\mathbf{A} \in \mathbb{R}^{n \times n}$, denote by $\lambda_i(\overline{\mathbf{A}})$, $i = 1,\ldots,n$, the eigenvalues of matrix $\overline{\mathbf{A}}$.*
a) Matrix $\overline{\mathbf{A}}$ has a real eigenvalue (simple or multiple), denoted by $\lambda_{\max}(\overline{\mathbf{A}})$, which fulfils the dominance condition

$$\Re e[\lambda_i(\overline{\mathbf{A}})] \leq \lambda_{\max}(\overline{\mathbf{A}}), \quad i = 1, \ldots, n. \tag{3.17}$$

b) The system of inequalities (3.16) is compatible (has a solution $\mathbf{d} > 0$) if and only if

$$\lambda_{\max}(\overline{\mathbf{A}}) \leq -\beta. \tag{3.18}$$

Consider the situation when system (3.7) exhibits a perturbed dynamics, described by:

$$\dot{\mathbf{x}}(t) = (\mathbf{A} + \mathbf{B}\Delta\mathbf{C})\mathbf{x}(t), \quad \mathbf{x}(t_0) = \mathbf{x}_0, \tag{3.19}$$

where $\mathbf{A} \in \mathbb{R}^{n \times n}$ is the same matrix as used in (3.7), $\mathbf{B} \in \mathbb{R}^{n \times m}$, $\mathbf{C} \in \mathbb{R}^{p \times n}$ and $\Delta \in \mathbb{R}^{m \times p}$ is a componentwise bounded matrix:

$$-\mathbf{D} \leq \Delta \leq \mathbf{D}, \ \mathbf{D} \geq 0. \tag{3.20}$$

Denote by $|\mathbf{B}| \in \mathbb{R}^{n\times m}$, $|\mathbf{C}| \in \mathbb{R}^{p\times n}$ the matrices obtained from $\mathbf{B}$ and $\mathbf{C}$, respectively, by replacing the original entries with their absolute values. Our robustness analysis relies on the following main result.

Theorem 3.11 *Assume system (3.7) is CWEAS and the spectrum of* $\overline{\mathbf{A}}$ *is dominated by* $\lambda_{\max}(\overline{\mathbf{A}})$ *in the sense formulated by Theorem 3.10. For an arbitrary* $\sigma \in \mathbb{R}$, *if* $\lambda_{\max}(\overline{\mathbf{A}}) < \sigma$, *then:*
a) $|\mathbf{C}| \cdot (\sigma \mathbf{I}_n - \overline{\mathbf{A}})^{-1} \cdot |\mathbf{B}| \cdot \mathbf{D}$ *is a nonnegative matrix;*
b) $\lambda_{\max}\left(|\mathbf{C}| \cdot (\sigma \mathbf{I}_n - \overline{\mathbf{A}})^{-1} \cdot |\mathbf{B}| \cdot \mathbf{D}\right) < 1$ *ensures* $\lambda_{\max}\left(\overline{\mathbf{A}+\mathbf{BC}}\right) < \sigma$ *for the perturbed system (3.19).*

Using the robustness property of eigenvalue location, let us first deal with the case when *unstructured perturbations* affect the dynamics of the linear continuous-time system. The preservation of componentwise asymptotic stability under this class of disturbances can be addressed within the framework built in Theorem 3.11, by searching an upper bound for $\|\mathbf{D}\|$, where $\| \;\; \|$ denotes an arbitrary matrix norm.

Theorem 3.12 *Assume system (3.7) is CWEAS and the spectrum of* $\overline{\mathbf{A}}$ *is dominated by* $\lambda_{\max}(\overline{\mathbf{A}})$ *in the sense formulated by Theorem 3.10. If* $\lambda_{\max}(\overline{\mathbf{A}}) < \sigma \leq 0$ *and if, for an arbitrary matrix norm* $\| \;\; \|$, *matrix* $\mathbf{D}$ *in (3.20) meets the condition:*

$$\|\mathbf{D}\| < 1/\left\| |\mathbf{C}| \cdot \left(\sigma \mathbf{I}_n - \overline{\mathbf{A}}\right)^{-1} \cdot |\mathbf{B}| \right\|, \tag{3.21}$$

then the perturbed system (3.19) is CWEAS for any $\beta > 0$, *with* $\lambda_{\max}(\overline{\mathbf{A}}) < -\beta < \sigma$ *and adequate* $\mathbf{d} > 0$ *satisfying* $(\overline{\mathbf{A}} + |\mathbf{B}| \cdot \mathbf{D} \cdot |\mathbf{C}|)\mathbf{d} \leq -\beta \mathbf{d}$.

Now, let us focus on the case when *structured perturbations* affect the dynamics of the linear continuous-time system (3.7). Assume that matrix $\mathbf{D}$ used in (3.20) for bounding the perturbation matrix Δ can be expressed in the following particular form:

$$\mathbf{D} = \Omega\omega, \quad \Omega \in \mathbb{R}^{m\times p}, \; 0 \leq \Omega, \; 0 < \omega, \tag{3.22}$$

where Ω is a known matrix, specifying the actions of a given class of structured perturbations. The preservation of componentwise asymptotic stability under this type of disturbances will be approached within the framework built in Theorem 3.11, by searching an upper bound for the positive constant ω occurring in expression (3.22).

Theorem 3.13 *Assume system (3.7) is CWEAS and the spectrum of* $\overline{\mathbf{A}}$ *is dominated by* $\lambda_{\max}(\overline{\mathbf{A}})$ *in the sense formulated by Theorem 3.11. If* $\lambda_{\max}(\overline{\mathbf{A}}) < \sigma \leq 0$ *and if* ω *in (3.22) meets the condition:*

$$\omega < 1/\lambda_{\max}\left(|\mathbf{C}| \cdot (\sigma \mathbf{I}_n - \overline{\mathbf{A}})^{-1} \cdot |\mathbf{B}| \cdot \Omega\right), \tag{3.23}$$

then the perturbed system (3.19) is CWEAS for any $\beta > 0$, *with* $\lambda_{\max}(\overline{\mathbf{A}}) < -\beta < \sigma$ *and adequate* $\mathbf{d} > 0$ *satisfying* $(\overline{\mathbf{A}} + |\mathbf{B}| \cdot \mathbf{D} \cdot |\mathbf{C}|)\mathbf{d} \le -\beta\mathbf{d}$.

Theorems 3.12 and 3.13 can also be regarded as providing conditions for the preservation of the componentwise asymptotic stability, whenever system (3.7) is used in a *closed-loop architecture*, with proportional output (or state) feedback, whose gain factors are constrained according to (3.20) or (3.22).

4. Free response analysis of interval matrix systems, [21], [29]

Consider the free response of an *interval matrix system* (IMS) described by the equation:

$$\dot{\mathbf{x}}(t) = \mathbf{A}^I\mathbf{x}(t); \quad \mathbf{x}(t_0) = \mathbf{x}_0; \quad t, t_0 \in T,\ t \ge t_0, \tag{4.1}$$

where the *interval matrix* $\mathbf{A}^I$ is defined as:

$$\mathbf{A}^I = \{\mathbf{A} \in \mathbb{R}^{n\times n} : \mathbf{A}^- \le \mathbf{A} \le \mathbf{A}^+\}. \tag{4.2}$$

Relying on the background laid out in sections 2 and 3, we are now interested to investigate the flow-invariance, $\mathrm{CWAS}_{\mathbf{h}}$ and CWEAS properties of IMS (4.1).

4.1. Flow-invariance of time-dependent rectangular sets

Consider IMS (4.1) and let $X^{\mathbf{h}}(t)$ be the *time-dependent rectangular set* (TDRS) defined by (3.2). For the beginning, assume that there exists no requirement of type (3.3) for vector function $\mathbf{h}(t)$, referring to its behaviour to infinity.

According to CCE concepts for the particular case of null inputs, TDRS (3.2) is *flow-invariant* (FI) w.r.t. IMS (4.1) if for any $t_0 \in T$ and any initial condition $\mathbf{x}(t_0) = x_0 \in X^{\mathbf{h}}(t_0)$, the whole state trajectory $\mathbf{x}(t)$ corresponding to $\mathbf{x}(t_0)$ remains inside the set $X^{\mathbf{h}}(t)$, i.e.

$$\forall t_0,\ \ t \in T,\ t \ge t_0,\ \ \forall \mathbf{x}(t_0) = \mathbf{x}_0 \in X^{\mathbf{h}}(t_0) : \mathbf{x}(t) \in X^{\mathbf{h}}(t). \tag{4.3}$$

This property is formulated in terms of the state-space trajectories of IMS (4.1), and our interest now focuses on a characterization based on the interval matrix $\mathbf{A}^I$ (4.2).

Theorem 4.1 *TDRS* $X^{\mathbf{h}}(t)$ *(3.2) is FI w.r.t. IMS (4.1) if and only if*

$$\forall t \in T : \dot{\mathbf{h}}(t) \ge \overline{\mathbf{A}}\mathbf{h}(t), \tag{4.4}$$

where $\overline{\mathbf{A}}$ *denotes a constant matrix constructed from the interval matrix* $\mathbf{A}^I$ *(4.2) as follows:*

$$\overline{a}_{ii} = \sup_{\mathbf{A}\in\mathbf{A}^I} \{a_{ii}\} = a_{ii}^+, \quad i = 1, ..., n, \tag{4.5–a}$$

$$\overline{a}_{ij} = \sup_{\mathbf{A}\in\mathbf{A}^I} \{|a_{ij}|\} = \max\{|a_{ij}^-|, |a_{ij}^+|\}, \quad i \neq j, \; i, j = 1, ..., n. \tag{4.5–b}$$

Matrix $\overline{\mathbf{A}}$ built according to (4.5) is *essentially nonnegative* (i.e. the off-diagonal entries are nonnegative) and it can be uniquely expressed as a sum of two real matrices:

$$\overline{\mathbf{A}} = \overline{\mathbf{A}}_D + \overline{\mathbf{A}}_E, \tag{4.6}$$

where $\overline{\mathbf{A}}_D$ is a diagonal matrix containing the diagonal entries of $\overline{\mathbf{A}}$, and $\overline{\mathbf{A}}_E \geq 0$ is a nonnegative matrix with zero diagonal entries, containing the off-diagonal entries of $\overline{\mathbf{A}}$.

In general, matrix $\overline{\mathbf{A}}$ does not belong to the interval matrix $\mathbf{A}^I$ (4.2), and for any $\mathbf{A} \in \mathbf{A}^I$ one can write the following matrix inequalities:

$$\forall \mathbf{A} = \mathbf{A}_D + \mathbf{A}_E \in \mathbf{A}^I : \; \mathbf{A}_D \leq \overline{\mathbf{A}}_D, \; |\mathbf{A}_E| \leq \overline{\mathbf{A}}_E, \tag{4.7}$$

where $\mathbf{A}_D$ and $\mathbf{A}_E$ are two real matrices containing the diagonal and off-diagonal entries of $\mathbf{A}$, respectively.

Inequalities presented in (4.7) show that matrix $\overline{\mathbf{A}}$ *dominates a symmetric convex set* of matrices $\mathbf{M}^C$, including the interval matrix $\mathbf{A}^I$ (4.2), which is defined as follows:

$$\mathbf{M}^C = \left\{\mathbf{M} =: \mathbf{M}_D + \mathbf{M}_E \in \mathbb{R}^{n\times n} : \; \mathbf{M}_D \leq \overline{\mathbf{A}}_D, \; |\mathbf{M}_E| \leq \overline{\mathbf{A}}_E\right\}, \tag{4.8}$$

where $\mathbf{M}_D$ and $\mathbf{M}_E$ are two real matrices containing the diagonal and off-diagonal entries of $\mathbf{M}$, respectively. In the space of coefficients, symmetry with respect to the space origin refers only to the extra-diagonal elements.

Based on this domination of matrix $\overline{\mathbf{A}}$, we will also say that the linear constant system:

$$\dot{\mathbf{x}}(t) = \overline{\mathbf{A}}\mathbf{x}(t), \quad \mathbf{x}(t_0) = \mathbf{x}_0; \; t, t_0 \in T, \; t \geq t_0, \tag{4.9}$$

dominates any system belonging to the family of systems generating IMS (4.1). Moreover, one can simply notice that linear constant system (4.9) dominates the whole family of systems generated with matrices belonging to the convex set $\mathbf{M}^C$ defined by (4.8). This fact should be interpreted as a key point in model construction when dealing with flow invariance, because $\overline{\mathbf{A}}$ ultimately characterizes an IMS with symmetric uncertainties in the sense formulated above, even if interval matrix $\mathbf{A}^I$ (4.2) does not exhibit such a symmetry.

Although system (4.9) is not necessarily a member of the family of systems that defines IMS (4.1), the state-transition matrix of constant system (4.9)

$$\overline{\Phi}(t_0, t) = e^{\overline{\mathbf{A}}(t-t_0)}; \quad t, t_0 \in T,\ t \geq t_0, \tag{4.10}$$

plays an important role in characterizing the flow-invariance with respect to IMS (4.1), as revealed by the next result.

Theorem 4.2 *TDRS $X^{\mathbf{h}}(t)$ (3.2) is FI w.r.t. IMS (4.1) if and only if*

$$\forall\, t_0, t \in T, \quad t \geq t_0 : \overline{\Phi}(t_0, t)\mathbf{h}(t_0) \leq \mathbf{h}(t). \tag{4.11}$$

A direct consequence of Theorems 4.1 or 4.2 is the fact that the flow-invariance of the set $X^{\mathbf{h}}(t)$ (3.2) with respect to IMS (4.1) guarantees the existence of this property for the whole *class* of rectangular sets $\tilde{X}^{\mathbf{h}}(t)$ *homothetic* to $X^{\mathbf{h}}(t)$, the time-dependence of which is defined by the vector functions:

$$\tilde{\mathbf{h}}(t) = c\mathbf{h}(t), \quad c > 0,\ t \in T. \tag{4.12}$$

4.2. $\mathbf{CWAS_h}$ of an IMS

Consider TDRS $X^{\mathbf{h}}(t)$ (3.2) with vector function $\mathbf{h}(t)$ meeting the supplementary condition (3.3) that refers to the behaviour to infinity. Thus, we can approach the $\mathrm{CWAS}_{\mathbf{h}}$ for IMS (4.1).

Theorem 4.3 *IMS (4.1) is $CWAS_{\mathbf{h}}$ if and only if the constant matrix $\overline{\mathbf{A}}$ built from the interval matrix $\mathbf{A}^I$ (4.2) according to (4.5) is Hurwitz stable.*

Now, once we have got this necessary and sufficient condition, we are interested to explore its algebraic background in terms of *matrix stability*. For the proposed development, the result formulated below plays a crucial role.

Theorem 4.4 *Let $\overline{\mathbf{A}}$ be a constant matrix built according to (4.5) from the interval matrix $\mathbf{A}^I$ (4.2). The eigenvalue $\lambda_{\max}(\overline{\mathbf{A}})$, which dominates the spectrum of $\overline{\mathbf{A}}$ in the sense formulated by Theorem 3.10, also dominates the spectrum of any arbitrary matrix $\mathbf{A} \in \mathbf{A}^I$ in the same sense, as follows:*

$$\forall\, \mathbf{A} \in \mathbf{A}^I : \Re e[\lambda_i(\mathbf{A})] \leq \lambda_{\max}(\overline{\mathbf{A}}), \quad i = 1, ..., n. \tag{4.13}$$

It is worth also noticing that the domination of the eigenvalue $\lambda_{\max}(\overline{\mathbf{A}})$ remains valid in the sense formulated by Theorem 4.4 for the whole symmetric convex set of matrices $\mathbf{M}^C$ defined by inequalities (4.8).

Now, resuming our discussion on the stability of matrix $\overline{\mathbf{A}}$ versus the stability of the interval matrix $\mathbf{A}^I$, we can simply formulate an immediate consequence of Theorem 4.4.

Corollary 4.1 *Let $\overline{\mathbf{A}}$ be a constant matrix built according to (4.5) from the interval matrix $\mathbf{A}^I$ (4.2). If $\overline{\mathbf{A}}$ is Hurwitz stable, then $\mathbf{A}^I$ is Hurwitz stable.*

The result formulated in Corollary 4.1 can be also found in [65]. However, the noticeable achievement of our work consists in revealing the *complete meaning* of the algebraic condition of stability for matrix $\overline{\mathbf{A}}$, based on the stronger concept of componentwise asymptotic stability. Thus, the stability of matrix $\overline{\mathbf{A}}$ operates only as a *sufficient criterion* for the asymptotic stability of IMS (4.1) – result already known, whereas it represents a *necessary and sufficient condition* for the componentwise asymptotic stability of IMS (4.1) – new result. Moreover, for $\overline{\mathbf{A}}$ stable, our recent insight demonstrates that the usage of the dominant eigenvalue $\lambda_{\max}(\overline{\mathbf{A}})$ as a stability margin for IMS (4.1) (proposed in [66] for standard asymptotic stability) fully characterizes only the situation when the margin refers to the componentwise asymptotic stability of IMS (4.1).

In the light of these comments, if IMS (4.1) is $\text{CWAS}_{\mathbf{h}}$, the dominant eigenvalue $\lambda_{\max}(\overline{\mathbf{A}})$ is expected to provide supplementary information with regard to the time-dependent evolution of those TDRSs $X^{\mathbf{h}}(t)$ (3.2) & (3.3) that are FI w.r.t. IMS (4.1).

Theorem 4.5 *Let IMS (4.1) be $\text{CWAS}_{\mathbf{h}}$ and consider a set $X^{\mathbf{h}}(t)$ (3.2) & (3.3) which is FI w.r.t. IMS (4.1). For any two arbitrary time instants t_0, $t \in T$, $t > t_0$, for which $\mathbf{h}(t) < \mathbf{h}(t_0)$, define the decreasing rate of $X^{\mathbf{h}}(t)$ between t_0 and t, denoted by $d(t_0, t)$, as the minimal subunitary value ensuring the fulfilment of the following inequality:*

$$\mathbf{h}(t) \leq d(t_0, t)(t_0). \tag{4.14}$$

The decreasing rate of $X^{\mathbf{h}}(t)$ has a lower bound that can be expressed as a time-dependent exponential function of parameter $\lambda_{\max}(\overline{\mathbf{A}})$:

$$e^{\lambda_{\max}(\overline{\mathbf{A}})(t-t_0)} \leq d(t_0, t). \tag{4.15}$$

Although standard asymptotic stability of IMS (4.1) represents only a necessary condition for the componentwise asymptotic stability, one can identify several classes of IMSs for which this condition is also sufficient. In this paragraph we restraint our presentation to three classes of IMSs, whose particular structures correspond to special types of interval matrices $\mathbf{A}^I$ (4.2), where the stability of $\mathbf{A}^I$ is equivalent to the stability of $\overline{\mathbf{A}}$.

(i) Assume that all the extra-diagonal entries of $\mathbf{A}^I$ (4.2) are nonnegative. IMS (4.1) is $\text{CWAS}_{\mathbf{h}}$ if and only if it is asymptotically stable.

(ii) Assume that $\mathbf{A}^I$ (4.2) is either lower- or upper-triangular. IMS (4.1) is $\text{CWAS}_{\mathbf{h}}$ if and only if it is asymptotically stable.

(iii) Assume that $\mathbf{V} = \mathbf{V}_D + \mathbf{V}_E$ is an extreme vertex of the hyperrectangle described in $\mathbf{R}^{n\times n}$ by $\mathbf{A}^I$ (4.2), such that $\mathbf{V}_D = \overline{\mathbf{A}}_D$, $|\mathbf{V}_E| = \overline{\mathbf{A}}_E$, with $\overline{\mathbf{A}}_D$, $\overline{\mathbf{A}}_E$ defined by (4.6), i.e. $\mathbf{V}$ is also a vertex of the convex set of matrices $\mathbf{M}^C$ defined by (4.8). Assume that $\mathbf{V}_E$ or $-\mathbf{V}_E$ is a Morishima matrix, [65]. IMS (4.1) is $\text{CWAS}_{\mathbf{h}}$ if and only if it is asymptotically stable.

4.3. CWEAS of an IMS

Consider TDRS $X^{\mathbf{h}}(t)$ (3.2) with vector function $\mathbf{h}(t)$ meeting the additional condition (3.10), which allows us to address the CWEAS for IMS (4.1). Since condition (3.10) is more restrictive than condition (3.3) that was assumed in the previous paragraph for $\text{CWAS}_{\mathbf{h}}$ analysis, it is perfectly reasonable to explore the relationship between the $\text{CWAS}_{\mathbf{h}}$ and CWEAS of IMS (4.1).

Theorem 4.6 *IMS (4.1) is* $\text{CWAS}_{\mathbf{h}}$ *if and only if it is CWEAS.*

For testing the componentwise (exponential) asymptotic stability of IMS (4.1), besides the investigation of Hurwitz stability for constant matrix $\overline{\mathbf{A}}$, built according to (4.5) from the interval matrix $\mathbf{A}^I$ (4.2), one can apply the equivalent criteria presented by Theorem 3.8 in the previous section.

5. Free response analysis for a class of nonlinear uncertain systems, [24] – [28], [67]

Consider the class of *nonlinear uncertain systems* (NUSs) defined as:

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}),\ \mathbf{x} \in \mathbb{R}^n,\ \mathbf{x}(t_0) = \mathbf{x}_0,\ t \geq t_0;$$

$$f_i(\mathbf{x}) = \sum_{j=1}^{n} a_{ij} x_j^{p_{ij}}, \quad p_{ij} \in \mathbb{N}, i = 1, ..., n, \tag{5.1}$$

where the *interval-type coefficients*:

$$a_{ij}^- \leq a_{ij} \leq a_{ij}^+ \tag{5.2}$$

are chosen to cover the inherent errors which frequently affect the accuracy of model construction.

Relying on the background laid out in sections 2 and 3, we are now interested to investigate the flow-invariance, $\text{CWAS}_{\mathbf{h}}$ and CWEAS properties of NUS (5.1).

5.1. Flow-invariance of time-dependent rectangular sets

Consider NUS (5.1) and let $X^{\mathbf{h}}(t)$ be the *time-dependent rectangular set* (TDRS) defined by (3.2). For the beginning, assume that there exists no requirement of type (3.3) for vector function $\mathbf{h}(t) =: [h_1(t) \ldots h_n(t)]^*$ (* means transposition), referring to its behaviour to infinity.

According to CCE concepts for the particular case of null inputs, TDRS (3.2) is *flow-invariant* (FI) *with respect to* (w.r.t.) NUS (5.1) if for any $t_0 \in T$ and any initial condition $\mathbf{x}(t_0) = \mathbf{x}_0 \in X^{\mathbf{h}}(t_0)$, the whole state trajectory $\mathbf{x}(t)$ corresponding to $\mathbf{x}(t_0)$ remains inside the set $X^{\mathbf{h}}(t)$, i.e.

$$\forall t_0,\ t \in T,\ t \geq t_0,\ \forall \mathbf{x}(t_0) = \mathbf{x}_0 \in X^{\mathbf{h}}(t_0) : \mathbf{x}(t) \in X^{\mathbf{h}}(t). \qquad (5.3)$$

This property is formulated in terms of the state-space trajectories of NUS (5.1), and our interest now focuses on a characterization based on the interval-type coefficients (5.2).

Theorem 5.1 *TDRS $X^{\mathbf{h}}(t)$ (3.2) is FI w.r.t. NUS (5.1) if and only if the following inequalities hold for $t \in [t_0, \theta)$, $\theta > t_0$:*

$$\dot{\mathbf{h}}(t) \geq \overline{\mathbf{g}}(\mathbf{h});\quad \overline{g}_i(\mathbf{h}) = \sum_{j=1}^{n} \overline{c}_{ij} h_j^{p_{ij}},\quad i = 1, ..., n, \qquad (5.4-\text{a})$$

$$\dot{\mathbf{h}}(t) \geq \tilde{\mathbf{g}}(\mathbf{h});\quad \tilde{g}_i(\mathbf{h}) = \sum_{j=1}^{n} \tilde{c}_{ij} h_j^{p_{ij}},\quad i = 1, ..., n, \qquad (5.4-\text{b})$$

where $\overline{c}_{ij}$, $\tilde{c}_{ij}$ have unique values, derived from the interval-type coefficients a_{ij} of NUS (5.1) as follows:

$$\overline{c}_{ii} = a_{ii}^{+},\ \text{for } p_{ii} \text{ odd or even};\quad \underset{i \neq j}{\overline{c}_{ij}} = \begin{cases} \max\left\{\left|a_{ij}^{-}\right|, \left|a_{ij}^{+}\right|\right\}, & \text{if } p_{ij} \text{ odd} \\ \max\left\{0, a_{ij}^{+}\right\}, & \text{if } p_{ij} \text{ even} \end{cases}; \qquad (5.5-\text{a})$$

$$\tilde{c}_{ii} = \begin{cases} a_{ii}^{+}, & \text{if } p_{ii} \text{ odd} \\ -a_{ii}^{-}, & \text{if } p_{ii} \text{ even} \end{cases};\quad \underset{i \neq j}{\tilde{c}_{ij}} = \begin{cases} \max\left\{\left|a_{ij}^{-}\right|, \left|a_{ij}^{+}\right|\right\}, & \text{if } p_{ij} \text{ odd} \\ \max\left\{0, -a_{ij}^{-}\right\}, & \text{if } p_{ij} \text{ even} \end{cases}. \qquad (5.5-\text{b})$$

Theorem 5.2 *There exist TDRSs (3.2) which are FI w.r.t. NUS (5.1) if and only if there exist common positive solutions (PSs) for the following differential inequalities (DIs):*

$$\dot{\mathbf{y}} \geq \overline{\mathbf{g}}(\mathbf{y}) \qquad (5.6-\text{a})$$

$$\dot{\mathbf{y}} \geq \tilde{\mathbf{g}}(\mathbf{y}). \qquad (5.6-\text{b})$$

Theorem 5.3 *There exist TDRSs (3.2) which are FI w.r.t. NUS (5.1) if and only if there exist PSs for the following DI:*

$$\dot{\mathbf{y}} \geq \mathbf{g}(\mathbf{y});\ g_i(\mathbf{y}) = \max_{\mathbf{y} \in \mathbb{R}^n} \left\{\overline{g}_i(\mathbf{y}), \tilde{g}_i(\mathbf{y})\right\},\ i = 1, ..., n. \qquad (5.7)$$

Let us study the family of TDRSs that are FI w.r.t. a given NUS (5.1). We start with the qualitative exploration of the solution of the following *differential equation* (DE):

$$\dot{\mathbf{z}} = \mathbf{g}(\mathbf{z}), \tag{5.8}$$

which is obtained from DI (5.7) by replacing "≥" with "=".

Lemma 5.1 *DE (5.8) with arbitrary* t_0 *and arbitrary initial condition* $\mathbf{z}(t_0) = \mathbf{z}_0$ *has a unique solution* $\mathbf{z}(t) = \mathbf{z}(t; t_0, \mathbf{z}_0)$ *defined on* $[t_0, \theta)$*, for some* $\theta > t_0$*.*

Lemma 5.2 *For any* t_0 *and any positive initial condition* $\mathbf{z}(t_0) = \mathbf{z}_0 > 0$*, the unique solution* $\mathbf{z}(t) = \mathbf{z}(t; t_0, \mathbf{z}_0)$ *of DE (5.8) remains positive for its maximal interval of existence* $[t_0, \theta)$*.*

One can easily see that Lemma 5.2 guarantees the existence of PSs for DI (5.7) in the particular case when "≥" is replaced by "=". However DI (5.7) might have PSs that do not satisfy DE (5.8) and, therefore, we further establish a connection between the PSs of DI (5.7) and the PSs of DE (5.8).

Lemma 5.3 *Let* $\mathbf{y}(t) > 0$ *be an arbitrary PS of DI (5.7), with the maximal interval of existence* $[t_0, \theta)$*. Denote by* $\mathbf{z}(t)$ *an arbitrary PS of DE (5.8), corresponding to an initial condition* $\mathbf{z}(t_0)$ *which satisfies the componentwise inequality:*

$$0 < \mathbf{z}(t_0) \leq \mathbf{y}(t_0). \tag{5.9}$$

Denote by $\mathbf{z}^{\mathbf{y}_0}(t)$ *the unique PS of DE (5.8) corresponding to the initial condition of* $\mathbf{y}(t)$*, i.e.*

$$\mathbf{z}^{\mathbf{y}_0}(t_0) \equiv \mathbf{y}(t_0). \tag{5.10}$$

For $t \in [t_0, \theta)$ *the following inequalities hold:*

$$0 < \mathbf{z}(t) \leq \mathbf{z}^{\mathbf{y}_0}(t) \leq \mathbf{y}(t). \tag{5.11}$$

We are now able to formulate a comparison between three different types of TDRSs, FI w.r.t. NUS (5.1), which are built (according to Theorem 5.3) by the help of the PSs of DI (5.7).

Theorem 5.4 *If* $X^{\mathbf{y}}(t)$*,* $X^{\mathbf{y}_0}(t)$ *and* $X^{\mathbf{z}}(t)$ *denote three TDRSs, FI w.r.t. NUS (5.1), generated by the following three types of PSs of DI (4.7):* $\mathbf{y}(t)$ *– arbitrary PS of DI (5.7);* $\mathbf{z}^{\mathbf{y}_0}(t)$ *– unique PS of DE (5.8), with* $\mathbf{z}^{\mathbf{y}_0}(t_0) = \mathbf{y}(t_0)$*;* $\mathbf{z}(t)$ *– arbitrary PS of DE (5.8), with* $\mathbf{z}(t_0) \leq \mathbf{y}(t_0)$*, then:*

$$X^{\mathbf{z}}(t) \subseteq X^{\mathbf{y}_0}(t) \subseteq X^{\mathbf{y}}(t) \quad \forall\, t \in [t_0, \theta), \tag{5.12}$$

where $[t_0, \theta)$ *denotes the maximal interval of existence for* $X^{\mathbf{y}}(t)$*.*

Given a TDRS which is FI w.r.t. NUS (5.1), we can also formulate a condition for the existence of other TDRSs, *strictly included* in the former one, which are FI w.r.t. NUS (5.1) too.

Theorem 5.5 *Denote by $X^{\mathbf{y}}(t)$ a TDRS, FI w.r.t. NUS (5.1) for its maximal interval of existence $[t_0, \theta)$. If there exist n functions $\delta_i(t) \in C^1$, non-decreasing, positive and subunitary $0 < \delta_i(t) < 1$, $i = 1, ..., n$, such that*

$$\mathbf{g}\left(\Delta(t)\mathbf{y}(t)\right) \leq \Delta(t)\mathbf{g}\left(\mathbf{y}(t)\right); \quad \Delta(t) =: \mathrm{diag}\left\{\delta_i(t), ..., \delta_n(t)\right\}, \qquad (5.13)$$

then the TDRS $X^{\Delta y}(t)$, generated by the vector function $\Delta(t)\mathbf{y}(t)$ is also FI w.r.t. NUS (5.1) and

$$X^{\Delta\mathbf{y}}(t) \subset X^{\mathbf{y}}(t), \quad t \in [t_0, \theta). \qquad (5.14)$$

A great interest for practice presents those TDRSs, FI w.r.t. NUS (5.1), which are defined on $[t_0, \infty)$ and *remain bounded* for any $t \in [t_0, \infty)$. Therefore our next theorem deals with the case of *infinite-time horizon*, for which it gives a necessary condition, formulated directly in terms of interval-type coefficients a_{ii} and exponents p_{ii} in NUS (5.1).

Theorem 5.6 *For the existence of TDRSs, FI w.r.t. NUS (5.1), which are bounded on $[t_0, \infty)$, it is necessary (but not sufficient) that a_{ii} and p_{ii} of NUS (5.1) meet the following requirement, for i = 1,...,n:*

$$(p_{ii} \text{ odd}, \ a_{ii}^{+} \leq 0) \ \text{ OR } \ (p_{ii} \text{ even}, \ a_{ii}^{-} = a_{ii}^{+} = 0). \qquad (5.15)$$

5.2. CWAS$_{\mathbf{h}}$ of a NUS

Consider TDRS $X^{\mathbf{h}}(t)$ (3.2) with vector function $\mathbf{h}(t)$ meeting the supplementary condition (3.3) that refers to the behaviour to infinity. Thus, we can approach the CWAS$_{\mathbf{h}}$ for NUS (5.1).

Theorem 5.7 *Equilibrium state ES $\{0\}$ of NUS (5.1) is CWAS$_{\mathbf{h}}$ if and only if there exist common PSs $\mathbf{h}(t) > 0$ for DI (5.6-a) & (5.6-b), with $\lim_{t\to\infty} \mathbf{h}(t) = 0$.*

Theorem 5.8 *ES $\{0\}$ of NUS (5.1) is CWAS$_{\mathbf{h}}$ if and only if there exist PSs $\mathbf{h}(t) > 0$ for DI (5.7), with $\lim_{t\to\infty} \mathbf{h}(t) = 0$.*

As the boundedness of TDRSs on $[t_0, \infty)$ introduces some restrictions for the exponents p_{ii} and interval-type coefficients a_{ii} of NUS (5.1) (formulated in Theorem 5.6), more restrictive conditions are expected when replacing boundedness of vector function $\mathbf{h}(t)$ with the stronger requirement (3.3).

Theorem 5.9 *A necessary condition for ES* $\{0\}$ *of NUS (5.1) to be* $CWAS_{\mathbf{h}}$ *is*

$$p_{ii} \text{ odd}, \quad a_{ii}^{+} < 0, \; i = 1, ..., n. \tag{5.16}$$

We now resume our qualitative analysis of the solutions of DI (5.7) and DE (5.8), in order to allow a refined interpretation of the result stated in Theorem 5.8.

Lemma 5.4 *Let* p_{ii} *be odd and* $a_{ii}^{+} < 0$ *for all* $i = 1, ..., n$. *Consider an arbitrary PS* $\mathbf{y}(t) > 0$ *of DI (5.7), with its maximal interval of existence* $[t_0, \theta)$. *If* $\mathbf{z}(t)$ *denotes an arbitrary solution of DE (5.8) corresponding to the initial condition* $\mathbf{z}(t_0)$ *which satisfies:*

$$-\mathbf{y}(t_0) \leq \mathbf{z}(t_0) \leq \mathbf{y}(t_0), \tag{5.17}$$

then the following inequalities hold for $t \in [t_0, T)$:

$$-\mathbf{y}(t) \leq \mathbf{z}(t) \leq \mathbf{y}(t). \tag{5.18}$$

This lemma completes the picture on the topology of the solutions (not only positive) for DE (5.8) in the vicinity of $\{0\}$, fact which permits revealing the link between condition (3.3) and the nature of ES $\{0\}$ for DE (5.8).

Theorem 5.10 *ES* $\{0\}$ *of NUS (5.1) is* $CWAS_{\mathbf{h}}$ *if and only if ES* $\{0\}$ *of DE (5.8) is asymptotically stable.*

For practice, it might be rather difficult to handle DE (5.8) in order to check its asymptotic stability. A more attractive approach is to find just a sufficient condition for $\text{CWAS}_{\mathbf{h}}$, based on an operator with a more tractable form than $\mathbf{g}$ in DE (5.8).

Theorem 5.11 *Consider the DE:*

$$\dot{\mathbf{z}} = \hat{\mathbf{g}}(\mathbf{z}); \quad \hat{g}_i(\mathbf{z}) = \sum_{j=1}^{n} \hat{c}_{ij} z_j^{p_{ij}}, \quad i = 1, ..., n, \tag{5.19}$$

where the coefficients $\hat{c}_{ij}$ *are defined by:*

$$\hat{c}_{ij} = \max\{\overline{c}_{ij}, \tilde{c}_{ij}\} \quad i, j = 1, ..., n. \tag{5.20}$$

(i) If ES $\{0\}$ *is asymptotically stable for DE (5.19), then ES* $\{0\}$ *is* $CWAS_{\mathbf{h}}$ *for NUS (5.1).*

(ii) In the particular case when the interval-type coefficients a_{ij} *of NUS (5.1) satisfy the inequalities given below, for each* i, $i = 1, ..., n$:

$$\text{IF } \underset{i \neq j}{p_{ij}} \text{ even THEN } (a_{ij}^{+} \geq -a_{ij}^{-} \text{ for all } j \text{ OR } a_{ij}^{+} \leq -a_{ij}^{-} \text{ for all } j), \tag{5.21}$$

the sufficient condition stated at (i) is also necessary for the ES $\{0\}$ *of NUS (5.1) to be* $CWAS_{\mathbf{h}}$.

5.3. CWEAS of a NUS

Consider TDRS $X^{\mathbf{h}}(t)$ (3.2) with vector function $\mathbf{h}(t)$ meeting the additional condition (3.10), which allows us to address the CWEAS for NUS (5.1). It is natural to understand this particularization of the function $\mathbf{h}(t)$ as a new restriction on the time-dependence of TDRS $X^{\mathbf{h}}(t)$ (3.2), which imposes a certain decreasing rate that should be followed by the state trajectories of NUS (5.1). Such a restriction is reflected by more severe conditions on the exponents p_{ii} and interval-type coefficients a_{ii} of NUS (5.1) than stated in Theorem 5.9.

Theorem 5.12 *A necessary condition for the ES* $\{0\}$ *of NUS (5.1) to be CWEAS is that*

$$p_{ii} = 1, \quad a_{ii}^{+} < 0, \quad i = 1, ..., n. \tag{5.22}$$

Theorem 5.13 *ES* $\{0\}$ *of NUS (5.1) is CWEAS if and only if the following nonlinear algebraic inequalities are compatible (have solutions* $d_i > 0$, *i = 1,...,n,* $r < 0$*):*

$$\sum_{j=1}^{n} \bar{c}_{ij} d_j^{p_{ij}} / d_i \leq r; \quad i = 1, ..., n, \tag{5.23-a}$$

$$\sum_{j=1}^{n} \tilde{c}_{ij} d_j^{p_{ij}} / d_i \leq r; \quad i = 1, ..., n. \tag{5.23-b}$$

Theorem 5.14 *ES* $\{0\}$ *of NUS (5.1) is CWEAS if and only if the following nonlinear algebraic inequalities are compatible (have solutions* $d_i > 0$, *i = 1,...,n):*

$$\sum_{j=1}^{n} \bar{c}_{ij} d_j^{p_{ij}} / d_i < 0; \quad i = 1, ..., n, \tag{5.24-a}$$

$$\sum_{j=1}^{n} \tilde{c}_{ij} d_j^{p_{ij}} / d_i < 0; \quad i = 1, ..., n. \tag{5.24-b}$$

Conditioning the existence of CWEAS to the values of p_{ii} in NUS (5.1) (as stated in Theorem 5.12) raises a direct question about the link between CWEAS and $\text{CWAS}_{\mathbf{h}}$.

Theorem 5.15 *For* $p_{ii} = 1$ *and* $a_{ii}^{+} < 0$, *i = 1,...,n, the ES* $\{0\}$ *of NUS (5.1) is* $CWAS_{\mathbf{h}}$ *if and only if it is CWEAS.*

The nonlinear algebraic inequalities (5.23) and (5.24) can be written compactly in a matrix form, using norm ∞, by considering the square matrices $\overline{\mathbf{M}}$, $\tilde{\mathbf{M}} \in \mathbb{R}^{n \times n}$ with the following entries:

$$(\overline{\mathbf{M}})_{ij} = \bar{c}_{ij} d_j^{p_{ij}} / d_i, \tag{5.25-a}$$

$$(\tilde{\mathbf{M}})_{ij} = \tilde{c}_{ij} d_j^{p_{ij}} / d_i, \tag{5.25−b}$$

and a positive real number:

$$s \geq |\hat{c}_{ii}| \quad i = 1, ..., n. \tag{2.26}$$

Theorem 5.16 *ES {0} of NUS (5.1) is CWEAS if and only if there exist $d_i > 0$, $i = 1,...,n$, and $r < 0$ such that*

$$\max \left\{ \left\| \overline{\mathbf{M}} + s\mathbf{I} \right\|_\infty , \left\| \tilde{\mathbf{M}} + s\mathbf{I} \right\|_\infty \right\} \leq r + s. \tag{5.27}$$

Theorem 5.17 *ES {0} of NUS (5.1) is CWEAS if and only if there exist $d_i > 0$, $i = 1,...,n$, and $r < 0$ such that*

$$\max \left\{ \left\| \overline{\mathbf{M}} + s\mathbf{I} \right\|_\infty , \left\| \tilde{\mathbf{M}} + s\mathbf{I} \right\|_\infty \right\} < s. \tag{5.28}$$

As already discussed in the general case of CWAS$_\mathbf{h}$, it might be preferable to use a sufficient condition generated from DE (5.19) in Theorem 5.11. Therefore, consider the square matrix $\hat{\mathbf{P}} \in \mathbb{R}^{n \times n}$, with the following entries:

$$(\hat{\mathbf{P}})_{ij} = \hat{c}_{ij} \varepsilon^{p_{ij}-1}, \quad \varepsilon > 0, \tag{5.29}$$

where $\hat{c}_{ij}$, $i, j = 1,...,n$, are defined by (5.20) in Theorem 5.11. Denote by $\lambda_{\max}(\hat{\mathbf{P}})$ the eigenvalue of $\hat{\mathbf{P}}$ (simple or multiple) with the greatest real part. As $\hat{c}_{ij} \geq 0$, $i \neq j$, $i, j = 1,...,n$, according to Theorem 3.10, $\lambda_{\max}(\hat{\mathbf{P}})$ is a real number.

Theorem 5.18 *If, for a given $\varepsilon > 0$, matrix $\hat{\mathbf{P}}$ is Hurwitz stable, then the ES {0} of NUS (5.1) is CWEAS for some $0 < d_i \leq \varepsilon$, $i = 1,...,n$, and $\lambda_{\max}(\hat{\mathbf{P}}) \leq r < 0$.*

According to Theorem 5.11, whenever inequalities (5.21) are satisfied, the existence of a positive $\varepsilon > 0$ for which matrix $\hat{\mathbf{P}}$ is Hurwitz stable represents a necessary and sufficient condition for the ES {0} of NUS (5.1) to be CWEAS.

5.4. CWAS$_\mathbf{h}$ and CWEAS in linear approximation

The linear approximation of NUS (5.1) is an uncertain system with interval matrix of form (4.1), preserving only those elements $a_{ij} x_j^{p_{ij}}$ in $f_i(\mathbf{x})$, $i, j = 1, ..., n$, for which $p_{ij} = 1$. Thus, it is necessary to have $p_{ii} = 1$, $i = 1,...,n$, because, otherwise, the linear approximation of NUS (5.1) cannot be CWAS$_\mathbf{h}$ (or equivalently CWEAS), according to Theorem 4.3. In other words CWAS$_\mathbf{h}$ and CWEAS in linear approximation are equivalent concepts, and the linear approximation inherits the CWEAS property from NUS (5.1) as shown below.

Theorem 5.19 *Let $p_{ii} = 1$ and $a_{ii}^{+} < 0$, $i = 1,\ldots,n$. ES $\{0\}$ of NUS (5.1) is CWEAS if and only if the linear approximation of NUS (5.1) is CWEAS.*

The equivalence stated by this theorem should be understood just in qualitative terms, around the ES $\{0\}$, since, for the IMS representing the linear approximation of NUS (5.1), CWEAS is a global property (as resulting from Section 4), whereas Theorem 5.19 refers to a local property of NUS (5.1).

6. Componentwise detectability and stabilizability

6.1. CWEAS state observer, [42]

Consider the linear time-invariant control system:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u}, \quad t \in \mathbb{R}_+, \ \mathbf{u} \in \mathbb{R}^m, \ \mathbf{x} \in \mathbb{R}^n, \tag{6.1}$$

$$\mathbf{y} = \mathbf{C}\mathbf{x}, \quad \mathbf{y} \in \mathbb{R}^p, \tag{6.2}$$

with $\mathbf{x}(t_0) = \mathbf{x}_0$, $t_0 \in \mathbb{R}_+$, where $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$ are appropriate matrices.

In many cases the state feedback control relies on an estimate $\hat{\mathbf{x}}$ (instead of $\mathbf{x}$ – not measurable) that can be obtained using the observer:

$$\dot{\hat{\mathbf{x}}} = (\mathbf{A} - \mathbf{L}\mathbf{C})\hat{\mathbf{x}} + \mathbf{B}\mathbf{u} + \mathbf{L}\mathbf{y}, \quad t \in \mathbb{R}_+, \ \hat{\mathbf{x}} \in \mathbb{R}^n, \tag{6.3}$$

with $\hat{\mathbf{x}}(t_0) = 0$. The matrix $\mathbf{L}$ (when it exists) must be determined such that:

$$\lim_{t\to\infty} [\mathbf{x}(t) - \hat{\mathbf{x}}(t)] = 0. \tag{6.4}$$

However, in practice the estimation error:

$$\mathbf{x}_\varepsilon =: \mathbf{x} - \hat{\mathbf{x}} \tag{6.5}$$

must converge to zero as quickly as possible, with not too many oscillations, and at the same time it must satisfy some prescribed constraints.

In order to use the results on CWEAS (see 3.2), let X_0 be the bounded set of all possible $\mathbf{x}_0$, i.e. the bounded set of all $\mathbf{x}_\varepsilon(t_0) = \mathbf{x}_0$. Under these circumstances there may be determined an appropriate constant hyperinterval $X^{\mathbf{d}} =: =: \{\mathbf{v} \in \mathbb{R}^n, |\mathbf{v}| \le \mathbf{d}\}$, with $\mathbf{d} > 0$, $\mathbf{d} \in \mathbb{R}^n$ such that $X_0 \subseteq X^{\mathbf{d}}$.

Definition 6.1 The observer (6.3) is called CWEAS w.r.t. $\mathbf{d}$ and to a prescribed scalar $\beta > 0$ if for each $(t_0, \mathbf{x}_\varepsilon(t_0))$ the estimation error satisfies:

$$|\mathbf{x}_\varepsilon(t)| \le \mathbf{d}e^{-\beta(t-t_0)}, \quad t \ge t_0. \ \square \tag{6.6}$$

Elimination of $\mathbf{x}$, $\hat{\mathbf{x}}$ and $\mathbf{y}$ between (6.1) – (6.3) and (6.5) yields:

$$\dot{\mathbf{x}}_\varepsilon = (\mathbf{A} - \mathbf{L}\mathbf{C})\mathbf{x}_\varepsilon, \quad t \in \mathbb{R}_+, \ \mathbf{x}_\varepsilon(t_0) = \mathbf{x}_0, \tag{6.7}$$

called the *equation of estimation error*, for which (6.6) is the CWEAS condition. According to Theorem 3.8-2° the characterization corresponding to Definition 6.1 is the following.

Theorem 6.1 *he observer (6.3) is CWEAS w.r.t.* $\mathbf{d}$ *and* β *if and only if there exists at least one gain matrix* $\mathbf{L}$ *such that:*

$$\overline{(\mathbf{A}-\mathbf{LC})}\mathbf{d} \le -\beta\mathbf{d}. \tag{6.8}$$

Further the existence problem of a matrix or in general of a set of matrices $\mathbf{L}$ that satisfy (6.8) can be dealt with by applying results on the *solvability* [68] of a system of inequalities or by exploiting the *convexity* [69] of the set of $\mathbf{L}$ matrices involved in (6.8). The first approach yields existence and rather computational results while the second one provides a better insight into the structure of both the set of observers and the observed system.

To present the second approach, assume that $\mathcal{L}$ is the set of all matrices $\mathbf{L}^*$ (transpose of $\mathbf{L}$) satisfying (6.8). It is a simple matter to prove the following.

Theorem 6.2 $\mathcal{L}$ *is a convex set.*

To check the no emptiness of $\mathcal{L}$, let us observe that (6.8) is equivalent to the assignment of the $\mathbf{d}$-Gershgorin's discs (see the sentence preceding Theorem 3.8) associated with the rows of matrix $\mathbf{F} =: \mathbf{A} - \mathbf{LC}$ in the complex half plane $\{\Re e\, s \le -\beta\}$ (see Theorem 3.8–7°) via an adequate choice of gain matrix $\mathbf{L}$ (if there exists). Thus, the spectrum $\Lambda =: \{\lambda_k, k = 1, ..., n\}$ of $\mathbf{F}$, i.e. of observer (6.3), satisfies $\Re e\lambda_k \le -\beta$, $k=1,...,n$. The evaluation of $\Im m\lambda_k$, $k=1,...,n$ (the eigenfrequencies of $\mathbf{F}$), is possible by the $\mathbf{d}$-Gershgorin's discs of $\mathbf{F} =: (f_{ij})$: $G_i(\mathbf{F_d}) =: \left\{ s\mathbb{C};\, |s - f_{ii}| \le r_i =: \frac{1}{d_i}\sum_{j=1,...,n\ j\ne i} |f_{ij}|d_j \right\}$, $i=1,...,n$. They have the property $\Lambda \subseteq \cup_{i=1,...,n} G_i(\mathbf{F_d})$ that allows the immediate evaluation: $\max_k |\Im m\lambda_k| \le \max_i r_i$. Clearly, if $\Im m\lambda_k$, $k=1,...,n$, are too great, then $\mathbf{x}_\varepsilon$ may be too oscillatory, even when condition (6.6) is fulfilled. Consequently, it is more natural to prescribe already from the beginning of the design a certain assignment region for the $\mathbf{d}$-Gershgorin's discs (implicitly for Λ) in $\{\Re e\, s \le -\beta\}$ (see Theorem 3.8-7°). Thus, the simplest possibility in this respect seems to be $r_i = 0$, $i=1,...,n$, which are met if and only if $\mathbf{F}$ satisfies $f_{ij} = 0$, $i, j = 1,...,n$, $i \ne j$. Then $\mathbf{F} = \mathrm{diag}\{f_{11}, ..., f_{nn}\}$ would result, but it remains to examine if $f_{ii} \le -\beta$, $i=1,...,n$, can be fulfilled, i.e. if there exists $\mathbf{L}^* \in \mathcal{L}$ such that $\mathbf{F} = \mathbf{A} - \mathbf{LC}$ is diagonal. A pertinent analysis in this respect can be achieved in the context of maximization of β.

An essential aspect in the design of a CWEAS observer is that of the adequate choice of β in order to establish a good vanishing speed for

$\mathbf{x}_\varepsilon$. In this respect it is natural to choose the maximum value of β for which there still exists a CWEAS observer for system (6.1), (6.2). Now, as suggested by Theorem 6.2, a convex optimisation problem is to solve.

In this respect one has to start by defining the convex functions:

$$f_i(\mathbf{z}) = \frac{1}{d_i}\mathbf{d}^*\overline{(\mathbf{a}_i^* - \mathbf{C}^*\mathbf{z})}, \quad \mathbf{z} \in \mathbb{R}^p, \; i = 1, ..., n. \tag{6.9}$$

where $\mathbf{d} =: (d_i)$ and $\mathbf{a}_i$, $i = 1...,n$, are the rows of matrix $\mathbf{A}$. For (6.9) the following set of clarifying statements may be formulated.

Theorem 6.3 *$\mathcal{L}$ is nonempty if and only if*

$$\max_i \min_{\mathbf{z}\in\mathbb{R}^p} f_i(\mathbf{z}) \le -\beta. \tag{6.10}$$

Definition 6.2 The observer (6.3) is called *β-maximal* if $\beta > 0$ from Definition 6.1 has the maximum possible value $\beta_{\max}$. □

To conveniently express the next result let us consider the equations:

$$\mathbf{C}^*_{(i)}\mathbf{z} = \mathbf{a}^*_{(i)}, \quad \mathbf{z} \in \mathbb{R}^p, \; i = 1, ..., n, \tag{6.11}$$

where $\mathbf{C}^*_{(i)}$ and $\mathbf{a}^*_{(i)}$ are to be obtained by deleting the row $\mathbf{c}_i^*$ from $\mathbf{C}^*$ and respectively the element a_{ii} from $\mathbf{a}_i^*$.

Theorem 6.4 *Assume that (6.6) hold for any $\mathbf{d} > 0$. Then the following statements are equivalent: 1° $\mathcal{L}$ is nonempty. 2° There exists a β-maximal observer.*

$$3^\circ \quad \operatorname{rank}\mathbf{C}^*_{(i)} = \operatorname{rank}\left[\mathbf{C}^*_{(i)}, \mathbf{a}^*_{(i)}\right], \quad i = 1, ..., n, \tag{6.12}$$

$$\max_i \inf_{\mathbf{z}\in Z_i} (a_{ii} - \mathbf{c}_i^*\mathbf{z}) \le -\beta, \tag{6.13}$$

where $Z_i \subseteq \mathbb{R}^p$, $i = 1,...,n$, are respectively the solution sets of (6.11).

The result (6.12), (6.13) shows that for any $\mathbf{d} > 0$ a β-maximal observer, individualized by some $\mathbf{L}_0^*$, has $\mathbf{F} = \operatorname{diag}\{f_{1\ \min}, .., f_{n\ \min}\}$, with

$$f_{i\ \min} =: \min_{\mathbf{z}\in\mathbb{R}^p} f_i(\mathbf{z}) = \inf_{\mathbf{z}\in Z_i} (\mathbf{a}_{ii} - \mathbf{c}_i^*\mathbf{z}), \quad i = 1, ..., n, \tag{6.14}$$

$$\beta_{\max} = -\max_i f_{i\ \min} \ge \beta. \tag{6.15}$$

The β-maximal observer for any $\mathbf{d} > 0$ is the best in the following sense: $\mathbf{d}$ does not play any role more in (6.6) and (6.8); the vanishing velocity of $\mathbf{x}_\varepsilon(t)$ is the greatest possible; $\mathbf{F}$ is diagonal. According to (6.11) and Theorem 6.4, the gain matrix can be determined as follows:

$$\mathbf{L}_0^* \in \left\{[\mathbf{z}_1\, \mathbf{z}_2 \ldots \mathbf{z}_n]; \; \mathbf{z}_i \in Z_i \subseteq \mathbb{R}^p, \; i = 1, ..., n, \; \max_i \inf_{\mathbf{z}\in Z_i} (a_{ii} - \mathbf{c}_i^*\mathbf{z} \le -\beta\right\}.$$

At the same time, the result (6.12), (6.13) refers only to the *observability* pair ($\mathbf{A}$, $\mathbf{C}$), reflecting after all a certain property of system (6.1), (6.2) which is defined and characterized by the following statements.

Definition 6.3 The system (6.1), (6.2) is called CWEAS-*detectable* if there exists an observer (6.3) that satisfies (6.6) with (6.5). □

Clearly, according to (6.6), a necessary condition that system (6.1), (6.2) be CWEAS-detectable is that the pair ($\mathbf{A}$, $\mathbf{C}$) be detectable.

Theorem 6.5 *Assume that (6.6) hold for any* $\mathbf{d} > 0$. *Then system (6.1), (6.2) is CWEAS detectable if and only if (6.12) and*

$$\max_i \inf_{\mathbf{z}\in Z_i} (\mathbf{a}_{ii} - \mathbf{c}_i^* \mathbf{z}) < 0 \tag{6.16}$$

are met.

6.2. CWEAS of state feedback control system

Consider for the control system (6.1), (6.2) the state feedback control:

$$\mathbf{u} = -\mathbf{K}\mathbf{x} + \mathbf{M}\mathbf{v}, \quad \mathbf{v} \in \mathbb{R}^q, \tag{6.17}$$

where $\mathbf{v}$ is the new control and $\mathbf{K}$ and $\mathbf{M}$ are appropriate matrices. By replacing (6.17) into (6.1) it results the equation of state feedback system:

$$\dot{\mathbf{x}} = (\mathbf{A} - \mathbf{B}\mathbf{K})\mathbf{x} + \mathbf{B}\mathbf{M}\mathbf{v}. \tag{6.18}$$

Concerning the CWEAS of system (6.18) w.r.t. $\mathbf{d}$, i.e. the hyper-interval $X^{\mathbf{d}} =: =: \{\mathbf{v} \in \mathbb{R}^n, |\mathbf{v}| \leq \mathbf{d}\}$ (including the bounded set X_0 of all possible x_0 of (6.18)), and a prescribed $\beta > 0$, Theorem 3.8-2° yields the following result.

Theorem 6.6 *The system (6.18) is CWEAS w.r.t.* $\mathbf{d}$ *and* β *if and only if*

$$\overline{(\mathbf{A} - \mathbf{B}\mathbf{K})}\mathbf{d} \leq -\beta\mathbf{d}. \tag{6.19}$$

By analogy with Theorem 6.2, observe that for the set $\mathcal{K}$ of all solution matrices $\mathbf{K}$ of (6.19), i.e. the set of all controllers for which system (6.18) is CWEAS, the following result may be derived.

Theorem 6.7 $\mathcal{K}$ *is a convex set.*

The no emptiness of $\mathcal{K}$ (the solvability of (6.19)) may be checked by duality according to 6.1. A more direct alternative is to simply examine if there exists $\mathbf{K} \in \mathcal{K}$ for which the matrix $\mathbf{G} =: \mathbf{A} - \mathbf{B}\mathbf{K}$, [40], is diagonal and satisfies (6.19) for any $\mathbf{d} > 0$. Then, Definition 6.3 and Theorems 6.3-6.5 will be used. To proceed conveniently, consider the equation:

$$\mathbf{B}\mathbf{K} = \mathbf{A} - \mathbf{G}, \tag{6.20}$$

where $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{G} =: (g_{ij})$ are given and $\mathbf{K}$ is unknown. Denote by $\mathbf{a}_i^c$, $\mathbf{g}_i^c$ and $\mathbf{k}_i^c$, $i=1,...,n$, the columns of $\mathbf{A}$, $\mathbf{G}$ and $\mathbf{K}$ respectively, and by $\mathbf{b}_i$, $i=1,...,n$, the rows of $\mathbf{B}$. Now, (6.20) may be equivalently rewritten as:

$$\mathbf{B}_{(i)}\mathbf{k}_i^c = \mathbf{a}_{(i)}^c - \mathbf{g}_{(i)}^c, \quad (\mathbf{g}_{(i)}^c = 0), \quad i = 1, ..., n, \tag{6.21}$$

$$g_{ii} = a_{ii} - \mathbf{b}_i\mathbf{k}_i^c, \quad i = 1, ..., n, \tag{6.22}$$

where $\mathbf{a}_{(i)}^c$, $\mathbf{g}_{(i)}^c$ and $\mathbf{B}_{(i)}$ are to be obtained by deleting the elements a_{ii}, g_{ii} from $\mathbf{a}_i^c$, $\mathbf{g}_i^c$ and respectively the row $\mathbf{b}_i$ from $\mathbf{B}$.

Theorem 6.8 *Assume that for system (6.18) condition (3.11) hold for any $d > 0$. Then $\mathcal{K}$ is nonempty if and only if*

$$\operatorname{rank}\mathbf{B}_{(i)} = \operatorname{rank}\left[\mathbf{B}_{(i)}, \mathbf{a}_{(i)}^c\right], \quad i = 1, ..., n, \tag{6.23}$$

$$\max_i \inf_{\mathbf{w}\in W_i} (a_{ii} - \mathbf{b}_i\mathbf{w}) \leq -\beta, \tag{6.24}$$

where $W_i \subseteq \mathbb{R}^m$, $i = 1,...,n$, are respectively the solution sets of (6.21).

From equation (6.21) and Theorem 6.8 it follows that matrix $\mathbf{K}_0$, for which $\mathbf{G}$ is diagonal, can be determined as follows: $\mathbf{K}_0 \in \{[\mathbf{w}_1\, \mathbf{w}_2 \ldots \mathbf{w}_n]; \mathbf{w}_i \in W_i \subseteq \mathbb{R}^m,\ i = 1, ..., n,\ \max_i \inf_{\mathbf{w}\in W_i}(a_{ii} - \mathbf{b}_i\mathbf{w}) \leq -\beta\}$.

Clearly, the result (6.23), (6.24) refers only to the controllability pair ($\mathbf{A}$, $\mathbf{B}$), reflecting a certain property of system (6.1) that is defined and characterized by the following statements.

Definition 6.4 The system (6.1) is called CWEAS-*stabilizable* if there exists a controller (6.17) such that system (6.18) satisfies (3.4) with (3.11). □

Clearly, according to (3.11), a necessary condition that system (6.1) be CWEAS-stabilizable is that the pair ($\mathbf{A}$, $\mathbf{B}$) be stabilizable.

Theorem 6.9 *The system (6.1) is CWEAS stabilizable for any* $\mathbf{d} > 0$ *if and only if (6.23) and*

$$\max_i \inf_{\mathbf{w}\in W_i} (a_{ii} - \mathbf{b}_i\mathbf{w}) < 0 \tag{6.25}$$

are met.

7. Design of sliding motion control, [50]

7.1. Preliminaries on sliding motion

Consider the nonlinear continuous-time control system (2.1), endowed with a state feedback discontinuous control and described by:

$$\dot{\mathbf{x}} = \mathbf{f}(t, \mathbf{x}, \mathbf{u}(t, \mathbf{x})) =: \mathbf{F}(t, \mathbf{x}), \quad t \in \mathbb{R}_+,\ \mathbf{x} \in \mathbb{R}^n,\ \mathbf{u} \in \mathbb{R}^m, \tag{7.1}$$

with $\mathbf{x} =: (x_i)$, and the control *switching hyperplane*:

$$S =: \{\mathbf{x} \in \mathbb{R}^n; \quad x_n = 0\}\,. \tag{7.2}$$

The achievement of *sliding motion* of $\mathbf{x}$ on S towards an equilibrium state in $\mathbb{R}^n$ consists in the synthesis of control $\mathbf{u}(t,\, \mathbf{x})$, discontinuous on S, such that the following three requirements are fulfilled:

1°. For every initial pair $(t_0,\, \mathbf{x}_0) \in \mathbb{R}_+ \times (\mathbb{R}^n \setminus S)$, with $\mathbf{x}(t_0) = \mathbf{x}_0$, system (7.1) evolves towards S, i.e. its state reaches S, in a *reaching point*, after a finite time interval $[t_0,\, \tau]$, $\tau > t_0$. This is the *reaching condition.*

2°. Since the *reaching instant* τ the state of system (7.1) remains to evolve on S, which is called the *ideal sliding motion* and S is the *ideal sliding domain* for system (7.1). This is the *ideal sliding motion condition.*

3°. The *ideal sliding motion* (on S) must be *asymptotically stable* towards an equilibrium state (usually or conventionally $\mathbf{x} = 0$) belonging to S. This is the *stability condition* of the ideal sliding motion.

Chronologically, conditions 1° and 2° must be successively fulfilled, while 2° and 3° must be simultaneously satisfied. This means that for each pair $(t_0,\, \mathbf{x}_0) \in\ \in \mathbb{R}_+ \times (\mathbb{R}^n \setminus S)$ the whole evolution of system (7.1), covers *two concatenated time intervals*: first, a finite one, $[t_0, \tau]$, according to condition 1°, followed by the second one, (τ, t_f), finite or not, according to conditions 2° and 3°. This essential and, as a matter of fact, natural concatenation of the reaching process followed by the ideal sliding motion can be approached in a unified manner by using adequately Theorem 2.2.

7.2. Some supporting results for the application of Theorem 2.2

Assume that $\mathbf{F} : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^n$ is a continuous and locally Lipschitzian function, i.e. for each $(t_0, x_0) \in \mathbb{R}_+ \times \mathbb{R}^n$ system (7.1) has a unique solution $\mathbf{x}(t)$, with $\mathbf{x}(t_0) = \mathbf{x}_0$, defined on a maximal time interval $(t_a, t_b) \subseteq \mathbb{R}_+$, with $t_0 \in (t_a, t_b)$.

The solutions $\mathbf{x}^-(t) = \mathbf{x}(t)$, $t \in (t_a, t_0]$, and $\mathbf{x}^+(t) = \mathbf{x}(t)$, $t \in [t_0, t_b)$, are respectively called *negative* (or to the left) and *positive* (or to the right) solutions through the point $(t_0, \mathbf{x}_0)$.

Definition 7.1 A set $X(t) \subseteq \mathbb{R}^n$, $t \in \mathbb{R}_+$, is called *negatively* or *positively flow-invariant* w.r.t. system (7.1) if for each $(t_0, \mathbf{x}_0) \in \mathbb{R}_+ \times \mathbb{R}^n$ the conditions: $\mathbf{x}^-(t) \in X(t)$, $t \in (t_a, t_0]$ or $\mathbf{x}^+(t) \in X(t)$, $t \in [t_0, t_b)$ are respectively met. □

Theorem 7.1 *$X(t)$ is positively flow-invariant w.r.t. system (7.1) if and only if $\mathbb{R}^n \setminus X(t)$ is negatively flow-invariant w.r.t. system (7.1).*

Consider now that $X(t)$ is defined by (2.6), with $\underline{\mathbf{a}}(t) =: (\underline{a}_i(t))$ and $\overline{\mathbf{a}}(t) =: =: (\overline{a}_i(t))$. Applying Theorem 2.2 for system (7.1), with $\mathbf{F} =: (F_i)$, ((2.8), (2.9) expressed componentwise) the following result may be derived.

Theorem 7.2 *$X(t)$ defined by (2.6) is positively flow-invariant w.r.t. system (7.1) if and only if:*

$$\begin{cases} F_i(t, x_1, ..., x_{i-1}, \underline{a}_i(t), x_{i+1}, ..., x_n) \geq \dot{\underline{a}}_i(t) \\ \\ F_i(t, x_1, ..., x_{i-1}, \overline{a}_i(t), x_{i+1}, ..., x_n) \leq \dot{\overline{a}}_i \end{cases} , i=1, ..., n, \; \forall \, (t, x) \in \mathbb{R}_+ \times X(t). \tag{7.3}$$

7.3. Flow structure pertaining to ideal sliding motion

Theorems 7.1 and 7.2 may be used for proving the flow structure of state space generated by the ideal sliding motion occurring in system (7.1).

As a matter of fact, system (7.1), with (7.2), has a variable structure and is constituted of two *switching subsystems* according to the equation:

$$\dot{\mathbf{x}} = \begin{cases} \mathbf{F}^-(t, \mathbf{x}), \;\; \mathbf{x} \in S^- \\ \\ \mathbf{F}^+(t, x), \;\; \mathbf{x} \in S^+ \end{cases} , \; t \in \mathbb{R}_+, \tag{7.4}$$

$S^- =: \{\mathbf{x} \in \mathbb{R}^n, \, x_n < 0\}$, $S^+ =: \{\mathbf{x} \in \mathbb{R}^n, \, \mathbf{x}_n > 0\}$ being the *functioning subsets.*

In keeping with system (7.4) one may also define the systems:

$$\dot{\mathbf{x}} = \mathbf{F}^-(t, x), \;\; (t, x) \in \mathbb{R}_+ \times \mathbb{R}^n, \tag{7.5}$$

$$\dot{\mathbf{x}} = \mathbf{F}^+(t, x), \;\; (t, x) \in \mathbb{R}_+ \times \mathbb{R}^n, \tag{7.6}$$

which are called the *adjacent systems* of (7.4) to S.

Definition 7.2 Switching surface S is an *ideal sliding motion domain* for system (7.4) if S does not contain any trajectory segments of adjacent systems (7.5), (7.6) and for each $\varepsilon > 0$ and each $\mathbf{x}^s \in S$ there exists a neighbourhood V^s of $\mathbf{x}^s$ such that for each fixed $\mathbf{x}_0 \in V^s \setminus S$ the state of system (7.4) evolves inside the domain $S_\varepsilon =: \{\mathbf{x} \in \mathbb{R}^n; \, |x_n| \leq \varepsilon\}$ for each $t \in [t_0, +\infty)$. □

To emphasize the state space flow structure of system (7.4), let us assume that $\mathbf{F}^\mp : \mathbb{R}_+ \times \mathbb{R}^n \to \mathbb{R}^n$ are continuous and locally Lipschitzian and $\mathbf{F}^\mp =: \left(\mathbf{F}_i^\mp\right)$.

Theorem 7.3 *Assume that adjacent systems (7.5), (7.6) have no trajectory segments on S. Then for system (7.4) the following statements are equivalent:*

1° S is the ideal sliding domain.

2° Each functioning subset is negatively flow-invariant w.r.t. its own switching subsystem.

3° The complementary of each functioning subset is positively flow-invariant w.r.t. the switching subsystem corresponding to the respective functioning subset.

4° $\lim_{h\downarrow 0}\inf h^{-1}d(\mathbf{x}+h\mathbf{F}^{\mp}(t,\mathbf{x});\, S\cup S^{\pm}) = 0,\ \forall\,(t,\mathbf{x}) \in \mathbb{R}_+\times(S\cup S^{\mp})$.

$$5^\circ \begin{cases} F_n^-(t,x_1 \ldots x_{n-1},0) \geq 0 \\ \\ F_n^+(t,x_1,\ldots,x_{n-1},0) \leq 0 \end{cases},\ \forall\,(t,[x_1,\ldots,x_{n-1}]^* \in \mathbb{R}_+ \times \mathbb{R}^{n-1}. \quad (7.7)$$

Besides the cardinal result stated by 5° (that allows to solve the ideal sliding motion problem (and control design) through conditions on S only, unlike the sufficient classical ones that must hold on a vicinity of S, [70], [71]), this theorem, by 3° and 4°, depicts the global flow structure of the state space of system (7.4) induced by its ideal sliding domain. Thus, precisely this flow structure allows deriving further results concerning the *reaching process* as a natural *flowing precursor* of the *ideal sliding motion*, i.e. the ideal sliding motion as a natural goal and as final (chronological) part of the reaching process.

7.4. Reaching process – flowing precursor of ideal sliding motion

The reaching process may be adequately dealt with Theorem 7.2, namely as suggested by Theorem 7.3-2°.

Theorem 7.4 *For each $(t_0,\mathbf{x}_0) \in \mathbb{R}_+\times(\mathbb{R}^n\setminus S)$ the state of system (7.4) reaches ideal sliding domain S if and only if there exists a differentiable function $r:\mathbb{R}_+\to\mathbb{R}$, depending on $(t_0,\mathbf{x}_0)$, that satisfies the following conditions:*

1° There exists $\tau \in (t_0,+\infty)$ such that $r(\tau)=0$.

$$2^\circ \begin{cases} \text{If } x_{0n} < 0 \text{ then } x_{0n} \geq r(t_0) \\ \\ \text{If } x_{0n} > 0 \text{ then } x_{0n} \leq r(t_0) \end{cases};\ x_{0n} =: x_n(t_0). \quad (7.8)$$

$$3^\circ \begin{cases} \text{If } x_{0n} < 0 \text{ then } F_n^-(t,x_1,\ldots,x_{n-1},r(t)) \geq \dot r(t) \\ \\ \text{If } x_{0n} > 0 \text{ then } F_n^+(t,x_1,\ldots,x_{n-1},r(t)) \leq \dot r(t) \end{cases}. \quad (7.9)$$

$$\forall\,(t,[x_1 ... x_{n-1}]^*) \in [t_0,\tau] \times \mathbb{R}^{n-1}.$$

Notice that this theorem is consistent w.r.t. conditions (7.7) because for $r(t) \equiv 0$ in (7.9) one obtains (7.7). Since conditions (7.7) are equivalent to the existence of the ideal sliding domain of system (7.4), it results that Theorem 7.4 also includes (unlike the classical approach, [70], [71]) the case of ideal sliding motion as a chronological subsequent part of the reaching process. In order to use Theorem 7.4 for the synthesis of sliding motion control, one has to appropriately choose the *reaching function* $r(t)$ that allows, to some extent, to prescribe the velocity of the reaching process.

7.5. Sliding motion control of a linear disturbed plant

Consider the linear time – invariant disturbed plant:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{b}u + \mathbf{D}\mathbf{z}, \quad t \in \mathbb{R}_+,\ \mathbf{x} \in \mathbb{R}^n,\ u \in,\ \mathbf{z} \in \mathbb{R}^q, \tag{7.10}$$

where the state $\mathbf{x}$ is available for measurement, u is the scalar control, and $\mathbf{z}$ is the disturbance; $\mathbf{A} =: (a_{ij})$, $\mathbf{b} =: (b_j)$, $\mathbf{D} = (d_{ij})$ are appropriate matrices.

Let us associate with system (7.10) the switching hyperplane:

$$S_c =: \{\mathbf{x} \in \mathbb{R}^n;\ s = \mathbf{c}^*\mathbf{x} = 0\}, \tag{7.11}$$

where $\mathbf{c} =: [c_1 ... c_{n-1}\, 1]^*$. Using the transformation $\tilde{\mathbf{x}} =: [x_1 ... x_{n-1}\, s]^* = \mathbf{P}\mathbf{x}$, where $\mathbf{P} =: \begin{bmatrix} \mathbf{I}_{n-1} & \vdots & 0 \\ \multicolumn{3}{c}{\dots\dots\dots} \\ \mathbf{c}^*_{(n)} & \vdots & 1 \end{bmatrix}$, $\mathbf{I}_{n-1}$ is the unit matrix of order $n-1$, and $\mathbf{c}_{(n)}$ is obtained by deleting $c_n = 1$ of $\mathbf{c}$, system (7.10) becomes:

$$\dot{\mathbf{x}}_{(n)} = \mathbf{E}\mathbf{x}_{(n)} + \mathbf{a}^c_{(n)}s + \mathbf{b}_{(n)}u + \mathbf{D}_{(n)}\mathbf{z}, \tag{7.12}$$

$$\dot{s} = \sum_{i=1}^{n} \mathbf{c}^*(\mathbf{a}^c_i - c_i\mathbf{a}^c_n)x_i + \mathbf{c}^*\mathbf{a}^c_n s + \mathbf{c}^*\mathbf{b}u + \mathbf{c}^*\mathbf{D}\mathbf{z}. \tag{7.13}$$

In (7.12), (7.13) $\mathbf{E} =: (a_{ij} - a_{in}c_j)$ is an $(n-1)\times(n-1)$ matrix, $\mathbf{A}_{(n)}$ and $\mathbf{D}_{(n)}$ denote the matrices obtained by deleting the n-th rows and columns of $\mathbf{A}$ and $\mathbf{D}$ respectively, $\mathbf{a}^c_i$ is the i-th column of $\mathbf{A}$; $\mathbf{x}_{(n)}$, $\mathbf{a}^c_{(n)}$ and $\mathbf{b}_{(n)}$ are obtained by deleting the n-th components of $\mathbf{x}$, $\mathbf{a}^c_n$ and $\mathbf{b}$, respectively.

According to Theorems 7.3, 7.4 it follows that for the control design only (7.13) is to be used. For $\mathbf{c}^*\mathbf{b} \neq 0$, consider the control algorithm:

$$u = -u_r(s) - u_s(\mathbf{x}_{(n)}) - u_z(\mathbf{x}_{(n)}), \tag{7.14}$$

$$u_r(s) = \rho s + \delta_r, \quad \delta_r = \begin{cases} \delta_0, & s < 0, \\ 0, & s = 0, , \\ -\delta_0, & s > 0, \end{cases} \tag{7.15}$$

$$u_s(\mathbf{x}_{(n)}) = \sum_{i \in J} \psi_i x_i, \quad \psi_i = \begin{cases} \beta_i, & x_i s < 0, \\ \alpha_i, & x_i s > 0, \end{cases} \tag{7.16}$$

$$u_z(\mathbf{x}_{(n)}) = \begin{cases} \beta_0, & s < 0, \\ \alpha_0, & s > 0, \end{cases} \tag{7.17}$$

where $u_r(s)$, $u_s(\mathbf{x}_{(n)})$ and $u_z(\mathbf{x}_{(n)})$ have to control the reaching process, the ideal sliding motion, and the disturbance rejection respectively; ρ, δ_0, α_i, β_i, $(i \in J =: =: \{i \in \{1, 2, ..., n-1\};\ \mathbf{c}^*(\mathbf{a}_i^c - c_i \mathbf{a}_n^c) \neq 0\})$, and α_0, β_0 are adjustable parameters. Using (7.7) for (7.13) the following result can be formulated.

Theorem 7.5 *S is the ideal sliding domain for system (7.10), (7.14), with,* $\mathbf{c}^*\mathbf{b} \neq 0$, *if and only if*

$$\alpha_i \mathbf{c}^*\mathbf{b} \geq \mathbf{c}^*(\mathbf{a}_i^c - c_i \mathbf{a}_n^c),\ \beta_i \mathbf{c}^*\mathbf{b} \leq \mathbf{c}^*(\mathbf{a}_i^c - c_i \mathbf{a}_n^c),\ i \in J, \tag{7.18}$$

$$\alpha_0 \mathbf{c}^*\mathbf{b} \geq \sup_t \mathbf{c}^*\mathbf{D}\mathbf{z}(t),\ \beta_0 \mathbf{c}^*\mathbf{b} \leq \inf_t \mathbf{c}^*\mathbf{D}\mathbf{z}(t).$$

In order to design the reaching process as a precursor of the ideal sliding motion, let us use the following reaching function:

$$r(t) = \begin{cases} \left[(|s_0| + \delta) e^{-\lambda(t-t_0)} - \delta\right] \operatorname{sgn}(s), & t \in [t_0, \tau], \\ 0, & t \in (\tau, t_f), \end{cases} \tag{7.20}$$

where $s_0 =: s(\mathbf{x}_0) = s(\mathbf{x}(t_0))$, and $\delta > 0$, $\lambda > 0$, $\tau > t_0$ are pre-assignable parameters. Using Theorem 7.4-3° (1° and 2° are satisfied with $r(\tau) = 0$, $\tau = = t_0 + \lambda^{-1} \ln(1 + |s_0|/\delta) > t_0$), with (7.15) – (7.17), under (7.18), (7.19), one can formulate the following result.

Theorem 7.6 *For each pair* $(t_0, \mathbf{x}_0) \in \mathbb{R}_+ \times (\mathbb{R}_n \setminus S)$ *the state of system (7.10), (7.14), with* $\mathbf{c}^*\mathbf{b} \neq 0$, *reaches ideal sliding domain S if and only if*

$$\rho \mathbf{c}^*\mathbf{b} \geq \mathbf{c}^*\mathbf{a}_n^c + \lambda, \quad \delta_0 \mathbf{c}^*\mathbf{b} \leq \mathbf{c}^*\mathbf{a}_n^c. \tag{7.21}$$

The equations (7.14) – (7.17) and the inequalities (7.18), (7.19) and (7.21) define the (variable) structure and the adjustable parameters of the controller in order to fulfil requirements 1° and 2° from 7.1.

By solving the equation $\dot{s} = 0$ (according to (7.13)) w.r.t. u and replacing the result into (7.12) ($s = 0$), one obtains the *ideal sliding equation*:

$$\dot{\mathbf{x}}_{(n)} = \mathbf{A}^1 \mathbf{x}_{(n)} + \mathbf{D}^1 \mathbf{z}, \tag{7.22}$$

where matrices $\mathbf{A}^1$ and $\mathbf{D}^1$ can be appropriately calculated. If $\operatorname{rank} \mathbf{b} = \operatorname{rank} [\mathbf{b}, \mathbf{D}]$, then $\mathbf{D}^1 = 0$ and the disturbance rejection is ensured.

The stability of system (7.22) depends only on $\mathbf{A}^1$ and may be improved (requirement 3° from 7.1) by an additional state feedback. Thus, the state trajectories of system (7.10), (7.14) evolve with pre-assignable velocity from each initial state $\mathbf{x}_0 \in \mathbb{R}_n \backslash S$ towards ideal sliding domain S and then towards the final equilibrium state $\mathbf{x} = 0$.

8. Conclusions

The flow-invariance method proves to be an efficient tool for a more subtle characterization of the temporal evolution of the dynamical systems. It certainly occasions a deeper insight into the system behavior, fact which becomes extremely relevant when the flow-invariant set considered in the state space is a time-dependent hyper-interval. In this case, a componentwise characterisation is available for the state variables. Such a characterisation is useful especially when the state variables present different importance for the normal system evolution, unlike the usual evaluation ensured by a norm-based characterisation (which is rather global and, consequently, it does not allow accurate distinctions, when necessary, between the individual dynamics of the state variables). Generally speaking, the corresponding results of the state componentwise characterisation have the form of necessary and sufficient conditions that analytically are expressed by differential or algebraic inequalities. These results are interesting not only at the theoretical level as yielding non-conventional approaches to system analysis and design, but also from the point of view of their applicability to large classes of *real* dynamical systems encountered in engineering (electric circuits and networks, neuronal networks, control systems), biology, ecology, pharmacokinetics etc. In such cases, meaningful and significant non-standard knowledge may be gained that accurately enlighten refined aspects of theoretical and practical interest for various applications.

In this context, the main conceptual sphere developed in the paper refers to the componentwise asymptotic stability, whose construction relies, as preliminaries, on the analysis of constrained evolution of dynamical systems, and later on, is able to accommodate synthesis problems covering observer and state feedback design, as well as sliding motion control. Thus, the overall approach is founded on the notion of $T\,\mathcal{U}\,X$ –

constrained evolution of a continuous-time dynamical system (Definition 2.1) which can be characterized in terms of flow invariance (Theorem 2.1), so as a noticeable particularization can be obtained when a special type of constraints, consisting in time-dependent hyper-intervals, is considered for the state and input variables (Theorem 2.2 - the general case, Theorem 2.3 - the linear, time-invariant case).

This background allows focusing on the free dynamics and defining the $\mathrm{CWAS}_{\mathbf{h}}$ and global $\mathrm{CWAS}_{\mathbf{h}}$ properties (Definitions 3.1 and 3.2, respectively) for the existence of which necessary and sufficient conditions are derived (Theorems 3.1, 3.2 – the general case, Theorems 3.3, 3.4 – the linear, time-invariant case). A refinement of the $\mathrm{CWAS}_{\mathbf{h}}$ requirements leads to the CWEAS and global CWEAS concepts (Definitions 3.3 and 3.4, respectively), which can be characterized algebraically (Theorems 3.5, 3.6 – the general case, Theorems 3.7, 3.8 – the linear, time-invariant case). By exploiting CWEAS, the componentwise absolute stability is introduced for a class of nonlinear systems (Definition 3.5) and a necessary and sufficient condition for its existence is given (Theorem 3.9). For the case of linear, time-invariant systems, two results exploring the eigenvalue location (Theorems 3.10, 3.11) are used to analyze the CWEAS robustness under unstructured and structured perturbations (Theorem 3.12 and Theorem 3.13, respectively). Remaining within the framework of linear dynamics, the study of the free response based on flow-invariance is extended for interval matrix systems (Theorems 4.1, 4.2), offering a nice and natural generalization of the results on $\mathrm{CWAS}_{\mathbf{h}}$ and CWEAS previously reported, in the case of constant coefficients (Theorems 4.3, 4.5 and Theorem 4.6, respectively). This study also allows a refined interpretation of Hurwitz stability for interval matrices (Theorems 4.4).

The usage of flow-invariance method in exploring free dynamics is further applied for a class of nonlinear uncertain systems, yielding characterizations expressed in terms of nonlinear differential inequalities with constant coefficients (Theorems 5.1, 5.2, 5.3); the nonlinearity of the dynamics involves a detailed discussion on the geometry of the flow-invariant time-dependent hyper-intervals (Theorems 5.4, 5.5, 5.6). Relying on some intermediary results (Theorems 5.7, 5.8, 5.9), it is shown that the $\mathrm{CWAS}_{\mathbf{h}}$ of the equilibrium state $\{0\}$ of the nonlinear uncertain system is equivalent to the standard asymptotic stability of the equilibrium state $\{0\}$ of a differential equation with constant coefficients (Theorem 5.10); along the same lines, a sufficient condition is also given (Theorem 5.11), presenting the advantage of a simpler manipulation. Unlike the linear dynamics previously investigated, the existence of CWEAS requires a special form for the state-space equations (Theorem 5.12), but for such special forms, CWEAS is equivalent to $CWAS_{\mathbf{h}}$ (Theorem 5.15), as in the linear case. Moreover, a similarity with the linear case can also be found with regard to the algebraic characterization of CWEAS (Theorems 5.13, 5.14, 5.16, 5.17, 5.18). A necessary and sufficient condition for CWEAS of the nonlinear uncertain system is CWEAS of the interval matrix system which represents the linear approximation of the former (Theorem 5.19).

CWEAS principles can be exploited to incorporate typical problems of linear system analysis/synthesis such as detectability and stabilizability redefined in the sense of componentwise constrained evolutions (Definitions 6.1,6.2, 6.3 and Definition 6.4, respectively), instead of the classical formulations based on global approaches in terms of vector norms. Thus CWEAS detectability and CWEAS stabilizability are studied in conjunction with the design of CWEAS observers (Theorems 6.1, 6.2, 6.3, 6.4, 6.5) and CWEAS state feedback control (Theorems 6.6, 6.7, 6.8, 6.9), respectively.

The flow-invariance method gives an appropriate framework to deal with sliding motion by using the notions of negatively and positively flow-invariant sets (Definition 7.1) and contextually defining the ideal sliding motion domain (Definition 7.2). Relying on two preparatory results (Theorems 7.1, 7.2), the approach first focuses on ideal sliding motions (Theorems 7.3, 7.4) and afterwards emphasis is placed on sliding motion control of disturbed plants with linear dynamics (Theorems 7.5, 7.6).

Far from scrutinizing the whole research potential offered by the applicability of flow invariance in control system analysis and design, this survey paper creates an overview of most noticeable results emerged from the main concept of componentwise asymptotic stability. Further investigations can be directed towards various objectives pertaining to linear and nonlinear system theory, such as: CWEAS preservation under linear transformations of the state-space variables, connections between CWEAS stabilizability/detectability and standard strategies based on eigenvalue assignment, links between CWEAS stabilizability/detectability and linear quadratic problems, componentwise asymptotic stability of different classes of nonlinear systems etc. Along the same lines, a special remark deserves the new orientation of the flow-invariance instruments towards the discrete-time dynamics, which was just simply mentioned in the introductory section, without any details in the text, as being beyond the scope of the current survey.

List of acronyms

AS	*asymptotic stability* (as a compound noun), **or** *asymptotically stable* (as a compound adjective)
CAS	*componentwise absolute stability* (as a compound noun), **or** *componentwise absolutely stable* (as a compound adjective)
CCE	*componentwise constrained evolution*
$\text{CWAS}_{\mathbf{h}}$	*componentwise asymptotic stability w.r.t.* **h** (as a compound noun), **or** *componentwise asymptotically stable w.r.t.* **h** (as a compound adjective)
CWEAS	*componentwise exponential asymptotic stability* (as a compound noun), **or** *componentwise exponential asymptotically stable* (as a compound adjective)
DE	*differential equation*
DI	*differential inequality*
ES	*equilibrium state*
FI	*flow-invariance* (as a compound noun), **or** *flow-invariant* (as a compound adjective)
IMS	*interval matrix system*
LE$\mathcal{C}$M	*linear elementwise $\mathcal{C}$-majorant*
NUS	*nonlinear uncertain system*
PS	*positive solution*
TDRS	*time-dependent rectangular set*
w.r.t.	*with respect to*

References

[1] Naugmo, M., *Über die Lage der Integralkurven gewöhnlicher Differentialgleichungen*, Proc. Phys. Math. Soc. Japan, **24**, pp. 551-559, 1942.

[2] Hukuhara, M., *Sur la théorie des equations differentieles ordinaries*, J. Fac. Sci. Univ. Tokio, Math. **7**, no. 8, pp. 483-510, 1954.

B̊REZIS, H., *On a characterization of flow-invariant sets*, Comm. Pure Appl. Math., **23**, pp. 261-263, 1970.

[3] Crandall, M. G., *A generalization of Peano's existence theorem and flow-invariance*, Proc. Amer. Math. Soc., **36**, pp. 151-155, 1972.

[4] Martin, R. H. jr., *Differential equations on closed subsets of a Banach space*, Trans. Amer. Math. Soc., **179**, pp 399-414, 1973.

[5] Pavel, H. N., *Differential Equations, Flow-invariance and Applications*, Pitman, Boston, 1984.

5M̊OTREANU, D., PAVEL, H. N., *Tangency, flow-invariance for differential equations and optimisation problems*, Marcel Dekker, New York, 1999.

[6] Voicu, M., *State constraints and asymptotic stability of linear constant dynamical systems*, Bul. Inst. Politehnic Iaşi, **XXVII (XXXI)**, s. III, pp. 57-60, 1981.

[7] Voicu, M., *Componentwise asymptotic stability of the linear constant dynamical systems*, (Al 2-lea Simp. Naţ. de Teoria Sist., Univ. din Craiova, vol. I, pp. 69-75, 1982); Preprint Series in Math. of "A. Myller" math. Seminar, University of Iaşi, **5**, pp. 1-14, 1982.

[8] Voicu, M., *Free response characterization via flow-invariance*, 9-th World Congress of Int. Fed. of Automatic Control, Budapest; Preprints, **5**, pp. 12-17, 1984.

[9] Voicu, M., *Componentwise asymptotic stability of linear constant dynamical systems*, IEEE Trans. on Aut. Control, **AC-29**, no. 10, pp. 937-939, 1984.

[10] Voicu, M., *Evolution on control and state hyper-intervals*, 6-th Internat. Conf. Control Syst. & Comp. Sci., Bucharest; Preprints, **1**, pp. 81-83, 1985.

[11] Voicu, M., *Structural properties of the spatial manipulating systems in connection with the state and control constraints*, IFAC Symposium on Robot Control 1985, Barcelona; pp. 425-428.

[12] Voicu, M., *Tehnici de analiză a stabilitătii sistemelor automate* (*Stability Analysis Techniques of the Automatic Control Systems*), Editura Tehnică, Bucureşti, 1986.

[13] Voicu, M., *On the application of the flow-invariance method in control theory and design*, 10-th World Congress of Int. Fed. of Aut. Control, Munich; Preprints, **8**, pp. 364-369, 1987.

[14] Voicu, M., *On the existence of limit cycles* (in Romanian), Al 2-lea Simp. "Structuri, algoritmi şi echipamente de conducere a proc. ind.", Inst. Polit. Iaşi; Preprints, pp. 39-43, 1989.

[15] Voicu, M., *Sufficient existence conditions of limit cycles*, Bul. Inst. Politehnic Iaşi, **XXXII (XXXVI)**, s. III, pp. 23-26, 1990.

[16] Voicu, M., *Flow-invariance method and componentwise absolute stability; applications in biotechnological systems*, Conferenece presented at Lab. d'Automatique de Grenoble, 1993.

V̊OICU, M., *Sufficient conditions for Hurwitz and Schur polynomials*, Bul. Inst. Politehnic Iaşi, **XLIV (XLVII)**, s. IV, pp. 1-6, 1998.

[17] Voicu, M., *Componentwise absolute stability of endemic epidemic systems*, Workshop on "Theories of Popov-type", *Romanian Academy*, Bucharest; Preprints, pp. 97-104, 1999.

[18] Păstrăvanu, O., Voicu, M., *Flow-invariant rectangular sets and componentwise asymptotic stability of interval matrix systems*, 5-th European Control Conf., Karlsruhe; CDROM, rubicon-Agentur für digitale Medien, Aachen, 16 p, 1999.

[19] Păstrăvanu, O., Voicu, M., *Robustness analysis of componentwise asymptotic stability*, 16-th IMACS World Congress, Lausanne; CDROM, imacs (ISBN 3 95222075 1 9), 2000.

[20] Păstrăvanu, O., Voicu, M., *Preserving componentwise asymptotic stability under disturbances*, Rev. Roum. Sci. Techn., ser. Electrot. Energet., **45**, no. 3, pp. 413-425, 2000.

[21] Păstrăvanu, O., Voicu, M., *Robustness of componentwise asymptotic stability for a class of nonlinear systems*, Proc. of Romanian Academy, ser. A, **1**, 1-2, 2001, 61-67.

[22] Păstrăvanu, O., Voicu, M., *Componentwise characterization of the free response for a class of nonlinear systems*, 13th International Conf. on Control Systems and Computer Science, *University Politehnica of Bucharest*; Proceedings, pp. 50-55, 2001.

[23] Păstrăvanu, O., Voicu, M., *Dynamics of a class of nonlinear systems under flow-invariance constraints*, 9-th IEEE Med. Conf. on Control and Aut. (MED'01), Dubrovnik; CD-ROM (ISBN 953 6037 35 1), Book of abstracts (ISBN 953 6037 34 3); ©KoREMA (Zagreb), 6p, 2001.

[24] Păstrăvanu, O., Voicu, M., *Componentwise asymptotic stability of a class of nonlinear systems.* 1st IFAC Symposium on System Structure and Control (SSSC 01), August 29-31, 2001, *Czech Techn. Univ. of Prague*; Book of Abstracts, p. 28; CD-ROM, 078, 6 p, 2001.

[25] Păstrăvanu, O., Voicu, M., *Flow-invariance in exploring stability for a class of nonlinear uncertain systems*, 6th European Control Conf., Porto, Sept. 4-7, 2001; CD-ROM, 6 p, 2001.

[26] Păstrăvanu, O., Voicu, M., *Interval matrix systems – flow-invariance and componentwise asymptotic stability*, J. Diff. & Int. Eq, Nov., 2002 (accepted for publication).

[27] Hmamed, A., *Componentwise stability of continuous-time delay linear systems*, Automatica, **32**, no. 4, pp. 651-653, 1996.

[28] Hmamed, A., *Componentwise stability of 1D and 2D linear discrete-time systems*, Automatica, **33**, no. 9, pp. 1759-1762, 1997.

[29] Hmamed A., Benzaouia A., *Componentwise stability of linear systems: a non-symmetrical case*, Int. Journal Robust Nonlin. **7**, no. 11, pp. 1023-1028, 1997.

[30] Chu, T. G., *Convergence in discrete-time neural networks with specific performance*, Phys. Revue E, 6305, pp. 1904- + Part 1, art. No. 051904, May 2001.

[31] Voicu, M., *State feedback matrices for linear constant dynamical systems with state constraints*, 4-th Internat. Conf. Control Syst. & Comp. Sci., University Politehnica of Bucharest; Preprints, **1**, pp. 110-115, 1981.

[32] Voicu, M., *Componentwise stabilization of constant linear dynamical systems* (in Romanian), 3-th MSIOPT Symp., *University of Galaţi*, Preprints, pp. 103-108, 1982.

[33] Voicu, M., *On the determination of linear state feedback matrix*, 5-th Internat. Conf. Control Syst. & Comp. Sci., University Politehnica of Bucharest; Preprints, **1**, pp. 119-123, 1983.

[34] Voicu, M., *Gerschgorinsche Kreise und die komponentenweise Stabilisierung*, Bul. Inst. Politehnic Iaşi, **XXXI (XXXV)**, s. III, pp. 45-50, 1985.

[35] Voicu, M., *Ein Anwendungsbeispiel der Komponentenweisen Stabilisierung*, Bul. Inst. Politehnic Iaşi, **XXXI (XXXV)**, s. III, pp. 57-60, 1985.

[36] Voicu, M., *A sufficient asymptotic stability condition for discrete systems*, Bul. Inst. Politehnic Iaşi, **XXXII (XXXVI)**, s. III, pp. 38-39, 1986.

[37] Voicu, M., *System matrix with prescribed off-diagonal entries obtained via state feedback*, Bul. Inst. Politehnic Iaşi, **XLIII (XLVII)**, s. IV, pp. 5-9, 1997.

[38] Voicu, M., Bulea, M., *Pole assignment in a region of complex plane via linear programming* (in Romanian), Conf. Naţională de Electronică, Telecomunicaţii, Automatică şi Calculatoare, Bucureşti; Preprints, **TS**, pp.153-165, 1982.

[39] Voicu, M., *Observing the state with componentwise decaying error*, Systems & Control Letters, **9**, pp. 33-42, 1987.

[40] Păstrăvanu O., Voicu, M., *State feedback design for componentwise exponential asymptotic stability*, Int. Symp. on Aut. Control and Comp. Sci. (SACCS 2001), *Techn. Univ. of Iaşi*; CD-ROM, 6 p, 2001.

[41] Voicu, M., Moroşan, B. I., *A theorem on structure-variable systems* (in Romanian), Al 2-lea Simpozion "Structuri, algoritmi şi echipamente de conducere a proceselor industriale", *Inst. Politehnic Iaşi*; Preprints, pp. 21-26, 1989.

[42] Voicu, M., Moroşan, B. I., *Existence problem of an ideal sliding domain* (in Romanian), Zilele academice ieşene, *Academia Română, Fliala Iaşi*, 7 p., 1989.

[43] Voicu, M., Moroşan, B. I., *State space flow structure induced by sliding motion control*, Bul. Inst. Politehnic Iaşi, **XXXV (XXXIV)**, s. III, pp. 25-29, 1989.

[44] Voicu, M., Moroşan, B. I., *Flow structure involving reaching of the sliding surface in variable structure systems*, Int. Conf. on El. Drives, *Univ. of Braşov*; Preprints, **G4**, 1991.

[45] Voicu, M., Moroşan, B. I., *Variable-structure controller for feedback positioning of d.c. electric motor*, Rev. Roum. Sci. Techn., ser. Electrot. Energet., **36**, no. 2, pp. 12-24, 1991.

[46] Voicu, M., Moroşan, B. I.,*Flow invariance based characterization of quasi-sliding motion*, International Workshop on Applied Automatic Control, *Czech Technical University of Prague*; Preprints, pp.86-89, 1993.

[47] Voicu, M., Moroşan, B. I., *Variable-structure control characterization and design via flow-invariance method*, Int. J. Automation Austria, **5**, no. 1, pp.12-24, 1997.

[48] Moroşan, B. I., *On the sliding conditions in sampled-data systems*, Bul. Inst. Politehnic Iaşi, **XXXVII (XLI)**, s. VI, pp. 18-22, 1991.

[49] Moroşan, B. I., *Quasi-sliding mode with integral control for a second order uncertain plant*, Int. Symp. Aut. Control & Comp. Sci., *Techn. Univ. of Iaşi*; Preprints, pp. 1-6, 1993.

[50] Moroşan, B. I., *Variable structure systems analysis with* "MATHEMATICA", Workshop on Computer Sci. Topics for Control Engrg. Educ., *Vienna University of Technology*; Proceedings SEFI, no. 10 (ISBN 2 87352 012 4), Bruxelles, 1993.

[51] Moroşan, B. I., *Discrete quasi-sliding mode control for coupled tanks system*, Workshop on Automation and Control Technology Educ. 2001, *Vienna University of Technology*; Proceedings SEFI, no. 2 (ISBN 2 87352 005 1), Bruxelles, 1993.

[52] Moroşan, B. I., *Fuzzy control in quasi-sliding mode systems*, Int. Workshop on Fuzzy Technology in Automation and Intelligent Systems – "Fuzzy Duisburg '94", *University of Duisburg*; Proceedings, pp. 155-161, 1994.

[53] Moroşan, B. I., *Digital variable-structure systems studied by flow-invariance method*, Doctoral thesis (in Romanian), *Technical University of Iaşi*, 1994.

[54] Moroşan, B. I., VOICU, M., *Flow-invariance method in general sliding mode system*, Technical Report, *Delft University of Technology*, 1993.

[55] Moroşan, B. I., VOICU, M., *General sliding mode systems analysis and design via flow-invariance method*, Studies in Informatics and Control, **3**, no. 4, pp.347-365, 1994.

[56] VOICU, M., PĂSTRĂVANU, O., *Flow-invariance method in systems theory* (in Romanian), Seminarul "Stabilitatea şi robusteţea sistemelor automate", *Academia Română, Filiala Iaşi*, 12 p., 21-22.01.2000.

[57] Ursescu, C., *Caratheodory solutions of ordinary differential equations on locally compact sets in Fréchet spaces*, Preprint Series in Math., *University of Iaşi*, 18, pp. 1-27, 1982.

[58] Voicu, M., PĂSTRĂVANU, O., *Componentwise asymptotic stability inducced by symmetrical polyhedral time-dependent constrants*, IFIP Conf. "Analysis and Optim. Diff. Syst.", Sept. 2002, Constanţa (accepted for presentation).

[59] Păstrăvanu, O., VOICU, M., *Infinity norm in analysis of componentwise asymptotic stability*, IFIP Conf. "Analysis and Optim Diff. Syst.", Sept. 2002, Constanţa (accepted for presentation).

[60] Šiljak, D. D., *When is a complex ecosystem stable?* Math. Biosci., **25**, pp. 25-50, 1975.

[61] Šiljak, D. D., *Connective stability of competitive equilibrium*, Automatica, **11**, pp. 389-400, 1975.

[62] Sezer M. E., Šiljak D. D., *On stability of interval matrices*, IEEE Trans. on Aut. Control, **AC-39**, pp. 368-371, 1994.

[63] Chen J., *Sufficient condition on stability of interval matrices: connections and new results*, IEEE Trans. on Aut. Control, **AC-37**, pp. 541-544, 1992.

[64] Păstrăvanu, O., VOICU, M., *Dynamics of a class of uncertain nonlinear systems under flow-invariance constraints*, Int. J. Math. and Math. Scie. (accepted for publication).

[65] Tschernikow, S. N., *Lineare Ungleichungen*, Deutscher Verlag der Wiss., Berlin, 1971.

[66] Barbu, V., Precupanu, Th., *Convexity and optimisation in Banach spaces*, Editura Academiei, Bucureşti, Reidel, Dordrecht, 1986.

[67] Utkin, V. I., *Sliding modes and their application in variable structure systems*, MIR, Moscow, 1978 (English transl. of Russian edition, Nauka, Moscow, 1974).

[68] Utkin, V. I., *Sliding mode in control optimisation*, Springer-Verlag, Berlin, 1992.

ADDENDUM

BRIEF HISTORY OF THE AUTOMATIC CONTROL DEGREE COURSE AT TECHNICAL UNIVERSITY "GH. ASACHI" OF IAŞI

Corneliu Lazar, Teohari Ganciu, Eugen Balaban, Ioan Bejan

Technical University "Gh. Asachi" of Iaşi

Dept. of Automatic Control and Industrial Informatics

e-mail: {clazar, tganciu, ebalaban}@ac.tuiasi.ro

Abstract: The Faculty of Automatic Control and Computer Engineering, one of the leading training and research centre in the field of automatic control and computer engineering from Romania, during the university year 2002 – 2003 has celebrated 40 years from the appearance of the first courses in the above mentioned field and 25 years from the emergence of the Automatic Control and Computer Engineering degree course. On this occasion, some of the most significant scientific achievements have been put together in this article, as a brief history of the Automatic Control degree course at Technical University "Gh.Asachi" of Iaşi.

1. Origin (the end of 50s– 1977)

Control engineering saw rapid development in many countries in the period immediately following the Second World War. Engineers and scientists concerned with control problems have formed new professional groupings and university courses dedicated to this subject have arisen. At the same time, research groups have been set up both in the industrial and in the academic communities.

In the above context, control engineering has started at the Technical University "Gh. Asachi" of Iaşi in the 50s. The Department of Electrical Drives from the Electrical Engineering Faculty was approaching issues in control engineering, which were introduced as chapters in the courses "Electrical Drives" and "Electromechanical Equipment". The first course actually entitled "Automation" was an optional course and it was initiated in the same period by professor Nicolae Boţan. As a result of their didactic and research activities, Nicolae Botan, Ioan Bejan and Eugen Balaban published the first book, named "Eletromechanical Drives and Automation", in the field of automatic control at "Editura didactică şi pedagogică", Bucureşti in 1962.

Other compulsory courses have been established prior to 1962, which were mainly focused on the requirements of electromechanical engineering and electrical power engineering programs. This was the case with the course of "Automation and Remote Control" (within the two degree courses) and the course on "Relay protection" (for the electrical power engineering program) taught by professors Leopold Sebastian and Ioan Bejan. In 1967, Ioan Bejan and Gherghina Balaban published the first course in control, entitled "Automation and Remote Control" at the publishing house of the University of Medicine and Pharmacy "Gr. T. Popa" of Iaşi.

In the 60s and the beginning of the 70s, the course of "Automation and Remote Control" has known a rapid development, e.g. the course "Automatic Control" (Leopold Sebastian) started at the electromechanical engineering program and the course "Automation of Electrical Power Systems" (Ioan Bejan) was introduced for the electrical power engineering program. New optional courses have also simultaneously appeared for electrical engineering, such as: "Logical Circuits and Sequential Control Systems" (Corneliu Hutanu), "Servomechanisms" (Iosif Olah), "Computer Controlled Processes" (Simona Caba) and "Advanced Automation" (Gherghina Balaban and Iosif Olah).

Under the supervision of professors Ioan Bejan and Leopold Sebastian extensive research have been carried out at the Department of Electrical Drives ranging from control theory problems (nonlinear systems, identification, adaptive and optimal control, controller tuning) to the application of control methods to areas of thermal processes, electrical drives, electrical power systems, relay protection systems, servomechanisms or machines tools control. The above mentioned professors have initiated PhD positions in Industrial Automation (Leopold Sebastian – 1966) and in Electrical Power System Automation (Ioan Bejan – 1972). Then, two research groups have emerged, headed by Leopold Sebastian for the electrical engineering program (Eugen Balaban, Corneliu Hutanu, Iosif Olah, Corneliu Botan, Teohari Ganciu and Simona Caba) and Ioan Bejan for the electromechanical power engineering programme, respectively (Gherghina Balaban, Ioan Titaru, Mihail Voicu, Cristea Pal and Dumitru Asandei).

Due to the fact that these groups were at that time the only ones offering courses in automatic control, one of the main tasks was from the very beginning to write manuals and monographs. In the 70s, several well-known books have been published in Romanian, e.g.:

- L. Sebastian, "Automatic Control", Editura didactică şi pedagogică, Bucureşti, 1975;

- I. Bejan, "Automation and remote control of electrical power systems", Editura didactică şi pedagogică, Bucureşti, 1976;
- I. Bejan, "Magnetic amplifier for control systems", Editura Tehnică, Bucureşti, 1972;
- N. V. Boțan, "Speed control of electrical drives", Editura Tehnică, Bucureşti, 1974;
- N. V. Boțan, "Electrical drives control", Editura Tehnică, Bucureşti, 1977.

In the above mentioned period, PhD degrees in the field of automatic control have been obtained by Mihail Voicu, Corneliu Botan, Gherghina Balaban and Iosif Olah.

Also in the same period, the members of the automatic control group were awarded the following prizes:

- the Ministry of Education prize - Ioan Bejan, Nicolae Botan, Leopold Sebastian, Eugen Balaban, Gherghina Balaban and Mihail Voicu;
- professor emeritus awarded by the Ministry of Education – Ioan Bejan, Eugen Balaban and Iosif Olah.

2. First automatic control program (1977 – 1990)

In 1977, a five years course entitled "Automatic Control and Computer Engineering" has been developed within the Faculty of Electrical Engineering. This came as a response to the demands from industry, which began to require well prepared engineers in the fields of control and computer engineering.

A study program of two years began in 1977. There were two groups of first-year students which took entrance examination at the new degree course and other two of second-year students transferred from the electrical and electronic engineering programs. Each group of students had separate curricula, one for Automatic Control within the Electrical Drives Department and the other for Computer Engineering functioning at the Electronic Department.

As it has been mentioned in the previous section, at the Electrical Drives Department there were already teaching staff and research laboratories in the field of control engineering. The control team elaborated the first Automatic Control curriculum that comprised courses of System theory, Digital control systems, Analogue control systems, System identification, Hydraulic and pneumatic control equipment, Control system design, Systems and equipment for process control and Optimal control. The scientific research of the automatic control staff has known a significant progress characterized by several research projects, the publishing of monographs and participation at international conferences.

In the 80s, the following new areas of research have appeared: flow-invariance in control theory, computer controlled processes, machine vision, pattern recognition, computer aided control engineering and robotics. In 1983 professor Corneliu Hutanu's book "Digital circuits and sequential control systems" was published at "Junimea" publishing house from Iaşi and professor Mihail Voicu published an important monograph at "Editura Tehnică, Bucureşti" in 1986 entitled "Stability analysis techniques for control systems".

The deep crisis experienced by Romania as a country in the 80s particularly affected higher education and inevitably, also the automatic control teaching staff, which faced serious problems. For example only three teaching assistants were admitted as PhD students in that period in the field of automatic control. It was extremely difficult, from an administrative point of view, to publish abroad or to participate at international conferences. However, under these severe conditions, 10 papers have still been published in important foreign journals by Mihail Voicu, Leopold Sebastian, Octavian Pastravanu and Teohari Ganciu and 9 papers appeared in the proceedings of international conferences. In this respect, the participations of professor Mihail Voicu at the 9th IFAC World Congress in Budapest (1984) and at the 10th IFAC World Congress in Munich (1987) can be considered as remarkable achievements. Also, Octavian Pastravanu has presented papers at the international conferences "Symposium on Systems Science IX" organized by University of Wroclaw (1986), "European Congress of Simulation" organized by Czechoslovak Academy of Science in cooperation with IMACS at Prague (1987) and the 4th International Symposium on Systems Analysis and Simulation organized by DDR Academy of Science in cooperation with IMACS at Berlin (1988), and Corneliu Lazar has presented a paper at "The 3rd International Conference on Automatic Image Processing" organized by Scientific Technological Society for Measurement and Automatic Control and DDR Academy of Science at Leipzig (1989).

The main difficulties in teaching and research activities were caused by the lack of computer facilities. Notable efforts have been done by the teaching assistants Octavian Pastravanu and Corneliu Lazar in the mid 80s, which were supported by the dean of the faculty, professor Ioan Bejan and by professor Mihail Voicu, in order to achieve proper computer equipment and software and to develop laboratories for computer aided control engineering.

Beginning with 1987, the control group formed within the teaching staff of the Electrical drives Department organized every two years the national scientific symposium "Structures, Algorithms and Equipment for Process Control".

In the mid 80s, professor Teohari Ganciu concentrated his efforts on the foundation of an important research centre – the Iaşi branch of the "Automation Design Institute".

3. The Faculty of Automatic Control and Computer Engineering

As a direct consequence of the profound changes experienced by the Romanian nation in 1989, the Romanian higher educational system has known a lively development. The Faculty of Electrical Engineering has split in 1990 in three faculties. One of these, the Faculty of Automatic Control and Computer Engineering has been founded at the initiative and due to the efforts of the teaching staff from the degree course "Automatic Control and Computer Engineering" of the former Electrical Engineering Faculty. Professor Corneliu Hutanu was the first dean of the faculty from 1990 to 1992. From the beginning, the faculty had two departments, Automatic Control and Industrial Informatics and Computer Engineering, each of them offering the following degree courses: Automatic Control and Industrial Informatics and Computer Engineering, respectively.

The first head of the Automatic Control and Industrial Informatics department, professor Mihail Voicu, together with the department staff began in 1990 to develop a new curriculum in Automatic Control and Industrial Informatics. It must be noted that this curriculum was also influenced by a consultative council of the Automatic Control professors from Romania in order to maintain certain compatibilities between similar curricula introduced in other university centers of the country. Starting with the beginning of the 90s, this curriculum has been changed and improved on a yearly basis, also based on the knowledge and the experience of other European universities, with which several contacts have been established in the framework of EU programs.

Due to the efforts of Mihail Voicu and Octavian Pastravanu, TEMPUS Joint European Projects (JEP) have been developed in collaboration with the Control Engineering Departments of other European universities, e.g. JEP 0886/1990 "Higher Education in Control Engineering", JEP 02011/1991 "Improvement in Automatic Control Technologies", JEP 07101/1994 "Development in Romania of Short-Term Higher Education in Computing, Centered on Distributed Processing and Its Application", MJEP 11467/1996 "EU Compatible Training in Industrial Automation" – COMPANION. The last TEMPUS project of the 90's, UM-JEP 13133/98 "Quality Management", has been managed by Alexandru Onea and it has resulted in a significant contribution to the implementation of the quality assurance system of our faculty.

These projects also offered a good opportunity for establishing relationships between our department and other European universities, which materialized in the participation of all the teaching staff to workshops organized by the JEP members and dedicated to Control Engineering Education and in the acquisition of modern laboratory setup. Due to the JEP framework, most of the teaching staff and especially young PhD students had the possibility to attend training stages at the partner universities.

Based on TEMPUS programs, the “traditional” approach towards teaching has been modified substantially. The faculty has accepted and implemented the European Credit Transfer System (ECTS) starting with the academic year 1998 – 1999 and has made our teaching procedures compatible with similarly oriented universities in the world. This makes possible the exchange of students and academic staff and mutual recognition of qualifications.

The European Credit Transfer System formed the basis of future collaborations that continued after the end of the TEMPUS program. Thus, with some of the partner universities, the program Socrates-Erasmus began to grow at the end of the 90s. In the framework of this program, each year students from our department worked on the diploma project at the following universities: University of Gent - Department of Electrical energy, Systems & Automation, University of Sheffield - Department of Automatic Control and System Engineering, University of Duisburg - Department of Measurement and Control Engineering, Technical University of Vienna, Université Joseph Fourier – Laboratoire d’Automatique de Grenoble. Within this program, an important number of MSc and PhD students had also training stages. Since 1999 till now, in the framework of the Socrates – Erasmus program, professor Robin De Keyser from the University of Gent has taught each year a module of the Predictive Control Systems course for the MSc program of our department. At the same time, professors Corneliu Lazar and Octavian Pastravanu have taught mini-courses on Predictive Control and Process Modelling Using Bond Graph, respectively, in the last two years at the University of Ghent.

Together with the development of a new curriculum for the Automatic control program, at the beginning of the 90s, new teaching staff has been recruited from the research institutes and young graduates. In 1992 professor Mihail Voicu, who had a strong experience in managing research and teaching staff, has become the dean of the faculty. Since the same year, professor Ganciu has been the head of the department of Automatic Control and Industrial Informatics.

Also, it must be mentioned that due to the changes that took place in Romania at the beginning of 90s, besides professors Bejan and Sebastian, the professors Voicu, Hutanu, Balaban, Botan and Olah also have received the right to be PhD supervisors in Automatic Control. Thus, new research areas have appeared and a greater number of graduates in Automatic Control have become PhD students, which ultimately led to an increased research activity. As a result of this, the number of scientific works published in journals and at international conferences and congresses has also increased considerably.

At the end of the 90s, what seemed to be a “natural” development took place, and the research groups from the department of Automatic Control and Industrial Informatics merged and they formed the Automatic Control and Applied Informatics (ACAI) Research Centre accredited by Ministry of Education and Research classified in the C category. ACAI is an interdisciplinary research centre managed by professor Mihail Voicu within the faculty of Automatic Control and Computer Engineering and its mission is to fulfill statements of the university and of the faculty within the field of Systems and Control Engineering by creating and sustaining a world class research group. The main directions of the scientific research are: System Theory, Robotics and CIM Optimal and Predictive Control, Artificial Intelligence in Process Control, Systems Identification and Fault Detection, Microprocessor Based Control Systems, CAD for Dynamic Systems. The activity of ACAI has been carried out in grants with the Ministry of Education and Research and industrial companies (grant directors: Mihail Voicu, Corneliu Botan, Corneliu Lazar). The ACAI research staff had a very productive period from 2000 to 2003, publishing 15 monographs and courses, 14 journal papers among which 16 in ISI journals and 121 conference papers.

Since 1998, the department of Automatic Control and Industrial Informatics offered a short cycle degree program of 3 years on Automation Equipment. The undergraduate program of 3 years curriculum represents in a way a compromise between the necessity for a graduate with a diploma from the Engineering College to be operational in his job and the possibility for him to go on with more advanced studies.

After graduating the long cycle degree program, the students who wish to continue their studies can choose to apply for master program – one year of specialization – which is organized in the area of Automatic Control, having the subjects: distributed parameter control systems, parallel programming algorithms and techniques, predictive control, parameter estimation, distributed control and artificial intelligence in control.

In September 1998, the faculty has been moved in a new building having 7700 m^2 useful area with 2 amphitheatres, 7 lecture rooms, 30 laboratories and 32 offices for the teaching staff. The two departments of the faculty develop their activity in the new building beside the library and the Communication Centre of Technical University “Gh. Asachi” of Iaşi. It is important to say that the building has been started off at the beginning of the 90s at the initiative of professor Mihail Voicu, supported by professor Ioan Bejan to take the initial necessary steps at the Ministry of Education.

Unfortunately, the 90’s economic decline of Romania created serious problems regarding higher education financing. As a result of this, several young people from the teaching staff left the department trying to fulfill their professional careers in more developed countries. At the same time serious difficulties have appeared in providing facilities for teaching and research work. However, the department members could adapt to the new forms of financing by having access to external funds, mainly from the following sources: European Trading Foundation – TEMPUS JEPs, Higher Education Financing Council of Ministry of Education, World Bank, Romanian Government and industrial companies.

Thus, during 1998 – 2003 the following five major projects have been financed by the World Bank and Romanian Government: System of Integrated Laboratories for Studying CIM (director Mihail Voicu – 193000 USD), Training Laboratory in the Field of Computer Aided Process Control (director Octavian Pastravanu – 150000 USD), Laboratory for Electrical Drives Control (director Corneliu Botan – 150000 USD), Integrated Laboratory for Studying, Designing and Implementation of Digital Structure for Process Control (director Teohari Ganciu – 150000 USD), Upgrading of the Short Cycle Degree Program on Automation Equipment (director Alexandru Onea – 35000 USD). There were also 2 individual projects only for the equipment acquisition: Microprocessor Based Control Systems (director Corneliu Hutanu – 5000 USD) and Digital Controller for Process Control (director – Corneliu Lazar – 5000 USD). These funds allowed the development of new laboratories and the update of most of the existent laboratories at our department with the following major facilities for teaching and research work:

- *Flexible Manufacturing System* containing two ABB robots (IRB 1400 and IRB 2400), machine tool (EMCO PC Mill), conveyor (FlexLink), computer vision system (OptiMaster), CAD system (8 PC stations); Robot Soccer System containing 8 MiaBot mobile robots and a computer vision system; Androtec mobile robot;
- *Process control setups*: FieldPoint Distributed Control System (National Instruments) for the distributed control of industrial processes, PROCON process control trainers for level, flow and temperature (Feedback), LEYBOLD and ELWE electrical drive control trainers, Twin Rotor MIMO System (Feedback), Moeller PLCs, Laboratory kits for teaching microprocessor based systems;
- *Laboratory installation for making printed circuit boards* (Lpkf Germany);
- *Computers*: 45 computers (IBM compatible) with operating systems and basic software, 5 analogue COMDYNA computers, 7 data acquisition cards – National Instruments ATMIO16E10 with starter Kit and related drivers;
- *Software*: CATIA V5R8 (CAD); Robot Studio (robot simulation software), Eclipse, RT++, AgentOCX, LPA Prolog, Flex, Agent Toolkit (artificial intelligence software), Sucosoft V5 (PLC software), Discovery computer control aided learning software, MATLAB- Simulink 6.0, HMI/SCADA software Lookout, Cadence software (OrCAD) for PCB design.

The above funds have also been used to purchase important textbooks in the field of automatic control for the faculty library. They also contributed in the publishing of monographs and university manuals, the development of new courses in Automatic Control degree program, the finalization of PhD thesis and the participation at international conferences.

Based on the experience accumulated from the last TEMPUS project, Alexandru Onea initiated and led other two projects having the same theme, quality assurance in higher education: MATRA project – "Developing the national strategy in the field of quality assurance in higher education in Romania, financed by EU through the Dutch Government and the Leonardo da Vinci project – "Training in quality management system for information technology in higher education".

In the period from 1990 until the present time the teaching staff of the department published 22 monographs, 31 manuals and courses and it largely participated in many international conferences both in the country and abroad. As a remarkable achievement, it can be mentioned the contribution to each edition of the European Control Conference from the beginning until now of professors Mihail Voicu and Octavian Pastravanu as well as the participation of the lecturers Letitia Mirea and Lavinia Ferariu at the 15th IFAC World Congress in Barcelona 2002.

After the foundation of the EU Control Association (EUCA), part of the research developed within our department became known for the scientific community of Automatic Control by a number of papers published in different editions of the European Control Conference (ECC). These papers focus on the following topics (i) AI techniques in identification and diagnosis (Marcu and Voicu - ECC'93, ECC'95; Marcu and Ferariu - ECC'99; Marcu and Matcovschi - ECC'99; Marcu and Mirea - ECC'01; Ferariu - ECC'03); (ii) Dynamics of systems with unknown parameters (Voicu and Pastravanu - ECC'95; Pastravanu and Voicu - ECC'97); (iii) Time-dependent invariant sets and componentwise stability (Pastravanu and Voicu - ECC'99, ECC'01, ECC'03; Matcovschi and Pastravanu - ECC'03).

Due to his outstanding contributions professor Mihail Voicu was elected correspondent member of the Romanian Academy of Technical Sciences in 1997 and from 1998 he becomes senior member of IEEE. Also, he has received the "Aurel Vlaicu" prize of the Romanian Academy for the year 1987 (awarded 1990) for the papers:

- M. Voicu, Observing the State with Componentwise Exponentially Decaying Error. Systems & Control Letters **9** (1987), pp. 33 – 42.
- M. Voicu, On the Application of the Flow-Invariance Method in Control Theory and Design. 10th World Congress of International Federation of Automatic Control, Munich, July 26–31, 1987. Preprints, vol. VIII, pp. 364–369.

Starting with 2001 the department offered a new degree course on Industrial Informatics and from 2002 the degree course on Automatic

Control and Industrial Informatics became Automatic Control and the short cycle degree course Automation Equipment was renamed Automation.

The numerous contacts with the western universities permitted the change of the national scientific symposium of the faculty in the international symposium on Automatic Control and Computer Science organized every two years, namely in 1993 and 1995 and every three years afterwards.

Since 1991, owing to the initiative of professor Mihail Voicu, the Bulletin of the Polytechnic Institute of Iaşi has a new fascicule (1-4) with Section 4 dedicated to Automatic Control and Computer Engineering, editors being from the beginning till now Octavian Pastravanu, Doru Panescu and Alexandru Onea. This journal allowed the specialists in the automatic control field to point out the results of their scientific research activity.

At the present time, the Automatic Control degree program at the Technical University "Gh. Asachi" of Iaşi is carried out by the following teaching staff, of the Automatic Control and Industrial Informatics Department from the Faculty of Automatic Control and Computer Engineering:

- *Professors*: Eugen Balaban, Ioan Bejan (honorary professor), Corneliu Botan, Teohari Ganciu (head of the Department), Corneliu Hutanu, Corneliu Lazar (scientific chancellor of the Faculty Council), Iosif Olah, Cristea Pal (head of the Iaşi branch of the Romanian Society of Automatic Control and Technical Informatics), Doru Panescu (deputy head of the Department), Octavian Pastravanu, Leopold Sebastian (honorary professor) Mihail Voicu (dean of the Faculty)
- *Associate Professors*: Stefan Dumbrava, Lucian Mastacan, Mihaela Matcovschi, Alexandru Onea (editor of the Automatic Control Section of the Bulletin of the Polytechnic Institute of Iaşi), Andrei Pricop, Gabriela Varvara
- *Lecturers*: Catalin Calistru, Lavinia Ferariu, Letitia Mirea, Florin Ostafi, Mihai Postolache
- *Teaching Assistants*: Alina Barabula, Catalin Braescu, Laurentiu Boboc, Vasile Dorin, Catalin Dosoftei, Claudiu Lefter, Laurentiu Marinovici, Bogdan Mustata, Cristina Tugurlan
- *Junior Teaching Assistants*: Sorin Carari, Marius Kloetzer, Mircea Lazar, Cristian Mahulea

SECS 754
1-4020-7607-X