An Introduction to the Special Issue on Time Delay Systems: Modelling, Identification, Stability, Control and Applications

LIBOR PEKAŘ [1], FILIPPO NERI [2]
[1] Department of Automation and Control Engineering, Faculty of Applied Informatics,
Tomas Bata University in Zlín, Zlín, CZECH REPUBLIC,
[2] Editor in Chief of WSEAS Transactions on Systems,
Department of Computer Science, University of Naples "Federico II", Naples, ITALY
si27v06@gmail.com, nerifil@gmail.com

The special Issue “Time Delay Systems: Modelling, Identification, Stability, Control and Applications” intents to collect high-erudite papers aiming theoretical and/or practical matters dealing with time delay systems.

In feedback control systems, delay as a generic part of many processes is the phenomenon which unambiguously deteriorates the quality of a control performance. Modern control theory has been dealing with this problem since its nascence – the well known Smith predictor has been known for longer than five decades. Systems with delays in technological and other processes are usually assumed to contain delay elements in input-output relations only, which results in shifted arguments on the right-hand side of differential equations. All the system dynamics is hence traditionally modelled by point accumulations in the form of a set of ordinary differential equations. However, this conception is a rather restrictive in effort to fit the real plant dynamics since inner feedbacks can be of time-distributed or delayed nature. Time delay (hereditary, anisochronic) models, contrariwise, offer a more universal dynamics description applying both derivatives and delay elements on the left-hand side of a differential equation, either in a lumped or distributed form.

Modelling, identification, stability analysis, stabilization, control, etc. of time delay systems are challenging and fascinating tasks in modern systems and control theory as well as in academic and industrial applications. Many related problems are unsolved and many questions remain unanswered.

The aim of this special issue is to highlight greatly significant recent developments on the topics of time delay systems, their estimation, modelling and identification, stability analysis, various (algebraic, adaptive and predictive) control strategies, relay-based autotuning and interesting academic and real-life applications. The papers included in the special issue are ordered from the most theoretical one to more directly applicable ones.

The state estimation or filtering problem is of great importance in both theory and application, and in the last decades, this problem has gotten extensive concern and many solution schemes have been proposed and successfully put into action. Among them, Kalman filtering, which minimizes the variance of the estimation error, is the most famous one. Observer design for linear time delay systems is the matter of the first paper of Mohammad Ali Pakzad, entitled “Kalman Filter Design for Time Delay Systems”. An easy way to compute least square estimation error of an observer for time delay systems is derived, where the time delay terms exist in the state and output of the system. Based on the least square estimation error an optimization algorithm to compute a Kalman filter for time delay systems is proposed. By employing the finite characterization of a Lyapunov functional equation, the existence of sufficient conditions for obtaining the right solution and guaranteeing the proper convergence rate of the estimation error is evaluated. It is shown that this finite characterization can be calculated by means of a matrix exponential function. The desirable performance of the proposed observer is demonstrated through the simulation of several numerical examples.

Fractional differential equations have gained considerable significance and most fractional systems may contain a delay term. The authors Mohammad Ali Pakzad and Sara Pakzad propose an exact method for the BIBO stability analysis of a large class of fractional order delay systems. In their paper entitled “Stability Map of Fractional Order Time-Delay Systems”, the stability robustness for linear time invariant fractional order systems with time delay against delay uncertainties is considered. The complexity arises due to the exponential type transcendental terms and fractional order in their characteristic equation. It is shown that this procedure numerically reveals all possible stability regions exclusively in the space of the delay. Using the approach presented in this study, all the locations where roots cross the imaginary axis are found. Finally, the concept of stability as a function of time delay is described for a general class of linear fractional order systems with multiple commensurate delays.

A revision and extension of the ring of retarded quasipolynomial meromorphic functions for description and control of time-delay systems is the aim of the paper “A Ring for Description and Control of Time-Delay Systems” by Libor Pekař. The new definition extends the usability to neutral systems and to those with distributed delays. First, basic algebraic notions useful for this paper are introduced. A concise overview of algebraic methods for time delay systems follows. The original and the revised definitions of the ring together with some its properties finish the contribution. Many illustrative examples that explain introduced terms and findings can be found throughout the paper.

Algebraic design of controllers for time delay systems having integrative or unstable properties is the topic of the paper entitled “Control of Unstable and Integrating Time Delay Systems Using Time Delay Approximations” by Petr Dostál, Vladimír Bobál and Zdeněk Babík. The proposed method is based on two methods of time delay approximations. The control system with two feedback controllers obtained via the polynomial approach and the linear quadratic technique is considered. Resulting continuous-time controllers ensure asymptotic tracking of step references as well as step disturbances attenuation.

The majority of processes in the industrial practice have stochastic characteristics and eventually they exhibit nonlinear behaviour. Traditional controllers with fixed parameters are often unsuitable for such processes because parameters of the process change. One possible alternative for improving the quality of control of such processes is application of adaptive control systems. The authors Vladimír Bobál, Petr Chalupa, Marek Kubalčík and Petr Dostál designed a toolbox in the MATLAB/SIMULINK environment for identification and self-tuning control of time delay systems in the paper entitled “Identification and Self-tuning Control of Time-delay Systems”. The control algorithms are based on modifications of the Smith predictor. The designed algorithms that are included in the toolbox are suitable not only for simulation purposes but also for implementation in real time conditions. Verification of the designed toolbox is demonstrated on a self-tuning control of a laboratory heat exchanger in simulation conditions.

A rather similar problem is the issue of the next paper entitled “Predictive Control of Higher Order Systems Approximated by Lower Order Time-Delay Models”. There often occur higher order processes in technical practice, the designed optimal controllers for which lead to complicated control algorithms. One of possibilities of solving the problem is their approximation a by lower-order model with dead time. The contribution of Marek Kubalčík and Vladimír Bobál is focused on a choice of a suitable experimental identification method and excitation input signals for an estimation of process model parameters with time-delay. One of the possible approaches to control of time-delay processes is the application of model-based predictive control methods. Design of an algorithm for predictive control of high-order processes which are approximated by second-order model of the process with time-delay then follows in the paper.

The last paper in the special issue: “Autotuning Principles for Time-delay Systems” by Roman Prokop, Jiří Korbel and Radek Matušů focuses single input–output principles for tuning of continuous-time controllers used in autotuning schemes. Autotuners represent a combination of relay feedback identification and some control design method. In this contribution, models with up to three parameters are estimated by means of a single asymmetrical relay experiment. Then a stable low order transfer function with a time delay term is identified by a relay experiment. Controller parameters are analytically derived from general solutions of Diophantine equations in the ring of proper and stable rational functions. This approach covers a generalization of PID controllers and enables to define a scalar positive parameter for further tuning of the control performance. The analytical simple rule is derived for aperiodic control response. Simulations are performed in the Matlab environment and a toolbox for automatic design was developed.

This special issue is created by mathematicians, system- and control-engineers and scientists who study various problems in analysis and control of dynamical systems including the very rich family of those with time delays. We are looking forward to hearing reactions and comments from you, the reader, as you engage in your design struggles and successes as well. We hope this special issue can support the designing of new communities and facilitating new groups of learners, engineers and scientists as well. We dare to claim that this issue made a little contribution to a better, feel-good, world.

Stability Map of Fractional Order Time-Delay Systems

MOHAMMAD ALI PAKZAD [1], SARA PAKZAD [2]
[1] Department of Electrical Engineering, Science and Research Branch, Islamic Azad University
Tehran, IRAN
m.pakzad@srbiau.ac.ir , http://kiau.ac.ir/~electronics
[2] Departments of Electrical Engineering, South Tehran Branch, Islamic Azad University
Tehran, IRAN
spakzad@pedc.ir

Abstract: - In this paper, the stability robustness is considered for linear time invariant (LTI) fractional order systems with time delay against delay uncertainties. The complexity arises due to the exponential type transcendental terms and fractional order in their characteristic equation. We show that this procedure numerically reveals all possible stability regions exclusively in the space of the delay .Using the approach presented in this study, we can find all the locations where roots cross the imaginary axis. Finally, the concept of stability as a function of time delay is described for a general class of linear fractional order systems with multiple commensurate delays. Several numerical examples are provided to demonstrate the effectiveness of the proposed methodology.

Key-Words: - Fractional order systems, time-delay systems, characteristic equation, stability, Root-Locus.

1 Introduction

It is known that the presence of delay many cause poor performance and/or instability in a dynamical system. Therefore, time-delayed systems play significant roles in theoretical as well as practical fields; and this influence can be observed in numerous research articles written on various problems that involve this class of systems [1-6]. Fractional differential equations have gained considerable importance due to their application in various sciences, such as viscoelasticity, electroanalytical chemistry, electric conductance of biological systems, modeling of neurons, diffusion processes, damping laws, rheology, etc. Fractional order differential equation is represented in continuous-time domain by differential equations of non integer-order. Moreover, time delay is often present in real processes due to transportation of materials or energy. Therefore, most fractional systems may contain a delay term, such as fractional order neutral systems or some other fractional order delay systems (see [20,22] and the references cited in it). The characteristic function of a fractional-delay system involves exponential type transcendental terms, so a fractional delay system has in general an infinite number of characteristic roots. This makes the stability analysis of fractional-delay systems a challenging task. However, for fractional order dynamic systems, it is difficult to evaluate the stability by simply examining its characteristic equation either by finding its dominant roots or by using other algebraic methods. At the moment, direct check of the stability of fractional order systems using polynomial criteria (e.g., Routh's or Jury's type) is not possible, because the characteristic equation of the system is, in general, not a polynomial but a pseudopolynomial function of fractional powers of the complex variable s.

The researchers of [7] and [8] may be the pioneer to consider stability of the fractional order time delay system with single-delay. They have developed the Ruth-Hurwitz criteria for analyzing the stability of some special delay systems to those involve fractional power $\sqrt{s}$.

One of the important and basic things about each of the dynamical systems is the stability investigation. With respect to systems with delay, we'd like to study that how this stability property could behave, if we increased the time delay. It is known that an interesting phenomenon, namely stability windows, might happen. There has been a large effort to deal with this problem, as can be seen by the large quantity of articles dealing with it for the standard case; see. [9, 10,21,23], and many others. Recently, [11] has used numerical methods to investigate this subject in fractional order delay systems.

Bonnet and Partington introduced necessary and sufficient conditions for BIBO stability of the retarded fractional order delay systems and sufficient conditions for some neutral types [12]. Also, they studied necessary and sufficient conditions for H_∞ -stability (a weaker notation than BIBO stability) for a significant case of neutral type fractional order delay systems [13]. From the numerical analysis point of view, the effective numerical algorithms to check BIBO stability of fractional order delay systems have been discussed in [11] and [14]. In [15], a numerical method with complicated calculations based on the Cauchy's integral theory has been proposed for testing the stability of such systems. Some numerical techniques based on the Lambert W function have been presented in [16, 17] for the investigation of stability of these systems, where the rightmost characteristic root on the first Riemann sheet can be expressed explicitly in terms of Lambert W function.

In the present work, an exact method for the BIBO stability analysis of a large class of fractional order delay systems has been proposed, which is able to determine all the possible stability regions in the parametric space of these systems as a function of time delay. The main idea of this strategy has been adopted from [11], whose authors have studied the stability of integer order time-delayed systems. The rest of the present article has been organized as follows. Section 2 includes the necessary explanations and assumptions and also the stability test of these systems. In section 3, an exact method which expresses the stability regions as a function of delay has been presented. In section 4, some examples that demonstrate the effectiveness of the proposed approach have been described; and section 5 concludes the article.

2 Problem Formulation

Standard notation has been used throughout the article. The set of natural numbers is denoted by $\mathbb{N}$, whereas $\mathbb{N}_N$ denotes the set of its first N elements (i.e., $\mathbb{N}_N = \{1,\ldots,N\}$). $\mathbb{C}\left(\mathbb{C}^+,\mathbb{C}^-\right)$ is the set of complex numbers (with strictly positive, and strictly negative real parts), and $j=\sqrt{-1}$ is the imaginary unit . For $z\in\mathbb{C}$, $\bar{z}$ denotes its complex conjugate, and $\angle z$, $\Re(z)$ and $\Im(z)$ define the argument (taken here from $\left(-\pi,\pi\right]$), the real part and the imaginary part of z . $\mathbb{R}\left(\mathbb{R}^+,\mathbb{R}^-\right)$ denotes the set of real numbers (larger or equal to zero, smaller or equal to zero).

Consider a fractional order system with the following characteristic equation:

$$C\left(s,\tau\right)=\sum_{k=0}^{N} p_{k}\left(s^{\alpha}\right)e^{-ks\tau} \tag{1}$$

where parameter τ is non-negative, such that $\tau\in\mathbb{R}^+$; $p_k\left(s^\alpha\right)$ for $k\in\mathbb{N}_N$ are polynomials in s^α (where $\alpha\in\left(0,1\right)$). Note that the zeros of characteristic equation (1) are in fact the poles of the system under investigation. We find out from [14] that the transfer function of a system with a characteristic equation in the form of (1) will be H_∞ stable if, and only if, it doesn't have any pole at $\Re(s)\geq 0$ (in particular, no poles of fractional order at $s=0$).

For fractional order systems, if a auxiliary variable of $\varsigma=s^\alpha$ is used, a practical test for the evaluation of stability can be obtained. By applying this substitution in characteristic equation (1), the following relation is obtained:

$$C_\varsigma\left(\varsigma,\tau\right)=\sum_{k=0}^{N} p_k\left(\varsigma\right)e^{-k\tau\varsigma^{1/\alpha}} \tag{2}$$

For this new variable, the stability region of the original system is not expressed as the right half-plane, but as the region described below:

$$\left|\angle\varsigma\right|\leq\frac{\pi\alpha}{2} \tag{3}$$

with $\varsigma\in\mathbb{C}$, which the stable region has been displayed by shaded regions in Figure (1).

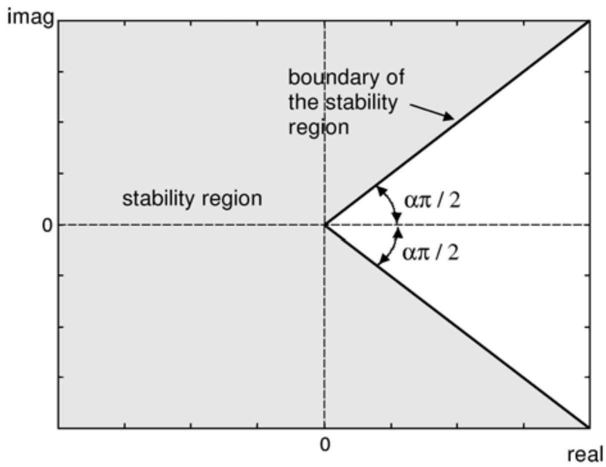

Fig. 1. The Stability Regions (Shaded) for linear fractional order systems

Note that under this transformation, the imaginary axis in the s-domain is mapped into the line

$$\angle\varsigma = \pm\frac{\pi\alpha}{2} \tag{4}$$

in the ς-domain, and therefore a solution $\varsigma^* = \left|\varsigma^*\right| \angle \pm \pi\alpha/2$ implies that the original system has a purely imaginary solution of the type

$$s^* = \pm j \left|\varsigma^*\right|^{\alpha^{-1}} \tag{5}$$

Let us assume that ; $s = \pm j\,\omega$ or in other words, $s = \omega\, e^{\pm j\pi/2}$ are the roots of characteristic equation (1) for a $\tau \in \mathbb{R}^+$. Then for the auxiliary variable, the roots are defined as follows:

$$\varsigma = s^{\alpha} = \omega^{\alpha}\, e^{\pm j\pi\alpha/2} \tag{6}$$

Therefore, with the auxiliary variable $\varsigma = s^{\alpha}$, there is a direct relation between the roots on the imaginary axis for the s-domain with the ones having argument $\pm\pi\alpha/2$ in the ς-domain.

3 Problem Solution

This is well known that one of the most important tools in the stability analysis of the systems is to use the location of the roots of system or root-locus method. About fractional-order delay systems, since there is exponential type transcendental terms in the characteristic equation and also because of having the fractional feature of the system, drawing the root-locus of the fractional delay systems makes the problem much more challenging compared to integer order systems. Some of the researchers have tried to present some algorithms to be able to draw the location of the roots. As an example, J.A. Tenreiro Machado in [19] and Fioravanti et al in [11]. But unfortunately, none of these algorithms could not properly display the root-locus of the fractional delay systems. We've tried to develop the existed algorithms, which are known in the resources, and also to draw the stability map and location of the roots correctly.

3.1 Crossing position

The main objective of this section is to present a new method for the evaluation of stability and determination of the unstable roots of a fractional order time delay system. A necessary and sufficient condition for the system to be asymptotically stable is that all the roots of the characteristic equation (1) lie in the left half of the complex plane.The proposed method eliminates the transcendental term of the characteristic equation without using any approximation or substitution and converts it into a equation without the transcendentality such that its real roots coincide with the imaginary roots of the characteristic equation exactly.

Based on the D-subdivision method, the number of unstable roots of a characteristic equation is invariant in some distinct regions of one-dimensional parameter space of time delay and the characteristic equation has at least one pair of purely imaginary roots at the boundary of these regions. After finding the boundaries and calculating the direction of imaginary roots variation, the number of unstable roots in each distinct interval is determined.

3.2 Single-Delay case

When there exists only a single delay in the system, the characteristic equation (1) becomes.

$$C\left(s^{\alpha},\tau\right) = p_0\left(s^{\alpha}\right) + p_1\left(s^{\alpha}\right)e^{-\tau s} \tag{7}$$

If for some finite $\tau, C\left(s^{\alpha},\tau\right)=0$ has root on the imaginary axis at $s = j\,\omega_c$ (where subscript c refers to "crossing" the imaginary axis), then the equation $C\left((-s)^{\alpha},\tau\right)=0$ must have the same root for the same value of τ because of the complex conjugate symmetry of roots. Therefore, looking for roots on the imaginary axis reduces to finding values of τ for which $C\left(s^{\alpha},\tau\right)=0$ and $C\left((-s)^{\alpha},\tau\right)=0$ have a common root. That is

$$\begin{aligned} C\left(s^{\alpha},\tau\right) &= p_0\left(s^{\alpha}\right) + p_1\left(s^{\alpha}\right)e^{-\tau s} = 0 \\ C\left((-s)^{\alpha},\tau\right) &= p_0\left((-s)^{\alpha}\right) + p_1\left((s)^{\alpha}\right)e^{\tau s} = 0 \end{aligned} \tag{8}$$

Let define variable $\overline{\varsigma}$ that is complex conjugate of auxiliary variable $\varsigma = \sqrt[\alpha]{\omega}e^{\,j\pi/2\alpha}$ as follows:

$$\varsigma = \sqrt[\alpha]{-s} = \sqrt[\alpha]{\omega} e^{-j\pi/2\alpha} = \left(e^{-j\pi/\alpha}\right)\varsigma \tag{9}$$

Where $\varsigma = \sqrt[\alpha]{\omega} e^{j\pi/2\alpha}$. Equation (8) can be written as follows

$$\begin{aligned} p_0\left(\varsigma^\alpha\right) + p_1\left(\varsigma\right)e^{-\tau\varsigma^\alpha} &= 0 \\ p_0\left(\bar{\varsigma}^\alpha\right) + p_1\left(\bar{\varsigma}^\alpha\right)e^{\tau\varsigma^\alpha} &= 0 \end{aligned} \tag{10}$$

By eliminating the exponential term in (10), we get the following polynomial:

$$p_0\left(\varsigma\right)p_0\left(\bar{\varsigma}\right) - p_1\left(\varsigma\right)p_1\left(\bar{\varsigma}\right) = 0 \tag{11}$$

Please note that transcendental characteristic equation with single delay given in (10) is now converted into a equation without transcendentality given by (11) and its positive real roots coincide with the imaginary roots of (7) exactly. The roots of this equation may easily be determined by standard methods. Depending on the roots of (11), the following situation may occur:

1. The equation of (10) does not have any positive real roots, which implies that the characteristic equation (6) does not have any roots on the $j\omega$-axis. In that case, the system is stable for all $\tau \geq 0$, indicating that the system is delay-independent stable.
2. The equation of (10) has at least one positive real root, which implies that the characteristic equation (6) has at least a pair complex eigenvalues on the $\tau \geq 0$-axis. In that case, the system is delay - dependent stable.

The roots of this equation may easily be determined by standard methods. For a positive real root ω_c, the corresponding value of delay margin τ can be easily obtained using (7) as:

$$\tau = \frac{1}{\omega_c}\left(\frac{\Im\left[p_0(\varsigma) \big/ p_1(\varsigma) \right]}{\Re\left[\left(-p_0(\varsigma)\right) \big/ p_1(\varsigma) \right]} \right)\Bigg|_{\varsigma = \sqrt[\alpha]{\omega}\, e^{\frac{j\pi}{2\alpha}}} + \frac{2k\pi}{\omega_c} \tag{12}$$

For the positive roots of (11), we also need to check if at $s = j\omega_c$, the root of characteristic equation (7) crosses the imaginary axis with increasing τ. This can be determined by the sign of $\Re\left[ds/d\tau\right]$. With these values in hand, we will be able to calculate the direction of crossing from the left half plane to the right one, which we will denote as a destabilizing crossing, or from the right to the left, meaning this is a stabilizing crossing. Notice that the use of the expressions destabilizing and stabilizing crossings means only that a pair of poles is crossing the imaginary axis in the defined direction, and not that the system is turning unstable or stable, respectively. For that, it is necessary to know the number of unstable poles before the crossings.

3.2 Commensurate-Delay case

The method given for the single-delay case could be easily extended to the stability analysis of the fractional order system with multiple commensurate time delays. The characteristic equation of such a system is given by (1). Similar to the single-delay case, if the characteristic equation (1) has a solution of $s = j\omega_c$ then $C\left((-s)^\alpha, \tau\right) = 0$ will have the same solution.

$$C\left((-s)^\alpha, \tau\right) = \sum_{k=0}^{N} p_k\left((-s)^\alpha\right)e^{ks\tau} \tag{13}$$

Characteristic equation (13) can be written in terms of the auxiliary parameter $\bar{\varsigma}$ as:

$$C\left(\bar{\varsigma}^\alpha, \tau\right) = \sum_{k=0}^{N} p_k\left(\bar{\varsigma}^\alpha\right)e^{k\varsigma^\alpha\tau} \tag{14}$$

Recall that in the single-delay case, the transcendental term and the time delay τ are eliminated. In this case, the purpose is the same. A recursive procedure should be developed to achieve that purpose. Therefore, let us define

$$\begin{aligned} &\sum_{k=0}^{n-1}\left[p_0\left(\bar{\varsigma}\right)p_k\left(\varsigma\right) - p_n\left(\varsigma\right)p_{n-k}\left(\bar{\varsigma}\right)\right]e^{k\varsigma^\alpha\tau} \\ &= C^{(1)}\left(\varsigma, \tau\right) = 0 \end{aligned} \tag{15}$$

Then, we have

$$\sum_{k=0}^{n-1}\left[p_0(\varsigma)p_k(\bar{\varsigma})-p_n(\bar{\varsigma})p_{n-k}(\varsigma)\right]e^{k\varsigma^{\alpha}\tau} = C^{(1)}(\bar{\varsigma},\tau)=0 \tag{16}$$

It should be observed from equations (15) and (16) that if $s = j\omega_c$ is the solution of equations (1) and (12) for some τ, then it must be a solution of the following augmented characteristic equations

$$\begin{aligned} C^{(1)}(\varsigma,\tau) &= \sum_{k=0}^{n-1} p_k^{(1)}(\varsigma)e^{k\varsigma^{\alpha}\tau}=0 \\ C^{(1)}(\bar{\varsigma},\tau) &= \sum_{k=0}^{n-1} p_k^{(1)}(\bar{\varsigma})e^{k\varsigma^{\alpha}\tau}=0 \end{aligned} \tag{17}$$

Where

$$p_k^{(1)}(\varsigma)=p_0^{(1)}(\bar{\varsigma})p_k^{(1)}(\varsigma)-p_n^{(1)}(\varsigma)p_{n-k}^{(1)}(\bar{\varsigma}) \tag{18}$$

Note that the characteristic equations (17) are of commensuracy degree of (n-1). We can easily repeat this procedure to eliminate commensuracy terms successively by defining a new polynomial

$$p_k^{(r+1)}(\varsigma)= p_0^{(r)}(\bar{\varsigma})p_0^{(r)}(\varsigma)-p_{n-r}^{(r)}(\varsigma)p_{n-r-k}^{(r)}(\bar{\varsigma}) \tag{19}$$

and an augmented characteristic equation

$$C^{(r)}(\varsigma,\tau)=\sum_{k=0}^{n-r} p_k^{(r)}(\varsigma)e^{-k\varsigma^{\alpha}\tau}=0 \tag{20}$$

By repeating this procedure n times, we eliminate the highest degree of commensuracy terms and obtain the following augmented characteristic equation

$$C^{(n)}(\varsigma)=p_0^{(n)}(\varsigma)=0 \tag{21}$$

Where

$$p_0^{(n)}(\varsigma)= p_0^{(n-1)}(\bar{\varsigma})p_0^{(n-1)}(\varsigma)-p_1^{(n-1)}(\varsigma)p_1^{(n-1)}(\bar{\varsigma})=0 \tag{22}$$

It should be emphasized that that if $s = j\omega_c$ is the solution of (1) for some τ, then it is also a solution of (20) since the imaginary roots of the original characteristic equation (1) are preserved during the manipulations. If we substitute $\bar{\varsigma}=\left(e^{-j\pi/\alpha}\right)\varsigma$ and $\varsigma=\sqrt[\alpha]{\omega}e^{j\pi/2\alpha}$ in (22), we get the following equation in ω

$$\begin{aligned} D(\omega)=\Big(&p_0^{(n-1)}\left(e^{-j\pi/\alpha}\varsigma\right)p_0^{(n-1)}(\varsigma) \\ &-p_1^{(n-1)}(\varsigma)p_1^{(n-1)}\left(e^{-j\pi/\alpha}\varsigma\right)\Big)\Big|_{\varsigma=\sqrt[\alpha]{\omega}e^{j\pi/2\alpha}}=0 \end{aligned} \tag{23}$$

One can easily notice that (22) is the generalization of (11) and allows us to determine the imaginary roots of charactristic equation (1), if there exists any. The corresponding value of time delay is then computed by

$$\tau=\frac{1}{\omega_c}\left(\frac{\Im\left[\dfrac{p_0^{(n-1)}(\varsigma)}{p_1^{(n-1)}(\varsigma)}\right]}{\Re\left[\dfrac{\left(-p_0^{(n-1)}(\varsigma)\right)}{p_1^{(n-1)}(\varsigma)}\right]}\right)\Bigg|_{\varsigma=\sqrt[\alpha]{\omega}\,e^{\frac{j\pi}{2\alpha}}}+\frac{2k\pi}{\omega_c} \tag{24}$$

The whole ω values, for which $s = j\omega$ is a root of equation (1) for some non-negative delays, is defined as the crossing frequency set.

$$\Omega=\left\{\omega\in\mathbb{R}^+ \,\middle|\, C(s,\tau)=0\,,\ \textit{for some}\ \ \tau\in\mathbb{R}^+\right\} \tag{25}$$

This class of systems exhibits only a finite number of possible imaginary characteristic roots for all $\tau\in\mathbb{R}^+$ at given frequencies. And this method detected all of them. Let us call this set

$$\{\omega_c\} = \{\omega_{c1}, \omega_{c2}, \ldots, \omega_{cn}\} \tag{26}$$

where subscript c refers to 'crossing' the imaginary axis. Furthermore to each ω_{cm}, $m = 1, \ldots, n$ correspond infinitely many, periodically spaced τ values. All this set

$$\{\tau_m\} = \{\tau_{m1}, \tau_{m2}, \ldots, \tau_{m\infty}\} \quad m = 1, \ldots, n \tag{27}$$

Where $\tau_{m,k+1} - \tau_{m,k} = 2\pi/\omega_m$ is the apparent period of repetition.

3.3 Direction of crossing

After the crossing points of characteristic equation (1) from the imaginary axis are obtained, the goal now is to determine whether each of these root crossings from the imaginary axis is a stabilizing cross or a destabilizing cross. As it was shown in [11], this is constant with respect to subsequential crossings , and therefore it is denoted as root tendency. Assume that (s, τ) is a simple root of $C(s, \tau) = 0$. The root sensitivities associated with each purely imaginary characteristic root $j\omega$, with respect to τ is defined as

$$S_\tau^s\Big|_{s=\omega j} = \left.\frac{ds}{d\tau}\right|_{s=\omega j} = -\left.\frac{\partial C/\partial \tau}{\partial C/\partial s}\right|_{s=j\omega}$$
$$= \left.\frac{s\sum_{k=1}^{N} \ell q_k\left(s^\alpha\right)e^{-\tau s k}}{p'\left(s^\alpha\right) + \sum_{k=1}^{N}\left(q_k'\left(s^\alpha\right) - \tau \ell q_k\left(s^\alpha\right)\right)e^{-\tau s k}}\right|_{s=j\omega} \tag{28}$$

Here, $p'(s^\alpha)$ and $q_\ell'(s^\alpha)$ denote the derivative of the polynomials $p(s^\alpha)$ and $q_\ell(s^\alpha)$ in s respectively. The root tendency is given by

It must be noted that the root tendency is independent of time delay τ. This implies that even though there is an infinite number of values of τ associated with each value of ω_c that make $C(j\omega_c, \tau) = 0$.

$$Root\ Tendency = sign\left(\Re\left(S_\tau^s\Big|_{s=j\omega_{cm}}\right)\right)$$
$$= sign\left(\Re\left(\frac{ds}{dv} \times \frac{d\varsigma}{d\tau}\right)\right)\Bigg|_{\substack{s=j\omega \\ \varsigma=\varsigma_{cm}=\sqrt[\alpha]{\omega}\, e^{\pm\frac{\pi}{2\alpha}}}}$$
$$= sign\left(\Re\left(\varsigma^{\alpha-1} \times \frac{d\varsigma}{d\tau}\right)\right)\Bigg|_{\substack{s=j\omega \\ \varsigma=\varsigma_{cm}=\sqrt[\alpha]{\omega}\, e^{\pm\frac{\pi}{2\alpha}}}}$$
$$= sign\left(\Re\left(je^{\frac{-j\pi}{2\alpha}} \times \frac{d\varsigma}{d\tau}\right)\right)\Bigg|_{\substack{s=j\omega \\ \varsigma=\varsigma_{cm}=\sqrt[\alpha]{\omega}\, e^{\pm\frac{\pi}{2\alpha}}}}$$
$$= sign\left(\Re\left(\left[\sin\left(\frac{\pi}{2\alpha}\right) + j\cos\left(\frac{\pi}{2\alpha}\right)\right] \times \frac{d\varsigma}{d\tau}\right)\right)\Bigg|_{\varsigma=\sqrt[\alpha]{\omega}\, e^{\pm\frac{\pi}{2\alpha}}}$$
$$= sign\left(\sin\left(\frac{\pi}{2\alpha}\right)\Re\left(\frac{d\varsigma}{d\tau}\right) - \cos\left(\frac{\pi}{2\alpha}\right)\Im\left(\frac{d\varsigma}{d\tau}\right)\right)\Bigg|_{\varsigma=\sqrt[\alpha]{\omega}\, e^{\pm\frac{\pi}{2\alpha}}}$$
$$= sign\left(\tan\left(\frac{\pi}{2\alpha}\right)\Re\left(\frac{d\varsigma}{d\tau}\right) - \Im\left(\frac{d\varsigma}{d\tau}\right)\right)\Bigg|_{\varsigma=\varsigma_{cm}=\sqrt[\alpha]{\omega}\, e^{\pm\frac{\pi}{2\alpha}}} \tag{29}$$

If it is positive, then it is a destabilizing crossing, whereas if it is negative, this means a stabilizing crossing. In case the result is 0, a higher order analysis is needed, since this might be the case where the root just touches the imaginary axis and returns to its original half-plane. Notice that root tendency represents the direction of transition of the roots at $j\omega_c$ as τ increases from $\tau_{mk-\varepsilon}$ to $\tau_{mk+\varepsilon}$, $0 < \varepsilon \le 1$.

$$sign\left(\Re\left(S_\tau^s\Big|_{\substack{s=j\omega \\ \tau}}\right)\right) \tag{30}$$

Is independent of k. Notice that for those given values, the exponential term

$$e^{-\tau s k} = e^{-j\theta\ell}e^{-j2\pi k\ell} = e^{-j\theta k} \tag{31}$$

is independent of ℓ.

Now having the cross points, their corresponding infinitely time delays, and root tendency of each time delay, the number of unstable roots in the regions subdivided by τ_{mk} can be calculated based on the D-subdivision method. To make a complete picture of the stability mapping for any fractional delay system, τ_{mk} is sorted from smallest to largest value. The number of unstable roots in each region between two successive time delays changes by $2RT$ if $\omega \neq 0$, and by RT if $\omega = 0$. Undoubtedly, after a specific value of time delay, the number of unstable roots cannot be zero and the calculation should be stopped.

4 Numerical Example

We present three example cases, which display all the features discussed in the text.

Example 1: This example has been taken from [7] and [15]. Consider the following linear time-invariant fractional order system with one delay:

$$C_1(s,\tau) = \left(\sqrt{s}\right)^3 - 1.5\left(\sqrt{s}\right)^2 + 4\left(\sqrt{s}\right) + 8 - 1.5\left(\sqrt{s}\right)^2 e^{-\tau s} \qquad (32)$$

This system has a pair of poles ($s = \pm 8j$) on the imaginary axis for $\tau = 0$. A very involved calculation scheme based on Cauchy's integral has been used in [15] to show that this system is unstable for $\tau = 0.99$, and stable for $\tau = 1$. Our objective in this example is to find all the stability windows based on the method described in this article for this system.

Applying the first part of the algorithm, we can see that

$$\begin{aligned} &\omega_1 = 8 \rightarrow \tau = 0.7854k \\ &\omega_2 = 6.6248 \rightarrow \tau = 0.0499 + 0.9485k \end{aligned} \qquad (33)$$

By applying the criterion expressed in the previous section, it is easy to find out that a destabilizing crossing of roots has occurred at $\tau = 0.7854k$ for $s = \pm 8j$ and a stabilizing crossing has taken place at $\tau = 0.0499 + 0.9485k$ for $s = \pm 6.6246j$, for all values of $k \in \{0,1,2,...\}$. Therefore, we will have 5 stability windows as follows: $0.0499 < \tau < 0.7854$, $0.9983 < \tau < 1.5708$, $1.9486 < \tau < 2.3562$, $2.8953 < \tau < 3.1416$ and $3.8437 < \tau < 3.9270$, which agree with the results presented by [7] and [11].

Note that at $\tau = 3.927$, an unstable pair of poles crosses toward the right half-plane, and before this unstable pole pair can turn to the left half-plane at $\tau = 4.7922$, another unstable pair of poles goes toward the right half-plane at $\tau = 4.7124$; and thus, the system cannot recover the stability.

In Table 1, the stability map of system defined via (32) is given. The number of unstable roots in each interval of unstable region has been determined as we

Table 1: Stability Regions (Shaded) for Example1

τ	$\omega\left(rad/s\right)$	Root Tendency	Number of unstable roots
0	8	+	2
0.0498	6.6248	-	
			0
0.7853	8	+	
			2
0.9938	6.6248	-	
			0
1.5707	8	+	
			2
1.9467	6.6248	-	
			0
2.3561	8	+	
			2
2.8952	6.6248	-	
			0
3.1415	8	+	
			2
3.8437	6.6248	-	
			0
3.9269	8	+	
			2
4.7123	8	+	
			4
4.7922	6.6248	-	
⋮	⋮		⋮

To get a better understanding of the properties of this system, its root-loci curve [19] has been plotted as a function of delay in Fig. 2. The color spectrum in the "color bar" indicates the selected τ; dark blue designates $\tau = 0$, and dark red is for $\tau = 3.9$. Since $\tau = 3.9$ is within the last stability window, we know in advance that the system will be stable for this amount of delay.

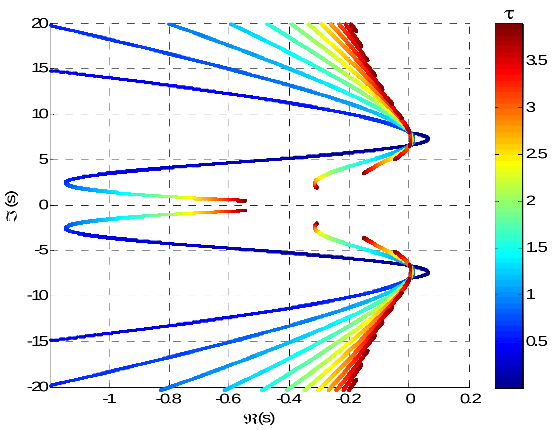

Fig. 2. Root-loci for $C_1(s,\tau)$ until τ=3.9

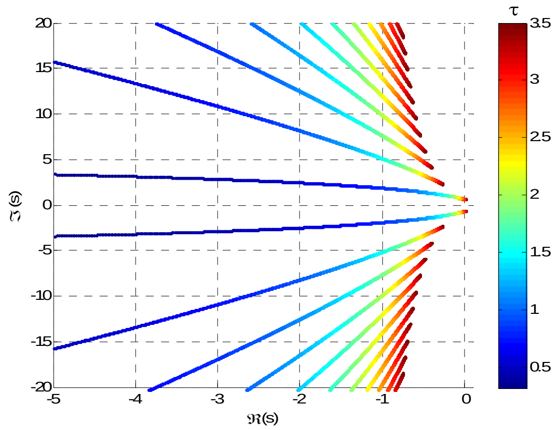

Fig. 3. Root-loci for $C_2(s,\tau)$ until τ=3.5

Example 2: This example comes from [18]. Consider the following fractional order system:

$$C_2(s,\tau) = 2\sqrt{s}\left(\sqrt{s}+1\right)\left(\sqrt{s}+10\right) + 5\left(\sqrt{s}+5\right)e^{-\tau s} \tag{34}$$

This system has no unstable pole for $\tau = 0$ (in fact, it has no pole in the physical Riemann sheet). By applying the previously described method, it is realized that the crossing through the imaginary axis occurs at $\tau = 3.2511598 + 9.9250867k$ and for $\omega = 0.633061$, which this crossing is a destabilizing crossing, and it means that the only stability window for this system is $0 < \tau < 3.2511598$. In Table 2, the stability map of system defined via (34) is given. The number of unstable roots in each interval of unstable region has been determined as we

Table 2: Stability Regions (Shaded) for Example2

τ	$\omega\,(rad/s)$	Root Tendency	Number of unstable roots
0			
			0
3.25116	0.63306	+	
			2
13.17624	0.63306	+	
			4
23.10132	0.63306	+	
⋮	⋮		⋮

To get a better understanding of the properties of this system, its root-loci curve has been plotted as a function of delay in Fig. 3.

The color spectrum indicating the changes of τ from $\tau = 0.322988$ (dark blue) to $\tau = 3.5$ (dark red) has been illustrated in the "color bar".

Example 3: Consider the following fractional order system with 2 delays [15]:

$$C(s) = s^{5/6} + \left(s^{1/2} + s^{1/3}\right)e^{-0.5s} + e^{-s} \tag{35}$$

It has been demonstrated in [15] that this system is stable. Let's change system (35) to the following form [13]:

$$C_3(s,\tau) = s^{5/6} + \left(s^{1/2} + s^{1/3}\right)e^{-\tau s} + e^{-2\tau s} \tag{36}$$

Now we evaluate the stability of this system for all the values of τ and also study the stability for $\tau = 0.5$.

By examining system (36) for $\tau = 0$, we find out that for this value, it has no pole in the physical Riemann sheet and thus, the system is stable without delay. In view of relations (12) and (36) we obtain

$$\begin{aligned} \omega_1 &= 1 \rightarrow \tau = 2.3562 + 6.2832k \\ \omega_2 &= 1 \rightarrow \tau = 2.6180 + 6.2832k \end{aligned} \tag{37}$$

As is observed, the crossing of the imaginary axis for $\tau = 2.3562 + 6.2832k$ and $\tau = 2.6180 + 6.2832k$ occurs at $s = \pm j$, and both of these are destabilizing crossings; and since the system is stable for $\tau = 0$, the only stability window for this system is $0 \le \tau < 2.3562$; and since $\tau = 0.5$ falls within this window, we can be sure that the original system $C(s)$ is stable.

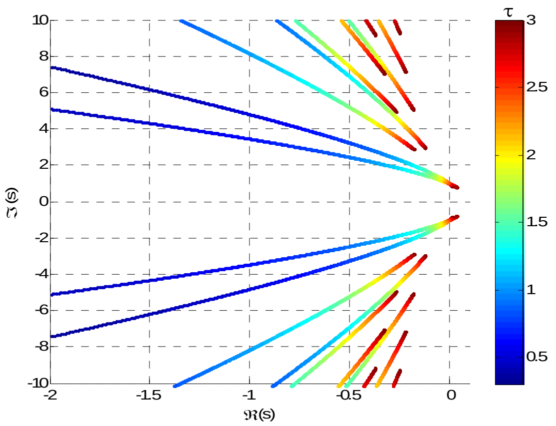

Fig. 4. Root-loci for $C_3(s,\tau)$ from τ=0.1 until τ=3

The obtained stability window matches the results presented in [15] and [11], which have used the numerical method to analyze this system.

In Fig. 4, the root-loci curve of this system for the changes of τ from $\tau = 0.1$ (dark blue) to $\tau = 3$ (dark red) has been presented for a better perception of the system.

5 Conclusion

In this paper, a new method for calculating stability windows and location of the unstable poles is proposed for a large class of fractional order time-delay systems. As the main advantages, we just deal with polynomials of the same order as that of the original system. and the use of auxiliary variable ς, the crossing points through the imaginary axis, and their direction of crossing, were determined. Then, system stability was expressed as a function of delay, based on the information obtained from the system. According to the infinitely countable time delays corresponding to each crossing point, the parametric space of τ is discretized to investigate stability in each interval. The number of unstable roots can be calculated with root tendency of each crossing point on the interval boundaries. Based on the proposed method, an upper bound for time delay is determined so that the system would not be stable any more for larger time delays. Finally, several examples were presented to highlight the proposed approach.

References:

[1] Y. Chang, S. Chen, Static Output-Feedback Simultaneous Stabilization of Interval Time-Delay Systems, *WSEAS Transactions on systems*, Vol. 7, No. 3, March 2008, pp. 185-194.

[2] R. Brcena, A. Etxebarria, Industrial PC-based real-time controllers applied to second-order and first-order plus time delay processes, *WSEAS Transactions on systems*, Vol. 7, No. 9, 2008, pp. 870-879.

[3] Y. Ke, C. Miao, Stability analysis of BAM neural networks with inertial term and time delay, *WSEAS Transactions on systems*, Vol. 10, No. 12, 2011, pp. 425-438.

[4] E. Gyurkovics, T. Takacs, Output Feedback Guaranteeing Cost Control by Matrix Inequalities for Discrete-Time Delay Systems, *WSEAS Transactions on systems*, Vol. 7, No. 7, 2008, pp. 645-654.

[5] L. Y. Chang, H. C. Chen, Tuning of Fractional PID Controllers Using Adaptive Genetic Algorithm for Active Magnetic Bearing System, *WSEAS Transactions on systems*, Vol. 8, No. 1, 2009, pp. 158-167.

[6] M. A. Pakzad, Kalman filter design for time delay systems, *WSEAS Transactions on Systems*, Vol.10, No.11, pp.551-560, 2012.

[7] N. Öztürk and A. Uraz, An analysis stability test for a certain class of distributed parameter systems with delays, *IEEE Trans. on Circuits and Systems*, vol. 34, no. 4, pp. 393-396, 1985.

[8] E.I. Jury and E. Zeheb, On a stability test for a class of distributed system with delays, *IEEE Trans. on Circuits and Systems*, vol. 37, no.10, pp. 1027-1028, 1986.

[9] N. Olgac and R. Sipahi, An exact method for the stability analysis of time delayed LTI systems, *IEEE Trans. on Automatic Control*, vol. 47, no.5, pp.793–797, 2002.

[10] N. Olgac and R. Sipahi, A practical method for analyzing the stability of neutral type LTI-time delayed systems, *Automatica*, vol. 40, pp.847–853, 2004.

[11] A.R. Fioravanti, C. Bonnet, H. Özbay and S.I. Niculescu, Stability windows and unstable root-loci for linear fractional time-delay systems, *The 18th IFAC World Congress*, Milan, Italy, August 28–September 02, pp. 12532–12537, 2011.

[12] C. Bonnet and J.R. Partington, Analysis of fractional delay systems of retarded and neutral type, *Automatica*, vol. 38, no. 7, pp. 1133-1138, 2002.

[13] C. Bonnet and J.R. Partington, “Stabilization of some fractional delay systems of neutral type, *Automatica*, vol. 43, no. 12, pp. 2047-2053, 2007.

[14] M. Buslowicz, Stability of linear continuous time fractional order systems with delays of the

retarded type, *Bull. Pol. Acad. Sci. Techn*., vol. 56, no. 4, pp. 319-324, 2008.

[15] C. Hwang and Y.C. Cheng, A numerical algorithm for stability testing of fractional delay systems, *Automatica,* vol. 42, no. 5, pp. 825-831, 2006.

[16] C. Hwang and Y.C. Cheng, A note on the use of the Lambert W function in the stability analysis of time-delay systems, *Automatica* , vol. 41, no. 11, pp. 1979–1985, 2005.

[17] Y. C. Cheng and C. Hwang, Use of the Lambert W function for time-domain analysis of feedback fractional delay systems, *IEE Proceedings - Control Theory and Applications*, vol. 153, No. 2, pp. 167 - 174, 2006.

[18] N. Öztürk , Stability Analysis of a Certain Class of Distributed Parameter Systems , *American Control Conference*, pp.415 - 416, 23-25 May 1990.

[19] J. A. Tenreiro Machado, Root locus of fractional linear systems, *Communications in Nonlinear Science and Numerical Simulation*, vol. 16, no.10, pp.3855-3862, 2011.

[20] J. A. Tenreiro Machado, V. Kiryakova and F. Mainardi, Recent history of fractional calculus, *Communications in Nonlinear Science and Numerical Simulation*, Vol. 16, no. 3, pp. 1140-1153, 2011.

[21] S. Pakzad, and M. A. Pakzad, Stability condition for discrete systems with multiple state delays, *WSEAS Transactions on Systems and Control*, Vol.6, No.11, 2011, pp. 417-426.

[22] J. Tenreiro Machado, A. Costa, M. Lima, A multidimensional scaling perspective of entropy analysis applied to musical compositions, *Nonlinear Dynamics*, vol. 65, no 4, 399–412, 2011.

[23] Pekař, L., Root locus analysis of a retarded quasipolynomial, *WSEAS Transactions on Systems and Control*, Vol.6, No.3, 2011, pp. 79-91.

Kalman Filter Design for Time Delay Systems

MOHAMMAD ALI PAKZAD
Department of Electrical Engineering
Science and Research Branch, Islamic Azad University
Tehran
IRAN
m.pakzad@srbiau.ac.ir

Abstract: - In this paper, an observer design is proposed for linear time delay systems. An easy way to compute least square estimation error of an observer for time delay systems is derived, where the time delay terms exist in the state and output of the system. Based on the least square estimation error an optimization algorithm to compute a Kalman filter for time delay systems is proposed. By employing the finite characterization of a Lyapunov functional equation, the existence of sufficient conditions for obtaining the right solution and guaranteeing the proper convergence rate of the estimation error has been evaluated. It will be shown that this finite characterization can be calculated by means of a matrix exponential function. The desirable performance of the proposed observer has been demonstrated through the simulation of several numerical examples.

Key-Words: - time delay system,state delay system, observer, Kalman filter

1 Introduction

As a dual of the control problem, the state estimation or filtering problem is of great importance in both theory and application, and in the last decades , this problem has gotten extensive concern and many solution schemes have been proposed and successfully put into action. Among them, Kalman filtering, which minimizes the variance of the estimation error, is the most famous one [11].

Kalman filtering is one of the most popular estimation approaches. This filtering method assures that both the state equation and output measurement are subjected to stationary Gaussian noises. The applications of the Kalman filtering theory may be found in a large spectrum of different fields ranging from various engineering problems to biology, geoscience, economics, and management, etc.

A dynamic system whose state variables are estimations of the state variables of another system is called the observer of that system. This expression was first introduced in 1963 into the theory of linear systems by Luenberger [1]. He showed that for every observable linear system, an observer can be designed whose estimation error (i.e. the difference between the real state of the system and the observer state) becomes zero at every considered speed. In fact, an observer is a dynamic system whose inputs are the process inputs and outputs, and whose outputs are the estimated state variables. It can be stated that an estimator of state is an indispensible member of the control systems theory, and it has important applications in feedback control, system supervision and in the fault diagnosis of dynamic systems.

In the control process, it is often assumed that the internal state vectors exist and are available in the measurement of the output; while in practice, this is not the case, and it is necessary to devise an observer in order to provide an estimation of state vectors. If the estimation and reconstruction of all the state variables is needed, the full-order observers, and if the estimation and reconstruction of a number of state variables is needed, the reduced-order observers are used. Time-delayed systems play significant roles in theoretical as well as practical fields; and this influence can be observed in numerous research articles written on various problems that involve this class of systems [2-8]. During the last decade, the theory of observer design for time delay systems has been widely contemplated [28-34]. The estimation of state variables is an important dynamic model, which adds to our knowledge of different systems and helps us analyze and design various controllers. Different approaches have been used for the designing of observers, including: the coordinate change approach [9], the LMI method [10], reducing transformation technique [11], factorization approach [12], polynomial approach [13], modal observer [14], reduced-order observer [15] and the output injection based observer [16]. In

[17], through an algebraic approach, an observer with delay-independent stability for systems with one output delay has been presented. In [18], an observer has been proposed that uses the H_∞ norm as the performance index. The H_∞ filter has been considered in [19], [20] by applying the delay independent stability conditions, in which the matrix inequality has been used. We also frequently encounter the issue of state delay in control problems and physical systems. In recent years, the systems with delay in state have attracted the attention of many researchers, and numerous approaches have been proposed for the evaluation of stability in these systems (see [21] and [22] and the references cited in them).

The goal of this article is to design an observer for time delay systems in which the time delay terms exist in the output and in the state variables, and also the inputs are mixed with noise and the system output accompanies measurement noises. An easy way to compute least square estimation error of an observer for time delay system is derived. This least square estimation error coincides with that of a Kalman filter when time delay is zero. Based on the least square estimation error, an optimization algorithm to compute an observer is proposed.

In the designing of this observer we have used the H_2 norm as the performance index. However, despite the usefulness of the H_2 norm, few observers have used it as the performance index. In [23] and [24] a method has been proposed for the calculation of the H_2 norm of time delay systems by means of the delay Lyapunov equation. In [25], an observer has been offered for time delay systems by applying the delay-independent stability conditions. It should be mentioned that delay-independent approaches are generally more conservative than delay-dependent ones. In this article, for the estimation of system states, a Kalman filter has been proposed whose design uses the delay-dependent stability conditions. Note that when there are no time delay terms, observer is a standard Kalman filter. The optimal observer will be designed by employing the finite characterization of a Lyapunov functional equation as a matrix exponential function and applying the unconstrained nonlinear optimization algorithm. Finally, the proposed observer in this article will be used to estimate the current states based on the time delay system, where the time delay terms exist in the state and in the output of the system.

This article has been organized in the following manner. In section 2, for the definition of the observer, the necessary mathematics has been presented. In section 3, the calculation of the H_2 norm has been offered for the state delay system. In section 4, the method of filter design has been described. In section 5, in order to test the practical usefulness of the proposed technique, it has been applied for solving the estimation problem of several linear systems with time delay. And finally, the summary and conclusion of the obtained results have been presented in the last section.

2 Problem Formulation And Assumptions

Consider linear time-invariant systems described by

$$\begin{aligned} \dot{x}(t) &= A_0 x(t) + A_1 x(t-h) + B_1 \omega(t) + B_2 u(t) \\ y(t) &= C_0 x(t) + C_1 x(t-h) + C_2 \nu(t) \end{aligned} \tag{1}$$

Where $x \in R^n$ is the state, $\omega \in R^p$ is the process noise, $u \in R^q$ is the input, $y \in R^r$ is the measurement, and $\nu \in R^r$ is the measurement noise. The h is constant known time delay in the states and the outputs.

It is assumed that ν and ω are uncorrelated white Gaussian processes, which satisfy

$$\begin{aligned} E\{\omega(t)\} = 0, E\{\omega(t)\omega(s)'\} = I\,\delta(t-s) \\ E\{\nu(t)\} = 0, E\{\nu(t)\nu(s)'\} = I\,\delta(t-s) \end{aligned} \tag{2}$$

The objective of this paper is to derive a Kalman filter for time delay system (l), where a filter has the following form:

$$\dot{\hat{x}}(t) = G\hat{x}(t) + Ky(t) + B_2 u(t) \tag{3}$$

Defining the estimation error $e(t)$ as

$$e(t) \triangleq x(t) - \hat{x}(t)$$

From (1) and (3) , we have

$$\begin{aligned} \dot{e}(t) = (A_0 - G - KC_0)x(t) + Ge(t) + \\ \left(A_1 - KC_1\right)x(t-h) + B_1\omega(t) - KC_2\nu(t) \end{aligned} \tag{4}$$

And the augmented system with (1) is given by

$$G_a : \begin{aligned} \dot{\eta}(t) &= \bar{A}_0 \eta(t) + \bar{A}_1 \eta(t-h) + B\,\zeta(t) \\ e(t) &= C\,\eta(t) \end{aligned} \tag{5}$$

where

$$\eta(t) \triangleq \begin{bmatrix} x(t) \\ e(t) \end{bmatrix} \quad , \quad \zeta(t) \triangleq \begin{bmatrix} \omega(t) \\ \upsilon(t) \end{bmatrix}$$

$$\bar{A}_0 \triangleq \begin{bmatrix} A_0 & 0 \\ A_0 - G - KC_0 & G \end{bmatrix} \quad , \quad \bar{A}_1 \triangleq \begin{bmatrix} A_1 & 0 \\ A_1 - KC_1 & 0 \end{bmatrix}$$

$$B \triangleq \begin{bmatrix} B_1 & 0 \\ B_1 & -KC_2 \end{bmatrix} \quad , \quad C \triangleq \begin{bmatrix} 0 & I \end{bmatrix}$$

The H_2 norm augmented system G_a is used as the performance index of estimation

$$\|G_a\|_2^2 = J(G, k, h) = \lim_{T \to \infty} E\left\{ \frac{1}{T} \int_0^T e'(t)\, e(t) dt \right\} \tag{6}$$

If there are no time delay terms (i.e., $A_1 = 0$ and $C_1 = 0$), then (1) becomes

$$\dot{x}(t) = A_0 x(t) + B_1 \omega(t) + B_2 u(t)$$
$$y(t) = C_0 x(t) + C_2 \nu(t)$$

and the filter, minimizing the H_2 norm (6) for this non-delayed system, is the standard Kalman filter. Thus we can call the proposed filter minimizing (6) a Kalman filter for time delay systems.

3 H_2 Norm Computatuon

The H_2 norm of G_a is expressed in terms of matrix function $P(s)$ in the next theorem.

Theorem 1: If is stable, then

$$\|G_a\|_2^2 = Tr(B'P(0)B) \tag{7}$$

Where $P(s), 0 \le s \le h$ is continuously differentiable and satisfies

$$P(0) = P'(0)$$
$$P(s) = \bar{A}_0' P(0) + \bar{A}_1' P(h-s), 0 \le s \le h \tag{8}$$
$$\dot{P}(0) + \dot{P}'(0) + C'C = 0$$

Remark 1: is related to the Lyapunov functional of state delay system (4). Let $V(\phi), \phi \in C[-h, 0]$ be defined by

$$V(\phi) \triangleq \phi'(0)P(0)\phi(0) + 2\phi'(0)\int_0^h P(r)\bar{A}_1\phi(-h+r)dr + \int_0^h \phi'(-h+r)\int_0^h \bar{A}_1' P(r-s)\bar{A}_1\phi(-h+r)dsdr \tag{9}$$

Where $P(s) \triangleq P'(-s)$ if $s < 0$. Equation (8) is derived from

$$\frac{d}{dt} V(x_t) = -x'(t)x(t) \tag{10}$$

Where

$$x_t(r) \triangleq x(t+r), \quad r \in [-h, 0]$$

Remark 2: If there are no time delay terms, the result in Theorem 1 becomes a standard H_2 norm computation. See, for example, Theorem 3.3.1 in [25]: the H_2 norm of a stable non-delay system is given by

$$\|G_a\|_2^2 = Tr(B'PB) \tag{11}$$

Where

$$\bar{A}_0' P + P\bar{A}_0 + C'C = 0$$

Note that conditions (7) are equivalent to those in (11) if $h = 0$.
The proof of Theorem 1 will be given using Lemma 1 and 2.

Lemma 1: If system G_a is stable, then

$$\|G_a\|_2^2 = \frac{1}{2\pi} \int_{-\infty}^{+\infty} Tr(G_a(j\omega)G_a'(-j\omega))d\omega \tag{12}$$

Proof: The result is standard (see Chap 3.3 in [25]).
Lemma 2: If G_a is stable and $P(s), 0 \le s \le h$ satisfies (8), then

$$P(0) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \Delta^{-1}(j\omega)'\Delta^{-1}(-j\omega)d\omega \tag{13}$$

Where

$$\Delta(j\omega) \triangleq j\omega I - \bar{A}_0 - \bar{A}_1 e^{-j\omega h} \tag{14}$$

Proof: See [26].
(Proof of Theorem 1) From Lemma 1,

$$Tr(B'P(0)B)$$
$$= Tr\left\{ \frac{1}{2\pi} \int_{-\infty}^{+\infty} B'\Delta^{-1}(j\omega)'\Delta^{-1}(-j\omega)Bd\omega \right\}$$
$$= \frac{1}{2\pi} \int_{-\infty}^{+\infty} Tr\left\{ B'\Delta^{-1}(j\omega)'\Delta^{-1}(-j\omega)B \right\} d\omega$$

Since $\int_{-\infty}^{+\infty} f(j\omega)d\omega = \int_{-\infty}^{+\infty} f(-j\omega)d\omega$, we have

$$Tr(B'P(0)B)$$
$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty} Tr\left\{B'\Delta^{-1}(-j\omega)'\Delta^{-1}(j\omega)B\right\}d\omega$$
$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty} Tr\left\{G_a'(-j\omega)'G_a(j\omega)\right\}d\omega$$

Since $Tr(AB) = Tr(BA)$ whenever AB and BA are square matrices, we have

$$Tr(B'P(0)B) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} Tr\left\{G_a(j\omega)'G_a'(-j\omega)\right\}d\omega = \|G_a\|_2^2$$

The last equality is from (12).
If G_a is stable, then $\|G_a\|_2^2$ can be computed from $P(0)$ in Theorem 1. How to check the stability of G_a will be considered later in Theorem 2; first we will consider how to compute $P(0)$ in the next lemma.

Notation: For a matrix $M \in \mathbb{C}^{n\times n}$ given by

$$M = \begin{bmatrix} m_{11} & m_{12} & \cdots & m_{1n} \\ m_{21} & m_{22} & \cdots & m_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ m_{n1} & m_{n2} & \cdots & m_{nn} \end{bmatrix}$$

M' denotes complex conjugate transpose of M
the column string csM is defined by

$$csM \triangleq \left[m_{11}\ m_{12}\cdots m_{1n} \mid m_{21}\ m_{22}\cdots m_{2n} \mid \right.$$
$$\left. \cdots \mid m_{n1}\ m_{n2}\cdots m_{nn}\right]' \in \mathbb{C}^{n^2\times 1}$$

How to compute $P(s), 0 \le s \le h$, is considered in the next lemma.

Lemma 3: If G_a is stable, then $P(0)$ and $P(h)$ satisfying (8) are given by

$$\begin{bmatrix} (I\otimes \bar{A}_0') + (\bar{A}_0'\otimes I) & (I\otimes \bar{A}_1')T + (\bar{A}_1'\otimes I) \\ R_1 & R_2 \end{bmatrix} \cdot \begin{bmatrix} csP(0) \\ csP(h) \end{bmatrix} = \begin{bmatrix} -csC'C \\ 0 \end{bmatrix} \tag{15}$$

Where

$$T := \left[T_1 | T_2 | \cdots | T_{n^2}\right], T_k \in R^{n^2\times 1}$$

Row vector $T_k, 1 \le k \le n^2$ is defined by

$$T_k, 1 \le k \le n^2\ T_{(i-1)n+j} := e_{(j-1)n+i},\ 1 \le i, j \le n^2$$

Where $e_k \in R^{n^2\times 1}, 1 \le k \le n^2$ is a row vector whose k-th element is 1 and all other elements are 0.
And

$$\left[R_1 \quad R_2\right] \triangleq \left[\Sigma_1 \quad 0\right]V^*$$

Matrices Σ_1 and V^* are from the singular value decomposition of the following

$$\left(I - J\exp(Hh)\right) = U\begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix}V^* \tag{16}$$

Where U and V are unitary matrices, and $\Sigma_1 \in R^{n^2\times n^2}$ is a diagonal matrix whose diagonal elements are nonzero singular values of $\left(I - J\exp(Hh)\right)$. Let T_{ij} denote an $n\times n$ matrix with (i, j)-entry equal to 1 and all other entries equal to zero, and let $T \in R^{n^2\times n^2}$ be the block matrix T, $\left[T_{ji}\right]$ (i.e., the (i, j)-block of T is T_{ji}). Matrices H and J are defined by

$$H \triangleq \begin{bmatrix} (I\otimes \bar{A}_0') & (I\otimes \bar{A}_1')T \\ -(I\otimes \bar{A}_1')T & -(I\otimes \bar{A}_0') \end{bmatrix}, \qquad J = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}$$

Proof: See [22].
Note that $P(0)$ can be computed from the matrix exponential (16) and a simple linear equation (15). Thus if G_a is stable, then we can easily compute H_2 norm: see (7).
Now the stability of G_e is considered in Theorem 2, where a stability condition for interval delay $h \in \left[0, \bar{h}\right)$ is provided.

Theorem 2: Suppose G_a is stable for $h = 0$. If H has imaginary eigenvalues $\left\{j\omega_1, \ \cdots \ , j\omega_k\right\}$ and their corresponding eigenvectors are given by

$$v_1 = \begin{bmatrix} v_{1,1} \\ v_{1,2} \\ \vdots \\ v_{1,2n^2} \end{bmatrix}, \cdots, v_k = \begin{bmatrix} v_{k,1} \\ v_{k,2} \\ \vdots \\ v_{k,2n^2} \end{bmatrix}$$

then G_a is stable for $h \in \left[0, \bar{h}\right)$ where $\bar{h}$ is defined by

$$\bar{h} \in \min_{1 \le i \le k} \left| \frac{1}{j\omega} \ln \left(\frac{v_{i,l}}{v_{i,l+n^2}} \right) \right| \tag{17}$$

where $v_{i,l}, 0 \le 1 \le n^2$ is any nonzero element of v_l . Theorem 2 is proved using Lemma 4 and 5. Lemma 4 is based on the fact that if G_a is stable for $h = 0$ and G_a does not have any imaginary poles for $h \in \left[0, \bar{h}\right)$, then G_a is stable for $h \in \left[0, \bar{h}\right)$.

Lemma 4: G_a is stable for $h \in \left[0, \bar{h}\right)$ if

- G_a is stable for $h = 0$
- The following equation does not have any roots for $h \in \left[0, \bar{h}\right)$:

$$\det(j\omega I - \bar{A}_0 - \bar{A}_1 e^{-j\omega h}) = 0 \tag{18}$$

Proof: See [27].
Stability of G_a for $h = 0$ can be easily checked from eigenvalues of $\bar{A}_0 + \bar{A}_1$. On the other hand, checking whether (18) has any roots for $h \in \left[0, \bar{h}\right)$ is not easy: (18) should be checked for all $0 \le \omega < \infty$ and $0 \le h < \bar{h}$ In the next lemma, it is shown that a root $j\omega$ of (18) (if any) is an eigenvalue of H .

Lemma 5: If (18) has a root ω, then it is an eigenvalue of H .

Proof: Suppose (18) has a root $j\omega$ for h ; then there exists $x\,(\in C^n) \ne 0$ such that

$$x'(j\omega I - \bar{A}_0 - \bar{A}_1 e^{-j\omega h}) = 0 \tag{19}$$

Taking the transpose (not complex conjugate), we Obtain

$$(j\omega I - \bar{A}_0 - \bar{A}_1 e^{-j\omega h})x = 0$$

Let $\alpha \in C^n$ be defined by

$$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} \triangleq x e^{-\frac{j\omega h}{2}} \tag{20}$$

where , $\alpha_i, 1 \le i \le n$ is a complex number. Let v be defined by ($\bar{u}$ is the complex conjugate of u)

$$v \triangleq \begin{bmatrix} u \\ \bar{u} \end{bmatrix} \tag{21}$$

Where

$$u = \begin{bmatrix} \bar{\alpha}_1 x \\ \bar{\alpha}_2 x \\ \vdots \\ \bar{\alpha}_n x \end{bmatrix} \in C^{n^2} \tag{22}$$

The theorem is proved if we show that this v ($v \ne 0$ from the construction) satisfies $(j\omega I - H)v = 0$: that is, $j\omega$ is an eigenvalue of H . From the definition of H , we obtain

$$\begin{aligned} &(j\omega I - H)v \\ &= \begin{bmatrix} j\omega I - (I \otimes \bar{A}_0') & -(I \otimes \bar{A}_1')T \\ (I \otimes \bar{A}_1')T & j\omega I + (I \otimes \bar{A}_0') \end{bmatrix} v \\ &= \begin{bmatrix} \left(j\omega I - (I \otimes \bar{A}_0')\right)u - (I \otimes \bar{A}_1')T\bar{u} \\ \left(j\omega I + (I \otimes \bar{A}_0')\right)\bar{u} + (I \otimes \bar{A}_1')Tu \end{bmatrix} \end{aligned} \tag{23}$$

Partition $(j\omega I - H)v$ into $2n$ complex vectors and let the i-th block of $(j\omega I - H)v$ be denoted by $r_i \in C^n$. Then $r_i, 1 \le i \le n$ is given by

$$r_i = (j\omega I - \bar{A}_0')\bar{\alpha}_i x - \bar{A}_1'(T_{1i}\bar{\alpha}_1 + T_{2i}\bar{\alpha}_2 + \cdots + T_{ni}\bar{\alpha}_n)\bar{x}$$

Noting the following relation

$$\begin{aligned} &(T_{1i}\bar{\alpha}_1 + T_{2i}\bar{\alpha}_2 + \cdots + T_{ni}\bar{\alpha}_n)\bar{x} \\ &= (T_{1i}\bar{\alpha}_1 + T_{2i}\bar{\alpha}_2 + \cdots + T_{ni}\bar{\alpha}_n)\bar{\alpha} e^{-\frac{j\omega h}{2}} \\ &= e^{-\frac{j\omega h}{2}} \bar{\alpha}_i \alpha \end{aligned}$$

We obtain

$$r_i = (j\omega I - \bar{A}_0')\bar{\alpha}_i \alpha e^{\frac{j\omega h}{2}} - \bar{\alpha}_i \bar{A}_1' \alpha e^{-\frac{j\omega h}{2}}$$
$$= \bar{\alpha}_i e^{\frac{j\omega h}{2}} \left(j\omega I - \bar{A}_0' - \bar{A}_1' e^{-j\omega h} \right) \alpha$$
$$= \bar{\alpha}_i \left(j\omega I - \bar{A}_0' - \bar{A}_1' e^{-j\omega h} \right) x = 0, \quad 1 \le i \le n$$

The last equality is from (19).

Since $r_{i+n} = -\bar{r}_i$, $1 \le i \le n$ (see (23)), we have $r_i = 0$, $n+1 \le i \le 2n$. Hence , $(j\omega I - H)v = 0$, where $v \ne 0$ (since $x \ne 0$).

Proof of Theorem 2: From the proof of Lemma 5, if (18) has a root ω_i for h_i $(1 \le i \le k)$, then ω_i is an eigenvalue of H. Furthermore, the corresponding eigenvector of H is of the form:

$$v_i = \Big[\bar{x}_1 x e^{\frac{j\omega_i h_i}{2}} \quad \bar{x}_2 x e^{\frac{j\omega_i h_i}{2}} \cdots \bar{x}_n x e^{\frac{j\omega_i h_i}{2}} \quad x_1 \bar{x} e^{-\frac{j\omega_i h_i}{2}}$$
$$x_2 \bar{x} e^{-\frac{j\omega_i h_i}{2}} \cdots x_n \bar{x} e^{-\frac{j\omega_i h_i}{2}} \Big]^T$$

Thus h_i can be computed as follows:

$$h_i = \left| \frac{1}{j\omega} \ln \left(\frac{v_{i,l}}{v_{i,l+n^2}} \right) \right|$$

Where $v_{i,l}$, $1 \le l \le n^2$ is any nonzero element of v_i .
If the minimum value of h_i $(1 \le i \le k)$ is h_i then (18) does not have a root for $h \in \left[0, \bar{h}\right)$. From Lemma 4, this proves the theorem.

Remark 3: Once (G, K) is determined, we can check the stability of the error system (4) (Theorem 2) and compute its H_2 norm (Theorem1).

4 Kalman Filter for Time Delay Systems: Synthesis

In this section, the synthesis algorithm of Kalman filter (3) is proposed, where the algorithm is formulated as a constrained nonlinear optimization problem and the output delay $h = h^*$.

When minimizing H_2 norm of G_a over (G, K) using Theorem 1, it should be guaranteed that G_a is stable. The approach presented here allows one to design linear observers for time delay systems (see Fig.1). If (G, K) is given, the stability of G_a can be checked using Theorem 2, which provides a upper stability bound $\bar{h}(k)$ (i.e., G_a is stable as long as $h < \bar{h}$). Thus finding (G, K), which stabilizes G_a and minimizes $\|G_a(G, K, h)\|_2$.
Kalman filter design problem can be formulated as follows:

$$\begin{aligned} &\min_{G,k} J_1(G, K, h^{/*}) \triangleq \left\| G_a(G, K, h^*) \right\|_2^2 \\ &\textit{subject to } h < \bar{h}(G, K) \end{aligned} \tag{24}$$

(24) is a constrained nonlinear optimization problem whose global solution is difficult to find. A suboptimal approach is proposed to compute (G, K) using penalty methods [26].

A penalty function is defined by

$$p(G, K) \triangleq \begin{cases} 0 & \textit{if} \quad h < \bar{h}(G, K) \\ \alpha(h^* - \bar{h})^2 & \textit{if} \quad h \ge \bar{h}(G, K) \end{cases}$$

where α is a constant and is chosen so that $p(G, K, h^*) \gg J(G, K, h^*)$ when $h^* \gg \bar{h}(G, K)$. With this penalty function, a constrained optimization problem (24) can be replaced by the following unconstrained optimization problem:

$$\min_{G,K} J_2(G, K, h^*) \triangleq \left\| G_a(G, K, h^*) \right\|_2^2 + p(G, K) \tag{25}$$

Note that if $h^* < \bar{h}(G, K)$ (i.e., G_a is stable), then $J_2(G, K, h^*) = J_1(G, K, h^*)$. Also note that if $h^* \ge \bar{h}(G, K)$,then $J_2(G, K, h^*)$ is dominated by the penalty function $p(G, K, h^*)$. Thus the penalty function $p(G, K, h^*)$ prevents unstable region searching when the H_2 norm is being minimized.
initial value of G and K can be chosen by minimizing $J(G, K, 0)$: the initial value corresponds to the Kalman filter gain for a non-delayed system. Minimization problem (25) can be solved, for example, using an unconstrained nonlinear optimization function fminunc in MATLAB optimization toolbox.

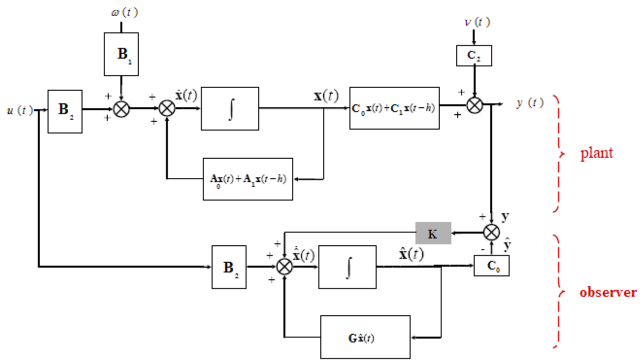

Fig. 1: The block diagram of observer

5 Numerical Example

In this section, the simulations have been performed by means of the MATLAB software.

Example 1: Consider the following first-order time delay system:

$$\begin{aligned} \dot{x}(t) &= -x(t) - 2x(t-h) + 0.5\omega(t) + u(t) \\ y(t) &= x(t) + x(t-h) + 0.5v(t) \end{aligned} \tag{26}$$

where $\omega(t)$ and $v(t)$ are the vectors of the input noise and measurement noise, respectively. It is assumed that these noises are Gaussian processes with an average of zero and that $\omega(t)$ and $v(t)$ are uncorrelated and they satisfy relation (2). In this example, $h = 0.5$.

The optimization problem (25) is solved by means of the Matlab optimization toolbox, and for this purpose, the optimization function "fminunc" in Matlab is used.

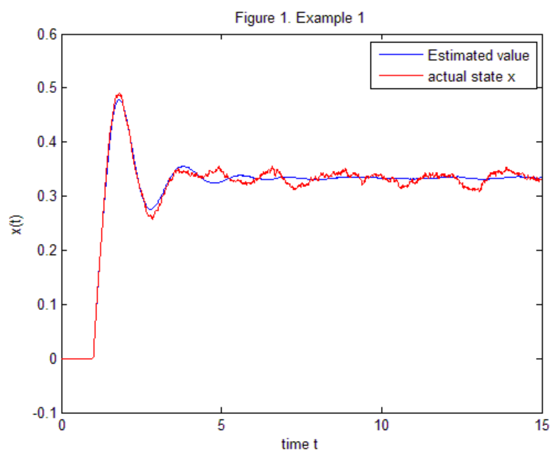

Fig. 2: Simulation result : true state and estimated value

By using $h = 0$, the initial value for (G, K) is obtained. The value of α in the penalty function has been adjusted at 200. The values calculated for $h = 0.5$ are as follows:

$$\|G_e(K,h)\|_2^2 = 0.0717$$

Using the computed filter gain, state estimation simulation was done, where a unit step signal was applied to the control input $u(t)$ at time 1s. The simulation result is given in Fig.2

it can be seen that the proposed Kalman filter estimates system states well.

To see how the time delay affects estimation performance, Kalman filters were designed for different h values.

As seen in Table 1, computed H_2 norm increases as time delay h increases.

Table 1. Time delay effects on estimation performance.

	$h = 0.1$	$h = 0.3$	$h = 0.5$	$h = 0.8$
$\|G_e(K,h)\|_2^2$	0.0399	0.0543	0.0717	0.1095
Variance of actual estimation error	0.000015	0.00035	0.00065	0.0009

Example 2: In this problem, the H_2 filter is designed for the second-order system given in the following relation.

$$\begin{aligned} \dot{x}(t) &= \begin{bmatrix} -2 & 1 \\ 0 & -1 \end{bmatrix} x(t) + \begin{bmatrix} -1 & 0 \\ -1 & -1 \end{bmatrix} x(t-h) \\ &\quad + \begin{bmatrix} 0.2 \\ 0.2 \end{bmatrix} \omega(t) + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u(t) \\ y(t) &= \begin{bmatrix} 0 & 1 \end{bmatrix} x(t) + \begin{bmatrix} 1 & 1 \end{bmatrix} x(t-h) + 0.5v(t) \end{aligned} \tag{27}$$

where $\omega(t)$ and $v(t)$ are zero-mean, uncorrelated white Gaussian processes satisfying (2). The time delay is set to be $h = 0.3$.

Optimization problem (25) was solved using Matlab optimization toolbox. The initial value of (G, K) is computed using $h = 0$, and α in the penalty function is set to 100. The computed values are as follows:

$$\bar{h} = 1.6309, \|G_e(K,h)\|_2^2 = 0.0240$$

Using the computed (G, K), state estimation simulation was done, where a unit step signal was applied to the control input $u(t)$ at time 1s. The simulation results are given in Fig.3 and Fig.4: it can be seen that the proposed Kalman filter estimates system states well.

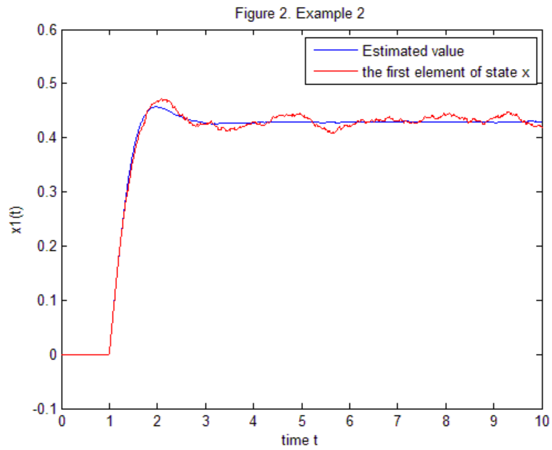

Fig. 3: Simulation result: true state (the first element of state x) and estimated value

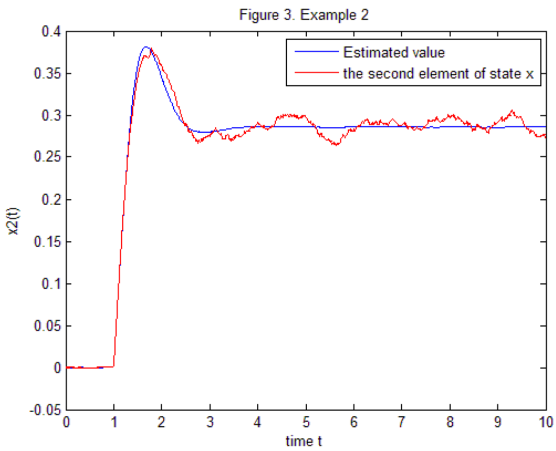

Fig. 4: Simulation result: true state (the second element of state x) and estimated value

To see how the time delay affects estimation performance, Kalman filters were designed for different h values.

As seen in Table 2, computed H_2 norm increases as time delay h increases. Variance of actual estimation error, which was computed from a simulation, also increases as time delay h increases. This verifies a common belief that the time delay adversely affects on estimation performance.

Table 2. Time delay effects on estimation performance.

	$h=0.1$	$h=0.3$	$h=0.5$	$h=0.7$
$\left\|G_e(K,h)\right\|_2^2$	0.0180	0.0243	0.0321	0.0424
Variance of actual estimation error	0.00088	0.00011	0.00013	0.00015

Example 3: Consider the following third-order system with delayed output and state:

$$\dot{x}(t)=\begin{bmatrix}-1 & 13.5 & -1\\ -3 & -1 & -2\\ -2 & -1 & -4\end{bmatrix}x(t)+\begin{bmatrix}-5.9 & 7.1 & -70.3\\ 2 & -1 & 5\\ 2 & 0 & 6\end{bmatrix}x(t-h)+\begin{bmatrix}0.2\\ 0.2\\ 0.2\end{bmatrix}\omega(t)+\begin{bmatrix}1\\ 1\\ 1\end{bmatrix}u(t) \tag{28}$$

$$y(t)=\begin{bmatrix}0 & 0 & 1\end{bmatrix}x(t)+\begin{bmatrix}1 & 1 & 1\end{bmatrix}x(t-h)+0.5v(t)$$

where $\omega(t)$ and $v(t)$ are the vectors of the input noise and measurement noise, respectively. In this example $h=0.06$.

The optimization problem (25) is solved by means of the Matlab optimization toolbox, and for this purpose, the optimization function “fminsearch” in Matlab is used.

By using $h=0$, the initial value for (G,K) is obtained. The value of α in the penalty function has been adjusted at 50. The values calculated for $h=0.06$ are as follows:

$$\bar{h}=0.1624,\ \left\|G_e(K,h)\right\|_2^2=1.3949$$

The simulation results are given in Fig.5, Fig.6 and Fig.7: it can be seen that the proposed H_2 filter estimates system states well.

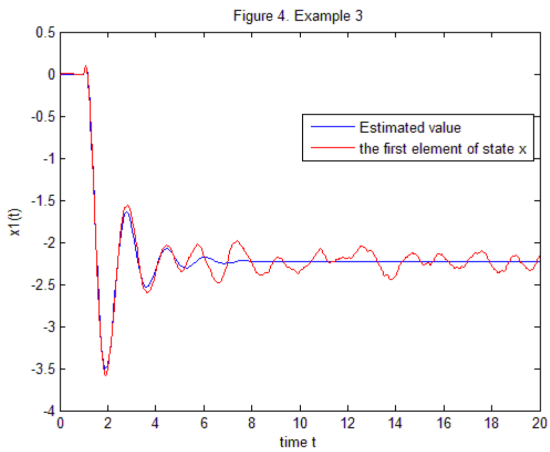

Fig. 5: Simulation result: true state (the first element of state x) and estimated value

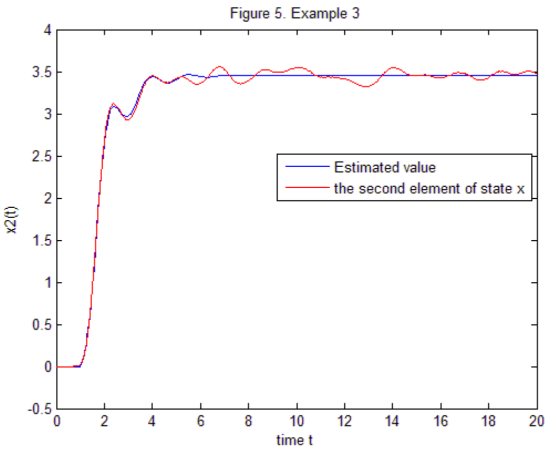

Fig. 6: Simulation result: true state (the second element of state x) and estimated value

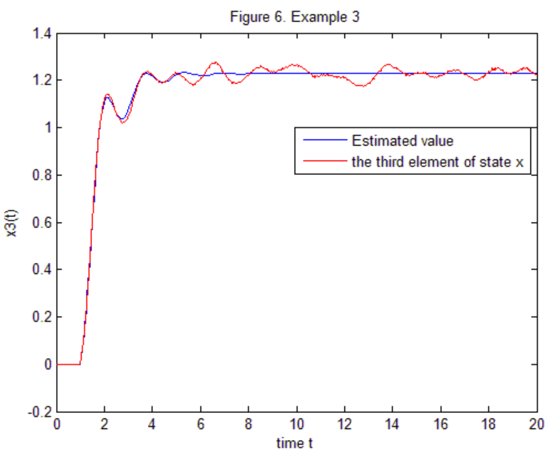

Fig. 7: Simulation result: true state (the third element of state x) and estimated value

As seen in Table 3, computed H_2 norm increases as time delay h increases.

Table 3. Time delay effects on estimation performance.

	$h = 0.01$	$h = 0.03$	$h = 0.06$	$h = 0.1$
$\|G_e(K,h)\|_2^2$	0.9765	1.0962	1.3949	2.0859
Variance of actual estimation error	0.000055	0.0001	0.0005	0.0007

As is observed, the increase of time delay has an opposite effect on the estimation performance, and with the increase of time delay, the estimation error variance also increases.

6 Conclusion

In this article, a method was proposed for the designing of Kalman filter for linear systems with time delay in the output and in state variables. By using the finite characterization of a Lyapunov functional equation, the existence of sufficient conditions for achieving the right solution and guaranteeing the proper convergence rate of the estimation error was evaluated. This observer provided satisfactory results in practical applications. Finally, by designing observers for three linear systems with time delays, the effectiveness of the proposed approach was demonstrated.

References:

[1] D. G. Luenberger, An Introduction to Observers, *IEEE Trans. Automat. Contr.*, vol. Ac-16, No. 6, 1971, pp. 596-602.

[2] Y. Chang, S. Chen, Static Output-Feedback Simultaneous Stabilization of Interval Time-Delay Systems, *WSEAS Transactions on systems*, Vol. 7, No. 3, March 2008, pp. 185-194.

[3] R. Brcena, A. Etxebarria, Industrial PC-based real-time controllers applied to second-order and first-order plus time delay processes, *WSEAS Transactions on systems*, Vol. 7, No. 9, 2008, pp. 870-879.

[4] Y. Ke, C. Miao, Stability analysis of BAM neural networks with inertial term and time delay, *WSEAS Transactions on systems*, Vol. 10, No. 12, 2011, pp. 425-438.

[5] E. Gyurkovics, T. Takacs, Output Feedback Guaranteeing Cost Control by Matrix Inequalities for Discrete-Time Delay Systems, *WSEAS Transactions on systems*, Vol. 7, No. 7, 2008, pp. 645-654.

[6] L. Y. CHANG, H. C. CHEN, Tuning of Fractional PID Controllers Using Adaptive Genetic Algorithm for Active Magnetic Bearing System, *WSEAS Transactions on systems*, Vol. 8, No. 1, 2009, pp. 158-167.

[7] D. Liu, G. Xu, N. E. Mastorakis, Reliability Analysis of a Deteriorating System with Delayed Vacation of Repairman, *WSEAS Transactions on systems*, Vol. 10, No. 12, 2011, pp. 413-424.

[8] L. Pekar, R. Prokop, P. Dostalek, Circuit Heating Plant Model with Internal Delays, *WSEAS Transactions on systems*, Vol. 8, No. 9, 2009, pp. 1093-1104.

[9] M. Hou, P. Zitek, and R. J. Patton, An observer design for linear time-delay systems, *IEEE Trans. on Automatic Control*, vol. 47, No. 1, 2002, pp. 121-125.

[10] M. Darouch, Linear functional observers for systems with delays in state variables, *IEEE Trans. on Automatic Control,* vol. 46, No. 3, 2001, pp. 491-496.
[11] A. E. Pearson and Y. A. Fiagbedzi, An observer for time lag systems, *IEEE Trans. on Automatic. Control*, vol. 34, No. 4, 1989, pp. 775-777.
[12] Y. X. Yao and Y. M. Zhang, Parameterization of observers for time delay systems and its application in observer design, *IEE Proc. Control Theory Appl.*, vol. 143, no. 3, 1996, pp. 225- 232.
[13] O. Sename, Unknown input robust observer for time delay system, *Proc. of the 36th IEEE Conference on Decision and Control*, pp. 1629-1630, 1997.
[14] J. Leyva-Ramos and A. E. Pearson, An asymptotic modal observer for linear autonomous time lag systems, *IEEE Trans. Automat. Contr.*, vol. 40, No. 7, 1995, pp. 1291-1294.
[15] M. Darouach, P. Pierrot, and E. Richard, Design of reduced-order observers without internal delays, *IEEE Trans. Automat. Contr.*, vol. 44, No. 9, 1999, pp. 1711-1713.
[16] M. Hou, P. Zitek, and R. J. Patton, An observer design for linear time-delay systems, *IEEE Trans. Automat. Contr.*, vol. 47, No. 1, 2002, pp. 121-125.
[17] P. M. Nia , R. Sipahi, An algebraic approach to design observers for delay-independent stability of systems with single output delay, *Proceedings of the American Control Conference*, (San Francisco), 2011, pp. 4231-4236
[18] A. Fattouh, O. Sename, and J. M. Dion, H_∞ observer design for time-delay systems, *Proc. of the 37th Conference on Decision and Control, (Florida, USA)*, 1998, pp. 4545-4546.
[19] E. Fridman, U. Shaked, and L. Xie, Robust $H\infty$ filtering of linear systems with time-varying delay, *IEEE Trans. Automat. Contr.*, vol. 48, No. 1, 2003, pp. 159-165.
[20] H. Gao and C. Wang, Delay-dependent robust $H\infty$ and $L2$ /$L\infty$ filtering for a class of uncertain nonlinear time-delay systems, *IEEE Trans. Automat. Contr.*, vol. 48, No. 9, 2003, pp. 1661-1666.
[21] S. Pakzad, and M. A. Pakzad, Stability condition for discrete systems with multiple state delays, *WSEAS Transactions on Systems and Control*, Vol.6, No.11, 2011, pp. 417-426.
[22] Y. S. Suh and S. Shin, Stability of State Delay Systems Based on Finite Characterization of a Lyapunov Functional, *Trans. of the Society of Instrument and Control Engineers*, Vol.35, No.9, 1999, pp. 1170-1175.
[23] E. Jarlebring, J. Vanbierviet, and W. Michiels., Explicit expressions for the H_2 norm of time-delay systems based on the delay Lyapunov equation, *in Proceedings of the 49st IEEE Control and Decision Conference*, (Atlanta, U.S.A), 2010, pp. 164-169.
[24] E. Jarlebring, J. Vanbierviet, and W. Michiels., Characterizing and computing the H_2 norm of time-delay systems based on the delay Lyapunov equation, *IEEE Trans. Automat. Contr.*, vol. 56, No. 4, 2011, pp. 814-825.
[25] M. S. Mahmoud, *Robust Control and Filtering For Time-Delay Systems*, Marcel Dekker, New York, 2000.
[26] D. Luenberger, *Optimization by Vector Space Methods*, Wiley, New York, 1969.
[27] M.Green and D.J.N.Limebeer, *Linear Robust Control*, Prentice-Hall, Englewood Cliffs, NJ, 1995
[28] Y. M. Fu, G. R Duan, and S. M. Song, Design of unknown input observer for linear time-delay systems , *International Journal of Control, Automation, Systems*, vol.2, no 4, 2004, pp 530-535.
[29] K. Subbarao, and P. Muralidhar, A State Observer for LTI Systems with Delayed Outputs: Time-Varying Delay, *Proceedings of the American Control Conference*, (Washington, U.S.A), 2008, pp. 3029-3033.
[30] C.-M. Zhang , G.-Y. Tang, M. Bai, State observer design for linear systems with delayed measurements, *Conference Proceedings - IEEE International Conference on Systems*, Man and Cybernetics, 2008, pp. 3222-3225.
[31] A. Fattouh, O. Sename, and J. M. Dion, An unknown input observer design for linear time-delay systems , *in Proceedings of the 38st IEEE Control and Decision Conference* ,(Arizona, USA) 1999., pp 4222-4227.
[32] Y. S. Suh, Kalman Filter for Output Delay Systems, *in Proceedings of the 41st IEEE Control and Decision Conference*,Vol. 3, (Las Vegas, U.S.A.), 2002, pp. 3033-3034.
[33] O. Sename, A. Fattouh, and J. M. Dion, Further results on unknown input observers design time-delay systems , *in Proceedings of the 40st IEEE Control and Decision Conference*, Florida, USA, 2001, pp4635-4636.
[34] Y. S. Suh, H. J. Kang, Y. S. Ro, H_2 Filter for Time Delay Systems, *International Journal of Control, Automation, Systems*, Vol. 4, No. 5, 2006, pp. 539-544

Autotuning Principles for Time-delay Systems

ROMAN PROKOP, JIŘÍ KORBEL, RADEK MATUŠŮ
Faculty of Applied Informatics
Tomas Bata University in Zlín
Nám. TGM 5555, 760 01 Zlín
CZECH REPUBLIC
prokop@fai.utb.cz http://www.fai.utb.cz

Abstract: This paper presents a set of single input – output (SISO) principles for tuning of continuous-time controllers used in autotuning schemes. The emphasis of designed autotuners is laid to SISO systems with time delay. Autotuners represent a combination of relay feedback identification and some control design method. In this contribution, models with up to three parameters are estimated by means of a single asymmetrical relay experiment. Then a stable low order transfer function with a time delay term is identified by a relay experiment. Controller parameters are analytically derived from general solutions of Diophantine equations in the ring of proper and stable rational functions R_{PS}. The generalization for a two degree of freedom (2DOF) control structure is performed. This approach covers a generalization of PID controllers and enables to define a scalar positive parameter for further tuning of the control performance. The analytical simple rule is derived for aperiodic control response and the scalar tuning parameter m>0. Autotuning principles of this contribution are applied to SISO systems with delays. Moreover, the Smith predictor scheme is applied for systems with a time delay term. The simulations are performed in the Matlab environment and a toolbox for automatic design and simulation was developed.

Key-Words: Algebraic control design, Relay experiment, Autotuning, Pole-placement problem, Smith predictor

1 Introduction

Proportional-Integral-Derivative (PID) controllers have survived changes in technology and they have been the most common way of using feedback in engineering systems [1], [2]. Yu in [3] refers that more than 97 % of control loops are of this type and most of them are actually under PI control. The practical advantages of PID controllers can be seen in a simple structure, in an understandable principle and in control capabilities. It is widely known that PID controllers are quite resistant to changes in the controlled process without meaningful deterioration of the loop behavior. The Ziegler – Nichols tuning rule has been glorified and vilified as well. However, there are many limitations, drawbacks and infirmities in the behavior of the Ziegler–Nichols setting. A solution for qualified choice of controller parameters can be seen in more sophisticated, proper and automatic tuning of PID controllers. Besides, the PID-based control loops are easy to simulate so no complex methods have to be used [4].

The development of various autotuning principles was started by a simple symmetrical relay feedback experiment proposed by Åström and Hägglund in [5] in the year 1984. The ultimate gain and ultimate frequency are then used for adjusting of parameters by original Ziegler-Nichols rules. During the period of more than two decades, many studies have been reported to extend and improve autotuners principles; see e.g. [6], [7], [11], [12]. The extension in relay utilization was performed in [3], [8], [10], [17] by an asymmetry and hysteresis of a relay. Over time, the direct estimation of transfer function parameters instead of critical values began to appear. Experiments with asymmetrical and dead-zone relay feedback are reported in [13]. Nowadays, almost all commercial industrial PID controllers provide the feature of autotuning.

In this paper, a new combination for autotunig method of PI and PID controllers with an aperiodic control rule is proposed and developed. The basic autotuning principle combines an asymmetrical relay identification experiment and a control design performed in the ring of proper and stable rational functions R_{PS}. The factorization approach proposed in [14] was generalized to a wide spectrum of control problems in [15], [18] - [23]. The pole placement problem in R_{PS} ring is formulated through a Diophantine equation and the pole is analytically tuned according to aperiodic response of the closed loop. The proposed method is compared by an equalization setting proposed in [16]. Naturally, there exist also different principles of control design syntheses which can be used for autotuning methods, e.g. [25], [31], [33].

This contribution deals with two simplest SISO linear dynamic systems with a delay term. The first model of the first order (stable) plus dead time (FOPDT) is supposed in the form:

$$G(s) = \frac{K}{Ts+1} \cdot e^{-\Theta s} \tag{1}$$

Similarly, the second order model plus dead time (SOPDT) is assumed in the form:

$$G(s) = \frac{K}{(Ts+1)^2} \cdot e^{-\Theta s} \tag{2}$$

The contribution is organized as follows. Section 2 represents a background of algebraic control design and the derivation for first and second order systems is derived. Section 3 deals with aperiodic tuning for a PI controller. Then the principle of the Smith predictor is introduced. Section 5 presents some facts about relay identification for autotuning principles. Then a Matlab program environment for design and simulations is described. Finally, section 7 presents a simulation results for three types of SISO systems.

2 Algebraic Control Design

The control design is based on the fractional approach; see e.g. [14], [15], [18]. Any transfer function $G(s)$ of a (continuous-time) linear system is expressed as a ratio of two elements of R_{PS}. The set R_{PS} means the ring of (Hurwitz) stable and proper rational functions. Traditional transfer functions as a ratio of two polynomials can be easily transformed into the fractional form simply by dividing, both the polynomial denominator and numerator by the same stable polynomial of the appropriate order.

Then all transfer functions can be expressed by the ratio:

$$G(s) = \frac{b(s)}{a(s)} = \frac{\dfrac{b(s)}{(s+m)^n}}{\dfrac{a(s)}{(s+m)^n}} = \frac{B(s)}{A(s)} \tag{3}$$

$$n = \max(\deg(a), \deg(b)), \quad m > 0 \tag{4}$$

Then, all feedback stabilizing controllers for the feedback system depicted in Fig. 1 are given by a general solution of the Diophantine equation:

$$AP + BQ = 1 \tag{5}$$

which can be expressed with Z free in R_{PS}:

$$\frac{Q}{P} = \frac{Q_0 - AZ}{P_0 + BZ} \tag{6}$$

In contrast of polynomial design, all controllers are proper and can be utilized.

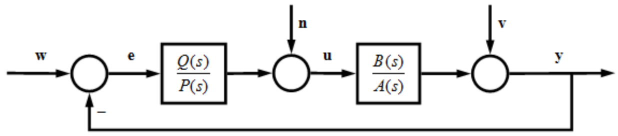

Fig. 1: One-degree of freedom (1DOF) control loop

The Diophantine equation for designing the feedforward controller depicted in Fig. 2 is:

$$F_w S + BR = 1 \tag{7}$$

with parametric solution:

$$\frac{R}{P} = \frac{R_0 - F_w Z}{P_0 + BZ} \tag{8}$$

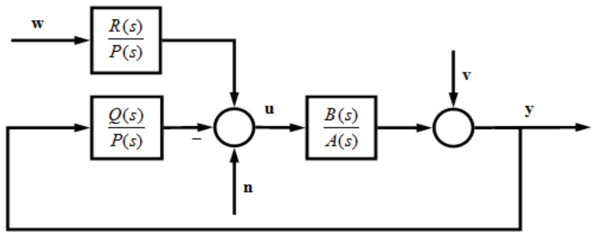

Fig. 2: Two-degree of freedom (2DOF) control loop

Asymptotic tracking is then ensured by the divisibility of the denominator P in (6) by the denominator of the reference $w = G_w / F_w$. The most frequent case is a stepwise reference with the denominator in the form:

$$F_w = \frac{s}{s+m}; \quad m > 0 \tag{9}$$

The similar conclusion is valid also for the load disturbance $d = G_d / F_d$. The load disturbance attenuation is then achieved by divisibility of P by F_d. More precisely, for tracking and attenuation in the closed loop according to Fig. 2 the multiple of AP must be divisible by the least common multiple of denominators of all input signals. The divisibility in R_{PS} is defined through unstable zeros and it can be achieved by a suitable choice of rational function Z in (6), see [14], [18] for details.

2.1 First order systems

Diophantine equation (5) for the first order systems (1) without the time delay term can be easily transformed into polynomial equation:

$$\frac{(Ts+1)}{s+m}p_0 + \frac{K}{s+m}q_0 = 1 \tag{10}$$

with general solution:

$$P = \frac{1}{T} + \frac{K}{s+m}\cdot Z$$
$$Q = \frac{Tm-1}{TK} - \frac{Ts+1}{s+m}\cdot Z \tag{11}$$

where Z is free in the ring R_{PS}. Asymptotic tracking is achieved by the choice:

$$Z = -\frac{m}{TK} \tag{12}$$

and the resulting PI controller is in the form:

$$C(s) = \frac{Q}{P} = \frac{q_1 s + q_0}{s} \tag{13}$$

where parameters q_1 a q_0 are given by:

$$q_1 = \frac{2Tm-1}{K} \qquad q_0 = \frac{Tm^2}{K} \tag{14}$$

The feedforward part of the 2DOF controller follows from (7):

$$\frac{s}{s+m}s_0 + \frac{K}{s+m}r_0 = 1 \tag{15}$$

with general solution:

$$P = \frac{1}{T} + \frac{K}{s+m}\cdot Z$$
$$R = \frac{m}{K} - \frac{s}{s+m}\cdot Z \tag{16}$$

The final PI like controller is given:

$$C_1(s) = \frac{R}{P} = \frac{r_1 s + r_0}{s} \tag{17}$$

with parameters

$$r_1 = \frac{Tm+m}{K} \qquad r_0 = \frac{Tm^2}{K} \tag{18}$$

2.2 Second order systems

The control synthesis for the SOPDT is based on stabilizing Diophantine equation (8) applied for the transfer function (4) without a time delay term. The Diophantine equation (5) takes the form:

$$\frac{(Ts+1)^2}{(s+m)^2}\cdot\frac{p_1 s + p_0}{s+m} + \frac{K}{(s+m)^2}\cdot\frac{q_1 s + q_0}{s+m} = 1 \tag{19}$$

and after equating the coefficients at like powers of s in (22) it is possible to obtain explicit formulas for *pi*, q_i:

$$p_1 = \frac{1}{T^2}; \qquad p_0 = \frac{3Tm-2}{T}$$
$$q_1 = \frac{1}{K}\left[3m^2 - \frac{1}{T^2}(1+3m-\frac{2}{T})\right]; \tag{20}$$
$$q_0 = \frac{1}{K}\left[m^3 - \frac{1}{T^2}(3m-\frac{2}{T})\right]$$

The rational function $P(s)$ has its parametric form (similar as in (14) for FOPDT):

$$P = \frac{p_1 s + p_0}{(s+m)} + \frac{K}{(s+m)^2}\cdot Z \tag{21}$$

with Z free in R_{PS}. Now, the function Z must be chosen so that P is divisible by the denominator of the reference which is (12). The required divisibility is achieved by $z_0 = -\frac{p_0 m}{K}$. Then, the particular solution for P, Q is

$$P = \frac{s\left[p_1 s + (p_1 m + p_0)\right]}{(s+m)^2}$$
$$Q = \frac{\tilde{q}_2 s^2 + \tilde{q}_1 s + \tilde{q}_0}{(s+m)^2}, \tag{22}$$

where

$$\tilde{q}_0 = q_0 + p_0 m, \quad \tilde{q}_1 = q_0 + q_1 m + 2Tp_0 m,$$
$$\tilde{q}_2 = q_1 + T^2 p_0 m. \tag{23}$$

The final (asymptotic tracking) controller has the transfer function:

$$C(s) = \frac{Q}{P} = \frac{\tilde{q}_2 s^2 + \tilde{q}_1 s + \tilde{q}_0}{s(p_1 s + (p_1 m + p_0))} \tag{24}$$

Also the feedforward part for the 2DOF structure can be derived for the second order system. For asymptotic tracking Diophantine equation takes the form:

$$\frac{s}{s+m}\frac{s_1 s + s_0}{(s+m)} + \frac{K}{(s+m)^2}r_0 = 1 \tag{25}$$

The 2DOF control law is only dependent upon the rational function R with general expression

$$R = \frac{m^2}{K} - \frac{s}{s+m}Z \tag{26}$$

also with Z free in R_{PS}. The final feedforward controller

$$C_1(s)=\frac{R}{P}=\frac{\dfrac{m^2}{K}(s+m)^2}{s\left[p_1s+(p_1m+p_0)\right]} \tag{27}$$

It is obvious that parameters of both parts of the controller (feedback and/or feedforward) depend on the tuning parameter $m>0$ in a nonlinear way. For both systems FOPDT and SOPDT the scalar parameter $m>0$ seems to be a suitable „tuning knob" influencing control performance as well as robustness properties of the closed loop system. Naturally, both derived controllers correspond to classical PI and PID ones. Equation (13) represents a PI controller:

$$u(t)=K_P\cdot\left(e(t)+\frac{1}{T_I}\cdot\int e(\tau)d\tau\right) \tag{28}$$

and the conversion of parameters is trivial. Relation (20) represents a PID in the standard four-parameter form [6]:

$$u(t)=K_P\cdot\left(e(t)+\frac{1}{T_I}\cdot\int e(\tau)d\tau+T_D y_f(t)\right)$$
$$\tau y_f'(t)+y_f(t)=y\ (t) \tag{29}$$

3 Aperiodic Tuning

There are many tuning principles and modifications of the Ziegler – Nichols rule developed from 1940s, see [6], [16], [25], [32]. Only in [25], more than 240 tuning rules are referred for PID and more than 100 rules for PI controllers.

A simple and attractive choice for the tuning parameter $m>0$ can be easily obtained analytically. In the R_{PS} expression, the closed-loop transfer function K_{wy} is for (1) and PI controller (13) given in a very simple form:

$$K_{wy}=\frac{BQ}{AP+BQ}=BQ=\frac{(2Tm-1)s+Tm^2}{(s+m)^2} \tag{30}$$

The step response of (30) can be expressed by Laplace transform:

$$h(t)=L^{-1}\left\{\frac{K_{wy}}{s}\right\}=L^{-1}\left\{\frac{k_1s+k_0}{s(s+m)^2}\right\}=$$
$$=L^{-1}\left\{\frac{A}{s}+\frac{B}{(s+m)}+\frac{C}{(s+m)^2}\right\}, \tag{31}$$

where A, B, C are calculated by comparing appropriate fractions in (31) and $k_1=2mT-1$, $k_0=Tm^2$.

The response $h(t)$ in time domain is then

$$h(t)=A+Be^{-mt}+Cte^{-mt} \tag{32}$$

The overshoot or undershoot of this response is characterized by the first derivative condition

$$h'(t)=-mBe^{-mt}+C(e^{-mt}-tme^{-mt})=0 \tag{33}$$

From (33) time of the extreme of response $h(t)$ is then easily calculated by the relation:

$$t_e=\frac{C-mB}{mC}=\frac{1}{m}-\frac{B}{C} \tag{34}$$

Since the aperiodic response means that the extreme does not exist for positive t_e, it implies $t_e<0$ and after substitutions of A, B, C, k_1, k_0 relation (34) takes the simple form

$$1<m\frac{B}{C}=\frac{1}{\dfrac{1}{Tm}-1} \tag{35}$$

The denominator of (35) must be positive and less than 1 and $m>0$ which implies the inequality:

$$\frac{1}{2T}<m<\frac{1}{T} \tag{36}$$

Any positive parameter m from (36) ensures aperiodic response. It is a question for further investigation and simulation what choice from interval (36) is the best. The time constant is always an estimation in the autotuning philosophy and then the middle value of (36) would be a reasonable choice in the form

$$m=\frac{3}{4\cdot T} \tag{37}$$

Also other tuning principles for aperiodic tuning certainly exist. For the mentioned algebraic synthesis, the equalization method developed by Gorez and Klán in [16]. The idea goes out from PI controller in the form (24). The tuning rule is very simple and it leads in relations:

$$K_P=\frac{1}{2K}\quad T_I=0.4\cdot T_u \tag{38}$$

where K is a process gain and T_u is the ultimate period obtained from the Ziegler-Nichols experiment. However, the fulfillment of (38) by unique value of $m>0$ is impossible, see [19]. The exact fulfillment of both relations in (38) could be obtained in the case of two distinct roots in denominator (30), so $(s+m_1)(s+m_2)$ instead of $(s+m)^2$.

4 Smith Predictors

The Smith predictor was designed in the late 1950s for systems with time delay, see e.g. [31], [32]. The basic classical interpretation of the Smith predictor is depicted in Fig. 3. The time delay term $e^{-\Theta s}$ has a negative influence to feedback stability which follows from frequency analysis. The feedback signal for the main controller $C(s)$ in Fig. 3 is a predicted value of the output. It means that the signal $y(t)$ inputs into the control error instead of the delayed $y(t-\Theta)$, it explains the name predictor. The Smith predictor launched the high development of *Internal Model Controllers* (*IMC*), where the plant model is present in the feedback loop (see [31], [33]). When the transfer function $G(s)$ is stable then the feedback systems in Fig. 3 is equivalent to the IMC version depicted in Fig. 4.

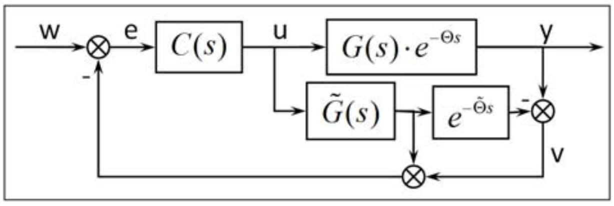

Fig. 3: Smith predictor – classical version

The main advantage of the Smith predictor is that the controller $C(s)$ can be designed according to delay-free part $G(s)$ of the plant. However, there are two main weak points in this sophisticated scheme. The first one is that the signal $v(t)$ is zero only in the case when the transfer function $G(s)$ is the same in the outer and inner loops in Fig. 3. The second weakness is that the transfer function must be stable. In the case of autotuning, always the approximated transfer function of the plant can be incorporated into the feedback.

Then the signal $v(t)$ in Fig. 3 and Fig. 4 is:

$$V(s) = G(s)e^{-\Theta s} - \tilde{G}(s)e^{-\tilde{\Theta}s} \tag{39}$$

In the case of discrepancy, this non-zero signal indeed negatively influences the control performance. Note that the nominal transfer function for control design is $\tilde{G}(s)$.

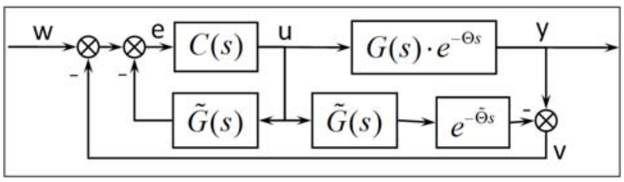

Fig. 4: Smith predictor – IMC version

5 Relay Feedback Estimation

The estimation of the process or ultimate parameters is a crucial point in all autotuning principles. The relay feedback test can utilize various types of relay for the parameter estimation procedure. The classical relay feedback test [5] was proposed for stable processes by symmetrical relay without hysteresis. Following sustained oscillation are then used for determining the critical (ultimate) values. The control parameters (PI or PID) are then generated in standard manner.

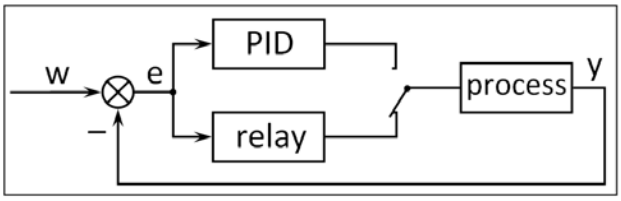

Fig. 5: Block diagram of an autotuning principle

Asymmetrical relays with or without hysteresis bring further progress [3], [17]. After the relay feedback test, the estimation of process parameters can be performed. A typical data response of such relay experiment is depicted in Fig. 6. The relay asymmetry is required for the process gain estimation (40) while a symmetrical relay would cause the zero division in the appropriate formula.

In this paper, an asymmetrical relay with hysteresis is used. This relay enables to estimate transfer function parameters as well as a time delay term. For the purpose of the aperiodic tuning the time delay is not exploited.

The process gain can be computed by the relation (see [13]):

$$K = \frac{\int_0^{iT_y} y(t)dt}{\int_0^{iT_y} u(t)dt}; \quad i = 1,2,3,.. \tag{40}$$

The time constant and time delay terms are then given by:

$$\begin{aligned} T &= \frac{T_y}{2\pi}\cdot\sqrt{\frac{16\cdot K^2\cdot u_0^2}{\pi^2\cdot a_y^2}-1} \\ \Theta &= \frac{T_y}{2\pi}\cdot\left[\pi - arctg\frac{2\pi T}{T_y} - arctg\frac{\varepsilon}{\sqrt{a_y^2-\varepsilon^2}}\right] \end{aligned} \tag{41}$$

where a_y and T_y are depicted in Fig. 6 and ε is the hysteresis.

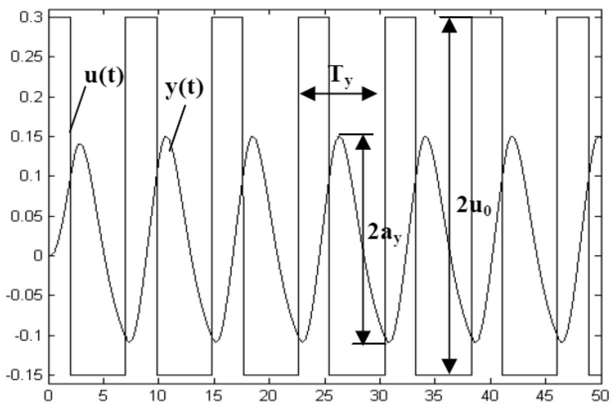

Fig. 6: Asymmetrical relay oscillation

The gain is given by (40), the time constant and time delay term can be estimated according to [13] by the relation:

$$
\begin{aligned}
T &= \frac{T_y}{2\pi} \cdot \sqrt{\frac{4 \cdot K \cdot u_0}{\pi \cdot a_y} - 1} \\
\Theta &= \frac{T_y}{2\pi} \cdot \left[\pi - 2arctg \frac{2\pi T}{T_y} - arctg \frac{\varepsilon}{\sqrt{a_y^2 - \varepsilon^2}} \right]
\end{aligned}
\qquad (42)
$$

6 Simulation and Program System

A Matlab program system was developed for engineering applications of auto-tuning principles. This program enables a choice for the identification of the controlled system of arbitrary order. The estimated model is of a first or second order transfer function with time delay. The user can choose three cases for the time delay term. In the first case the time term is neglected, in the second one the term is approximated by the Pade expansion and the third case utilizes the Smith predictor control structure. The program is developed with the support of the Polynomial Toolbox. The Main menu window of the program system can be seen in Fig. 7.

In the first phase of the program routine, the controlled transfer function is defined and parameters for the relay experiment can be adjusted. Then, the experiment is performed and it can be repeated with modified parameters if necessary. After the experiment, an estimated transfer function in the form of (1) or (2) is performed automatically and controller parameters are generated after pushing of the appropriate button. Parameters for experimental adjustment are defined in the upper part of the window.

The second phase begins with the "Design controller parameters" button and the actual control design is performed. According to above mentioned methodology and identified parameters, the controller is derived and displayed. The control scheme depends on the choice for the 1DOF or 2DOF structure and on the choice of the treatment with the time delay term.

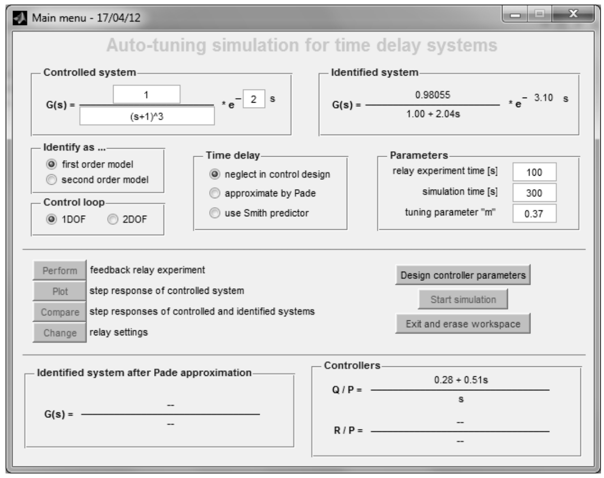

Fig. 7: Main Menu

During the third phase, after pushing the "Start simulation" button, the simulation routine is performed and required outputs are displayed. The simulation horizon can be prescribed as well as tuning parameter m, other simulation parameters can be specified in the Simulink environment. In all simulation a change of the step reference is performed in the second third of the simulation horizon and a step change in the load is injected in the last third. A typical control loop of the case with the Smith predictor in Simulink is depicted in Fig. 8.

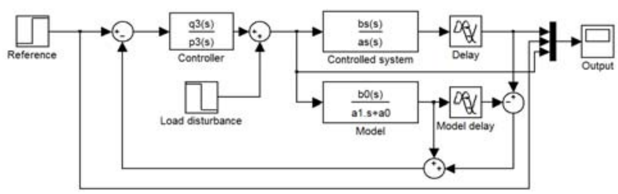

Fig. 8: Control loop in Simulink

Also the step responses can be displayed and the comparison of the controlled and estimated systems can be depicted. Another versions of the similar program systems were developed and they are referred in e.g. [19], [20].

7 Examples and Simulations

The following examples illustrate the situation where the estimated model is in the form (1) or (2) with a time delay term. The controllers are designed according to Part 2 with neglecting of the time delays.

<u>Example 1:</u> A second order controlled system with time delay with the transfer function:

$$G(s) = \frac{1}{(2s+1)^2} \cdot e^{-2s} \tag{43}$$

was identified by the relay experiments as a first and second order system. The results give the following transfer functions:

$$\tilde{G}(s) = \frac{0.98}{3.46s+1} \cdot e^{-2.77s}, \qquad \tilde{\tilde{G}}(s) = \frac{0.98}{3.41s^2+3.69s+1} \cdot e^{-2.49s} \tag{44}$$

The first controller was designed for the identified system with neglecting of the time delay term and the tuning parameter m = 0.22 was derived from the aperiodic condition (36). The PID for the second order estimation (44) was designed for the tuning parameter m = 0.41. The final controllers are governed by the transfer functions:

$$C_1(s) = \frac{\tilde{Q}(s)}{\tilde{P}(s)} = \frac{0.51s+0.17}{s}, \quad C_2(s) = \frac{\tilde{\tilde{Q}}(s)}{\tilde{\tilde{P}}(s)} = \frac{0.71s^2+0.70s+0.18}{1.85s^2+s}, \quad C_3(s) = \frac{\tilde{\tilde{R}}(s)}{\tilde{\tilde{P}}(s)} = \frac{1.06s^2+0.86s+0.18}{1.85s^2+s} \tag{45}$$

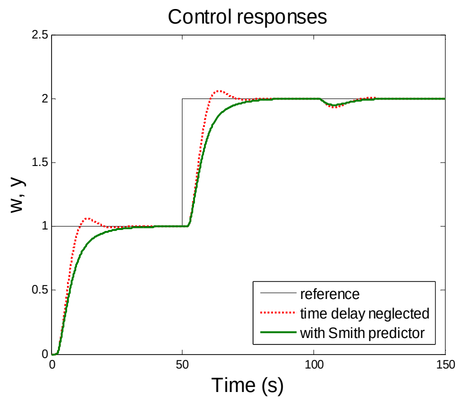

Fig. 9: Control responses 1DOF first order

The original system $G(s)$ from (43) was controlled by (44) in two different control ways. The simple control response in the sense of 1DOF is depicted in Fig. 9 by dashed line while the Smith predictor scheme represents an aperiodic response in the same figure. Fig. 10 displays the same simulation for the second order controller C_2 in (45).

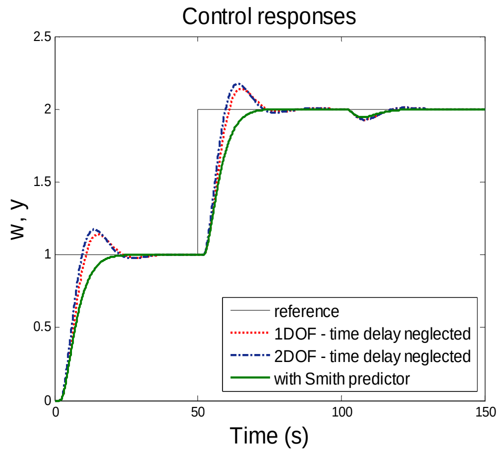

Fig. 10: Control responses second order

<u>Example 2:</u> A fifth order system with time delay $G(s)$ was identified in the form of a first order transfer function with time delay:

$$G(s) = \frac{3}{(2s+1)^5} \cdot e^{-5s} \tag{46}$$

The first and second order estimation results in the following transfer functions:

$$\tilde{G}(s) = \frac{2.99}{5.88s+1} \cdot e^{-10.35s}, \qquad \tilde{\tilde{G}}(s) = \frac{2.99}{11.19s^2+6.69s+1} \cdot e^{-8.49s} \tag{47}$$

Then controllers were designed for the identified models (47) with time delay terms neglected. The PI controller was derived for the value of m = 0.13 and the PID one was derived for m = 0.22. Both controllers in the 1DOF structure have the transfer functions:

$$C_1(s) = \frac{0.17s+0.03}{s}, \qquad C_2(s) = \frac{0.42s^2+0.23s+0.03}{3.35s^2+s} \tag{48}$$

The control responses for the first order approximation and design are depicted in Fig. 11. In this case the difference of responses between neglecting the time delay term and with the use of the Smith predictor is remarkably stronger. While standard feedback control response is quite poor and oscillating then the response with Smith predictor in the loop is smooth and aperiodic.

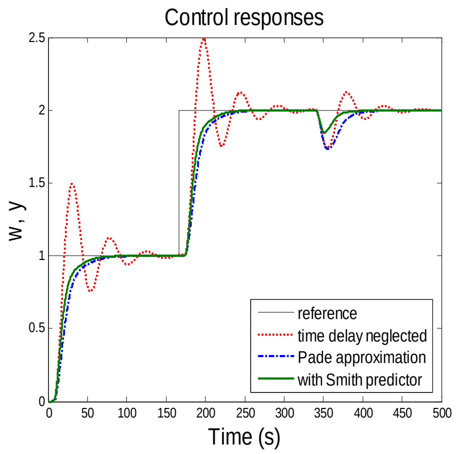

Fig. 11: Control responses 1DOF - first order

Almost the same situation is illustrated in Fig. 12 where the second order approximation and synthesis were utilized. However, comparison of Fig. 11 and Fig. 12 shows that the first order synthesis is sufficient and the second order is redundant.

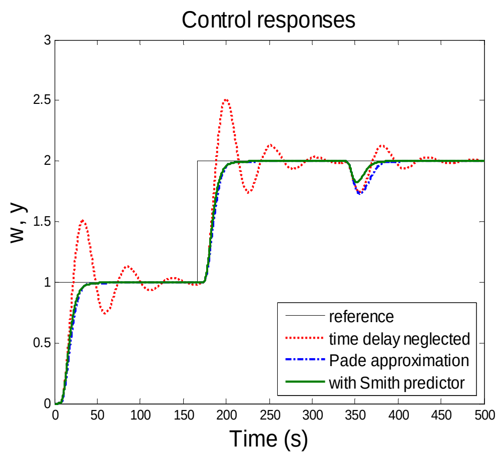

Fig. 12: Control responses 1DOF - second order

Example 3: This example represents a case of higher order system without delay approximated by a law order system with a time delay term. A higher order system (8th order) with transfer function $G(s)$ is supposed:

$$G(s) = \frac{3}{(s+1)^8} \tag{49}$$

After the relay experiment, a first order and second estimation gives the following transfer functions:

$$\tilde{G}(s) = \frac{2.96}{4.22s+1} \cdot e^{-4.96s}$$
$$\tilde{\tilde{G}}(s) = \frac{2.96}{4.83s^2+4.40s+1} \cdot e^{-4s} \tag{50}$$

The step responses of systems (49) and (50) are shown in Fig. 13.

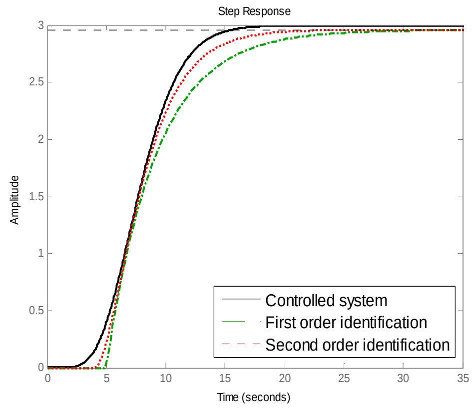

Fig. 13: Step responses of systems (49)

Naturally, both step responses of the estimated systems are quite different from the original system $G(s)$.

Again, PI controllers are derived from (10), (11) and the tuning parameter $m>0$ can influence the control behaviour. Since the difference of controlled and estimated systems is considerable, it can be expected that not all values of and some of $m>0$ represent acceptable behaviour.

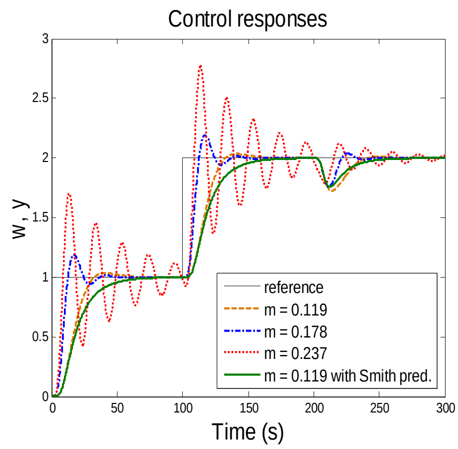

Fig. 14: Control responses 1DOF first order

With respect of (36), three responses are shown in Fig. 14. Generally, larger values of $m>0$ implicate larger overshoots and oscillations. As a consequence, for inaccurate relay identifications, lower values of $m>0$ in interval (36) can be recommended. The PI controller for $m = 0.18$ has the form

$$C(s) = \frac{0.17s + 0.05}{s} \tag{51}$$

The control responses for (49) and (51) with and without the Smith predictor are shown in Fig. 14.

The second order identification and synthesis of example 3 for $m = 0.34$ gives the PID controller:

$$C(s) = \frac{0.28s^2 + 0.23s + 0.05}{2.20s^2 + s} \tag{52}$$

The higher order system (49) was controlled by (52) and two responses are depicted in Fig. 15. The first one represents neglecting of a time delay term in (50) while the second one utilizes the Smith predictor structure. It is obvious that the Smith predictor brings a significant improvement of overshoots.

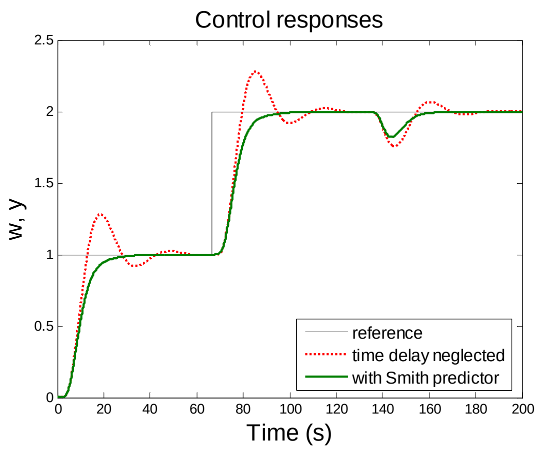

Fig. 15: Control responses 1DOF second order

8 Conclusion

This contribution gives some rules for autotuning principles with a combination of relay feedback identification and a control design method.

The estimation of a low order transfer function parameters is performed from asymmetric limit cycle data, see [13]. The control synthesis is carried out through the solution of a linear Diophantine equation according to [14], [15], [18]. This approach brings a scalar tuning parameter which can be adjusted by various strategies. A first order estimated model generates PI-like controllers while a second order model generates a class of PID ones. The aperiodic tuning through the parameter $m>0$ is proposed by the analytic derivation, more details in [20]. In both cases also the Smith predictor influence was compared with neglecting of time delay terms. The methodology is illustrated by several examples of various orders and dynamics. The results of all simulations prove that the Smith predictor structure brings a significant improvement of the aperiodic responses. The price for the improvement is a more complex structure of the feedback control system.

Acknowledgements

This work was supported by the European Regional Development Fund under the project CEBIA-Tech No. CZ.11.05./2.1.00 /03.0089.

References:

[1] D. H. Kim, J. H. Cho, Robust Tuning of PID Controller with Disturbance Rejection Using Bacterial Foraging Based Optimization, *WSEAS Transactions on Systems*, Vol. 3, No. 11, 2004, pp. 2834-2840.

[2] P. C. Patic, L. Duta, Comparative Study Regarding the Optimization and Analysis of Automatic Systems using PID, LQR and PI-MIMO, *WSEAS Transactions on Systems*, Vol. 9, No. 7, 2010, pp. 754-763.

[3] Ch.Ch. Yu, *Autotuning of PID Controllers. Springer,* London, 1999.

[4] F. Neri, Software Agents as a Versatile Simulation Tool to Model Complex Systems. *WSEAS Transactions on Information Science and Applications*, WSEAS Press (Wisconsin, USA), Vol. 7, No. 5, 2010, pp. 609-618.

[5] K.J. Åström and T. Hägglund, Automatic tuning of simple regulators with specification on phase and amplitude margins. *Automatica*, Vol.20, 1984, pp. 645-651.

[6] K.J. Åström and T. Hägglund, *PID Controllers: Theory, Design and Tuning.* Research Triangle Park, NC: Instrumental Society of America, 1995.

[7] R.F. Garcia and F.J.P. Castelo, A complement to autotuning methods on PID controllers, In: *Preprints of IFAC Workshop PID'00*, 2000, pp. 101-106.

[8] R.R. Pecharromán and F.L. Pagola, Control design for PID controllers auto-tuning based on improved identification, In: *Preprints of IFAC Workshop PID'00,* pp. 89-94, 2000.

[9] C.C. Hang, K.J. Åström and Q.C. Wang, Relay feedback auto-tuning of process controllers – a tutorial review, *Journal of Process Control*, Vol. 12, No. 6, 2002.

[10] T. Thyagarajan and Ch.Ch. Yu, Improved autotuning using shape factor from relay feedback, In: *Preprints of IFAC World Congres*, 2002.

[11] S. Majhi and D.P. Atherton, Autotuning and controller design for unstable time delay processes, In*: Preprints of UKACC Conf an Control,* 1998, pp. 769-774.

[12] F. Morilla, A. Gonzáles and N. Duro, Auto-tuning PID controllers in terms of relative damping , In: *Preprints of IFAC Workshop PID'00,* 2000, pp. 161-166.

[13] M. Vítečková, and A. Víteček, Plant identification by relay methods. In: *Engineering the future (edited by L. Dudas)*. Sciyo, Rijeka, 2010, pp. 242-256.

[14] M. Vidyasagar*, Control System Synthesis: A Factorization Approach*. MIT Press, Cambridge, M.A., 1987.

[15] V. Kučera, Diophantine equations in control - A survey, Automatica, Vol. 29, No. 6, 1993, pp. 1361-75.

[16] R. Gorez and P. Klán, Nonmodel-based explicit design relations for PID controllers, In: *Preprints of IFAC Workshop PID'00*, 2000, pp. 141-146.

[17] I. Kaya and D.P. Atherton, Parameter estimation from relay autotuning with asymmetric limit cycle data, *Journal of Process Control*, Vol. 11, No4, 2001, pp. 429-439.

[18] R. Prokop and J.P. Corriou, Design and analysis of simple robust controllers*, Int. J. Control,* Vol. 66, 1997, pp. 905-921.

[19] R. Prokop, J. Korbel and Z. Prokopová, Relay feedback autotuning – A polynomial approach, In: *Preprints of 15th IFAC World Congress*, 2010.

[20] R. Prokop, Korbel, J. and Prokopová, Z., Relay based autotuning with algebraic control design, In: *Preprints of the 23rd European Conf. on modelling and Simulation*, Madrid, 2009, pp. 531-536.

[21] R. Matušů and R. Prokop, Robust Stabilization of Interval Plants using Kronecker Summation Method. In: *Last Trends on Systems, 14th WSEAS International Conference on Systems*, Corfu Island, Greece, 2010, pp. 261-265.

[22] L. Pekař and R. Prokop, Non-delay depending stability of a time-delay system. In*: Last Trends on Systems, 14th WSEAS International Conference on Systems*, Corfu Island, Greece, 2010, pp. 271-275.

[23] L. Pekař and R. Prokop, Control of Delayed Integrating Processes Using Two Feedback Controllers: RMS Approach, In*: Proceedings of the 7th WSEAS International Conference on System Science and Simulation in Engineering*, Venice, 2008, pp. 35-40.

[24] R. Prokop, J. Korbel and O.Líška, A novel principle for relay-based autotuning. *International Journal of Mathematical Models and Methods in Applied Science,* 2011, Vol. 5, No. 7, pp. 1180-1188.

[25] A. O´Dwyer, *Handbook of PI and PID controller tuning rules.* London: Imperial College Press, 2003.

[26] L. Pekař and R. Prokop: Algebraic Control of integrating Processes with Dead Time by Two Feedback Controllers in the Ring RMS. *Int. J. of Circuits, Systems and Signal Processing*, Vol. 2, No. 4, 2008, pp. 249-263.

[27] P. Dostálek, J. Dolinay, V. Vašek and L. Pekař. Self-tuning digital PID controller implemented on –bit Freescale microcontroller. *International Journal of Mathematical Models and Methods in Applied Sciences,* Vol. 4, No. 4, 2010, pp. 274-281.

[28] L. Pekař, R. Prokop and R. Matušů. Stability conditions for a retarded quasipolynomial and their applications*. International Journal of Mathematics and Computers in Simulations*, Vol. 4, No. 3, 2010, pp. 90-98

[29] R. Matušů and R. Prokop. Experimental verification of design methods for conventional PI/PID controllers. *WSEAS Trans. on Systems and Control*, Vol. 5, Issue 5, 2010, pp.269-280.

[30] Q. Zhohg. *Robust Control of Time-delay systems.* Springer, London, 2006.

[31] K.J. Åström and R.M. Murray, *Feedback Systems.* Research Triangle Park, NC: Instrumental Society of America, 1995.

[32] O.J.M. Smith, *Feedback Control Systems.* McGraw-Hill Book CompanyInc., 1958.

[33] M. Morari and E. Zafiriou, *Robust Process Control.* Prentice Hall, New Jersey, 1989.

A Ring for Description and Control of Time-Delay Systems

LIBOR PEKAR
Department of Automation and Control Engineering, Faculty of Applied Informatics
Tomas Bata University in Zlin
nam. T. G. Masaryka 5555, 76001 Zlin
CZECH REPUBLIC
pekar@fai.utb.cz http://www.fai.utb.cz

Abstract: - This contribution aims a revision and extension of the ring of retarded quasipolynomial meromorphic functions (R_{MS}) for description and control of time-delay systems (TDS). The original definition has some significant drawbacks – especially, it does not constitute a ring. Our new definition extends the usability to neutral TDS and to those with distributed delays. As first, basic algebraic notions useful for this paper are introduced. A concise overview of algebraic methods for TDS follows. The original and the revised definitions of the ring together with some its properties finish the contribution. There are many illustrative examples that explain introduced terms and findings throughout the paper.

Key-Words: - Time-delay systems, Algebraic description and control, Ring, Coprime factorization, Bézout domain

1 Introduction

Algebraic structures in their charming and attractive elegance proved to be suitable and effective tools for system dynamics description and control system design. Modern control theory has been adopting algebraic approaches and parlance, which are based on TDS description in a suitable field, ring or module and the subsequent operation in the algebraic structure, for decades.

The aim of this paper is to introduce a revise the definition and some basic properties of the R_{MS} ring for description and control of TDS in input-output space, unlike some other methods using state-space domain which prevail. R_{MS} structure was originally introduced in [1]; however, the genesis of the idea can be view already in works of Vidyasagar [2] and Kucera [3] for delayless systems and/or in [4] for TDS. Nevertheless, it has been pointed out in [5] that the structure does not constitute ring. In addition to that, the structure is applicable to retarded systems only and it brings problems when comprising models with distributed delays.

The revised and extended structure can useful when analysis and control of neutral TDS and those with distributed delays. Some stability notions are also discussed and taken into account. Basic properties of the revisited R_{MS} are given for the record as well. To illuminate the ideas and statements, many illustrative examples are introduced throughout the paper.

The paper is organized as follows. Section 2 provides an overview of algebraic notions useful for uninitiated readers to comprehend the rest of the contribution. A non-exhaustive introduction to algebraic structures and methods used in description, analysis and control of TDS can be found in Section 3. The original and a revised definition of R_{MS} are the contents of Section 4. Section 5 includes a list of selected properties of the now conception supported by some examples. Section 6 concludes the paper and outlines the usability of the R_{MS} ring.

2 Basic Algebraic Notions

Prior to a brief overview of particular algebraic structures utilized by some authors when analysis (and/or synthesis) of TDS, it is convenient to introduce some basic algebraic notions being used in this paper and their elementary properties if useful, see e.g. [6], [7].

A *group*, G, is an algebraic structure with binary operation $\cdot$ satisfying:

a) For each $a,b \in G$, it holds that $a \cdot b \in G$.

b) For all $a,b,c \in G$, $(a \cdot b) \cdot c = a \cdot (b \cdot c) \in G$ (associativity).

c) There exists an element $e \in G$, such that for every element $a \in G$, it holds that $a = a \cdot e = e \cdot a \in G$ (*identity element, neutral element*).

d) For each $a \in G$, there exists an element $b \in G$ such that $a \cdot b = b \cdot a = e \in G$ (*inverse element*).

A set satisfying a) and b) only from the definition above, i.e. without a necessity of identity and inverse elements, is called a *semigroup*. If one requires the existence of an identity element, so-called *monoid* is obtained. A group with the commutative property, i.e.

e) For each $a,b \in G$, $a \cdot b = b \cdot a \in G$

is called a commutative (abelian) group.

A *ring*, *R*, is a set with two binary operations +, · (generally interpreted summation and addition) for which it holds true the following:

a) *R* is a commutative group under addition with an identity element denotes as 0.

b) For any $a,b,c \in R$, $(a+b) \cdot c = a \cdot b + b \cdot c \in R$ and $c \cdot (a+b) = c \cdot a + c \cdot b \in R$ (left and right distributivity).

c) For every $a,b,c \in R$, it holds that $(a \cdot b) \cdot c = a \cdot (b \cdot c) \in R$ (Associativity of multiplication).

Some authors add another property of a ring as:

d) There exists $1 \in R$ such that for every $a \neq 0 \in R$, $a \cdot 1 = 1 \cdot a \in R$ (multiplicative identity).

If d) holds, then a ring is a commutative group under + and a commutative monoid under ·, together with distributivity. In a *commutative ring*, the commutative property holds also for multiplication.

A *unit* of the ring (or an *invertible element*) is $a \neq 0 \in R$, for which there exists $a^{-1} \in R$, such that $a \cdot a^{-1} = a^{-1} \cdot a = 1$. If all elements of a ring are units, the ring is called a *field*.

It is said that $b \in R$ *divides* $a \in R$ (i.e. $b \mid a$) if there exists $q \in R$, such that $a = q \cdot b$. Two elements $a,b \in R$ are *associated* if $b \mid a$ and $a \mid b$.

Let *R* be a commutative ring and $a,b \in R$. A *common divisor* $c \in R$ of *a*, *b* is an element of the ring, for which $c \mid a$ and $c \mid b$. $d \in R$ is the *greatest common divisor* (GCD) of *a*, *b* if for every common divisor $c \in R$ of $a,b \in R$ it holds that $c \mid d$. The CGD is determined unambiguously except for associativity.

A nonzero noninvertible element a of a commutative ring R is called *irreducible* if it is divisible solely by a unit or any element associated with a. In some rings, so-called *prime elements* generalizing prime numbers are introduced. A prime elements is a nonzero noninvertible $a \in R$, such that if $a \mid (b \cdot c)$ for some $b,c \in R$, then always $a \mid b$ or $a \mid c$. Every prime element is irreducible, the converse is not true in general.

A ring *R* in which every nonzero noninvertible $a \in R$ can be uniquely decomposed in a (finite) product of irreducible or prime elements (except for the ordering and associativity) is called a unique factorization ring (UFR).

A commutative ring with identity (under multiplication) such that for any two elements $a \neq 0 \in R$ and $b \neq 0 \in R$ it holds that $a \cdot b \neq 0$ is called an *integral domain*. An URF which is an integral domain is labeled as a unique factorization domain (UFD).

A *field of fractions* of an integral domain *R* (at least with one element) is the "smallest" field containing *R*, such that necessary elements satisfying the divisibility (by a nonzero element) are added. An element c of this field can be expresses in the form $c = a / b$ where $a,b \in R$, $b \neq 0$.

An *ideal I* (of the ring *R*) is a subset of *R* with the following properties:

a) For every $a,b \in I$, it holds that $a + b \in I$.

b) For each $a \in I$ and $r \in R$, $a \cdot r \in I$.

It holds that an intersection of ideals is an ideal as well. Let $M = \{a_1, a_2, \ldots a_n\} \subseteq R$, then an intersection of all ideals of *R* containing *M* is called an ideal *generated by M*. It is also the "smallest" ideal including *M*. Ideals of the form $aR = \{a \cdot r \mid r \in R\}$, i.e. those generated by (the only one) element *a* are called *principal*.

If every ideal of an integral domain is principal, so-called *principal ideal domain* (PID) is obtained. It holds true that every PID is UFD; however, the converse is not true in general.

A *Noetherian* ring *R* is primarily defined as that satisfying the so-called finite ascending chain condition. Equivalently, it is possible to circumscribe the term as follows: A ring *R* is Noetherian if its every ideal is finitely generated, i.e. $n = |M|$ is a finite number.

A (left) *module* (or *R-module*) *M* over the ring *R* is a commutative group satisfying:

a) For every $r \in R$, $a,b \in M$, it holds that $r \cdot (a+b) = r \cdot a + r \cdot b \in M$.

b) For every $r,s \in R$, $a \in M$, $(r+s) \cdot a = r \cdot a + s \cdot a \in M$.

c) For every $r,s \in R$, $a \in M$, $(r \cdot s) \cdot a = r \cdot (s \cdot a) \in M$.

d) If there exists a multiplicative identity $1 \in R$, and $a \in M$, then $1 \cdot a = a \in M$

Modules are similar to vector spaces, yet in modules, coefficients are taken from rings, not from fields. A *free module* is that with a basis. For instance, since nonzero elements in a ring are not necessarily invertible, a relation $\sum_{i=1}^{n} r_i \cdot a_i = 0, r_i \in R, a_i \in M$, where M is a free module, does not imply that each r_i is the linear combination of the remaining ones (Conte and Perdon, 2000).

A *partially ordered set* (poset) is defined as an ordered pair $P = (S, \preceq)$ where S is called the ground set of P and $\preceq$ is the partial order of P. A relation $\preceq$ is a poset on S if:

a) For all $a \in S$, $a \preceq a$ (reflexivity)

b) For $a, b \in S$, if $a \preceq b$ and $b \preceq a$, then $a \equiv b$ (antisymmetry)

c) For $a, b, c \in S$, $a \preceq b$ and $b \preceq c$ implies $a \preceq b$ (transitivity)

From a PID, a *Bézout domain* is distinguished in which every *finitely* generated ideal is principal. In a Bézout domain, PID is UFD and viceversa. Thus, a PID admits the existence of an infinitely generated ideal which is principal.

In a Bézout domain *R*, for every pair $a, b \in R$ (or generally for a finite set of elements) there exists the GCD which meets the *Bézout identity* (or more generally a *linear Diophantine equation*)

$$a \cdot x + b \cdot y = \mathrm{GCD}(a, b),\ x, y \in R \tag{1}$$

A solution $x, y \in R$ is not determined uniquely but (an infinitely many) solutions of (1) are given by the parameterization

$$x = x_0 \pm z \cdot \frac{b}{\mathrm{GCD}(a,b)},\ y = y_0 \mp z \cdot \frac{a}{\mathrm{GCD}(a,b)} \tag{2}$$

where $\{x_0,\ y_0\}$ is a particular solution of (1) and $z \in R$

If (1) is solved for any $c \in R$ on the right-hand side instead of $\mathrm{GCD}(a,b)$, it is necessary to verify whether there exists $\mathrm{GCD}(a,b)$ (especially in a ring which is not Bézout or PID) for which $\mathrm{GCD}(a,b) \mid c$.

The Bézout identity can be solved e.g. using a *generalized* (extended) *Euclidean algorithm* which can be described as follows. Let a, b be given and the task is to find $d = \mathrm{GCD}(a,b)$ and a pair x, y according to (1). The iterative procedure can be written as follows

$$\begin{aligned} & r_i = r_{i-2} - \lfloor q_i \rfloor \cdot r_{i-1},\ r_{i-2} \ge r_{i-1} \ge r_i \\ & i = 3, 4, \ldots, n \end{aligned} \tag{3}$$

i.e. the current reminder r_i of the division can be expressed by preceding reminders r_{i-1}, r_{i-2} and using the whole quotient q_i.

In every step of the algorithm, it is possible to write the following identity

$$r_i = a \cdot x_i + b \cdot y_i \tag{4}$$

where x_i, y_i are from the ring. The first two reminders are chosen as

$$\begin{aligned} & r_1 = a = a \cdot 1 + b \cdot 0 \\ & r_2 = b = a \cdot 0 + b \cdot 1 \end{aligned} \tag{5}$$

The desired $d = \mathrm{GCD}(a,b)$ then equals the last nonzero reminder, $r_n \ne 0,\ n < \infty$.

The whole procedure can be expressed in a table (matrix) form as follows

$$\left[\begin{array}{cc|c} 1 & 0 & a \\ 0 & 1 & b \end{array}\right] \sim \begin{array}{c}\text{elementary} \\ \text{matrix} \\ \text{operations}\end{array} \sim \left[\begin{array}{cc|c} v & t & 0 \\ x & y & d \end{array}\right] \tag{6}$$

The result is determined by two Diophantine equations

$$\begin{aligned} & a \cdot v + b \cdot t = 0 \\ & a \cdot x + b \cdot y = d \end{aligned} \tag{7}$$

In the case when (1) is solved for any fixed $c \in R$ on the right-hand side instead of $d = \mathrm{GCD}(a,b)$ it is possible (if a solution exists) to use the extended Euclidean algorithm again in the following two possibilities:

1) To use scheme (6) for $c \in R$ instead of $d = \mathrm{GCD}(a,b)$. Generally, it is not necessary to achieve the zero element on the upper right matrix corner.

2) Obviously

$$\begin{aligned} & a \cdot x + b \cdot y = \mathrm{GCD}(a,b) \quad / \frac{c}{\mathrm{GCD}(a,b)} \\ & a \frac{xc}{\mathrm{GCD}(a,b)} + b \frac{yc}{\mathrm{GCD}(a,b)} = c \\ & a x_1 + b \cdot y_1 = c \end{aligned} \tag{8}$$

Hence, $\mathrm{GCD}(a,b)$, *x,y* are found using (6) first, and subsequently, the following substitution is used

$$x_1 = x\frac{c}{\mathrm{GCD}(a,b)},\ y_1 = y\frac{c}{\mathrm{GCD}(a,b)} \qquad (9)$$

to get the desired solution.

For the necessity and comprehension of the further text, some basic notions from the complex functions analysis ought to be introduced as well.

A *holomorphic function* is a complex-valued function of a single (or multiple) complex variable defined on a region $\mathbb{D} \subseteq \mathbb{C}$ which is infinitely complex differentiable (i.e. there exists all complex derivatives) at any point $z_0 \in \mathbb{D}$.

The term holomorphic function is often used interchangeably with or compared to an *analytic function* which is generally a complex-valued function of a single (or multiple) complex variable defined on a region $\mathbb{D} \subseteq \mathbb{C}$, in which the Taylor series expansion exists at every point $z_0 \in \mathbb{D}$. That is, a series $T(z) = \frac{1}{i!}\sum_{i=0}^{\infty} f^{(i)}(z_0)(z-z_0)^i$ converges to $f(z)$ for every point z from a neighborhood of z_0. For complex functions, a holomorphic function implies an analytic function. A function holomorphic on all $\mathbb{C}$ is called *entire*.

An *isolated singularity* of a complex function $f(z)$ is a point z_0, in which the function is not differentiable; however, there exists an open disk D centered at z_0 such that $f(z)$ is holomorphic on the disk excluding z_0. There are several types of isolated singularities. A *pole* is an isolated singularity z_0 of $f(z)$ such that $f(z)$ converges uniformly to infinity for $z \to z_0$. Thus, if there exists the improper limit $\lim_{z\to z_0} f(z) = \infty$, then there exists also $n \in \mathbb{N}$, so that $\lim_{z\to z_0} (z-z_0)^n f(z) < \infty$. A *removable singularity* is another type of an isolated one for which $\lim_{z\to z_0} f(z) \neq \infty$. In this case, it is possible to define $f(z_0) = \lim_{z\to z_0} f(z)$, so that $f(z)$ becomes holomorphic. An *essential singularity* represents the last type of an isolated singularity which evinces "peculiar" behavior within the neighborhood of the singularity, and it holds that the limit $\lim_{z\to z_0} f(z)$ does not exist here.

A *meromorphic function* is a complex-valued function of a complex variable which is holomorphic on an open subset $\mathbb{D} \subseteq \mathbb{C}$ except a set of poles. The function can be expressed as a ratio of two holomorphic functions.

3 Fields, Rings and Modules for Description and Control of TDS

The nascence of algebraic methods in description of TDS is connected with fields, namely with systems over fields [9], which can be written in the (retarded) state-space form

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)\\ \mathbf{y}(t) &= \mathbf{C}\mathbf{x}(t)\end{aligned} \qquad (10)$$

where elements of $\mathbf{A},\mathbf{B},\mathbf{C}$ are from a fixed field and $\dot{\mathbf{x}}(t) = \frac{\mathrm{d}\mathbf{x}(t)}{\mathrm{d}t}$.

The next step was to further generalize the concept of linear systems, to include the case in which coefficients belong to a ring. The first, general, in-depth research into the properties of systems over rings was constituted in [10], [11]. One of the primordial attempts to utilize ring theory to infinite-dimensional linear systems was made by Kamen [12] where an operator theory was presented, the particular case of systems defined via rings of distributions. Namely, the ring Θ generated by the entire functions $\theta_\sigma(s)$ defined as

$$\begin{aligned}&\varphi_\sigma(s) = 0.5(\theta_\sigma + \theta_{\bar{\sigma}}), \psi_\sigma(s) = 0.5\mathrm{j}(\theta_\sigma - \theta_{\bar{\sigma}})\\ &\theta_\sigma(s) = \frac{1-\exp(-\tau(s-\sigma))}{s-\sigma}, \sigma \in \mathbb{C}\end{aligned} \qquad (11)$$

and their derivatives and 1 was introduced there. Ring models for TDS with lumped delays was published in [13].

In [14], linear systems over commutative rings, especially TDS, were intensively studied. The author i.a. presented the simplest TDS over rings, those with commensurate delays where the introduction of the operator $\delta x(t) := x(t-\tau)$, where τ represents the smallest delay, yields state matrix entries in the ring of polynomials $\mathrm{R}[\delta]$. In more details, let the model be

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \sum_{k=0}^{N} \mathbf{A}_k \mathbf{x}(t-k\tau) + \mathbf{B}_k \mathbf{u}(t-k\tau)\\ \mathbf{y}(t) &= \sum_{k=0}^{N} \mathbf{C}_k \mathbf{x}(t-k\tau)\end{aligned} \qquad (12)$$

then state and output matrices in (10) read

$$\mathbf{A} = \sum_{k=0}^{N} \mathbf{A}_k \delta^k, \mathbf{B} = \sum_{k=0}^{N} \mathbf{B}_k \delta^k, \mathbf{C} = \sum_{k=0}^{N} \mathbf{C}_k \delta^k \qquad (13)$$

Using a substitution $\delta^k \to \exp(-k\tau s)$, one can obtain the Laplace transform form of the state model for TDS with commensurate delays. If delays are not commensurate, we need to define a finite set of delay operators $\delta_1, \delta_2, ..., \delta_N$ resulting in a ring $\mathbb{R}[\delta_1, \delta_2, ..., \delta_N]$. Some authors, e.g. Youla [15], introduced the field $\mathbb{R}(\delta_1, \delta_2, ..., \delta_N)$ of rational functions in $\mathbb{R}[\delta_1, \delta_2, ..., \delta_N]$ in order to study networks with transmission lines (i.e. delayed system). Reachability and observability of a general system with coefficients over a ring are analyzed in [14] as well.

Conte and Perdon in [16] further studied the realization of such systems. These authors also developed the geometrical approach to the study of dynamical systems with coefficients over a ring concerning TDS. The overview of the methodology was presented in [8]. In this framework, the main tool lies in the view that $\mathbf{x}(t), \mathbf{u}(t), \mathbf{y}(t)$ in (10) are free *R*-modules.

Concerning input-output maps, which are substantive for the aim of this paper, the conception of 2-D systems which naturally arises from the transfer function of a TDS with commensurate delays over a ring (12), (13) was introduced in [13], [14]. Translating the state-space description into the transfer function results in a rational function in s and $\exp(-\tau s)$. This expresses that two operators are used here, i.e. the integrator and the delay operator, which are algebraically independent (due to the fact that the exponential term is a transcendental function) in the meaning of that there is no nontrivial linear combination of s and $\exp(-\tau s)$ over real numbers equals to zero. Thus, the ring $\mathbb{R}[s, \exp(-\tau s)]$ of quasipolynomials, which is isomorphic to the ring of real polynomials in two variables (a so-called 2-D polynomial) $\mathbb{R}[s,z]$, is obtained. Quasipolynomials defined here are connected with commensurate delays. This concept was further studied and developed e.g. in [17], [18].

However, some authors pointed out that the use quasipolynomials does not permit to effectively handle some stabilization and control tasks, thus other rings based on quasipolynomials for TDS with commensurate delays were introduced.

For instance, in [4], [19] there were established the following rings: A ring $\mathcal{R} = \Theta \cup \mathbb{R}[\exp(-\tau s)] = \Theta[\exp(-\tau s)]$ of all linear combinations, with real coefficients, of distributed delays from Θ and lumped delays, and a ring $\mathcal{E} = \mathrm{P}[s] = \Theta \cup \mathbb{R}[s, \exp(-\tau s)]$ of so-called *pseudopolynomials* which consists of Laplace transforms of operators that are generated using derivatives, lumped and distributed delays. Any element $T(s) \in \mathcal{E}$ can be written in the (coprime) form $T(s) \in N(s, \exp(-\tau s))/D(s)$, $N(s, \exp(-\tau s)) \in \mathbb{R}[s, \exp(-\tau s)]$, $D(s) \in \mathbb{R}[s]$. Two pseudopolynomials are coprime if and only if there are neither their common zeros nor factors in the form $\exp(-k\tau s)$. Ring $\mathcal{R}[s]$ is not isomorphic to $\mathcal{R}[x]$, which means that the variables are not algebraically independent (transcendental) over $\mathcal{R}$, see an example in [4]. Moreover, it is a Bézout domain, yet not a Euclidean ring nor a Noetherian ring nor a UFD. Notice that $\mathcal{E}$ and $\mathbb{R}[s, \exp(-\tau s)]$ share the same field of fractions, i.e. $\mathbb{R}(s, \exp(-\tau s))$. The transfer function can then be expresses as a fraction of two pseudopolynomials.

Behavioral approach, as it was introduced for dynamical systems in [20], was presented by [21] for TDS, again with commensurate delays. In contrast to above mentioned works, the author considered systems in the behavioral point of view instead of systems over rings. A behavior is the kernel of a delay-differential operator. More precisely, consider equations in the scalar case in the form

$$\sum_{j=0}^{L}\sum_{i=0}^{N} p_{ij} x^{(i)}(t-j) = 0 \tag{14}$$

where $p_{ij}, t \in \mathbb{R}$, $x^{(i)}(t)$ denotes the *i*-th derivative of the function $x(t)$: $\mathbb{R} \to \mathbb{R}$. Behaviors $\mathcal{B}$ are those functions $x(t)$ satisfying (14). Alternatively, $\mathcal{B} = \ker \widetilde{P}$ where $P = \sum_{j=0}^{L}\sum_{i=0}^{N} s^i z^j \in \mathbb{R}[s,z]$ and $\widetilde{P}$ denotes the associated delay-differential operator, i.e. $\widetilde{P}x(t) = \sum_{j=0}^{L}\sum_{i=0}^{N} p_{ij} x^{(i)}(t-j)$. It is stated in [21] that it is algebraically more adequate to consider the ring $\mathbb{R}[s, z, z^{-1}]$ instead of $\mathbb{R}[s,z]$. There is also defined the ring

$$\mathcal{H} := \{p \in \mathbb{R}(s)[z, z^{-1}] \mid p(s,z) \in H_C \tag{15}$$

as the appropriate domain in order to translate relations between behaviors, lying between $\mathbb{R}[s, z, z^{-1}]$ and $\mathbb{R}(s)[z, z^{-1}]$, where the latter means the ring of polynomials in z, z^{-1} with the coefficients in rational functions in s with real parameters, and H_C is the set of all entire functions.

It was proved that $\mathcal{H}$ is not UFD and not a Noetherian ring; however, it is a Bézout ring.

However, delays are naturally real-valued and thus the limitation to commensurate delays is rather restrictive for real applications [22]. Dealing with rings for input-output maps of TDS with even non-commensurate delays, it is crucial for this paper to mention here the family of approaches (originally developed for delayless systems) utilizing a field of fractions where the transfer function is expressed as a ratio of two coprime (or relatively prime) elements of a suitable ring [2], [3], [23]. The process of finding such coprime pair is called a *coprime factorization*.

One of such rings for continuous-time systems is a ring of stable and proper rational functions, R_{PS} [3], [24]. An element of this ring is defined as a ratio of two polynomials in s over $\mathbb{R}$ where the denominator polynomial is Hurwitz stable (i.e. free of roots located in the closed right-half plane including imaginary axis) and, moreover, the ratio is proper (i.e. the s-degree of the numerator is less or equal to the denominator). Alternatively, the element of R_{PS} is analytic and bounded for $\operatorname{Re} s \geq 0$ including infinity, i.e. it lies in $H_\infty(\mathbb{C}^+)$. Such a definition is, however, not sufficient for TDS since e.g. the Laplace form of a stable system including in $H_\infty(\mathbb{C}^+)$ can have an unstable denominator.

The utilization of R_{PS} in description (and control) of TDS requires a rational approximation of a general meromorphic transfer function as a first step of a coprime factorization, for instance, by a substitution of the exponential terms, $\exp(-\tau s) \approx X(s) \in \mathbb{R}(s)$, see e.g. [25], [26]. A similar idea, yet over $\mathbb{R}[s]$ was presented e.g. in [27].

An example of a coprime factorization in R_{PS} follows.

Example 1. Consider a stable TDS with distributed delays governed by the transfer function

$$G(s) = \frac{Y(s)}{U(s)} = \frac{1 - \exp(1)\exp(-s)}{s - 1} \tag{16}$$

Use of, e.g., the first order Padé rational approximation results in

$$G(s) = \frac{Y(s)}{U(s)} \approx \frac{0.5s(1 + \exp(1)) + 1 - \exp(1)}{(s - 1)(0.5s + 1)} = \frac{b(s)}{a(s)} \tag{17}$$

where $a(s)$, $b(s) \in \mathbb{R}[s]$. Notice that the common root $s = -1$ (removable singularity) characterizing the delay distribution in this example vanished after the rationalization. An addition, although the relative order of the transfer function is preserved, the absolute one has increased. To establish coprime factors $A(s) = a(s)/m(s)$, $B(s) = b(s)/m(s)$, $m(s) \in \mathbb{R}[s]$ (with no zero in $\mathbb{C}^+$), $A(s) \in R_{PS}$, $B(s) \in R_{PS}$, one has to realize the divisibility condition in R_{PS}: Any $A(s) \in R_{PS}$ divides $B(s) \in R_{PS}$ if and only if all unstable zeros (including $s \to \infty$) of $A(s)$ are those of $B(s)$. Inclusion of infinity in the condition gives rise to the requirement $\deg m(s) = \deg a(s) = 2$, and moreover, there is no s with $\operatorname{Re} s \geq 0$ satisfying $m(s) = 0$. ■

The main drawback of the ring, i.e. the necessity of a rational approximation, induces the idea of introduction a similar, yet rather different, ring avoiding this operation.

4 R_{MS} Ring

4.1 Original definition

The original definition of the ring of proper and stable retarded quasipolynomial (RQ) meromorphic functions, R_{MS}, is the subject of this subsection [1]. The basic idea for its introduction proceeds from the following ideas. First, as mentioned above in the previous section, a rational approximation of the transfer function in the form of a ratio of two quasipolynomials is required for the use of the ring R_{PS}. This operation brings a loss of system dynamics information, as can be seen from Example 1. Second, from the practical point of view, there is no reason to be limited to commensurate delays in a model, thus, a more universal description ought to be introduced. Third, authors took into account the fact that two variables, z and s, are not independent from the functional point of view, thus, a one-dimensional (1-D) instead of 2-D approach can be used. Last but not least, as stated above, quasipolynomials in the transfer function do not permit to effectively handle some stabilization and control tasks such as impulse-free stability and controller properness and parameterization.

Definition 1 (R_{MS} ring – original). An element $T(s) \in R_{MS}$ is represented by a proper fraction of two quasipolynomials

$$T(s) = \frac{y(s)}{x(s)} \tag{18}$$

where a denominator $x(s)$ is a quasipolynomial of degree n and a numerator can be factorized as

$$y(s) = \tilde{y}(s)\exp(-\tau s) \tag{19}$$

where $\tilde{y}(s)$ is a quasipolynomial of degree l and $\tau \geq 0$. $x(s)$ is stable, which means that there is no zero of $x(s)$, s_0, such that $\operatorname{Re} s_0 \geq 0$. Moreover, the ratio is proper, i.e. $l \leq n$. ■

Obviously, the condition $\tau > 0$ is too restrictive (or more likely a misprint); the inequality $\tau \geq 0$ would be more natural instead. The original definition of R_{MS} has some drawbacks; especially, it does not constitute a ring, which requires making some changes in the definition. Namely, although the retarded structure of TDS is considered only, the minimal ring conditions require the use of neutral quasipolynomials at least in the numerator of $T(s)$. Moreover, the original definition brings problems when comprising models with distributed delays and handling a coprime factorization.

4.2 H_∞ and BIBO stability

To comprehend the revisited definition, notion of H_∞, BIBO, formal and strong stability have to be briefly introduced first.

A system is H_∞ stable if its transfer function $G(s)$ lies in the space $H_\infty(\mathbb{C}^+)$ of functions analytic and bounded in the right-half complex plane, i.e. providing the finite norm

$$\|G\|_\infty := \sup\{G(s) : \operatorname{Re} s \geq 0\} < \infty \tag{20}$$

see e.g. [27]. That is, the system has finite $L_2(0,\infty)$ to $L_2(0,\infty)$ gain where $L_2(0,\infty)$ norm of an input or output signal $h(t)$ is defined as

$$\|h(t)\|_2 := \sqrt{\int_0^\infty |h(t)|^2 \, \mathrm{d}t} \tag{21}$$

Notice, for instance, that a transfer function having no pole in the right-half complex plane but a sequence of poles with real part converging to zero can be H_∞ unstable due to unbounded gain at the imaginary axis [28].

The notion of *BIBO* (Bounded Input Bounded Output) stability is stronger than the preceding one and usually more difficult to analyze. A single-input single-output (SISO) TDS is BIBO stable if a bounded input $|u(t)| < M_1$, $t < 0$, $M_1 \in \mathbb{R}$ produces a bounded output $|y(t)| < M_2$, $t < 0$, $M_2 \in \mathbb{R}$; in other words, it has a finite L_∞ gain. It holds that the system is BIBO stable if its transfer function is an element of a commutative Banach algebra $\Lambda(L_1 + \mathbb{R}\delta)$ of Laplace transforms of functions of the form

$$h(t) = h_a(t) + \sum_{i=1}^{\infty} h_i \delta(t - \tau_i), t \geq 0 \tag{22}$$

where $h_a(t) \in L_1(0,\infty)$, i.e.

$$\int_0^\infty |h_a(t)| \mathrm{d}t < \infty \tag{23}$$

$h_i \in \mathbb{R}$, $\tau_0 = 0, \tau_i > 0$, for $i > 0$, $\delta(t)$ stands for the Dirac delta function, and

$$\sum_{i=1}^{\infty} |h_i| < \infty \tag{24}$$

BIBO stability implies H_∞ stability [29], [30].

Formal stability of neutral TDS is defined in the state-space domain and this theory is going beyond the topic of this paper. However, it can be formulated simply as follows: formal stability means that the system has only a finite number of poles in the right-half complex plane [31]. In other words, the rightmost vertical strip of poles does not reach or cross the imaginary axis.

The feature of a neutral TDS that the position of the rightmost vertical strip is not continuous in real axis is not continuous [32] gives rise to another (yet a germane) stability notion. *Strong* stability means that the strip remains in $\mathbb{C}_0^-$ when subjected to small variations in delays (i.e. a TDS remains formally stable). Although this stability notion is defined in state-space domain, the following input-output test can be performed

$$\sum_{j=1}^{h_n} |m_{nj}| < 1 \tag{25}$$

see e.g. [33], [34] where m_{nj} are coefficients for the highest s-power in the characteristic quasipolynomial (transfer function denominator)

$$m_0(s) = s^n + \sum_{i=0}^{n} \sum_{j=1}^{h_i} m_{ij} s^i \exp(-s\eta_{ij}), \eta_{ij} \geq 0 \tag{26}$$

4.3 Revised definition

The following simple example shows that the original definition does not constitute a ring.

Example 2. Consider two elements of R_{MS}

$$T_1(s)=\frac{s}{s+2},T_2(s)=\frac{(s+1)\exp(-s)}{s+2} \tag{27}$$

Yet, a sum of them

$$\begin{aligned}T(s)&=T_1(s)+T_2(s)\\&=\frac{s(1+\exp(-s))+\exp(-s)}{s+2}\notin R_{MS}\end{aligned} \tag{28}$$

since the numerator is a neutral (even formally unstable) quasipolynomial, which is inconsistent with the original ring definition. ■

The above introduced example indicates that it is necessary to include neutral terms in the definition.

The second drawback comes from the requirement of stable denominator. The transfer function of a stable TDS with distributed delays has common numerator and denominator root from the right-half plane; however, there is no reason to consider it as unstable in any sense, see e.g. stable system (16). Rephrased, an element of the ring should include a removable singularity in $\mathbb{C}^+$ (but not poles). Analogously, spectral stabilizability can be viewed in the similar manner [35].

Because of this, $H_\infty(\mathbb{C}^+)$ seems to be a suitable candidate for the ring definition (as for R_{PS} ring).

However, there are some troubles with neutral systems, namely, although a formally unstable neutral TDS with a vertical strip of poles tending to the imaginary axis from left (for $\operatorname{Im} s_0\to\infty$) can be BIBO (and hence $H_\infty(\mathbb{C}^+)$) stable, it does not permit the so called Bézout factorization, [28], [30]. Any two elements $A(s),B(s)\in H_\infty(\mathbb{C}^+)$ form a Bézout (coprime) factorization if and only if

$$\inf_{\operatorname{Re}s\geq 0}\left(|A(s)|+|B(s)|\right)>0 \tag{29}$$

i.e. there exist (a stabilizing coprime pair) $Q(s),P(s)\in H_\infty(\mathbb{C}^+)$, such that

$$A(s)P(s)+B(s)Q(s)=1 \tag{30}$$

Example 3. A TDS of neutral type has a transfer function

$$G(s)=\frac{Y(s)}{U(s)}=\frac{b(s)}{a(s)}=\frac{1}{(1-\exp(-s))(s+1)} \tag{31}$$

Clearly, a pair

$$B(s)=\frac{1}{s+2},A(s)=\frac{(1-\exp(-s))(s+1)}{s+2} \tag{32}$$

has no nontrivial (non-unit) common factor, i.e. it is coprime. However, $|A(\pm k2\pi j)|=0,k\in\mathbb{N}$, and $\lim_{k\to\infty}|B(\pm k2\pi j)|=0$, hence (29) does not holds true and the system is not Bézout coprime nor BIBO stabilizable. ■

As stated in [35] for neutral-type TDS, a system that is not formally stable is not BIBO stable nor stabilizable. However, this is not true exactly, as shown in [28].

Since formal stability is not given in input-output relation (transfer function), consider a rather more strict notion – strong stability – given by condition (25) instead. Formal stability is hence required; however, its testing by strong stability condition (25) could not be included in the ring definition since it may lead to strong instability when algebraic operations on ring elements.

The following short examples demonstrate and clarify the above ideas.

Example 4. Let be given three neutral delayed systems (plants) governed by transfer functions

$$\begin{aligned}G_1(s)&=\frac{1}{s+s\exp(-s)+1},\\G_2(s)&=\frac{1}{(s+s\exp(-s)+1)(s+1)},\\G_3(s)&=\frac{1}{(s+s\exp(-s)+1)(s+1)^4}\end{aligned} \tag{33}$$

All the systems have poles located in the “stable” half-plane $\mathbb{C}_0^-$, except for $\operatorname{Im} s_0\to\infty$ where the asymptotic pole lies on the imaginary axis, see Fig. 1, where displayed poles (blue asterisks) are -0.4011, -0.0379 + 3.4264j, -0.0054 + 9.5293j, -0.0020 + 15.7713j, -0.0010 + 22.0365j, -0.0006 + 28.3096j, -0.0004 + 34.5864j, -0.0003 + 40.8652j, -0.0002 + 47.1451j. (2.67)

However, although there is no pole (except the asymptotic case) in $\mathbb{C}^+$, neutral systems (33) can not be considered as asymptotically stable since the is no positive α satisfying $\operatorname{Re} s_0\leq-\alpha$ for all s_0, which is necessary for stability of neutral TDS.

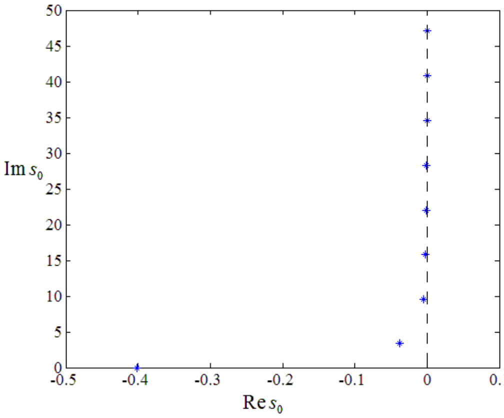

Fig. 1. Root loci of the rightmost poles of $G_1(s)$ from (33)

Moreover, these systems are neither strongly nor formally stable, simply, the chain of poles reaches the imaginary axis. Nevertheless, other stability notions are more attractive. An easy test on $G_1(\mathrm{j}\omega)$, $G_2(\mathrm{j}\omega)$, $G_3(\mathrm{j}\omega)$ shows that $\|G_1\|_\infty = \infty$, $\|G_2\|_\infty = 2$, $\|G_3\|_\infty = 1$, hence $G_1 \notin H_\infty(\mathbb{C}^+)$, $G_2, G_3 \in H_\infty(\mathbb{C}^+)$. As proved in [27], G_1 and G_2 are not BIBO stable, yet G_3 is BIBO stable. This means that formal instability does not automatically implies H_∞ or BIBO instability which makes problems when decision about the inclusion of the system into an algebraic structure (or set). ■

Example 5. This example demonstrates the necessity of formal stability in the definition of R_{MS} ring, not only for elements of R_{MS} but also for their inversions.

Consider a coprime factorization in $H_\infty(\mathbb{C}^+)$ of system $G_2(s)$ from (33), i.e.

$$G(s) = \frac{\dfrac{1}{(s+2)^2}}{\dfrac{[(1+\exp(-s)s+1)(s+1)]}{(s+2)^2}} = \frac{B(s)}{A(s)} \tag{34}$$

More information about (Bézout) coprime factorization can be found in Section 5. Notice that the factorization (34) is coprime yet not Bézout.

As stated above, the system $G(s)$ is formally unstable but from $H_\infty(\mathbb{C}^+)$, i.e. $B(s)/A(s) \in H_\infty(\mathbb{C}^+)$. However, one can verify that $1/A(s) \notin H_\infty(\mathbb{C}^+)$. This yields a mismatch in the ring definition since there is not an unambiguous answer whether $A(s)$ is invertible (a unit) or not. If both terms were not coprime, it would not pose a problem since such situations are natural also in R_{PS} ring. If $G(s)$ was formally stable, it would hold that $1/A(s) \in H_\infty(\mathbb{C}^+)$. As a conclusion, a set $H_\infty(\mathbb{C}^+)$ is not a sufficient candidate for R_{MS} ring. ■

Hence, there seem to be two possibilities for the ring definitions regarding formal stability. Either to include the requirement of formal stability of the quasipolynomial numerator in the ring definition and thus to exclude the existence of (Bézout) coprime factorization for formally unstable systems, or to take it into consideration in ring divisibility conditions. Naturally, we decided to choose the latter option, since it is not possible to avoid a formal unstable numerator in ring elements as demonstrated in Example 2.

Example 6. The aim of this example is to show that strong stability could not be included in the ring definition; however, the necessity of formal stability has been already proved in Example 5.

Consider a formally and strongly stable element from $H_\infty(\mathbb{C}^+)$

$$T(s) = \frac{1}{(1+\exp(-0.8s))s+1} \tag{35}$$

Now make a multiplication

$$T_2(s) = T(s)T(s) = \frac{1}{[(1+\exp(-0.8s))s+1]^2} = \frac{1}{(1+\exp(-1.6s)+2\exp(-0.8s))s^2+2(1+\exp(-0.8s))s+1} \tag{36}$$

which is obviously strongly unstable, yet formally stable, since $T(s)$ and $T_2(s)$ have the same spectrum (except for poles multiplicity). Hence, this algebraic operation (multiplication) preserves formal yet not strong stability. Recall, however, that formal stability will be tested by verification of strong stability, so there is some kind of conservativeness. ■

The crucial part of this section, the R_{MS} ring proposal, as a revisited and extended definition to the original one, follows.

Definition 2 (R_{MS} ring – a revision). An element $T(s)$ of R_{MS} ring is represented by a ratio of two (quasi)polynomials $y(s)/x(s)$ where the denominator is a (quasi)polynomial of degree n and the numerator can be factorized as

$$y(s) = \tilde{y}(s)\exp(-\tau s) \tag{37}$$

where $\tilde{y}(s)$ is a (quasi)polynomial of degree l and $\tau \geq 0$. Note that the degree of a quasipolynomial means its highest s-power.

The element lies in the space $H_\infty(\mathbb{C}^+)$, i.e. it is analytic and bounded in $\mathbb{C}^+$, particularly, there is no pole s_0 such that $\operatorname{Re} s_0 \geq 0$ for a retarded denominator or $\operatorname{Re} s_0 \geq -\varepsilon, \varepsilon > 0$ for a neutral one. If the term includes distributed delays, all roots of $x(s)$ in $\mathbb{C}^+$ are those of $y(s)$ (i.e. removable singularities). Moreover, $T(s)$ is formally stable. The strong stability condition (25) for (quasi)polynomial $x(s)$ is a sufficient but not necessary condition guaranteeing that.

In addition, the ratio is proper, i.e. $l \leq n$. More precisely, there exists a real number $R > 0$ for which holds that

$$\sup_{\operatorname{Re} s > 0, |s| \geq R} |T(s)| < \infty \tag{38}$$

see [28]. ■

5 Basic Properties of the Ring

5.1 Coprime factorization and Bézout identity

A basic operation on the quasipolynomial transfer function of TDS is coprime factorization by which the transfer function is decomposed into a coprime (or relatively prime) pair of ring elements. Since, in controller design, the intention is to use coprime factors in the Bézout equation (30), the factorization should also be Bézout, i.e. there must exists a stabilizing solution of (30) satisfying (29).

When dealing with coprime factorization, the divisibility condition has to be stated.

Lemma 1. (Divisibility in R_{MS}). Any $A(s) \in R_{MS}$ divides $B(s) \in R_{MS}$ if and only if all unstable zeros (including $s \to \infty$) of $A(s)$ are those of $B(s)$, and moreover, the numerator of $A(s)$ is formally stable. ■

Notice that zeros mean the roots of the whole term of the ring, not only those of the numerator.

Again, problems appear when dealing with neutral TDS or with those including distributed delays. An example of coprime, yet not Bézout, factorization of formally unstable neutral TDS was demonstrated in Example 3 and Example 5.

The following two examples demonstrate a typical coprime factorization over R_{MS} and a specific problem with distributed delays, respectively.

Example 7. The system is governed by the transfer function

$$G(s) = \frac{b(s)}{a(s)} = \frac{s + \exp(-s)}{s^2 + (2 + \exp(-s))s + 1} \exp(-2s) \tag{39}$$

which is a stable retarded TDS. Coprime factorization of (39) over R_{MS} can be performed e.g. as follows

$$G(s) = \frac{\frac{b(s)}{m(s)}}{\frac{a(s)}{m(s)}} = \frac{\frac{(s + \exp(-s))\exp(-2s)}{m(s)}}{\frac{s^2 + (2 + \exp(-s))s + 1}{m(s)}} = \frac{B(s)}{A(s)} \tag{40}$$

where $A(s)$, $B(s) \in R_{MS}$ and $m(s)$ stands for a stable (quasi)polynomial of degree 2. Its degree must equal 2; otherwise, elements would not be proper or coprime. ■

Example 8. Consider a simple system with distributed delays with transfer function (16) and suggest a factorization

$$G(s) = \frac{1 - \exp(1)\exp(-s)}{s - 1} = \frac{\frac{1 - \exp(1)\exp(-s)}{m(s)}}{\frac{s-1}{m(s)}} = \frac{B(s)}{A(s)} \tag{41}$$

In this case, the common denominator (quasi)polynomial $m(s)$ could not be stable since it would lead to prime elements in R_{MS}. Indeed, let, for instance, $m(s) = s + 1$, then there exists a term $T(s) \in R_{MS}$ that is a non-zero non-invertible common divisor of both $A(s), B(s)$ (which are then reducible), e.g.

$$\begin{aligned} A(s) &= T(s)A_0(s) = \frac{s-1}{s+2} \frac{s+2}{s+1} \\ B(s) &= T(s)B_0(s) = \frac{s-1}{s+2} \frac{1 - \exp(1)\exp(-s)}{s-1} \end{aligned} \tag{42}$$

The solution of this problem is read as follows: The common denominator $m(s)$ must include all common zeros s_0 of $a(s), b(s)$ with $\operatorname{Re} s_0 \geq 0$ (even asymptotic ones tending to the imaginary axis). Thus, the coprime factorization (41) should read

$$G(s)=\frac{1-\exp(1)\exp(-s)}{s-1}=\frac{\frac{1-\exp(1)\exp(-s)}{s-1}}{\frac{s-1}{s-1}}=\frac{B(s)}{A(s)} \tag{43}$$

■

The notion of coprime factorization is closely related to the existence of a solution of the Bézout identity. As stated e.g. in Example 3, for formally unstable TDS such solution in $H_\infty(\mathbb{C}^+)$ (an thus not in R_{MS}) does not exist – we can obtain coprime yet not Bézout coprime factors.

If a pair $A(s), B(s) \in R_{MS}$ is Bézout coprime, it is possible to solve the Bézout identity (or to find the GCD) using the extended Euclidean algorithm. Prior to the implementation of the extended Euclidean algorithm to R_{MS} ring, an ordering of ring elements has to be defined, so that a poset is obtained. Thus, define $P=(R_{MS},\preceq)$ as

a) $A(s)\preceq B(s)$ iff $A(s)\,|\,B(s)$.

b) $A(s)\equiv B(s)$ iff $A(s)\,|\,B(s)$ and $B(s)\,|\,A(s)$, or equivalently, $A(s)$ is associated with $B(s)$.

c) $A(s)$ is not related to $B(s)$ iff $A(s)\nmid B(s)$ and $B(s)\nmid A(s)$.

The procedure of finding the $\mathrm{GCD}(A(s),B(s))$ can be characterized as follows. Assume these three situations:

a) If $A(s)\equiv B(s)$, the GCD of both is simply either $A(s)$ or $B(s)$.

b) If $A(s)\succeq B(s)$, keep the following scheme

$$\left[\begin{array}{cc|c} 1 & 0 & A(s) \\ 0 & 1 & B(s) \end{array}\right] \sim \left[\begin{array}{cc|c} 1 & -\frac{A(s)}{B(s)} & 0 \\ 0 & 1 & B(s) \end{array}\right] \tag{44}$$

hence, $B(s)$ is the GCD of $A(s)$ and $B(s)$, according to (2.46). If $B(s)\succeq A(s)$, the procedure is analogous with $\mathrm{GCD}(A(s),B(s))= A(s)$.

c) Let $A(s)$ and $B(s)$ be not related to each other. In this case, follow the scheme (45).

Here, the GCD of $A(s)$ and $B(s)$ equals $A(s)X(s)+B(s)Y(s)$. In scheme (45), it is supposed that there can be found quotients $X(s), Y(s)$ such that the element $A(s)X(s)+B(s)Y(s)$ divides $A(s), B(s)$. Since $A(s), B(s)$ are Bézout coprime, $A(s)X(s)+B(s)Y(s)$ must be a unit of the ring.

$$\begin{aligned}
&\left[\begin{array}{cc|c} 1 & 0 & A(s) \\ 0 & 1 & B(s) \end{array}\right] \sim \left[\begin{array}{cc|c} X(s) & 0 & A(s)X(s) \\ 0 & 1 & B(s) \end{array}\right] \\
&\sim \left[\begin{array}{cc|c} X(s) & Y(s) & A(s)X(s)+B(s)Y(s) \\ 0 & 1 & B(s) \end{array}\right] \\
&\sim \left[\begin{array}{cc|c} 0 & 1 & B(s) \\ X(s) & Y(s) & A(s)X(s)+B(s)Y(s) \end{array}\right] \\
&\sim \left[\begin{array}{cc|c} \frac{-B(s)X(s)}{A(s)X(s)+B(s)Y(s)} & \frac{A(s)X(s)}{A(s)X(s)+B(s)Y(s)} & 0 \\ X(s) & Y(s) & A(s)X(s)+B(s)Y(s) \end{array}\right]
\end{aligned} \tag{45}$$

In other words, the objective is to find structures of $X(s), Y(s)$ and to set zeros and poles of $A(s)X(s)+B(s)Y(s)$ such that divisibility conditions as in Lemma 1 are satisfied or the element is invertible. This task can be troublesome; however, if formally unstable neutral TDS were avoided being included, every numerator/denominator quasipolynomial would have only a finite number of unstable zeros, which would make possible to find the $\mathrm{GCD}(A(s),B(s))$.

If the task is to solve the Bézout identity (30) itself instead of the $\mathrm{GCD}(A(s),B(s))$, one can use scheme (9) where $c=1$. This yields these results, respectively

$$\text{a)}\quad \begin{array}{l} P(s)=\frac{1}{A(s)},\ Q(s)=0 \ \text{and/or} \\ P(s)=0,\ Q(s)=\frac{1}{B(s)} \end{array} \tag{46}$$

$$\text{b)}\quad \begin{array}{l} P(s)=\frac{1}{A(s)},\ Q(s)=0 \ \text{or} \\ P(s)=0,\ Q(s)=\frac{1}{B(s)} \end{array} \tag{47}$$

$$\text{c)}\quad \begin{array}{l} P(s)=\frac{X(s)}{A(s)X(s)+B(s)Y(s)} \\ Q(s)=\frac{Y(s)}{A(s)X(s)+B(s)Y(s)} \end{array} \tag{48}$$

The following examples elucidate the whole procedure.

Example 9. Assume coprime factorization (43) and find $\mathrm{GCD}(A(s),B(s))$ first. Since $A(s)$ divides $B(s)$, it holds that $B(s)\succeq A(s)$, hence

$$\mathrm{GCD}\left(A(s),B(s)\right)=A(s)=\frac{s-1}{s-1}=1 \tag{49}$$

according to (44).

The Bézout identity (30) then has the solution given by (47) as

$$P(s)=\frac{1}{A(s)}=1,\ Q(s)=0 \tag{50}$$

Example 10. Now let the factorization be given by (40) with $m(s)=(s+1)^2$. In this case, the both elements $A(s)$ and $B(s)$ are associated, thus $A(s)\equiv B(s)$ and scheme (45) can be used when solving $\mathrm{GCD}\left(A(s),B(s)\right)$. This scheme yields e.g.

$$\begin{aligned}&X(s)=Y(s)=1\\&\Rightarrow A(s)X(s)+B(s)Y(s)\\&=\frac{s^2+(2+\exp(-s))s+1+(s+\exp(-s))\exp(-2s)}{(s+1)^2}\\&=\frac{s^2+(2+\exp(-s)+\exp(-2s)+\exp(-3s))s+1}{(s+1)^2}\\&=\mathrm{GCD}(A(s),B(s))\end{aligned} \tag{51}$$

where $X(s),Y(s)$ are chosen for the simplicity.

Then the solution of the Bézout identity according to (48) reads

$$\begin{aligned}&P(s)=Q(s)\\&=\frac{(s+1)^2}{s^2+(2+\exp(-s)+\exp(-2s)+\exp(-3s))s+1}\end{aligned} \tag{52}$$

In case of asymptotically stable systems, i.e. $A(s)$ is invertible (a unit), it is possible to use also a simple procedure when solving the Bézout identity

$$Q(s)=1\Rightarrow P(s)=\frac{1-B(s)}{A(s)} \tag{53}$$

By applying this rule to the example, the following solution is obtained

$$P(s)=\frac{(s+1)^2-(s+\exp(-s))\exp(-2s)}{s^2+(2+\exp(-s))s+1} \tag{54}$$

This scheme has some advantages in controller design (this topic is out of the aim of this paper). ■

5.2 Ring properties

Follow now terms introduced in Section 2 and try to match some of them with R_{MS} ring.

Lemma 2. A set R_{MS} introduced in Definition 2 constitutes a commutative ring. ■

Proof. A sketch of proof that R_{MS} meets ring conditions follows.

Clearly, R_{MS} is closed under addition with associativity and the neutral element $E=0$. The inverse element $B(s)\in R_{MS}$ under addition of $A(s)\in R_{MS}$ is simply $B(s)=-A(s)$. Since $A(s)+B(s)=B(s)+A(s)\in R_{MS}$, it is a commutative group.

The closure under multiplication with associativity is also evident since the numerator and denominator of any $A(s)\in R_{MS}$ are composed of quasipolynomial factors – retarded ones and formally stable neutral ones, respectively. Since the operation of multiplication is commutative, left and right distributivity hold as well. In case of distributed delays, it is not possible to obtain more unstable denominator zeros then numerator ones of any $A(s)\in R_{MS}$ under multiplication. The multiplicative identity element equals 1. □

Lemma 11. An element $A(s)\in R_{MS}$ is a unit (invertible element) iff $A(s)$ has zero relative order and has the (asymptotically and formally) stable numerator. ■

The proof of Lemma 11 is evident (e.g. the necessity can be proved by the negation of the right hand side of the lemma) with the aid of Lemma 1. Note that stable numerator means that is has only stable zeros in the appropriate meaning.

Lemma 12. An element $A(s)\in R_{MS}$ is irreducible iff its numerator is formally stable and

$$O_R+N_U\le 1 \tag{55}$$

where O_R is the relative order and N_U stands for the number of real zeros $s_{U,i}, i=1,2,...N_U$ or conjugate pairs $s_{U,i},\bar{s}_{U,i}, i=1,2,...N_U$ with $\mathrm{Re}\, s_{U,i}\ge 0$ and $\mathrm{Re}\,\bar{s}_{U,i}\ge 0$ of $A(s)$, respectively. ■

Proof. Necessity. Consider the following three cases

a) $O_R=0$, $N_U=1$

b) $O_R=1$, $N_U=0$

c) $O_R \geq 2$

Use an indirect proof. First, let a) is not valid; hence, $O_R = 0$, $N_U > 1$. Consider a (quasi)polynomial $c(s)$ with only one unstable zero (or a pair of unstable zeros), say $c(s_{U,1}) = 0$ (or $c(s_{U,1}) = c(\bar{s}_{U,1}) = 0$) and an arbitrary stable (quasi)polynomial $b(s)$ of the same order (i.e. first or second one). Then

$$A(s) = \frac{a_{num}(s)}{a_{den}(s)} = \frac{a_{num}(s)b(s)}{a_{den}(s)c(s)} \frac{c(s)}{b(s)} = A_1(s)A_2(s) \quad (56)$$

where $A_1(s)$ and $A_2(s)$ are neither associated with $A(s)$ nor units.

Now, let b) is not valid, i.e. $O_R = 1$, $N_U > 0$, and assume a stable (quasi)polynomial $d(s)$ of the first order. Then follow the scheme

$$A(s) = \frac{a_{num}(s)}{a_{den}(s)} = \frac{a_{num}(s)d(s)}{a_{den}(s)} \frac{1}{d(s)} = A_1(s)A_2(s) \quad (57)$$

Again, $A_1(s)$ and $A_2(s)$ are neither associated with $A(s)$ nor units.

Finally, let c) holds. Then it is possible to write e.g. scheme (57).

Sufficiency. Consider the three cases introduced above again.

If a) holds and the numerator is formally stable (even asymptotically), scheme (56) fails, since $A_1(s)$ is a unit and $A_2(s)$ is associated with $A(s)$. Moreover, there is not possible to find another "reducible" scheme.

Similarly, if b) holds and is formally stable, $A_1(s)$ is a unit and $A_2(s)$ is associated with $A(s)$ in scheme (57); hence, $A(s)$ is irreducible. □

Lemma 13. R_{MS} ring does not constitute UFR. ■

Proof. Consider the following element of the ring

$$\frac{1 - \exp(-\tau s)}{s} \quad (58)$$

Nonzero zeros of the numerator of (58) are

$$s_k = \frac{2k\pi}{\tau} \mathrm{j}, \bar{s}_k = -\frac{2k\pi}{\tau} \mathrm{j},\ k \in \mathrm{N} \quad (59)$$

Define polynomials

$$P_k(s) = (s - s_k)(s - \bar{s}_k) \quad (60)$$

Then the factorization

$$\frac{1 - \exp(-\tau s)}{s} = \frac{[1 - \exp(-\tau s)](s + m_0)^2}{sP_1(s)} \frac{P_1(s)}{(s + m_0)^2} =$$
$$= \frac{[1 - \exp(-\tau s)](s + m_0)^4}{sP_1(s)P_2(s)} \frac{P_1(s)P_2(s)}{(s + m_0)^4} =$$
$$\ldots \quad (61)$$

where $m_0 > 0$ is infinite and thus the R_{MS} ring is not a UFR, and none of left-hand factors in (61) is irreducible and none of all factors is a unit. □

Lemma 14. R_{MS} is an integral domain. ■

Proof. Consider $A(s), B(s) \in R_{MS}$ where $A(s)$ is a unit. Let $A(s)B(s) = 0$ and multiply the whole equation by $1/A(s)$. It yields $B(s) = 0$ and we have a contradiction. □

Hence, Lemma 13 and Lemma 14 imply that R_{MS} is UFD.

Lemma 15. R_{MS} does not constitute PID. ■

Proof. Simply, it holds that every PID is UFD. Since R_{MS} is not UFD according to Lemma 13, it is not PID. □

Lemma 16. R_{MS} does not constitute a Bézout domain. ■

Proof. It is sufficient to show that there exists a pair $A(s), B(s) \in R_{MS}$ which does not give a solution pair $Q(s), P(s) \in R_{MS}$ of (30). Indeed, as mentioned above, coprime factorization of formally unstable TDS does not have a stabilizing solution of the Bézout identity in $H_\infty(\mathrm{C}^+)$, i.e. condition (29) does not hold. Since $H_\infty(\mathrm{C}^+) \supset R_{MS}$, which is evident from Definition 2, such solution does not exist in R_{MS} as well. □

The decision whether R_{MS} is a Noetherian ring is not successfully solved. Typically, a ring is a Bézout domain yet not PID, i.e. there exists an infinitely generated ideal which is not principal. In such cases, the ring is not Noetherian, see e.g. ring $\mathcal{E}$ of pseudopolynomials or ring $\mathcal{H}$, see Section 3.

6 Conclusions

The presented paper has introduced the original and a revised (alternative) definition of a special algebraic structure (ring) of quasipolynomial meromorphic functions. After offering an acquaintance with basic algebraic notions, an overview of some algebraic analytic and control structures and methods has been given. The original definition of R_{MS} has followed and some its disadvantages have been mentioned. Thus, a proposition of a revised definition has been then introduced, which is the crucial part of this contribution. The most involved part of the paper, i.e. (Bézout) coprime factorization, issues about the solution of the Bézout identity in the ring and selected algebraic properties, has followed.

As mentioned above several times, the ring can be used not only for TDS description but primarily for algebraic controller design satisfying asymptotic and formal stability of a control feedback system, reference tracking, asymptotic load disturbance rejection, etc., see e.g. [36], [37]. To comprehend this broad topic, some preliminary and supporting problems had to be analyzed and solved, for instance [38]-[42].

The natural limitation of the methodology is that formally unstable neutral TDS can not be stabilized in the sense of the ring. A detailed description of this control approach in the revised ring will be the matter of a future paper.

Acknowledgements

The author kindly appreciates the financial support which was provided by the European Regional Development Fund under the project CEBIA-Tech No. CZ.1.05/2.1.00/03.0089.

References:

[1] P. Zitek and V. Kucera, Algebraic Design of Anisochronic Controllers for Time Delay Systems, *International Journal of Control*, Vol.76, No.16, 2003, pp. 1654-1665.

[2] M. Vidyasagar, *Control System Synthesis: A Factorization Approach*. MIT Press, Cambridge, M. A., 1985.

[3] V. Kucera, Diophantine Equations in Control - A Survey, *Automatica*, Vol.29, No.6, 1993, pp. 1361-1375.

[4] D. Brethe and J. J. Loiseau, An Effective Algorithm for Finite Spectrum Assignment of Single-Input Systems with Delays, *Mathematics and Computers in Simulation*, Vol.45, No.3-4, 1998, pp. 339-348.

[5] L. Pekar and R. Prokop, Some Observations About the RMS Ring for Delayed Systems, In *Proceedings of the 17th International Conference on Process Control '09*, Strbske Pleso, Slovakia, 2009, pp. 28-36.

[6] J. Rosicky, *Algebra I.*, Masaryk University in Brno, Brno, Czech Republic, 1994 (in Czech).

[7] E. W. Weisstein, *MathWorld – A Wolfram Web Resource* [internet], Wolfram Research, 1995 [updated Jun 16, 2012; cited Jun 22, 2012], Available from: http://mathworld.wolfram.com/.

[8] G. Conte and A. M. Perdon, Systems over Rings: Geometric Theory and Applications, *Annual Reviews in Control*, Vol.24, 2000, pp. 113-124.

[9] R. E. Kalman, P. L. Falb and M. A. Arbib, *Topics in Mathematical System Theory*, McGraw-Hill, 1969.

[10] Y. Rouchaleau, *Linear, Discrete Time, Finite Dimensional Dynamical Systems over Some Classes of Commutative Rings*, Ph.D. Thesis, Stanford University, 1972.

[11] Y. Rouchaleau, B. F. Wyman and R. E. Kalman, Algebraic Structure of Linear Dynamical Systems. III. Realization Theory over a Commutative Ring, *Proceedings of the National Academy of Sciences of the United States of America*, Vol.69. No.11, 1972, pp. 3404-3406.

[12] E. W. Kamen, On the Algebraic Theory of Systems Defined by Convolution Operations, *Mathematical Systems Theory*, Vol.9, 1975, pp. 57-74.

[13] A. S. Morse, Ring Models for Delay-Differential Systems, *Automatica*, Vol.12, No.5, 1976, pp. 529-531.

[14] E. D. Sontag, Linear Systems over Commutative Rings: A Survey, *Richerche di Automatica*, Vol.7, 1976, pp. 1-34.

[15] D. C. Youla, The of Networks Containing Lumped and Distributed Elements, Part I., *Network and Switching Theory*, Vol.11, 1968, pp. 73-133.

[16] G. Conte and A. M. Perdon, Systems over a Principal Ideal Domain: A Polynomial Model Approach, *SIAM Journal of Control and Optimization*, Vol.20, 1982, pp. 112-124.

[17] E. Fornasini and G. Marchesini, State-Space Realization Theory of Two-Dimensional Filters, *IEEE Transactions on Automatic Control*, Vol.21, No.4, 1976, pp. 484-492.

[18] M. Morf, B. C. Levy and S.-Y. Kung, New Results in 2-D Systems Theory, Part I: 2-D Polynomial Matrices, Factorization and Coprimeness, *Proceedings of the IEEE*, Vol.65, No.6, 1977, pp. 861-872.

[19] J. J. Loiseau, Algebraic Tools for the Control and Stabilization of Time-Delay Systems,

Annual Reviews in Control, Vol.24, 2000, pp. 135-149.

[20] J. C. Willems, Models for Dynamics, *Dynamics Reported*, Vol.2, 1989, pp. 171-269.

[21] H. Gluesing-Lueerssen, A Behavioral Approach to Delay-Differential Systems, *SIAM Journal of Control and Optimization*, Vol.35, 1997, pp. 480-499.

[22] W. Michiels and T. Vyhlidal, An Eigenvalue Based Approach for the Stabilization of Linear Time-Delay Systems of Neutral Type, *Automatica*, Vol.41, No.6, 2005, pp. 991-998.

[23] C. A. Desoer, R. W. Liu, J. Murray and R. Seaks, Feedback System Design: The Fractional Representation Approach to Analysis and Synthesis, *IEEE Transactions on Automatic Control*, Vol.25, No.3, 1980, pp. 399-412.

[24] R. Prokop and J. P. Corriou, Design and Analysis of Simple Robust Controllers, *International Journal of Control*, Vol.66, No.6, 1997, pp. 905-921.

[25] J. R. Partington, Some Frequency-Domain Approaches to the Model Reduction of Delay Systems, *Annual Reviews in Control*, Vol.28, No.1, 2004, pp. 65-73.

[26] L. Pekar, E. Kureckova, Does the Higher Order Mean the Better Internal Delay Rational Approximation?, *International Journal of Mathematics and Computers in Simulation*, Vol.6, No.1, 2012, pp. 153-160.

[27] P. Dostal, V. Bobal and M. Sysel, Design of Controllers for Integrating and Unstable Time Delay Systems using Polynomial Method, in *Proceedings of the 2002 American Control Conference*, Anchorage, Alaska, USA, 2002, pp. 2773-2778.

[28] J. R. Partington and C. Bonnet, H_∞ and BIBO Stabilization of Delay Systems of Neutral Type, *Systems & Control Letter*s, Vol.52, No.3-4, 2004, pp. 283-288.

[29] C. A. Desoer, M. Vidyasagar, *Feedback Systems: Input-Output Properties*, Academic Press, New York, 1975.

[30] J. J. Loiseau, M. Cardelli and X. Dusser, Neutral-Type Time-Delay Systems That Are Not Formally Stable Are Not BIBO Stabilizable, *IMA Journal of Mathematical Control and Information*, Vol.19, No.1-2, 2002, pp. 217-227.

[31] L. S. Pontryagin, On the Zeros of Some Elementary Transcendental Functions, *Izvestiya Akademii Nauk SSSR*, Vol.6, 1942, pp. 115-131.

[32] J. K. Hale and S. M. Verduyn Lunel, Strong Stabilization of Neutral Functional Differential Equations, *IMA Journal of Mathematical Control and Informatio*n, Vol.19, No.1-2, 2002, pp. 5-23.

[33] J. K. Hale and S. M. Verduyn Lunel, Introduction to Functional Differential Equations, in *Applied Math. Sciences*, Vol.99, Springer-Verlag, New York, 1993.

[34] P. Zitek and T. Vyhlidal, Argument-Increment Based Stability Criterion for Neutral Time Delay Systems, in *Proceedings of the 16th Mediterranean Conference on Control and Automation*, Ajaccio, France, 2008, pp. 824-829.

[35] J. J. Loiseau, Algebraic Tools for the Control and Stabilization of Time-Delay Systems, *Annual Reviews in Control*, Vol.24, 2000, pp. 135-149.

[36] L. Pekar and R. Prokop, Control of Delayed Integrating Processes Using Two Feedback Controllers – RMS Approach, in *Proceedings of the 7th WSEAS International Conference on System Science and Simulation in Engineering*, Venice, Italy, 2008, pp. 35–40.

[37] R. Prokop, L. Pekar and J. Korbel, Autotuning for Delay Systems using Meromorphic Functions, in *Proceedings of the 9th IFAC Workshop on Time Delay Systems*, 2010, Prague. FP-PR-333 [DVD-ROM].

[38] R. Matusu and R. Prokop, Control of Periodically Time-Varying Systems with Delay: An Algebraic Approach vs. Modified Smith Predictors, *WSEAS Transactions on Systems*, Vol.9, No.6, 2010, pp. 689-702.

[39] L. Pekar, R. Prokop and R. Matusu, A Stability Test for Control Systems with Delays Based on the Nyquist Criterion, *International Journal of Mathematical Models in Applied Sciences*, 2011, Vol.5, No.6, pp. 1213-1224.

[40] L. Pekar, Root Locus Analysis of a Retarded Quasipolynomial, *WSEAS Transaction on Systems and Control*, 2011, Vol.6, No.3, pp. 79-91.

[41] L. Pekar, On Finite-Dimensional Transformations of Anisochronic Controllers Designed by Algebraic Means: A User Interface, *Matlab / Book 2*, V. N. Katsikis (ed.), InTech, Rijeka, Croatia, 2012, accepted.

[42] F. Neri, Software Agents as A Versatile Simulation Tool to Model Complex Systems. *WSEAS Transactions on Information Science and Applications*, Vol.7, No.5, 2010, pp. 609-618.

Control of Unstable and Integrating Time Delay Systems Using Time Delay Approximations

PETR DOSTÁL, VLADIMÍR BOBÁL, and ZDENĚK BABÍK
Tomas Bata University in Zlin
Department of Process Control, Centre of Polymer Systems
nam. T.G. Masaryka 5555, 760 01 Zlin
CZECH REPUBLIC
{dostalp, bobal, babik}@fai.utb.cz http://www.fai.utb.cz/

Abstract: - The paper deals with design of controllers for time delay systems having integrative or unstable properties. The proposed method is based on two methods of time delay approximations. The control system with two feedback controllers is considered. For design of controllers, the polynomial approach is used. Resulting continuous-time controllers obtained via polynomial equations and the LQ control technique ensure asymptotic tracking of step references as well as step disturbances attenuation. Simulation results are presented to illustrate the proposed method.

Key-Words: - Time delay system, Time delay approximation, Polynomial method, LQ control.

1 Introduction

Different classes of technological processes include a time delay in their input-output relations. Plants with a time delay cannot often be controlled by conventional controllers designed without consideration of the dead-time. The control responses using such controllers are often of a poor quality or even can tend to destabilize the closed-loop system.

A part of time delay processes can be unstable or having integrating properties. Typical examples of such processes are e.g. pumps, liquid storing tanks, distillation columns and some types of chemical or biochemical reactors. A control of such processes represents a difficult problem especially for processes containing also other stable or unstable parts with the integrative term.

For control design of unstable and also integrating processes several ways exist. Some methods are based on several modifications of the Smith predictor which was originally developed for stable time delay systems. Such modified Smith predictors were published e.g. in [1] – [4]. Other group of methods employ PID control strategies [5] – [8], the robust control methods [9] and [10] or methods based on the ring of quasipolynomials, e.g. [11]. A solution of differential equations describing the time delay systems can be found e.g. in [12]. Other simulation possibilities are described e.g. in [13].

This paper presents one method of the controller design for unstable and integrating time delay systems and also for its combination with a stable or an unstable first order system. The presented procedure is based on approximations of the time delay term by the first order Taylor numerator expansion (TNE) and by the first order Padé approximation (PA). The control system with two feedback controllers is considered, see, e.g. [14], [15]. The controllers are derived using the polynomial approach published e.g. in [16]. For tuning of controller parameters, the pole assignment method exploiting the LQ control technique is used, see, e.g. [17]. The resulting proper and stable controllers obtained via polynomial Diophantine equations and spectral factorization techniques ensure the asymptotic tracking of step references as well as step disturbances attenuation.

The structures of developed controllers together with analytically derived formulas for computation of their parameters are presented for five typical plants of time delay systems: the unstable first order time delay system (UFOTDS), the unstable second order time delay system (USOTDS), integrating time delay system (ITDS), and, the stable and unstable first order plus integrating time delay system (SFOPITDS, UFOPITDS).

Presented simulation results obtained by both approximations document usefulness of the proposed method providing stable control responses of a good quality.

2 Approximate Transfer Functions

The transfer functions in the sequence UFOTDS, USOTDS, ITDS, SFOPITDS and UFOPITDS have forms

$$G_1(s) = \frac{K}{\tau s - 1} e^{-\tau_d s} \tag{1}$$

$$G_2(s) = \frac{K}{(\tau_1 s - 1)(\tau_2 s + 1)} e^{-\tau_d s} \tag{2}$$

$$G_3(s) = \frac{K}{s} e^{-\tau_d s} \tag{3}$$

$$G_4(s) = \frac{K}{s(\tau s \pm 1)} e^{-\tau_d s} \tag{4}$$

2.1 TN expansion

In the first case, the time delay terms in (1) – (4) are approximated by the TN expansion

$$e^{-\tau_d s} \approx 1 - \tau_d s \,. \tag{5}$$

Then, approximate transfer functions relating to (1) – (4) have forms

$$G_{1N}(s) = \frac{K(1 - \tau_d s)}{\tau s - 1} = \frac{b_0 - b_1 s}{s - a_0} \tag{6}$$

where

$$b_0 = \frac{K}{\tau},\ b_1 = \frac{K\tau_d}{\tau},\ a_0 = \frac{1}{\tau} \tag{7}$$

for the UFOTDS,

$$G_{2N}(s) = \frac{K(1 - \tau_d s)}{(\tau_1 s - 1)(\tau_2 s + 1)} = \frac{b_0 - b_1 s}{s^2 + a_1 s + a_0} \tag{8}$$

where

$$b_0 = \frac{K}{\tau_1 \tau_2},\ b_1 = \frac{K\tau_d}{\tau_1 \tau_2},\ a_0 = \frac{1}{\tau_1 \tau_2},\ a_1 = \frac{\tau_1 - \tau_2}{\tau_1 \tau_2} \tag{9}$$

for the USOTDS,

$$G_{3N}(s) = \frac{K(1 - \tau_d s)}{s} = \frac{b_0 - b_1 s}{s} \tag{10}$$

where

$$b_0 = K,\ \ b_1 = K\tau_d \tag{11}$$

for the ITDS, and,

$$G_{4,5N}(s) = \frac{K(1 - \tau_d s)}{s(\tau s \pm 1)} = \frac{b_0 - b_1 s}{s^2 + a_1 s} \tag{12}$$

where

$$b_0 = \frac{K}{\tau},\ \ b_1 = \frac{K\tau_d}{\tau},\ \ a_1 = \pm\frac{1}{\tau} \tag{13}$$

for the SFOPITDS and UFOPITDS.

2.2 Padé approximation

In the second case, the time delay terms in (1) – (4) are approximated by the by the first order Padé approximation

$$e^{-\tau_d s} \approx \frac{2 - \tau_d s}{2 + \tau_d s} \tag{14}$$

Now, approximate transfer functions in the same sequence take forms

$$G_{1P}(s) = \frac{K(2 - \tau_d s)}{(\tau s - 1)(2 + \tau_d s)} = \frac{b_0 - b_1 s}{s^2 + a_1 s - a_0} \tag{15}$$

where

$$b_0 = \frac{2K}{\tau\tau_d},\ b_1 = \frac{K}{\tau},\ a_0 = \frac{2}{\tau\tau_d},\ a_1 = \frac{2\tau - \tau_d}{\tau\tau_d} \tag{16}$$

for the UFOTDS,

$$G_{2P}(s) = \frac{K(2 - \tau_d s)}{(\tau_1 s - 1)(\tau_2 s + 1)(2 + \tau_d s)} = \frac{b_0 - b_1 s}{s^3 + a_2 s^2 + a_1 s - a_0} \tag{17}$$

where

$$b_0 = \frac{2K}{\tau_1 \tau_2 \tau_d},\ b_1 = \frac{K}{\tau_1 \tau_2},\ a_0 = \frac{2}{\tau_1 \tau_2 \tau_d}$$
$$a_1 = \frac{2(\tau_1 - \tau_2) - \tau_d}{\tau_1 \tau_2 \tau_d},\ a_2 = \frac{2\tau_1\tau_2 + \tau_1\tau_d - \tau_2\tau_d}{\tau_1 \tau_2 \tau_d} \tag{18}$$

for the USOTDS,

$$G_{3P}(s) = \frac{K(2 - \tau_d s)}{s(2 + \tau_d s)} = \frac{b_0 - b_1 s}{s^2 + a_1 s} \tag{19}$$

where

$$b_0 = \frac{2K}{\tau_d},\ b_1 = K,\ a_1 = \frac{2}{\tau_d} \tag{20}$$

for the ITDS, and,

$$G_{4,5P}(s) = \frac{K(2 - \tau_d s)}{s(\tau s \pm 1)(2 + \tau_d s)} = \frac{b_0 - b_1 s}{s^3 + a_2 s^2 + a_1 s} \tag{21}$$

where

$$b_0 = \frac{2K}{\tau\tau_d},\ b_1 = \frac{K}{\tau},\ a_1 = \pm\frac{2}{\tau\tau_d},\ a_2 = \frac{2\tau \pm \tau_d}{\tau\tau_d} \qquad (22)$$

for the SFOPITDS and UFOPTDS.

Remark: For the UFOTDS and the UFOPITDS the conditions $\tau_d \neq \tau$ in (6) and (12), $\tau_d \neq \tau_1$ in (8), and, $\tau_d \neq 2\tau$ in (15) and (21) must be fulfilled.

All approximate transfer functions have the form

$$G_A(s) = \frac{b(s)}{a(s)} \qquad (23)$$

where b and a are coprime polynomials in s that fulfill the inequality $\deg b \leq \deg a$.

3 Control Design

The control system with two feedback controllers is depicted in Fig.1.

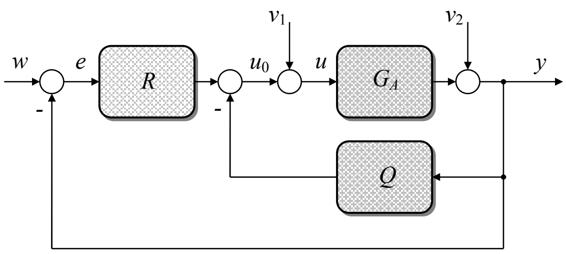

Fig.1. Control system.

In the scheme, w is the reference signal, v_1, v_2 are input and output disturbances, e is the tracking error, u_0 is the controller output, y is the controlled output and u is the control input. The reference w and both disturbances v_1 and v_2 are considered to be step functions with transforms

$$W(s) = \frac{w_0}{s},\ V_1(s) = \frac{v_{10}}{s},\ V_2(s) = \frac{v_{20}}{s} \qquad (24)$$

The transfer function G_A represents a proper approximate transfer function in the general form (23).

The transfer functions of controllers are

$$Q(s) = \frac{\tilde{q}(s)}{\tilde{p}(s)},\ R(s) = \frac{r(s)}{\tilde{p}(s)} \qquad (25)$$

where $\tilde{q}, r$ and $\tilde{p}$ are coprime polynomials in s.

3.1 Application of Polynomial Method

The controller design described in this section follows from the polynomial approach. The general conditions required to govern the control system properties are formulated as follows:

- Strong stability of the control system (in addition to the control system stability, also the stability of a controller is required).
- Internal properness of the control system.
- Asymptotic tracking of the reference.
- Attenuation of disturbances.

The procedure to derive admissible controllers can be carried out as follows:

Transforms of the controlled output and the tracking error take the form (for simplification, the argument s is in some equations omitted)

$$Y(s) = \frac{1}{d}\left[br\,W(s) + b\tilde{p}\,V_1(s) + a\tilde{p}V_2(s)\right] \qquad (26)$$

$$E(s) = \frac{1}{d}\left[(a\tilde{p} + b\tilde{q})W(s) - b\tilde{p}V_1(s) - a\tilde{p}V_2(s)\right] \qquad (27)$$

where

$$d(s) = a(s)\,\tilde{p}(s) + b(s)\big(r(s) + \tilde{q}(s)\big) \qquad (28)$$

is the characteristic polynomial with roots as poles of the closed-loop.

Establishing the polynomial t as

$$t(s) = r(s) + \tilde{q}(s) \qquad (29)$$

and substituting (29) into (28), the condition of the control system stability is ensured when polynomials $\tilde{p}$ and t are given by a solution of the polynomial Diophantine equation

$$a(s)\tilde{p}(s) + b(s)t(s) = d(s) \qquad (30)$$

with a stable polynomial d on the right side.

With regard to (24), asymptotic tracking and both disturbances attenuation are provided by divisibility of both terms $a\tilde{p} + b\tilde{q}$ and $\tilde{p}$ in (27) by s. This condition is fulfilled for polynomials $\tilde{p}$ and $\tilde{q}$ in the form

$$\tilde{p}(s) = s\,p(s),\ \ \tilde{q}(s) = s\,q(s). \qquad (31)$$

Subsequently, the transfer functions of controllers take forms

$$Q(s) = \frac{q(s)}{p(s)},\ R(s) = \frac{r(s)}{s\,p(s)}. \qquad (32)$$

A stable polynomial $p(s)$ in denominators of (32) ensures the stability of ontrollers.

The control system satisfies the condition of internal properness when the transfer functions of all its components are proper. Consequently, the degrees of polynomials q and r must fulfill inequalities

$$\deg q \le \deg p\,, \quad \deg r \le \deg p + 1\,. \qquad (33)$$

Now, the polynomial t can be rewritten to the form

$$t(s) = r(s) + s\,q(s)\,. \qquad (34)$$

Taking into account solvability of (30) and conditions (33), the degrees of polynomials in (30) and (32) can be easily derived as

$$\deg t = \deg r = \deg a\,,\ \deg q = \deg a - 1$$
$$\deg p = \deg a - 1\,,\ \ \deg d = 2\deg a\,. \qquad (35)$$

Denoting deg $a = n$, polynomials t, r and q have the form

$$t(s) = \sum_{i=0}^{n} t_i s^i\ ,\ r(s) = \sum_{i=0}^{n} r_i s^i\ ,\ q(s) = \sum_{i=1}^{n} q_i s^{i-1} \qquad (36)$$

and among of their coefficients equalities

$$r_0 = t_0\,,\ \ r_i + q_i = t_i\ \text{ for } i = 1, \dots, n \qquad (37)$$

hold. Since by a solution of the polynomial equation (30) only coefficients t_i can be calculated, unknown coefficients r_i and q_i can be obtained by a choice of selectable coefficients $\beta_i \in \langle 0,1 \rangle$ such that

$$r_i = \beta_i t_i\,,\ \ q_i = (1-\beta_i)\,t_i\ \text{ for } i = 1, \dots, n\,. \qquad (38)$$

The coefficients β_i distribute a weight between numerators of transfer functions Q and R. With respect to the transform (26), it may be expected that higher values of β_i speed up control responses to step references.

Remark: If $\beta_i = 1$ for all i, the control system in Fig. 1 demotes to the 1DOF control configuration. If $\beta_i = 0$ for all i and the reference and both disturbances are step functions, the control system corresponds to the 2DOF control configuration.

The controller parameters then follow from solutions of the polynomial equation (30) and depend upon coefficients of polynomial d. The next problem here means to find a stable polynomial d that enables to obtain the acceptable stabilizing and stable controllers.

3.2 Pole Assignment

In this paper, the polynomial d is considered as a product of two stable polynomials g and m in the form

$$d(s) = g(s)\,m(s) \qquad (39)$$

where the polynomial g is a monic form of the polynomial h obtained by spectral factorization

$$\left[s\,a(s)\right]^* \varphi \left[s\,a(s)\right] + b^*(s) b(s) = h^*(s)\,h(s) \qquad (40)$$

where $\varphi > 0$ is the weighting coefficient.

Remark: In the LQ control theory, the spectral factorization (40) is used in a procedure of minimization of the quadratic cost function

$$J = \int_0^\infty \left\{ e^2(t) + \varphi \dot{u}^2(t) \right\} dt \qquad (41)$$

where $e(t)$ is the tracking error and $\dot{u}(t)$ is the control input derivative.

The polynomials h and derived formulas for their parameters calculation have forms

$$h(s) = h_2 s^2 + h_1 s + h_0 \qquad (42)$$

for the UFOTDS and ITDS with the TN expansion where

$$h_0 = |b_0|\,,\ \ h_2 = \sqrt{\varphi}\ ,\ \ h_1 = \sqrt{\varphi a_0^2 + b_1^2 + 2h_0 h_2} \qquad (43)$$

and $a_0 = 0$ for the ITDS,

$$h(s) = h_3 s^3 + h_2 s^2 + h_1 s + h_0 \qquad (44)$$

for the UFOTDS and ITDS with the Padé approximation, and, for the USOTDS, SFOPITDS and UFOPITDS with the TN expansion where

$$h_0 = |b_0|, h_3 = \sqrt{\varphi}, h_1 = \sqrt{\varphi a_0^2 + b_1^2 + 2h_0 h_2}$$
$$h_2 = \sqrt{\varphi\,(a_1^2 - 2\,a_0) + 2h_1 h_3} \qquad (45)$$

and $a_0 = 0$ for the ITDS, SFOPITDS and UFOPITDS, and,

$$h(s) = h_4 s^4 + h_3 s^3 + h_2 s^2 + h_1 s + h_0 \qquad (46)$$

for the USOTDS, SFOPITDS and UFOPITDS with the Padé approximation where

$$h_0 = |b_0|, h_4 = \sqrt{\varphi}, h_1 = \sqrt{\varphi a_0^2 + b_1^2 + 2h_0 h_2}$$

$$h_2 = \sqrt{\varphi(a_1^2 - 2\,a_0\,a_2) + 2h_1 h_3 - 2h_0 h_4} \qquad (47)$$

$$h_3 = \sqrt{\varphi(a_2^2 - 2a_1) + 2h_2 h_4}$$

and $a_0 = 0$ for both SFOPITDS and UFOPITDS.
For calculation of d, polynomials (42), (44) and (46)

are arranged to monic forms $g(s)$ (with unit coefficients by the highest power of s) such that

$$g_j = h_j / h_n \quad j = 0,1,\dots,n \tag{48}$$

where $n = \deg h$.

The second polynomial m ensuring properness of the controller is chosen as

$$m(s) = 1 \tag{49}$$

for both UFOTDS and ITDS with the TN expansion,

$$m(s) = s + \frac{2}{\tau_d} \tag{50}$$

for both UFOTDS and ITDS with the Padé approximation,

$$m(s) = s + \frac{1}{\tau_2} \tag{51}$$

for the USOTDS with the TN expansion,

$$m(s) = \left(s + \frac{1}{\tau_2}\right)\left(s + \frac{2}{\tau_d}\right) \tag{52}$$

for the USOTDS with the Padé approximation,

$$m(s) = s + \frac{1}{\tau} \tag{53}$$

for both SFOPITDS and UFOPITDS with the TN expansion, and,

$$m(s) = \left(s + \frac{2}{\tau_d}\right)\left(s + \frac{1}{\tau}\right) \tag{54}$$

for both UFOPITDS and SFOPITDS with the Padé approximation.

The above forms of m lead to the polynomial d with coefficients containing only the selectable parameter φ with all other coefficients depending on parameters of polynomials b and a. Consequently, a location of the closed loop poles can be affected by the selectable parameter φ.

The transfer functions of controllers with degrees of polynomials in their numerators and denominators given by (35) are

$$Q(s) = \frac{q_1}{p_0}, \quad R(s) = \frac{r_1 s + r_0}{p_0 s} \tag{55}$$

for both UFOTDS and ITDS with the TN expansion,

$$Q(s) = \frac{q_2 s + q_1}{s + p_0}, \quad R(s) = \frac{r_2 s^2 + r_1 s + r_0}{s(s + p_0)} \tag{56}$$

for both UFOTDS and ITDS with the Padé approximation, and, for the USOTDS, SFOPITDS and UFOPITDS with the TN expansion. Further,

$$\begin{aligned} Q(s) &= \frac{q_3 s^2 + q_2 s + q_1}{s^2 + p_1 s + p_0} \\ R(s) &= \frac{r_3 s^3 + r_2 s^2 + r_1 s + r_0}{s(s^2 + p_1 s + p_0)} \end{aligned} \tag{57}$$

for the USOTDS, SFOPITDS and UFOPITDS with the Padé approximation.

In all cases, the parameters q in numerators of controllers are computed from parameters t according to (37).

For clarity, derived formulas for computation of parameters p_0 and t the controller derived for all considered cases together with conditions of the controllers' stability are introduced in the form of tables.

Table 1. Controller parameters for UFOTDS

TN expansion
$p_0 = \dfrac{\tau\tau_d(g_1 + \tau_d g_0) + \tau}{\tau - \tau_d}$ $t_0 = \dfrac{\tau}{K} g_0, \quad t_1 = \dfrac{1}{K}\dfrac{\tau}{\tau_d}(p_0 - 1)$
$p_0 > 0$ for $\tau_d < \tau$
Padé approximation
$p_0 = \dfrac{\tau\left[2g_2 + \tau_d(g_1 + \frac{\tau_d}{2} g_0)\right] + 2}{2\tau - \tau_d}, \quad t_0 = \dfrac{\tau}{K} g_0$ $t_1 = \dfrac{1}{K}\left[p_0 + \tau(g_1 + \tau_d g_0)\right]$, $t_2 = \dfrac{1}{K}\left[\tau(p_0 - g_2) - 1\right]$
$p_0 > 0$ for $\tau_d < 2\tau$

Table 2. Controller parameters for ITDS

TN expansion
$p_0 = 1 + \tau_d(g_1 + K\tau_d)$ $t_0 = r_0 = \dfrac{1}{K} g_0, \quad t_1 = \dfrac{1}{K}(g_1 + \tau_d g_0)$
$p_0 > 0$ for all τ_d
Padé approximation

$p_0 = g_2 + \frac{\tau_d}{4}(2g_1 + \tau_d g_0)\,,\;\; t_0 = r_0 = \frac{1}{K} g_0$ $t_1 = \frac{1}{K}(g_1 + \tau_d g_0)\,,\;\; t_2 = \frac{\tau_d}{4K}(2g_1 + \tau_d g_0)$
$p_0 > 0$ for all τ_d

Table 3. Controller parameters for USOTDS

TN expansion
$p_0 = \frac{\tau_1(g_2 + \tau_d g_1 + \tau_d^2 g_0) + 1}{\tau_1 - \tau_d}$ $t_0 = \frac{\tau_1}{K} g_0\,,\;\; t_1 = \frac{1}{K}\left[p_0 + \tau_1 g_1 + \tau_1(\tau_2 + \tau_d) g_0\right]$ $t_2 = \frac{1}{K}\frac{\tau_1 \tau_2}{\tau_d}\left[p_0 - g_2 - \frac{1}{\tau_1}\right]$
$p_0 > 0$ for $\tau_d < \tau_1$
Padé approximation
$p_0 = \frac{2g_3 + \tau_1\left[2g_2 + \tau_d\left(g_1 + \frac{\tau_d}{2} g_0\right)\right] + \frac{2}{\tau_1}}{2\tau_1 - \tau_d}$ $p_1 = g_3 + \frac{1}{\tau_1}$ $t_0 = \frac{\tau_1}{K} g_0\,,\;\; t_1 = \frac{1}{K}\left[p_0 + \tau_1\left(g_1 + (\tau_2 + \tau_d) g_0\right)\right]$ $t_2 = \frac{1}{K}\left[\left(\frac{4\tau_1\tau_2}{\tau_d} + \tau_1 - \tau_2\right) p_0 - \right.$ $\left. -\left(\frac{4\tau_2}{\tau_d} + 1\right)\left(g_3 + \tau_1 g_2 + \frac{1}{\tau_1}\right) - \tau_1 \tau_2 g_1\right]$ $t_3 = \frac{\tau_2}{K}\left[\tau_1(p_0 - g_2) - g_3 - \frac{1}{\tau_1}\right]$
$p_0 > 0$ for $\tau_d < 2\tau_1$

Table 4. Controller parameters for SFOPITDS

TN expansion
$p_0 = g_2 + \tau_d(g_1 + \tau_d g_0)\,,\;\; t_0 = \frac{1}{K} g_0$ $t_1 = \frac{1}{K}[g_1 + (\tau + \tau_d) g_0]\,,\;\; t_2 = \frac{\tau}{K}(g_1 + \tau_d g_0)$
$p_0 > 0$ for all τ_d
Padé approximation
$p_0 = g_2 + \frac{\tau_d}{4}(2g_1 + \tau_d g_0)\,,\;\; p_1 = g_3$ $t_0 = \frac{1}{K} g_0\,,\;\; t_1 = \frac{1}{K}\left[g_1 + (\tau + \tau_d) g_0\right]$ $t_2 = \frac{1}{4K}\left[(2\tau + \tau_d)(2g_1 + \tau_d g_0) + 2\tau\tau_d g_0\right]$ $t_3 = \frac{\tau\tau_d}{4K}(2g_1 + \tau_d g_0)$
$p_1 > 0$ for all τ_d, $p_0 > 0$ for all τ_d

Table 5. Controller parameters for UFOPITDS

TN expansion
$p_0 = \frac{(\tau + \tau_d)\left[g_2 + \tau_d(g_1 + \tau_d g_0)\right] + 2\tau}{\tau - \tau_d}$
$t_0 = \frac{1}{K} g_0\,,\;\; t_1 = \frac{1}{K}\left[g_1 + (\tau + \tau_d) g_0\right]$ $t_2 = \frac{1}{K}\frac{2\tau g_2 + \tau(\tau + \tau_d)(g_1 + \tau_d g_0) + 2}{\tau - \tau_d}$
$p_0 > 0$ for $\tau_d < \tau$
Padé approximation
$p_0 = \frac{4g_3 + (2\tau + \tau_d)\left(g_2 + \frac{\tau_d}{2} g_1 + \frac{\tau_d^2}{4} g_0\right) + \frac{4}{\tau}}{2\tau - \tau_d}$ $p_1 = g_3 + \frac{2}{\tau}$ $t_0 = \frac{1}{K} g_0\,,\;\; t_1 = \frac{1}{K}\left[g_1 + (\tau + \tau_d) g_0\right]$ $t_2 = \frac{1}{K}\left[\left(\frac{4\tau}{\tau_d} - 1\right) p_0 - \frac{8}{\tau_d} g_3 - \left(\frac{4\tau}{\tau_d} + 1\right) g_2 - \right.$ $\left. -\tau g_1 - \frac{8}{\tau\tau_d}\right]$ $t_3 = \frac{1}{K}\left[\tau(p_0 - g_2) - 2g_3 - \frac{2}{\tau}\right]$
$p_1 > 0$ for all τ_d , $p_0 > 0$ for $\tau_d < 2\tau$

4 Simulation Results

All simulations were performed by MATLAB-Simulink tools. In all cases, the unit step reference w was introduced at the time $t = 0$ and the step disturbances v_1 and v_2 were subsequently injected after settling of the control responses.

4.1 UFOTDS

The parameters in the transfer function (1) has been chosen as $K = 1$ and $\tau = 4$.

The responses in Fig.2 document applicability of the TNE for the UFOTDS with a small value of τ_d. Further, the responses illustrate necessity of a higher value of φ to achieving of an aperiodic character of responses. Smaller values of φ lead to their oscillatory character. An effect of the parameter β_1 can be seen in Fig.3. Its increasing value speeds the control but causes expressive overhoots.

A preference of the PA in comparison with the TN is evident from the controlled output responses in Fig.4 computed under the same conditions. Moreover, the PA enables a use also for higher values of τ_d as shown in Fig.5.

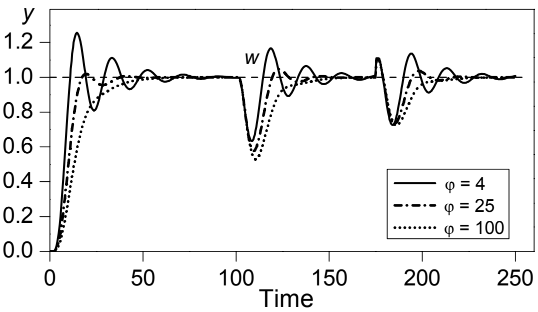

Fig.2. UFOTDS - TNE: Controlled output for various φ ($\tau_d = 2$, $\beta_1 = 0$, $v_1 = -0.2$, $v_2 = 0.1$).

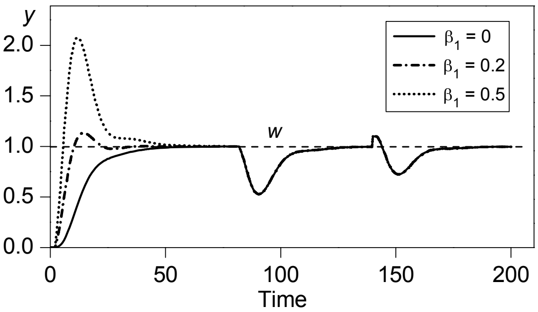

Fig.3. UFOTDS - TNE: Controlled output for various β_1 ($\tau_d = 2$, $\varphi = 100$, $v_1 = -0.2$, $v_2 = 0.1$).

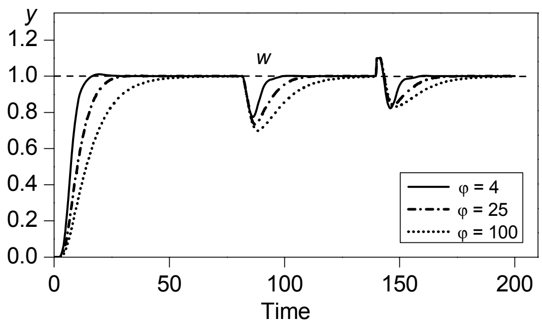

Fig.4. UFOTDS - PA: Controlled output for various φ ($\tau_d = 2$, $\beta_{1,2} = 0$, $v_1 = -0.2$, $v_2 = 0.1$).

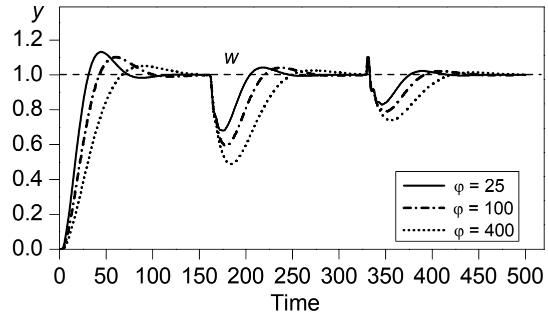

Fig.5. UFOTDS - PA: Controlled output for various φ ($\tau_d = 4$, $\beta_{1,2} = 0$, $v_1 = -0.2$, $v_2 = 0.1$).

4.2 USOTDS

The parameters in the transfer function (2) were chosen as $K = 1$, $\tau_1 = 4$, $\tau_2 = 2$.

Also in this case, an application of the TNE is possible for smaller values of the time delay and for higher values of φ. A higher value of τ_d needs a use of the PA. The simulation results can be seen in Figs.6 and 7.

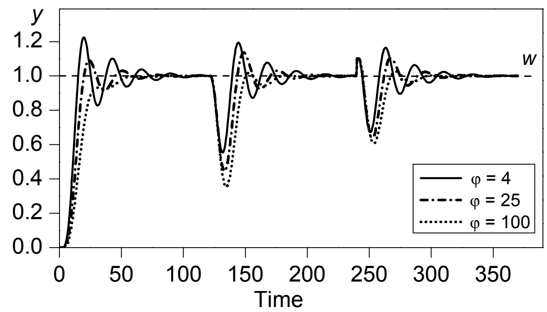

Fig.6. USOTDS - TNE: Controlled output for various φ ($\tau_d = 2$, $\beta_{1,2} = 0$, $v_1 = -0.2$, $v_2 = 0.1$).

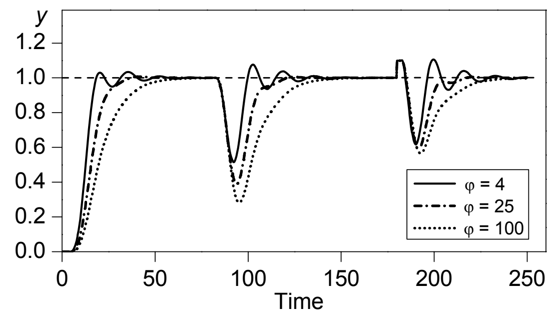

Fig.7. USOTDS - PA: Controlled output for various α ($\tau_d = 3$, $\beta_{1,2,3} = 0$, $v_1 = -0.2$, $v_2 = 0.1$).

The responses in Fig.8 demonstrate their high sensitivity to parameters β. Evidently, on behalf of acceleration of the control, only small values β should be chosen. Their higher values lead to expresive overshoots at the start of the tracking interval.

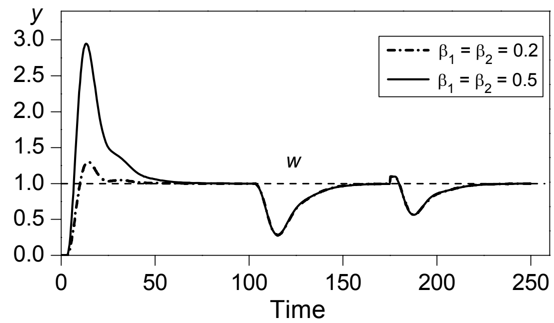

Fig.8. USOTDS - PA: Controlled output for various β_1, β_2 ($\tau_d = 3$, $\varphi = 100$, $\beta_3 = 0$, $v_1 = -0.2$, $v_2 = 0.1$).

4.3 ITDS

In this case, the parameter in (3) has been chosen as $K = 0.2$.

The responses in Fig.9 document applicability of the TNE for the ITDS with smaller values of τ_d. There is not a significant difference in comparison with utilization of the PA as shown in Fig.10. Here, also a selection of the parameter φ is not very important.

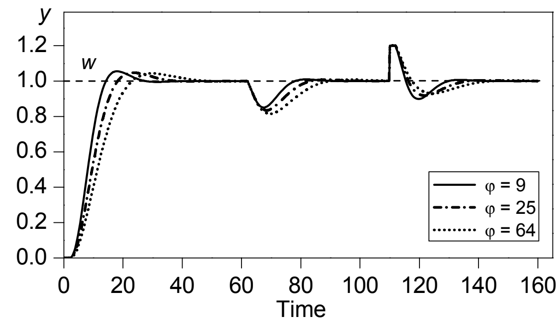

Fig.9. ITDS - TNE: Controlled output for various φ ($\tau_d = 2$, $\beta_1 = 0$, $v_1 = -0.2$, $v_2 = 0.2$).

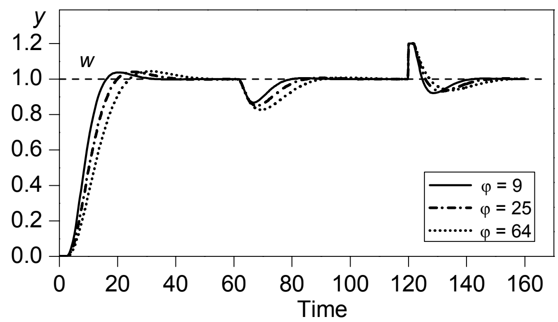

Fig.10. ITDS - PA: Controlled output for various φ ($\tau_d = 2$, $\beta_1 = 0$, $v_1 = -0.2$, $v_2 = 0.2$).

An effect of the parameter β_1 on the controlled output responses can be seen in Fig.11. A reasonable choice of this parameter can accelerate the control responses keeping their apperiodic character.

A difference between both approximations appears for higher values of τ_d as it can be seen in Figs.12, 13 and 14. There, a priority of the PA is evident. It is also clear that a higher value of τ_d requires a use of a higher value of φ.

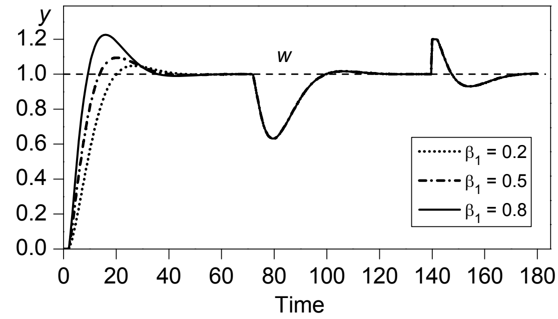

Fig.11. ITDS - TNE: Controlled output for various β_1 ($\tau_d = 5$, $\varphi = 25$, $v_1 = -0.4$, $v_2 = 0.2$).

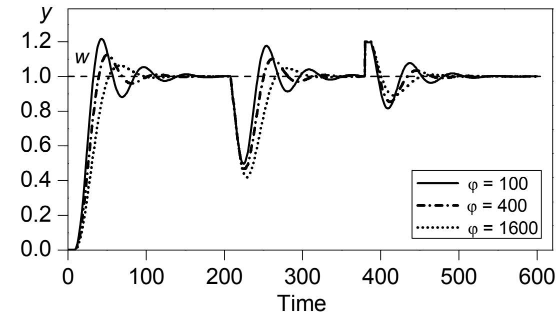

Fig.12. ITDS - TNE: Controlled output for various φ ($\tau_d = 8$, $\beta_1 = 0$, $v_1 = -0.2$, $v_2 = 0.2$).

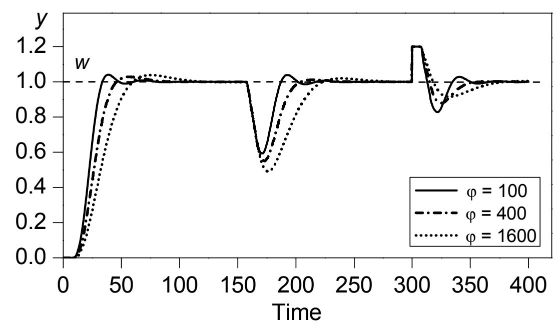

Fig.13. ITDS - PA: Controlled output for various φ ($\tau_d = 8$, $\beta_1 = \beta_2 = 0$, $v_1 = -0.2$, $v_2 = 0.2$).

4.4 SFOPITDS

For this model (and, also for the UFOPITDS), the parameters in (2) have been chosen as $K = 0.2$ and $\tau = 4$. The controlled output responses for various φ are shown in Figs.15 and 16, a comparison between application of the TNE and PA can be seen in Fig.17. The presented results clearly prove a better control quality obtained by the PA. It should be noted that for both SFOPITDS and UFOPITDS zero parameters β were chosen equivalent to the 2DOF control structure. This choice gave best control results.

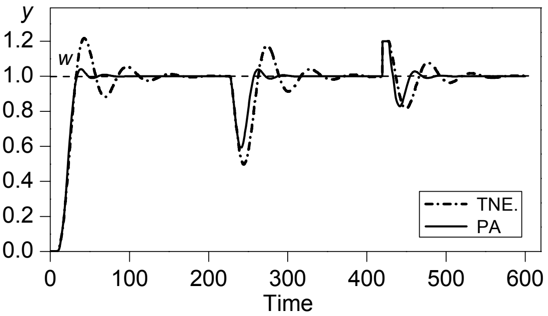

Fig.14. ITDS – Comparison of controlled outputs for TNE and PA ($\tau_d = 8$, $\varphi = 100$, $\beta_1 = \beta_2 = 0$, $v_1 = -0.2$, $v_2 = 0.2$).

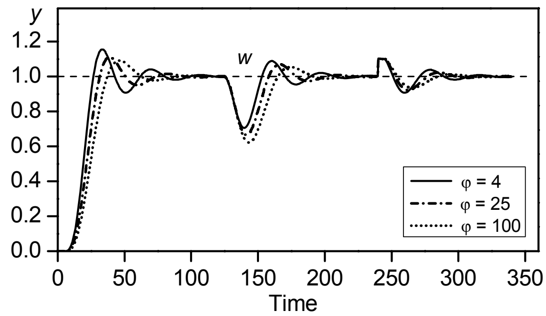

Fig.15. SFOPITDS - TNE: Controlled output for various φ ($\tau_d = 5$, $v_1 = -0.2$, $v_2 = 0.1$).

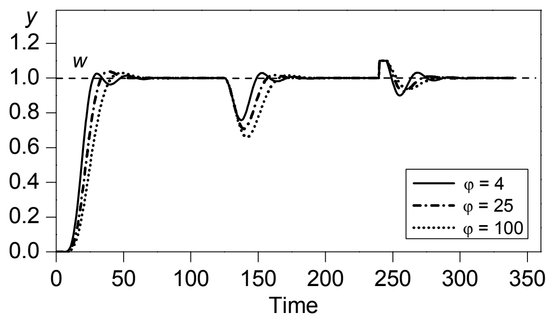

Fig.16. SFOPITDS - PA: Controlled output for various φ ($\tau_d = 5$, $v_1 = -0.1$, $v_2 = 0.1$).

4.5 UFOPITDS

With regard to a presence of both integrating and unstable parts, the UFOPITDSs belong to hardly controllable systems. However, the control responses in Fig.18 document usability of both TNE and PA for smaller value of τ_d. Higher values of τ_d require a selection of higher values of φ as shown for the PA in Fig.19 However, for higher values of φ, the TNE is unsuitable, as documented in Fig.20.

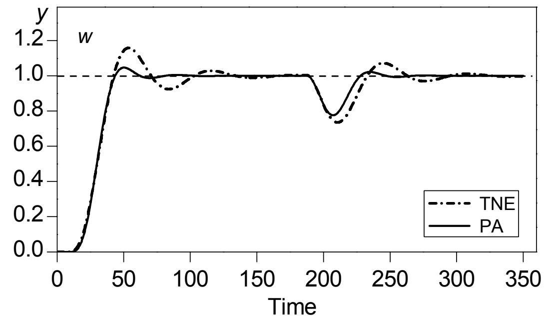

Fig.17. SFOPITDS - Comparison of controlled outputs for TNE and PA ($\tau_d = 8$, $\varphi = 100$, $v_1 = -0.1$, $v_2 = 0.2$)

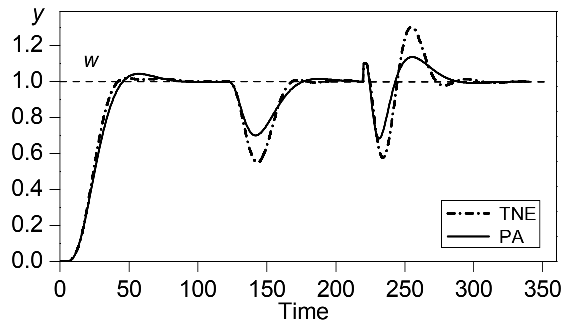

Fig.18. UFOPITDS - Comparison of controlled outputs for TNE and PA ($\tau_d = 2$, $\varphi = 400$, $v_1 = -0.05$, $v_2 = 0.1$).

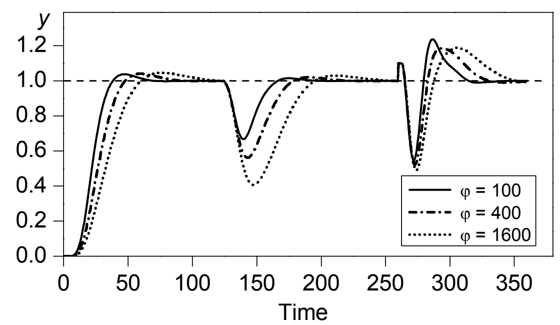

Fig.19. UFOPITDS - PA: Controlled output for various φ ($\tau_d = 3$, $v_1 = -0.05$, $v_2 = 0.1$).

5 Conclusions

The problem of control design for unstable and integrating time delay systems has been solved and analysed. The proposed method is based in two ways of the time delay approximation. The controller design uses the polynomial synthesis and the controller setting employs the results of the LQ control theory. The presented procedure provides satisfactory control responses in the tracking of a step reference as well as in step disturbances attenuation. The presented results have demonstrate the usability of the method and the control of a good quality also for relatively high ratio of the time delay to the time constant. The procedure makes possible a tuning of the controller parameters by two types of selectable parameters. Using derived formulas, the controller parameters can be automatically computed. From this reason, the method could also be used for an adaptive control.

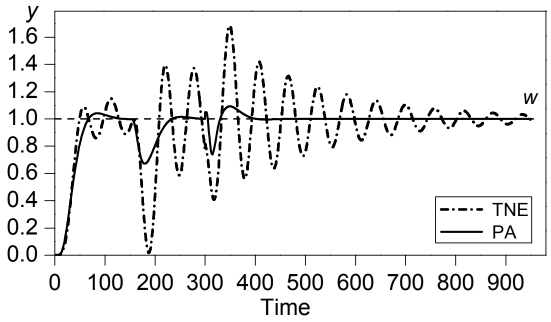

Fig.20. UFOPITDS – Comparison of controlled outputs for TNE and PA ($\tau_d = 3$, $\varphi = 2500$, $v_1 = -0.05$, $v_2 = 0.1$).

Acknowledgments

This article was written in part with support of Operational Program Research and Development for Innovations co-funded by the European Regional Development Fund (ERDF) and national budget of Czech Republic, within the framework of project Centre of Polymer Systems (reg. number: CZ.1.05/2.1.00/03.0111) and with support of the grant IGA/FAI/2012/020.

References:

[1] M. De Paor, A modified Smith predictor and controller for unstable processes with time delay, *Int. J. Control*, Vol. 58, 1995, pp. 1025-1036.

[2] S. Majhi, S., and D.P. Atherton, Modified Smith predictor and controller for processes with time delay, *IEE Proc. Control Theory Appl.*, Vol. 146, 1999, pp. 359-366.

[3] T. Liu, Y.Z. Cai, D.Y. Gu, and W.D. Zhang, New modified Smith predictor scheme for integrating and unstable processes with time delay, *IEE Proc. Control Theory Appl.*, Vol. 152, 2005, pp. 238-246.

[4] M.R. Matausek, and A.D. Micic, A modified Smith predictor for controlling a process with an integrator ang long dead-time. *IEEE Trans. Aut. Control*, Vol. 41, 1996, pp. 1199-1203.

[5] L. Wang, L., and W.R. Cluett, Tuning PID controllers for integrating processes. *IEE Proc. Control Theory Appl.*, Vol. 144, 1997, pp. 385-392.

[6] T. Emami, and J. M. Watkins, A unified approach for sensitivity design of PID controllers in the frequency domain, *WSEAS Transactions on Systems and Control*, Vol. 4, 2009, pp. 221-231.

[7] J.H. Park, S.W. Sung, and I. Lee, An enhanced PID control strategy for unstable processes. *Automatica*, Vol. 34, 1998, pp. 751-756.

[8] G.J. Silva, A. Datta, and S.P. Bhattacharyya, *PID controllers for time-delay systems*, Birkhäuser, Boston, 2005.

[9] R. Prokop, and J.P. Corriou, Design and analysis of simple robust controller. *Int. J. Control*, Vol. 66, 1997, pp. 905-921.

[10] R. Matušů, R. Prokop, and L. Pekař, Uncertainty modelling in time-delay systems: Parametric vs. unstructured approach, in *Recent Researches in Automatic Control* (*Proc. of the 13th WSEAS International Conference on Automatic Control, Modelling & Simulation*), Lanzarote, Spain, 2011, pp. 310-313.

[11] L. Pekař, Root locus analysis of a retarded quasipolynomial, *WSEAS Transactions on Systems and Control*, Vol. 6, 2010, pp. 79-91.

[12] Y. Zhang, Spectrum of a class of delay differential equations and its solution expansion, *WSEAS Trans. Mathematics*, Vol. 10, 2011, pp. 169-180.

[13] F. Neri, Software agents as a versatile simulation tool to model complex systems, *WSEAS Transactions on Information Science and Applications*, Vol. 7, 2010, pp.609-618.

[14] P. Dostál, F. Gazdoš, V. Bobál, and J. Vojtěšek, Adaptive control of a MIMO process by two feedback controllers, *AT&P Journal PLUS2*, 2007, No.1, pp. 93-98.

[15] P. Dostál, F. Gazdoš, V. Bobál, and J. Vojtěšek, Adaptive control of a continuous stirred tank reactor by two feedback controllers, in *9th IFAC Workshop Adaptation and Learning in Control and Signal Processing*, Saint Petersburg, Russia, 2007,

[16] V. Kučera, Diophantine equations in control – A survey. *Automatica*, Vol. 29, 1993, pp. 1361-1375.

[17] P. Dostál, V. Bobál, and M. Sysel, Design of controllers for integrating and unstable time delay systems using polynomial method, in *2002 American Control Conference*, Anchorage, Alaska, USA, 2002, 2773-2778.

Identification and Self-tuning Control of Time-delay Systems

VLADIMÍR BOBÁL, PETR CHALUPA, MAREK KUBALČÍK, PETR DOSTÁL
Tomas Bata University in Zlín, Faculty of Applied Informatics
Nám. T. G. Masaryka 5555, 7605 Zlín
CZECH REPUBLIC
bobal@fai.utb.cz http://web.fai.utb.cz/cs/docs/CV_english.pdf

Abstract: - Time-delays (dead times) occur in many processes in industry. A Toolbox in the MATLAB/SIMULINK environment was designed for identification and self-tuning control of such processes. The control algorithms are based on modifications of the Smith Predictor (SP). The designed algorithms that are included in the toolbox are suitable not only for simulation purposes but also for implementation in real time conditions. Verification of the designed Toolbox is demonstrated on a self-tuning control of a laboratory heat exchanger in simulation conditions.

Key-Words: - Time-delay; Smith predictor; Process identification; ARX model; Self-tuning control; PID control; Pole assignment; Time-delay Toolbox; Heat exchanger

1 Introduction

The majority of processes in the industrial practice have stochastic characteristics and eventually they exhibit nonlinear behaviour. Traditional controllers with fixed parameters are often unsuitable for such processes because parameters of the process change. One possible alternative for improving the quality of control of such processes is application of adaptive control systems. Different approaches were proposed and utilized. One of the successful approaches is self-tuning control (STC) [1] – [5].

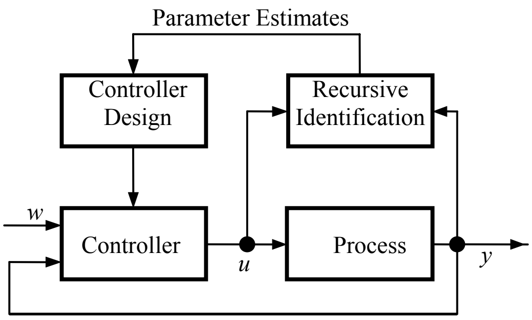

Fig. 1. Self-tuning control system

The block diagram of an STC is shown in Fig. 1, where *y*, *u* and *w* are the process output, the control signal and the reference signal. The main idea of the STC is based on combination of a recursive identification procedure and a particular controller synthesis. The self-tuning strategy was applied for design of control of time-delay systems.

Time-delays appear in many processes in industry and other fields, including economical and biological areas. They are caused by some of the following phenomena [6]:

- the time needed to transport mass, energy or information,
- the accumulation of time lags in a great numbers of low order systems connected in series,
- the required processing time for sensors, such as analyzers; controllers that need some time to implement a complicated control algorithms or processes.

Consider a continuous time dynamical linear SISO (single input $u(t)$ – single output $y(t)$) system with time-delay T_d. The transfer function of a pure transportation lag is $e^{-T_d s}$ where s is complex variable. Overall transfer function with time-delay is in the form

$$G_d(s) = G(s)e^{-T_d s} \tag{1}$$

where $G(s)$ is the transfer function without time-delay. Processes with significant time-delay are difficult to control using standard feedback controllers. When a high performance of the control process is desired or the relative time-delay is very large, a predictive control strategy must be used. The predictive control strategy includes a model of the process in the structure of the controller. The first time-delay compensation algorithm was proposed by Smith 1957 [7]. This control algorithm known as the Smith Predictor (SP) contained a dynamic model of the time-delay process and it can be considered as the first model predictive algorithm.

Although time-delay compensators appeared in the mid 1950s, their implementation with analog technique was very difficult and these were not used in industry. Since 1980s digital time-delay compensators can be implemented. In spite of the fact that all these algorithms are implemented on digital platforms, most works analyze only the continuous case. The digital time-delay compensators are presented e.g. in [8], [9], [10]. Two STC modifications of the digital Smith Predictors (STCSP) are designed in [11] and implemented into MATLAB/SIMULINK Toolbox [12].

The paper is organized in the following way. The principle of the digital Smith Predictor is described in Section 2. Section 3 contains description of the off-line and on-line (recursive) identification procedure. Two modifications of digital controllers that are used for self-tuning versions SPs are proposed in Section 4. The designed Toolbox is briefly described in Section 5. An example of the real-time identification and simulation control of the laboratory heat exchanger contains Section 6. Section 7 concludes the paper.

2 Digital Smith Predictors

The discrete versions of the SP and its modifications are suitable for time-delay compensation in industrial practice. Most of authors designed the digital SP using discrete PID controllers with fixed parameters. However, the SP is more sensitive to process parameter variations and therefore requires an auto-tuning or adaptive approach in many practical applications.

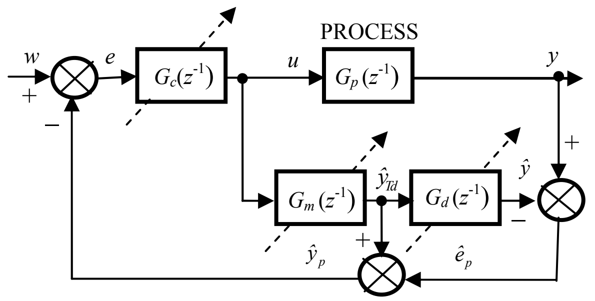

Fig. 2. Block Diagram of a Digital Smith Predictor

The block diagram of a digital SP [13], [14] is shown in Fig. 2. The function of the digital version is similar to the classical analog version. The block $G_m\left(z^{-1}\right)$ represents process dynamics without the time-delay and is used to compute an open-loop prediction. The difference between the output of the process y and the model including time delay $\hat{y}$ is the predicted error $\hat{e}_p$ as shown in Fig. 2, whereas e and e_s are the error and the noise, respectively and w is the reference signal. If there are no modelling errors or disturbances, the error between the current process output y and the model output $\hat{y}$ will be null. Then the predictor output signal $\hat{y}_p$ will be the time-delay-free output of the process. Under these conditions, the controller $G_c\left(z^{-1}\right)$ can be tuned, at least in the nominal case, as if the process had no time-delay. The primary (main) controller $G_c\left(z^{-1}\right)$ can be designed by different approaches (for example digital PID control or methods based on algebraic approach). The outward feedback-loop through the block $G_d\left(z^{-1}\right)$ in Fig. 2 is used to compensate for load disturbances and modelling errors. The dash arrows indicate the self-tuned parts of the Smith Predictor.

Most industrial processes can be approximated by a reduced order model with a pure time-delay. Consider the following second order linear model with a time-delay

$$G\left(z^{-1}\right)=\frac{B\left(z^{-1}\right)}{A\left(z^{-1}\right)}z^{-d}=\frac{b_1z^{-1}+b_2z^{-2}}{1+a_1z^{-1}+a_2z^{-2}}z^{-d} \tag{2}$$

for demonstration of some approaches to the design of the adaptive Smith Predictor. The term z^{-d} represents the pure discrete time-delay. The time-delay is equal to dT_0 where T_0 is the sampling period. Model (2) is used in control algorithms of the designed Toolbox.

3. Identification Procedure

3.1 Identification of Time-delay

In this paper, the time-delay models are obtained separately from an off-line identification using the least squares method (LSM) [15]. The measured process output $y(k)$ is generally influenced by noise. These nonmeasurable disturbances cause errors $\boldsymbol{e}$ in the determination of model parameters and therefore real output vector is in the form

$$\boldsymbol{y}=\boldsymbol{F\Theta}+\boldsymbol{e} \tag{3}$$

It is possible to obtain the LSM expression for calculation of the vector of the parameter estimates

$$\hat{\boldsymbol{\Theta}}=\left(\boldsymbol{F}^T\boldsymbol{F}\right)^{-1}\boldsymbol{F}^T\boldsymbol{y} \tag{4}$$

The matrix $\boldsymbol{F}$ has dimension (N-n-d, $2n$), the vector $\boldsymbol{y}$ (N-n-d) and the vector of parameter model estimates $\hat{\boldsymbol{\Theta}}$ ($2n$). N is the number of samples of measured input and output data, n is the model order [16].

Equation (4) serves for calculation of the vector of the parameter estimates $\hat{\boldsymbol{\Theta}}$ using N samples of measured input-output data. The individual vectors and matrices in equations (3) and (4) have the form

$$\boldsymbol{y}^T = \begin{bmatrix} y(n+d+1) & y(n+d+2) & \cdots & y(N) \end{bmatrix} \quad (5)$$

$$\boldsymbol{e}^T = \begin{bmatrix} \hat{e}(n+d+1) & \hat{e}(n+d+2) & \cdots & \hat{e}(N) \end{bmatrix} \quad (6)$$

$$\hat{\boldsymbol{\Theta}}^T = \begin{bmatrix} \hat{a}_1 & \hat{a}_2 & \cdots & \hat{a}_n & \hat{b}_1 & \hat{b}_2 & \cdots & \hat{b}_n \end{bmatrix} \quad (7)$$

$$\boldsymbol{F} = \begin{bmatrix} -y(n+d) & -y(n+d-1) & \cdots & -y(d+1) & u(n) & u(n-1) & \cdots & u(1) \\ -y(n+d+1) & -y(n+d) & \cdots & -y(d+2) & u(n+1) & u(n) & \cdots & u(2) \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ -y(N-1) & -y(N-2) & \cdots & -y(N-n) & u(N-d-1) & u(N-d-2) & \cdots & u(N-d-n) \end{bmatrix} \quad (8)$$

It is obvious that the quality of time-delay systems identification is very dependent on the choice of a suitable input exciting signal $u(k)$. Therefore the MATLAB function from the System Identification Toolbox

$$u = \text{idinput}(N, type, band, levels)$$

was used. This MATLAB code generates input signals u of different kinds, which are typically used for identification purposes. N determines the number of generated input data. *Type* defines the type of input signal to be generated. This argument takes one of the following values [17]:

type = 'rgs': Gives a random, Gaussian signal.
type = 'rbs': Gives a random, binary signal. This is the default.
type = 'prbs': Gives a pseudorandom, binary signal.
type = 'sine': Gives a signal that is a sum of sinusoids.

The frequency contents of the signal is determined by the argument *band*. For the choices type = 'rs', 'rbs', and 'sine', this argument is a row vector with two entries *band* = [wlow, whigh] that determine the lower and upper bound of the passband. The frequencies wlow and whigh are expressed in fractions of the Nyquist frequency. A white noise character input is thus obtained for *band* = [0 1], which is also the default value.
For the choice type = 'prbs', *band* = [0, B], where B is such that the signal is constant over intervals of length 1/B (the clock period). In this case the default is *band* = [0 1].

The argument *levels* defines the input level. It is a row vector *levels* = [minu, maxu] such that the signal u will always be between the values minu and maxu for the choices type = 'rbs', 'prbs', and 'sine'. For type = 'rgs', the signal level is such that minu is the mean value of the signal, minus one standard deviation, while maxu is the mean value plus one standard deviation. Gaussian white noise with zero mean and variance one is thus obtained for levels = [-1, 1], which is also the default value. The example of exciting input signals of the "idinput" function are depicted in the Fig. 3.

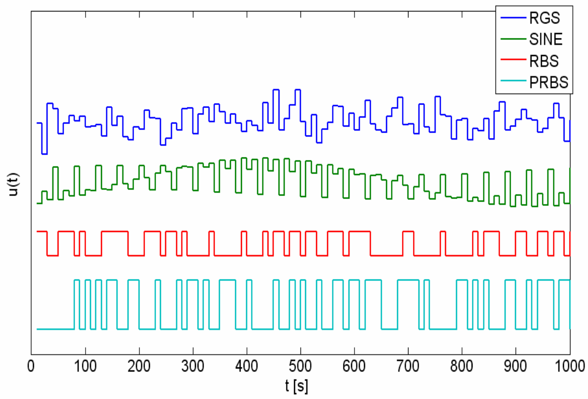

Fig. 3. Example of exciting input signals of "idinput" function

Consider that model (2) is the deterministic part of the stochastic process described by the ARX (regression) model

$$\begin{aligned} y(k) = &-a_1 y(k-1) - a_2 y(k-2) + \\ &+ b_1 y(k-1-d) + b_2 y(k-2-d) + e_s(k) \end{aligned} \quad (9)$$

where $e_s(k)$ is the random nonmeasurable component. The vector of parameter model estimates is computed by solving equation (4)

$$\hat{\boldsymbol{\Theta}}^T(k) = \begin{bmatrix} \hat{a}_1 & \hat{a}_2 & \hat{b}_1 & \hat{b}_2 \end{bmatrix} \quad (10)$$

and is used for computation of the prediction output

$$\begin{aligned} \hat{y}(k) = &-\hat{a}_1 y(k-1) - \hat{a}_2 y(k-2) + \\ &\hat{b}_1 u(k-1-d) + \hat{b}_2 u(k-2-d) \end{aligned} \quad (11)$$

The quality of identification can be considered according to error, i.e. the deviation

$$\hat{e}(k) = y(k) - \hat{y}(k) \tag{12}$$

In this paper, the error was used for suitable choice of the time-delay dT_0. The LSM algorithm (4) – (8) is computed for several time-delays dT_0 and the suitable time-delay is chosen according to quality of identification based on the prediction error (12).

For the off-line process identification the MATLAB function from the Optimization Toolbox

$$x = \text{fminsearch}('\mathit{name\_fce}', x_0)$$

was also used. This function find minimum of unconstrained multivariable function using derivative-free method. Algorithm "fminsearch" uses the simplex search method of [18]. This is a direct search method that does not use numerical or analytic gradients.

3.2 Recursive Identification Algorithm

The regression (ARX) model of the following form

$$y(k) = \boldsymbol{\Theta}^T(k)\boldsymbol{\Phi}(k) + e_s(k) \tag{13}$$

is used in the identification part of the designed controller algorithms, where

$$\boldsymbol{\Theta}^T(k) = \begin{bmatrix} a_1 & a_2 & b_1 & b_2 \end{bmatrix} \tag{14}$$

is the vector of model parameters and

$$\boldsymbol{\Phi}^T(k-1) = \left[-y(k-1) - y(k-2)\, u(k-d-1)\, u(k-d-2)\right] \tag{15}$$

is the regression vector. The non-measurable random component $e_s(k)$ is assumed to have zero mean value $E[e_s(k)] = 0$ and constant covariance (dispersion)
$R = E[e_s^{\,2}(k)]$.

Both digital adaptive SP controllers use the algorithm of identification based on the Recursive Least Squares Method (RLSM) extended to include the technique of directional (adaptive) forgetting. Numerical stability is improved by means of the LD decomposition [5], [19]. This method is based on the idea of changing the influence of input-output data pairs to the current estimates. The weights are assigned according to amount of information carried by the data.

When using the self-tuning principle, the model parameter estimates must approach the true values right from the start of the control. This means that as the self-tuning algorithm begins to operate, identification must be run from suitable conditions – the result of the possible *a priori* information. The role of suitable initial conditions in recursive identification is often underestimated.

4 Controller Algorithms

4.1 Digital PID Smith Predictor

Hang *et al.* [13], [14] used the Dahlin PID algorithm [20] for the design of the main controller $G_c(z^{-1})$. This algorithm is based on the desired close-loop transfer function in the form

$$G_e(z^{-1}) = \frac{1 - e^{-\alpha}}{1 - z^{-1}}; \;\; \alpha = \frac{T_0}{T_m} \tag{16}$$

where T_m is a desired time constant of the first order closed-loop response. It is not practical to set T_m to be small since it will demand a large control signal $u(k)$ which may easily exceed the saturation limit of the actuator. Then the individual parts of the controller are described by the transfer functions

$$G_c(z^{-1}) = \frac{(1 - e^{-\alpha})}{(1 - z^{-1})} \frac{\hat{A}(z^{-1})}{\hat{B}(1)}; \; G_m(z^{-1}) = \frac{z^{-1}\hat{B}(1)}{\hat{A}(z^{-1})}$$

$$G_d(z^{-1}) = \frac{z^{-d}\hat{B}(z^{-1})}{z^{-1}\hat{B}(1)} \tag{17}$$

where $B(1) = \hat{B}(z^{-1})\Big|_{z=1} = \hat{b}_1 + \hat{b}_2$.

Since $G_m(z^{-1})$ is the second order transfer function, the main controller $G_c(z^{-1})$ becomes a digital PID controller having the following form:

$$G_c(z^{-1}) = \frac{U(z)}{E(z)} = \frac{q_0 + q_1 z^{-1} + q_2 z^{-2}}{1 - z^{-1}} \tag{18}$$

where $q_0 = \gamma, q_1 = \hat{a}_1\gamma, q_2 = \hat{a}_2\gamma$ using by the substitution $\gamma = (1 - e^{-\alpha})/\hat{B}(1)$. The PID controller output is given by

$$u(k) = q_0 e(k) + q_1 e(k-1) + q_2 e(k-2) + u(k-1) \tag{19}$$

Some simulation experiments using this digital PID SP are presented in [11].

4.2 Digital Pole Assignment (PA) Smith Predictor

The digital pole assignment SP was designed using a polynomial approach in [11]. Polynomial control theory is based on the apparatus and methods of linear algebra (see e.g. [21] - [24]). The design of the controller algorithm is based on the general block scheme of a closed-loop with two degrees of freedom (2DOF) according to Fig. 4 [25].

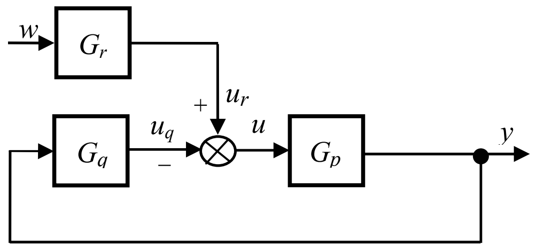

Fig. 4. Block Diagram of a Closed Loop 2DOF Control System

The controlled process is given by the transfer function in the form

$$G_p(z^{-1}) = \frac{Y(z)}{U(z)} = \frac{B(z^{-1})}{A(z^{-1})} \tag{20}$$

where A and B are the second order polynomials. The controller contains the feedback part G_q and the feedforward part G_r. Then the digital controllers can be expressed in the form of discrete transfer functions

$$G_r\left(z^{-1}\right) = \frac{R\left(z^{-1}\right)}{P\left(z^{-1}\right)} = \frac{r_0}{1 + p_1 z^{-1}} \tag{21}$$

$$G_q\left(z^{-1}\right) = \frac{Q\left(z^{-1}\right)}{P\left(z^{-1}\right)} = \frac{q_0 + q_1 z^{-1} + q_2 z^{-2}}{\left(1 + p_1 z^{-1}\right)\left(1 - z^{-1}\right)} \tag{22}$$

According to the scheme presented in Fig. 3 and Equations (20) – (22) it is possible to derive the characteristic polynomial

$$A(z^{-1})P(z^{-1}) + B(z^{-1})Q(z^{-1}) = D(z^{-1}) \tag{23}$$

where

$$D\left(z^{-1}\right) = 1 + d_1 z^{-1} + d_2 z^{-2} + d_3 z^{-3} + d_4 z^{-4} \tag{24}$$

The feedback part of the controller is given by solution of the polynomial Diophantine equation (23). The procedure leading to determination of controller parameters in polynomials Q, R and P (21) and (22) is in [5]. The asymptotic tracking is provided by the feedforward part of the controller given by solution of the polynomial Diophantine equation

$$S(z^{-1})D_w\left(z^{-1}\right) + B(z^{-1})R(z^{-1}) = D(z^{-1}) \tag{25}$$

For a step-changing reference signal value $D_w\left(z^{-1}\right) = 1 - z^{-1}$ holds and S is an auxiliary polynomial which does not enter into controller design and it is possible to solve Equation (25) by substituting $z = 1$

$$R\left(z^{-1}\right) = r_0 = \frac{D(1)}{B(1)} = \frac{1 + d_1 + d_2 + d_3 + d_4}{b_1 + b_2} \tag{26}$$

The 2DOF controller output is given by

$$\begin{aligned} u(k) &= r_0 w(k) - q_0 y(k) - q_1 y(k-1) - \\ &- q_2 y(k-2) + (1 + p_1) u(k-1) + p_1 u(k-2) \end{aligned} \tag{27}$$

The control quality is very dependent on the pole assignment of the characteristic polynomial

$$D(z) = z^4 + d_1 z^3 + d_2 z^2 + d_3 z + d_4 \tag{28}$$

inside the unit circle. The simple method for choice of individual poles is based on the following approach. Consider 1DOF control loop where controlled process (20) with second-order polynomials A and B is controlled using PID controller which is given by transfer function

$$G_q\left(z^{-1}\right) = \frac{Q\left(z^{-1}\right)}{P\left(z^{-1}\right)} = \frac{q_0\left(1 + a_1 z^{-1} + a_2 z^{-2}\right)}{\left(1 - z^{-1}\right)} \tag{29}$$

Substitution of polynomials A, B, Q, P into Equation (23) yields the following relation

$$\begin{aligned} &\hat{A}(z^{-1})\left(1 - z^{-1}\right) + \hat{B}(z^{-1}) q_0 \hat{A}(z^{-1}) = \\ &= \hat{A}(z^{-1})\left[\left(1 - z^{-1}\right) + \hat{B}(z^{-1}) q_0\right] = D(z^{-1}) \end{aligned} \tag{30}$$

where

$$\hat{A}\left(z^{-1}\right) = 1 + \hat{a}_1 z^{-1} + \hat{a}_2 z^{-2}; \quad \hat{B}\left(z^{-1}\right) = \hat{b}_1 z^{-1} + \hat{b}_2 z^{-2} \tag{31}$$

are polynomials with model parameter estimates.

From Equation (30) it is obvious that polynomial

$$A(z) = z^2 + a_1 z + a_2 \tag{32}$$

which have two different real poles α, β, is included in polynomial $D(z)$ (28). Its parameter estimates are known from process identification. Two possibilities are likely to solve using the Time-delay Toolbox.

Pole assignment with user-defined multiple pole (PAMP) method:

Polynomial (24) has two different real poles α, β and user-defined multiple pole γ. Then polynomial (24) has the form

$$D(z) = (z - \alpha)(z - \beta)(z - \gamma)^2$$

and it is possible to express its individual parameters as:

$$\begin{aligned} d_1 &= -(2\gamma + \alpha + \beta) \\ d_2 &= 2\gamma(\alpha + \beta) + \alpha\beta + \gamma^2 \\ d_3 &= -(2\alpha\beta\gamma + \gamma^2(\alpha + \beta)) \\ d_4 &= \alpha\beta\gamma^2 \end{aligned} \tag{33}$$

Pole assignment with user-defined different real poles (PADP) method:
Polynomial (24) has two different real poles α, β and user-defined real poles γ, δ. Then polynomial (24) has the form

$$D(z) = (z-\alpha)(z-\beta)(z-\gamma)(z-\delta)$$

and it is possible to express its individual parameters as:

$$\begin{aligned} d_1 &= -(\alpha + \beta + \gamma + \delta) \\ d_2 &= \alpha\beta + \gamma\delta + (\alpha + \beta)(\gamma + \delta) \\ d_3 &= -\left[(\alpha + \beta)\gamma\delta + (\gamma + \delta)\alpha\beta\right] \\ d_4 &= \alpha\beta\gamma\delta \end{aligned} \tag{34}$$

5 Toolbox Functions

The Toolbox [11] contains three main scripts (**start_PAMP.m, start_PADP.m** and **start_PID.m**) and other programs functions, models and scripts) that are called by these main scripts. These scripts perform similar sequence of operations:

- definition of the controlled system (transfer function, time delay), sample time and controller parameters,
- off-line identification of the controlled system,
- pole assignment control or PID control of the system.

Toolbox files are summarized in Table 1. The detailed instructions for use of the Toolbox are introduced in the User's Guide [12].
A typical control scheme used is depicted in Fig. 5. This scheme is used for systems with time-delay of two sample steps. Individual blocks of the SIMULINK scheme correspond to blocks of the general control scheme presented in Fig. 1. The green blocks represent the controlled system. Constants bc0, ac2, ac1, and ac0 are parameters of a continuous-time system. Blocks Compensator 1 and Compensator 2 are parts of the Smith Predictor and they correspond to $G_m\left(z^{-1}\right)$ and $G_d\left(z^{-1}\right)$ blocks of Fig. 2 respectively. The control algorithm is encapsulated in Main Pole Assignment Controller which corresponds to $G_c\left(z^{-1}\right)$ Fig. 2 block. The Identification block performs the on-line identification of a controlled system and outputs the estimates of the 2nd order ARX model (a1, b1, a2, b2) parameters.

Table 1. Toolbox Files

File	Description
start_PAMP.m	top-level script for pole assignment control (multiple pole γ)
start_PADP.m	top-level script for pole assignment control (poles γ,δ)
start_PID.m	top-level script for PID control
LSM_2or2td.m	off-line identification
Sm_adapt_pp2i.m	computation of control value in pole assignment control scheme SmP_ad_PA.mdl.
sid.m	on-line identification s-function used by both control schemes (SmP_ad_PA.mdl and SmP_ad_PID.mdl)
Ident_c_LSM.mdl	Simulink scheme used to collect data for off-line identification
SmP_ad_PA.mdl	Simulink control scheme of pole assignment control
SmP_ad_PID.mdl	Simulink control scheme of PID control

6 Experimental results

The experimental identification methods and use of the Time-delay Toolbox is demonstrated on a control of laboratory heat exchanger in simulation conditions. The laboratory heat exchanger [26], [27], {28] is based on the principle of transferring heat from a source through a piping system using a heat transferring media to a heat-consuming appliance. A scheme of the laboratory heat exchanger is depicted in Fig. 6.

The heat transferring fluid (e. g. water) is transported using a continuously controllable DC pump (6) into a flow heater (1) with max. power of 750 W. The temperature of a fluid at the heater output T_1 is measured by a platinum thermometer. Warmed liquid then goes through a 15 meters long insulated coiled pipeline (2) which causes the significant delay in the system. The air-water heat exchanger (3) with two cooling fans (4, 5) represents a heat-consuming appliance. The speed of the first fan can be continuously adjusted, whereas the second one is of on/off type. Input and

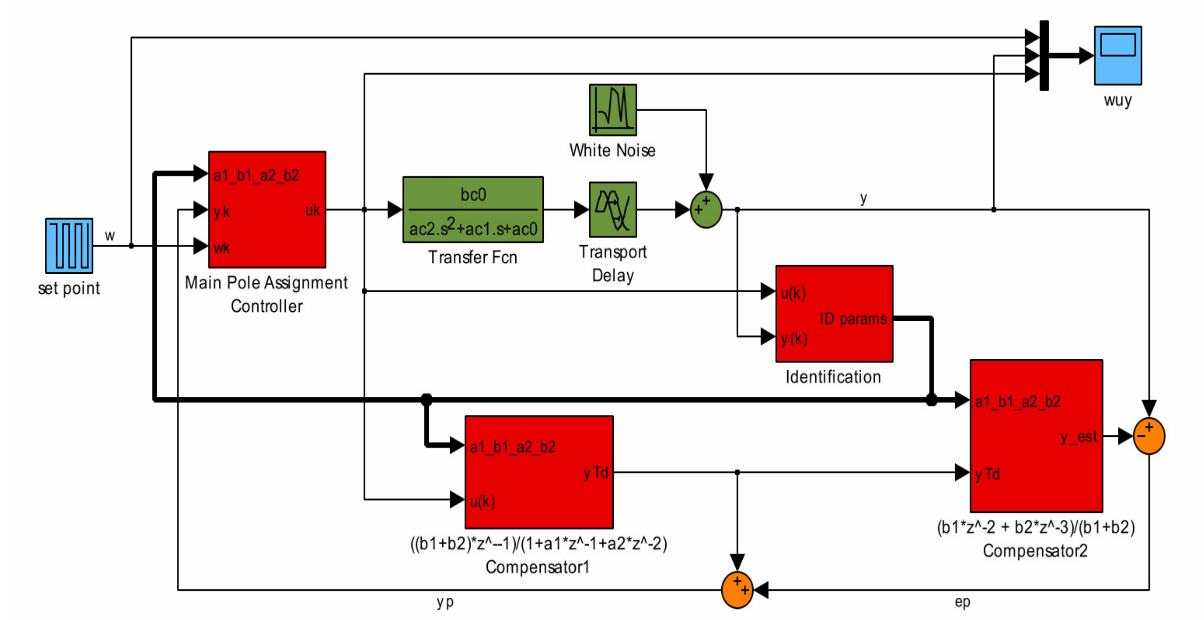

Fig. 5. Simulink control scheme

output temperatures of the cooler are measured again by platinum thermometers as T_2, respective T_3. The laboratory heat exchanger is connected to a standard PC via technological multifunction I/O card. For all monitoring and control functions the MATLAB/SIMULINK environment with Real Time Toolbox.

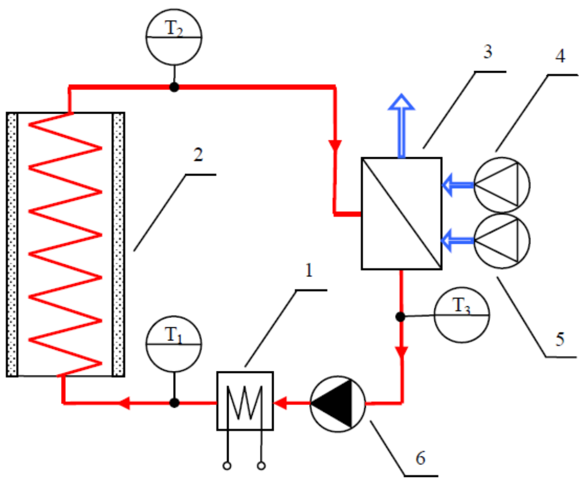

Fig. 6. Scheme of laboratory heat exchanger

6.1 Real-time Identification Experiments

The dynamic model of the laboratory heat exchanger was obtained from processed input (the power of a flow heater P [W]) and output (the temperature of a T_2 [^{o}C]) of the cooler) data. The input signal $u(k)$ was generated using the MATLAB function "idinput" and discrete parameter estimates of model (2) for sampling period T_0 = 100 s and time delay T_d = 200 s were computed using off-line LSM and MATLAB function "fminsearch" (see Paragraph 3.1).

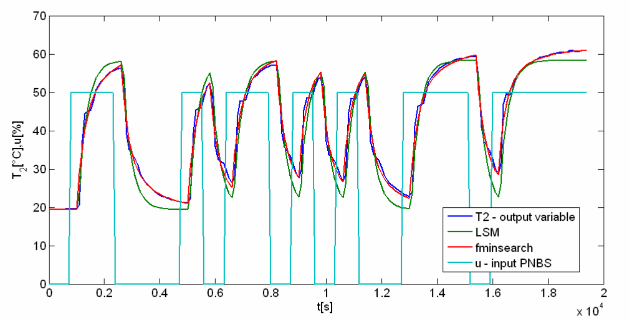

Fig. 7. Identification results: input PNBS

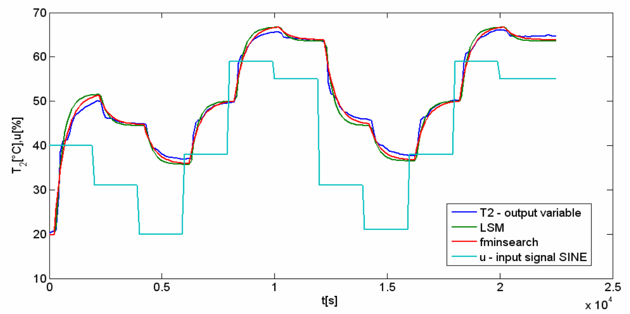

Fig. 8. Identification results: input signal SINE

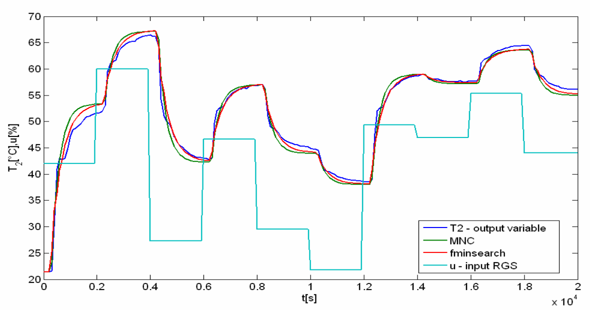

Fig. 9. Identification results: input signal RGS

The graphical variable courses of individual identification experiments are shown in Figs. 7 – 9. The discrete models which were obtained from individual experiments and criterions of identification quality are presented in Tab. 2. From comparison of the real output variable T_2 and the modelled output variables it is obvious that the criterion of identification quality

$$S_y = \frac{1}{N}\sum_{k=1}^{N}\left[y(k)-\hat{y}(k)\right]^2 \tag{35}$$

and the estimate of the static gain

$$\hat{K}_g = \frac{\hat{b}_1+\hat{b}_2}{1+\hat{a}_1+\hat{a}_2} \tag{36}$$

are relatively very good. This fact is confirmed also from courses of unit step responses in Figs. 10 and 11.

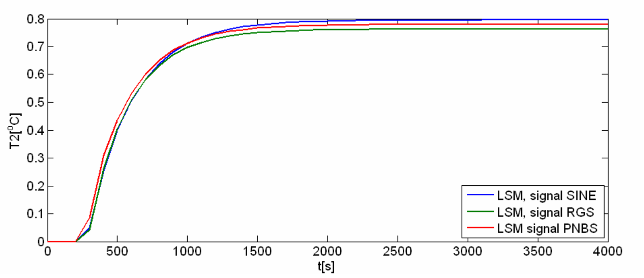

Fig. 10. Comparison of unit step responses, LSM identification

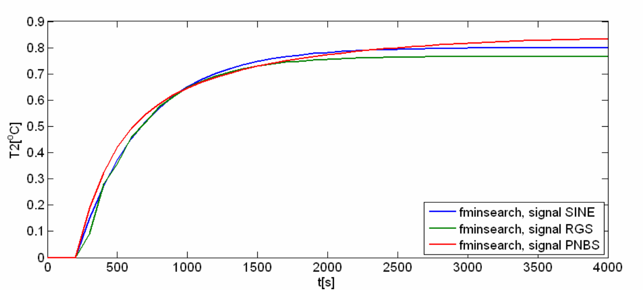

Fig. 11. Comparison of unit step responses, “fminsearch” identification

6.2 Simulation of Closed Control Loops

Simulation is a useful tool for the synthesis of control systems, allowing us not only to create mathematical models of a process but also to design virtual controllers in a computer [29]. The provided mathematical models are close enough to a real object and simulation can be used to verify the dynamic characteristics of control loops when the structure or parameters of the controller have changed. The models of the processes may also be excited by various random noise generators which can simulate the stochastic characteristics of the processes noise signals with similar properties as disturbance signals measured in the machinery can be directly used. The simulation results are valuable for an implementation of a chosen controller (control algorithm) under laboratory and industrial conditions. It must be borne in mind, however, that the practical application of a controller verified by simulation can not be taken as a routine event. Obviously simulation and laboratory conditions can be quite different from those in real plants, and therefore we must verify its practicability with regard to the process dynamics and the required standard of control quality (for example maximum sufferable overshoot, accuracy, settling time, etc.).

For the simulation verification of the proposed control algorithms was chosen the model (43) - see Tab. 2.

$$G(z^{-1}) = \frac{0.1494z^{-1}+0.028z^{-2}}{1-0.6376z^{-1}-0.1407z^{-2}}z^{-2} \tag{37}$$

Its qualitative identification parameters are the best (see Tab. 2). The followed simulation conditions were chosen for all control experiments: sampling period T_0 = 100 s, time-delay T_d = 200 s, as a random disturbance signal was used the white noise with the mean value μ=0 and the variance σ^2=0.01.

Table 2. Comparison of identification methods and input signals

Identification method	Input signal $u(k)$	Model $G(z^{-1})$	Static gain $\hat{K}$ [°C/%]	Criterion quality S_y
LSM	PNBS	$G(z^{-1}) = \frac{0.0862z^{-1}+0.1811z^{-2}}{1-0.4934z^{-1}-0.1636z^{-2}}z^{-2}$ (38)	0.7794	10.2726
	RGS	$G(z^{-1}) = \frac{0.0424z^{-1}+0.1917z^{-2}}{1-0.6445z^{-1}-0.0484z^{-2}}z^{-2}$ (39)	0.7626	1.8761
	SINE	$G(z^{-1}) = \frac{0.0493z^{-1}+0.1691z^{-2}}{1-0.7063z^{-1}-0.0191z^{-2}}z^{-2}$ (40)	0.7849	1.6583
fminsearch	PNBS	$G(z^{-1}) = \frac{0.1885z^{-1}-0.1647z^{-2}}{1-1.586z^{-1}-0.6151z^{-2}}z^{-2}$ (41)	0.8197	6.0292
	RGS	$G(z^{-1}) = \frac{0.0907z^{-1}+0.1708z^{-2}}{1-0.19z^{-1}-0.4689z^{-2}}z^{-2}$ (42)	0.7676	1.3249
	SINE	$G(z^{-1}) = \frac{0.1494z^{-1}+0.028z^{-2}}{1-0.6376z^{-1}-0.1407z^{-2}}z^{-2}$ (43)	0.7901	1.0963

The initial model parameter estimates were chosen using *a priori* information from previous off-line identification experiment

$$\hat{\boldsymbol{\Theta}}^T(0) = [-0.65 \quad -0.15 \quad 0.15 \quad 0.03].$$

6.2.1 PID Smith Predictor

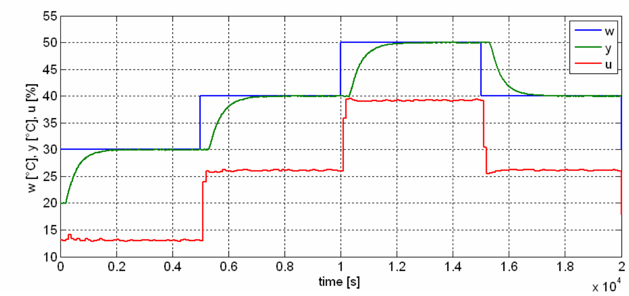

Fig. 11. PID Control of the model (43), $T_m = 400$ s

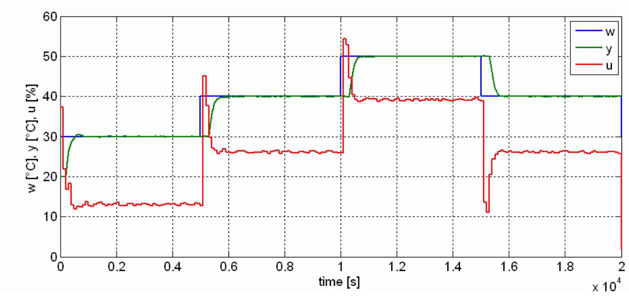

Fig. 12. PID Control of the model (43), $T_m = 100$ s

The simulation verification of the control model (43) using the main PID controller (19) is shown in Figs. 11 and 12. From these Figs. it is obvious that user-defined time constant T_m influences a speed of the step response and an overshoot of the controller output $u(k)$.

6.2.1 Pole Assignment (PA) Smith Predictor

The simulation verification of the control model (43) using the main PA controller (27) is shown in Figs. 13 and 14.

Simulated control responses when parameters of characteristic polynomial (28) were computed using equations (33) are shown in Fig. 13 and 14. Characteristic polynomials and individual poles are:

Fig. 13:

$$D(z) = z^4 - 1.25z^3 + 0.33z^2 + 0.0315z - 0.0135$$

$\alpha = 0.8111$; $\beta = -0.1735$; $\gamma = \delta = 0.3$.

Fig. 14:

$$D(z) = z^4 - 0.6576z^3 - 0.1278z^2 + 0.0028z$$

α=0.8111; β=-0.1735; $\gamma = \delta$=0.01. It is obvious that a small multiple pole leads to $d_4 = 0$.

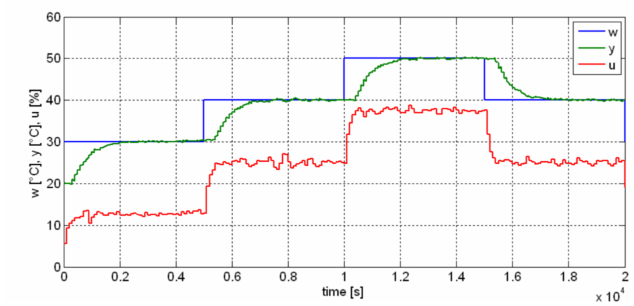

Fig. 13. PAMP Control of the model (43), $\gamma = \delta = 0.3$

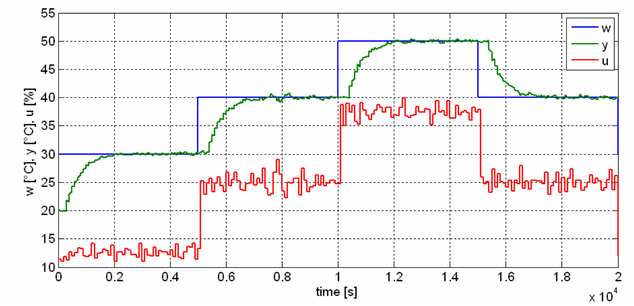

Fig. 14. PAMP Control of the model (43), $\gamma = \delta = 0.01$

Simulated control responses when parameters of characteristic polynomial (28) were computed using equations (34) are shown in Fig. 15. Characteristic polynomial has the form

$$D(z) = z^4 - 0.8876z^3 + 0.0337z^2 + 0.0256z - 0.0021$$

with poles α=0.8111; β=-0.1735; γ=0.1; δ=0.15.
The control quality using main PA controller is very dependent on the pole assignment in the characteristic polynomial (28). The simulation experiments proved that except poles of polynomial $A(z)$ it is suitable to choose next two real positive poles near the centre coordinates. But very small real poles can cause an oscillatory behaviour of the controller output $u(k)$ – see Fig. 14.

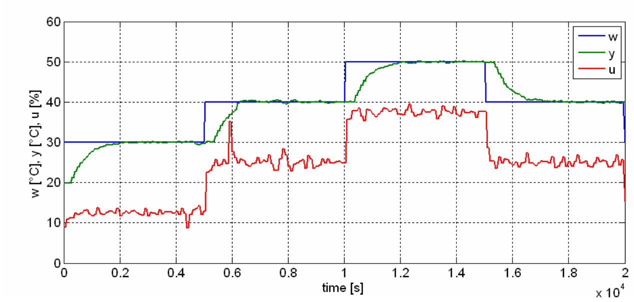

Fig. 15. PADP Control of the model (44), $\gamma = 0.1$, $\delta = 0.15$

7 Conclusion

Two methods for off-line identification with combination of several input exciting signals suitable for time-delay systems were analyzed.

These methods were experimentally verified by an identification of the laboratory heat exchanger in real-time conditions. The best discrete experimental model from view point of the identification quality was used for design of adaptive controllers, which are included in the *MATLAB Toolbox for CAD and Verification of Digital Adaptive Control Time-Delay Systems* [12]. This Toolbox is available free of charge from the Tomas Bata University Zlín Internet site. Both controllers were derived purposely by analytical way (without utilization of numerical methods) to obtain algorithms with easy implementability in industrial practice. The identification part of the adaptive controllers uses the regression ARX model, recursive identification is solved by the Least Squares Method with directional (adaptive) forgetting. Both controllers were successfully verified not only by simulation but also in real-time laboratory conditions for control of the heat exchanger. Very good results were achieved by implementation of the Adaptive Model Predictive Controller in simulation and real-time conditions [30].

Acknowledgment:
This work was supported in part by the Operational Programme Research and Development for Innovations co-funded by the Research Development Fund (ERDF) and national budget of Czech Republic within framework of the Centre of Polymer Systems project (reg. number: CZ. 1.05/2.1.00/03.0111) and the Ministry of Education of the Czech Republic under grant 1M0567.

References:

[1] K. J. Åström and B. Wittenmark, *Adaptive Control*, Massachusetts, Addison-Wesley Publishing Company, 1995.

[2] R. Isermann, K. Lachmann and D. Matko, *Adaptive Control Systems,* Prentice Hall, Inc., Englewood Clils, New Jersey, 1991.

[3] P. E. Welstead and M. B. Zarrop, *Self-Tuning Systems: Control and Signal Processing,* John Wiley, Chichester, 1991.

[4] D. Landau, R. Lozano and M'Saad, *Adaptive Control,* Springer-Verlag, London, 1998.

[5] V. Bobál, J. Böhm J. Fessl and J. Macháček, *Digital Self-tuning Controllers: Algorithms, Implementation and Applications,* Springer-Verlag, London, 2005.

[6] J. E. Normey-Rico, and E. F. Camacho, *Control of Dead-time Processes,* London, Springer-Verlag, 2007.

[7] J. G. Ziegler and N. B. Nichols, Optimum settings for automatic controllers, *Trans. ASME*, Vol. 64, pp. 759-768, 1942.

[8] E. F. Vogel, and T. F. Edgar, A new dead time compensator for digital control, In: *Proc. ISA Annual Conference*, pp. 29-46, Houston, 1980.

[9] Z. J. Palmor, and Y. Halevi, Robustness properties of sampled-data systems with dead time compensators, *Automatica*, Vol. 26, pp. 637-640, 1990.

[10] J. E. Normey-Rico and E. F. Camacho, Dead-time compensators: A unified approach, In: *Proc. of IFAC Workshop on Linear Time-Delay Systems (LDTS'98)*, Grenoble, 1998, pp. 141-146.

[11] V. Bobál, P. Chalupa, P. Dostál and M. Kubalčík, Design and simulation verification of self-tuning Smith Predictors, *International Journal of Mathematics and Computers in Simulation*, Vol. 5, No. 4, pp. 342-351, 2011.

[12] V. Bobál, P. Chalupa, and J. Novák, *Toolbox for CAD and Verfication of Digital Adaptive Control Time-Delay Systems,* Tomas Bata University in Zlín, Faculty of Applied Informatics, 2011. Available from http://nod32.fai.utb.cz/promotion/Software_OBD/Time_Delay_Tool.zip.

[13] C. C Hang, K. W. Lim and B. W. Chong, A dual-rate adaptive digital Smith predictor, *Automatica*, Vol. 20, No. 1, pp. 1-16, 1989.

[14] C. C Hang, H. L. Tong, and K. H. Weng, *Adaptive Control*, North Carolina, Instrument Society of America, 1993.

[15] L. Ljung, *System Identification: Theory for the User,* Englewood Cliffs, NJ, Prentice-Hall, 1987.

[16] V. Bobál, M. Kubalčík, P. Chalupa and P. Dostál, Identification and digital control of higher-order processes, In: *Proc. of the 26th European Control Conference on Modelling and Simulation*, Koblenz, Germany, 2012, pp. 426-433.

[17] L. Ljung, *System Identification Toolbox^{TM}*, User's Guide, The MathWorks, Natick, MA, 2012.

[18] J. C Lagaris, J. A Reeds, M. H. Wright and P. E. Wright, Convergence properties of the Nelder-Mead simplex method in low dimensions, *SIAM Journal of Optimization*, Vol. 9, No. 1, pp. 112-147, 1998.

[19] R. Kulhavý, Restricted exponential forgetting in real time identification, *Automatica*, Vol. 23, pp. 586-600, 1987.

[20] D. B. Dahlin, Designing and tuning digital controllers, *Inst. Control Systems*, Vol. 42, pp.77-73, 1968.
[21] V. Kučera, *Discrete Linear Control: the Polynomial Equation Approach,* Chichester, John Wiley, 1979.
[22] M. Vidyasagar, *Control System Synthesis: A Factorization Approach,* Cambridge M.A., MIT Press, 1985.
[23] V. Kučera, *Analysis and Design of Discrete Linear Control Systems,* London, Prentice Hall, 1991.
[24] V. Kučera, Diophantine equations in control – a survey, *Automatica*, Vol. 29, pp. 1361-1375, 1993.
[25] T. Kawabe, Robust 2DOF PID controller design of time-delay systems based on evolutionary computation, In: *Proceedings of 4th International Conference on ELACTRONICS, CONTROL and SIGNAL PROCESSING*, Miami, Florida, USA, 2005, pp. 144-149.
[26] Dostálek, P., Vašek, V. and J. Dolinay, Design and implementation of portable data acquisition unit in process control and supervision applications, *WSEAS Transactions on Systems and Control*, Vol. 3, No. 9, 2008, pp. 779-788.
[27] L. Pekař, R. Prokop and P. Dostálek, Circuit heating plant model with internal delays, *WSEAS Transactions on Systems and Control*, Vol. 8, No. 9, 2009, pp. 1093-1104.
[28] L. Pekař, R. Prokop and P. Dostálek, An anisochronic model of a laboratory heating system. In: *Proceedings of 13th WSEAS International Conference on Systems*, Rhodes, Greece, 2009, pp. 165-172.
[29] F. Neri, *Agent Based Modeling under Partial and Full Knowledge Learning Settings to Simulate Financial Markets*, AI Communications, IOS Press, 2012 (in print).
[30] V. Bobál, M. Kubalčík, P. Dostál and J. Matějíček, Adaptive predictive control of time-delay systems, In: Nostradamus - Conference on Prediction, Modeling and Analysis of Complex System Dynamics, Ostrava, Czech Republic, 2012 (accepted).

Predictive Control of Higher Order Systems Approximated by Lower Order Time-Delay Models

MAREK KUBALČÍK, VLADIMÍR BOBÁL
Tomas Bata University in Zlín, Faculty of Applied Informatics
Nám. T. G. Masaryka 5555, 7605 Zlín
CZECH REPUBLIC
kubalcik@fai.utb.cz, bobal@fai.utb.cz
http://web.fai.utb.cz/

Abstract: - In technical practice often occur higher order processes when a design of an optimal controller leads to complicated control algorithms. One of possibilities of control of such processes is their approximation by lower-order model with time-delay (dead time). The contribution is focused on a choice of a suitable experimental identification method and a suitable excitation input signals for an estimation of process model parameters with time-delay. One of the possible approaches to control of time-delay processes is application of model-based predictive control (MPC) methods. The further contribution is design of an algorithm for predictive control of high-order processes which are approximated by second-order model of the process with time-delay. The controller was tested and verified by control of several simulation models and a model of a laboratory heat exchanger.

Key-Words: - predictive control, time-delay systems, digital control, higher order systems, simulation

1 Introduction

Some technological processes in industry are characterized by high-order dynamic behaviour or large time constants and time-delays. Time-delay in a process increases the difficulty of controlling it. However using the approximation of higher-order process by lower-order model with time-delay provides simplification of the control algorithms. Let us consider a continuous-time dynamical linear SISO (single input $u(t)$ – single output $y(t)$) system with time-delay T_d. The transfer function of a pure transportation lag is $e^{-T_d s}$ where s is a complex variable. Overall transfer function with time-delay is in the form

$$G_d(s) = G(s)e^{-T_d s} \quad (1)$$

where $G(s)$ is the transfer function without time-delay. Methods and applications of control of time-delay systems are for example in [1], [2], [3].

Processes with time-delay are difficult to control using standard feedback controllers. One of the possible approaches to control processes with time delay is predictive control [4], [5], [6]. The predictive control strategy includes a model of the process in the structure of the controller. The first time-delay compensation algorithm was proposed by [8]. This control algorithm known as the Smith Predictor (SP) contains a dynamic model of the time-delay process and it can be considered as the first model predictive algorithm. An alternative method implemented to analyze heat diffusion system with time–delay, are the integer and fractional order controllers with a Smith Predictor controller [9].

Model Predictive Control (MPC) or only Predictive Control is one of the control methods which have developed considerably over a few past years. Predictive control is essentially based on discrete or sampled models of processes. Computation of appropriate control algorithms is then realized namely in the discrete domain.

The term Model Predictive Control designates a class of control methods which have common particular attributes [10], [11].

- Mathematical model of a system is used for prediction of future systems output.
- The input reference trajectory in future is known.
- A computation of the future control sequence includes minimization of an appropriate objective function (usually quadratic one) with the future trajectories of control increments and control errors.
- Only the first element of the control sequence is applied and the whole procedure of the objective function minimization is repeated in the next sampling period.

The principle of MPC is shown in Fig. 1, where $u(t)$ is the manipulated variable, $y(t)$ is the process output and $w(t)$ is the reference signal, N_1, N_2 and N_u are called minimum, maximum and control horizon. This principle is possible to define as follows:

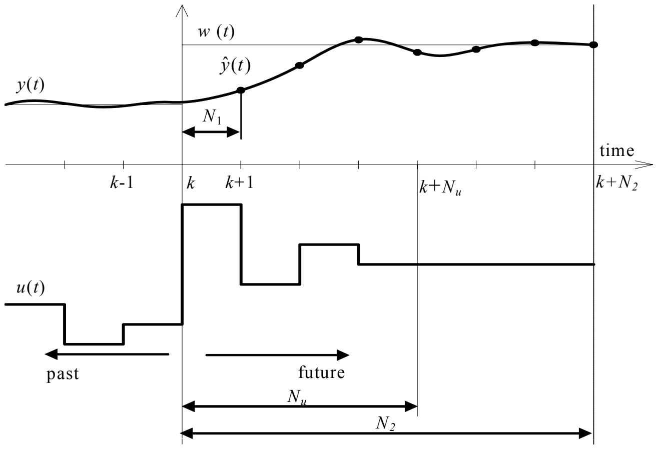

Fig. 1. Principle of MPC

1. The process model is used to predict the future outputs $\hat{y}(t)$ over some horizon *N*. The predictions are calculated based on information up to time *k* and on the future control actions that are to be determined.
2. The future control trajectory is calculated as a solution of an optimisation problem consisting of an objective function and constraints. The cost function comprises future output predictions, future reference trajectory, and future control actions.
3. Although the whole future control trajectory was calculated in the previous step, only first element $u(k)$ is actually applied to the process. At the next sampling time the procedure is repeated. This is known as the *Receding Horizon* concept.

Theoretical research in the area of predictive control has a great impact on the industrial world and there are many applications of predictive control in industry. Its development has been significantly influenced by industrial practice. At present, predictive control with a number of real industrial applications belongs among the most often implemented modern industrial process control approaches. First predictive control algorithms were implemented in industry as an effective tool for control of multivariable industrial processes with constraints more than twenty five years ago. The use of predictive control was limited on control of namely rather slow processes due to the amount of computation required. At present, with the computing power available today, this is not an essential problem. A fairly actual and extensive surveys of industrial applications of predictive control are presented in [12], [13], [14].

High-order processes are largely approximated by the FOTD (first-order-time-delay) model. The aim of the paper is implementation of a predictive controller for control of high-order processes which are approximated by second-order model with time delay of two steps. This model approximates the higher order dynamics more accurately than the first order time delay model whilst design of control algorithms is still quite simple. The designed controller was tested and verified by control of several simulation models and a model of a laboratory heat exchanger.

The paper is organized as follows: section 2 describes identification of time-delay processes; section 3 presents design and implementation of predictive control; section 4 introduces computation of predictor for time-delay systems; section 5 gives the simulation results; section 6 contains experimental results and finally section 7 concludes the paper.

2 Identification of Time-Delay Processes

In this paper, the time-delay model is obtained separately from an off-line identification using the least squares method (LSM). The measured process output $y(k)$ is generally influenced by noise. These nonmeasurable disturbances cause errors $\boldsymbol{e}$ in the determination of model parameters and therefore real output vector is in the form

$$\boldsymbol{y} = \boldsymbol{F\Theta} + \boldsymbol{e} \tag{2}$$

It is possible to obtain the LSM expression for calculation of the vector of the parameter estimates

$$\hat{\boldsymbol{\Theta}} = \left(\boldsymbol{F}^T\boldsymbol{F}\right)^{-1}\boldsymbol{F}^T\boldsymbol{y} \tag{3}$$

The matrix $\boldsymbol{F}$ has dimension (N-n-d, $2n$) and rank 2, the vector $\boldsymbol{y}$ (N-n-d) and the vector of parameter model estimates $\hat{\boldsymbol{\Theta}}\,(2n)$. N is the number of samples of measured input and output data, n is the model order, d is a number of time-delay steps.

Equation (3) serves for calculation of the vector of the parameter estimates $\hat{\boldsymbol{\Theta}}$ using N samples of measured input-output data. The individual vectors and matrices in Equations (2) and (3) have the form

$$\hat{\boldsymbol{\Theta}}^T = \begin{bmatrix} \hat{a}_1 & \hat{a}_2 & \cdots & \hat{a}_n & \hat{b}_1 & \hat{b}_2 & \cdots & \hat{b}_n \end{bmatrix} \tag{4}$$

$$\boldsymbol{y}^T = \begin{bmatrix} y(n+d+1) & y(n+d+2) & \cdots & y(N) \end{bmatrix} \tag{5}$$

$$\boldsymbol{e}^T = \begin{bmatrix} \hat{e}(n+d+1) & \hat{e}(n+d+2) & \cdots & \hat{e}(N) \end{bmatrix} \tag{6}$$

$$\boldsymbol{F} = \begin{bmatrix} -y(n+d) & -y(n+d-1) & \cdots & -y(d+1) & u(n) & u(n-1) & \cdots & u(1) \\ -y(n+d+1) & -y(n+d) & \cdots & -y(d+2) & u(n+1) & u(n) & \cdots & u(2) \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ -y(N-1) & -y(N-2) & \cdots & -y(N-n) & u(N-d-1) & u(N-d-2) & \cdots & u(N-d-n) \end{bmatrix} \tag{7}$$

Most of higher-order industrial processes can be approximated by a model of reduced order with pure time-delay. Let us consider the following second order linear model with a time-delay

$$G_d\left(z^{-1}\right) = \frac{B\left(z^{-1}\right)}{A\left(z^{-1}\right)} z^{-d} = \frac{b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}} z^{-d} \tag{8}$$

The term z^{-d} represents the pure discrete time-delay. The time-delay is equal to dT_0 where T_0 is the sampling period.

Our experience proved that quality of system identification when the higher-order process is identified by the lower-order model is very dependent on the choice of an input excitation signal $u(k)$. The best results were achieved using a Random Gaussian Signal (RGS).

Let us consider that model (8) is the deterministic part of the stochastic process described by the ARX (regression) model

$$\begin{aligned} y(k) = &-a_1 y(k-1) - a_2 y(k-2) + \\ &+ b_1 y(k-1-d) + b_2 y(k-2-d) + e_s(k) \end{aligned} \tag{9}$$

where $e_s(k)$ is the non-measurable random component. The vector of parameter model estimates is computed by solving equation (3)

$$\hat{\boldsymbol{\Theta}}^T(k) = \begin{bmatrix} \hat{a}_1 & \hat{a}_2 & \hat{b}_1 & \hat{b}_2 \end{bmatrix} \tag{10}$$

and is used for computation of the prediction output.

$$\begin{aligned} \hat{y}(k) = &-\hat{a}_1 y(k-1) - \hat{a}_2 y(k-2) + \\ &\hat{b}_1 u(k-1-d) + \hat{b}_2 u(k-2-d) \end{aligned} \tag{11}$$

The quality of identification can be considered according to error, i.e. the difference between the measured and modeled value of the systems output

$$\hat{e}(k) = y(k) - \hat{y}(k) \tag{12}$$

In this paper, a suitable choice of the number of time-delay steps was performed according to the error. The LSM algorithm (3) – (7) is computed for several numbers of time-delays steps and a suitable time-delay is chosen according to quality of identification based on the prediction error (12).

2.1 Stable Process

Let us consider the following stable fifth order linear system

$$G_A(s) = \frac{2}{(s+1)^5} = \frac{2}{s^5 + 5s^4 + 10s^3 + 10s^2 + 5s + 1} \tag{13}$$

The system (13) was identified by the discrete model (11) using off-line LSM (3) – (6) for different numbers of time-delay steps. As the input signal was used the Random Gaussian Signal (RGS). A criterion of the identification quality is based on sum of squares of error

$$J_{\hat{e}^2}(d) = \sum_{k=1}^{N} \hat{e}^2(k) \tag{14}$$

This criterion evaluates accuracy of the identification process. From Fig 2. , it is obvious that value of the criterion (14) decreases with increasing number of time-delay steps d. This is caused by the fact that the increasing of the number of time-delay steps improves estimation of the static gain

$$\hat{K}_g = \frac{\hat{b}_1 + \hat{b}_2}{1 + \hat{a}_1 + \hat{a}_2} \tag{15}$$

The difference between estimates of the static gain $\hat{K}_g$ of the discrete model (8) and the continuous-time model (13) plays an important role for the quality of identification because the identification time was relatively long (300 s) with regard to the response time (about 15 s).

The system was identified by the following model

$$G_A\left(z^{-1}\right)=\frac{-0.0424z^{-1}+0.0296z^{-2}}{1-1.6836z^{-1}+0.7199z^{-2}}z^{-d} \tag{16}$$

Comparisons of step responses of continuous-time (13) and discrete models (16) with sampling period $T_0 = 0.5$ s for different numbers of time-delay steps d are shown in Figs. 3-5, where yc is the step response of the model (13) and yd are step responses of the discrete model (16) for individual numbers of time-delay steps d.

From Figs. 2-5 it results that a suitable model for the design of the predictive controller is the model (13) with d = 2. Its structure is simple and it relatively well approximates the dynamic behaviour of the continuous-time model (16).

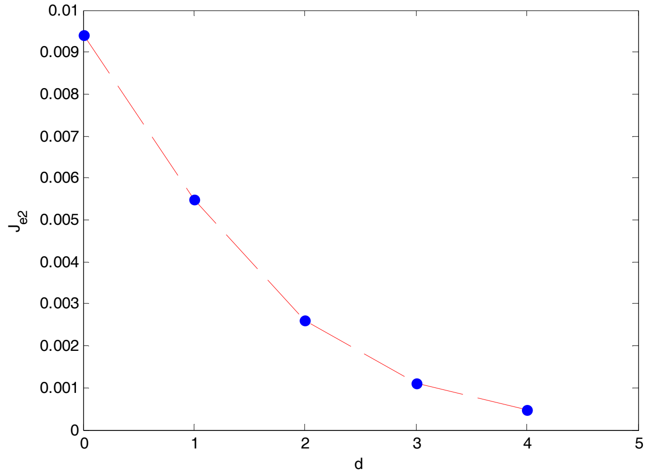

Fig. 2. Criterion of quality identification for $d \in [0,5]$

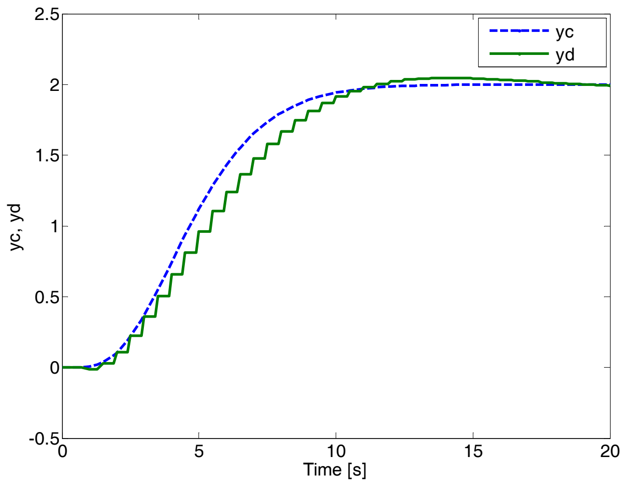

Fig. 3. Comparison of step responses yc, yd for d =0 (process (13))

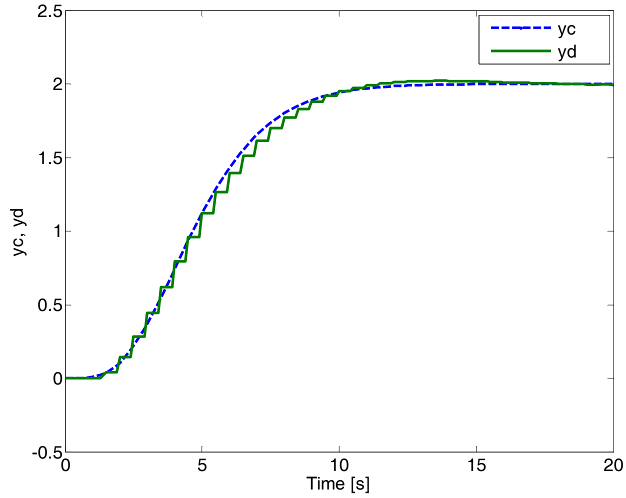

Fig. 4. Comparison of step responses yc, yd for d =2 (process (13))

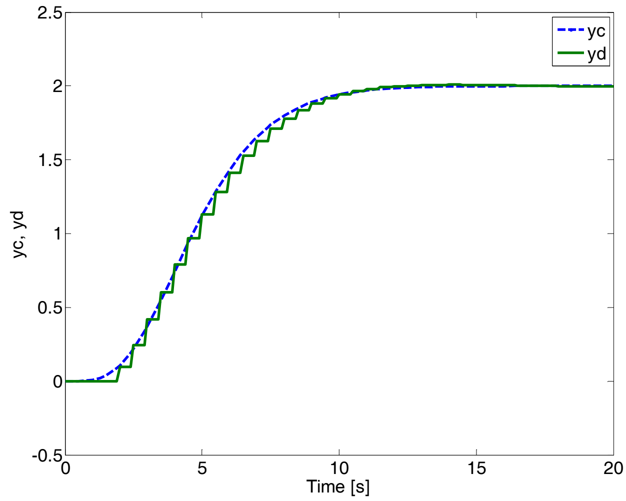

Fig. 5. Comparison of step responses yc, yd for d =3 (process (13))

2.2 Stable Non-Minimum Phase Process

Let us consider the following fifth-order linear system with non-minimum phase

$$G_B(s)=\frac{2(1-5s)}{s^5+5s^4+10s^3+10s^2+5s+1} \tag{17}$$

The process (17) was identified by the model (8) with a time-delay d=2 and sampling period $T_0 = 0{,}5s$. The discrete model which was obtained from the model (17) by Z-transform is in the following form

$$G_B\left(z^{-1}\right)=\frac{-0.7723z^{-1}+0.8514z^{-2}}{1-1.6521z^{-1}+0.8514z^{-2}}z^{-2} \tag{18}$$

The comparison of the step responses of the continuous-time model (17) and the discrete model (18) is shown in Fig. 6.

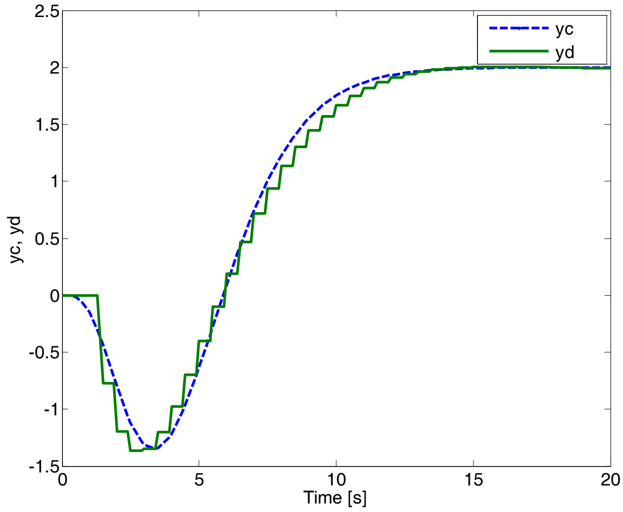

Fig. 6. Comparison of step responses yc, yd for d =2 (process (17))

3 Implementation of Predictive Control

In this Section, GPC (General predictive control) will be briefly described. The GPC method is in principle applicable to both SISO and MIMO processes and is based on input-output models. The standard cost function used in GPC contains quadratic terms of (possible filtered) control error and control increments on a finite horizon into the future

$$J = \sum_{i=N_1}^{N_2} \delta(i)\left[\hat{y}(k+i) - w(k+i)\right]^2 + \sum_{i=1}^{N_u}\left[\lambda(i)\Delta u(k+i-1)\right]^2 \quad (19)$$

where $\hat{y}(k+i)$ is the process output of i steps in the future predicted on the base of information available upon the time k, $w(k+1)$ is the sequence of the reference signal and $\Delta u(k+i-1)$ is the sequence of the future control increments that have to be calculated.

Implicit constraints on Δu are placed between N_u and N_2 as

$$\Delta u(k+i-1) = 0, \qquad N_u < i \le N_2 \quad (20)$$

The parameters $\delta(i)$ and $\lambda(i)$ are sequences which affect future behaviour of the controlled process. Generally, they are chosen in the form of constants or exponential weights, according to our requirements on control.

3.1 Calculation of the Optimal Control

The objective of predictive control is a computation of a sequence of future increments of the manipulated variable $[\Delta u(k), \Delta u(k+1), \ldots]$ so that the criterion (19) was minimized. For further computation, it is necessary to transform the criterion (19) to a matrix form.

The output of the model (predictor) is computed as the sum of the free response $\boldsymbol{y}_0$ and the forced response of the model $\boldsymbol{y}_n$

$$\hat{\boldsymbol{y}} = \boldsymbol{y}_n + \boldsymbol{y}_0 \quad (21)$$

It is possible to compute the forced response as the multiplication of the matrix $\boldsymbol{G}$ (Jacobian of the model) and the vector of future control increments $\Delta\boldsymbol{u}$, which is generally a priori unknown

$$\boldsymbol{y}_n = \boldsymbol{G}\Delta\boldsymbol{u} \quad (22)$$

where

$$\boldsymbol{G} = \begin{bmatrix} g_1 & 0 & 0 & \cdots & 0 \\ g_2 & g_1 & 0 & \cdots & 0 \\ g_3 & g_2 & g_1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ g_{N_2} & g_{N_2-1} & g_{N_2-2} & \cdots & g_{N_2-N_u+1} \end{bmatrix} \quad (23)$$

is a matrix containing values of the step sequence.

It follows from equations (21) and (22) that the predictor in a vector form is given by

$$\hat{\boldsymbol{y}} = \boldsymbol{G}\Delta\boldsymbol{u} + \boldsymbol{y}_0 \quad (24)$$

and the cost function (19) can be modified to the form below

$$\begin{aligned} J &= (\hat{\boldsymbol{y}} - \boldsymbol{w})^T(\hat{\boldsymbol{y}} - \boldsymbol{w}) + \lambda\Delta\boldsymbol{u}^T\Delta\boldsymbol{u} = \\ &= (\boldsymbol{G}\Delta\boldsymbol{u} + \boldsymbol{y}_0 - \boldsymbol{w})^T(\boldsymbol{G}\Delta\boldsymbol{u} + \boldsymbol{y}_0 - \boldsymbol{w}) + \lambda\Delta\boldsymbol{u}^T\Delta\boldsymbol{u} \end{aligned} \quad (25)$$

where $\boldsymbol{w}$ is the vector of future reference trajectory.

Minimisation of the cost function (25) now becomes a direct problem of linear algebra. The solution in an unconstrained case can be found by setting partial derivative of J with respect to $\Delta\boldsymbol{u}$ to zero and yields

$$\Delta\boldsymbol{u} = -\left(\boldsymbol{G}^T\boldsymbol{G} + \lambda\boldsymbol{I}\right)^{-1}\boldsymbol{G}^T(\boldsymbol{y}_0 - \boldsymbol{w}) \quad (26)$$

where the gradient $\boldsymbol{g}$ and Hessian $\boldsymbol{H}$ are defined as

$$\boldsymbol{g}^T = \boldsymbol{G}^T(\boldsymbol{y}_0 - \boldsymbol{w}) \quad (27)$$

$$\boldsymbol{H} = \boldsymbol{G}^T\boldsymbol{G} + \lambda\boldsymbol{I} \quad (28)$$

Equation (26) gives the whole trajectory of the future control increments and such is an open-loop strategy. To close the loop, only the first element $\boldsymbol{u}$, e. g. $\Delta u(k)$ is applied to the system and the whole algorithm is recomputed at time k+1. This strategy is called the *Receding Horizon Principle* and is one of the key issues in the MBPC concept.

If the first row of the matrix $\left(\boldsymbol{G}^T\boldsymbol{G}+\lambda\boldsymbol{I}\right)^{-1}\boldsymbol{G}^T$ is denoted as $\boldsymbol{K}$ then the actual control increment can be calculated as

$$\Delta u(k)=\boldsymbol{K}\left(\boldsymbol{w}-\boldsymbol{y}_0\right) \tag{29}$$

4 Computation of Predictor

An important task is computation of predictions for arbitrary prediction and control horizons. Dynamics of most of processes requires horizons of length where it is not possible to compute predictions in a simple straightforward way. Recursive expressions for computation of the free response and the matrix $\boldsymbol{G}$ in each sampling period had to be derived. There are several different ways of deriving the prediction equations for transfer function models. Some papers make use of Diophantine equations to form the prediction equations (e.g. [14]). In [10] matrix methods are used to compute predictions. We derived a method for recursive computation of both the free response and the matrix of the dynamics [15].

Computation of the predictor for the time-delay system can be obtained by modification of the predictor for the corresponding system without a time-delay. At first we will consider the second order system without time-delay and then we will modify the computation of predictions for the time-delay system.

4.1 Second Order System without Time-Delay

The model is described by the transfer function

$$G\left(z^{-1}\right)=\frac{b_1z^{-1}+b_2z^{-2}}{1+a_1z^{-1}+a_2z^{-2}}=\frac{B\left(z^{-1}\right)}{A\left(z^{-1}\right)} \tag{30}$$

$$A\left(z^{-1}\right)=1+a_1z^{-1}+a_2z^{-2};\; B\left(z^{-1}\right)=b_1z^{-1}+b_2z^{-2} \tag{31}$$

The model can be also written in the form

$$A\left(z^{-1}\right)y(k)=B\left(z^{-1}\right)u(k) \tag{32}$$

A widely used model in general model predictive control is the CARIMA model which we can obtain from the nominal model (32) by adding a disturbance model

$$A\left(z^{-1}\right)y(k)=B\left(z^{-1}\right)u(k)+\frac{C\left(z^{-1}\right)}{\Delta}n_c(k) \tag{33}$$

where $n_c(k)$ is a non-measurable random disturbance that is assumed to have zero mean value and constant covariance and the operator delta is $1-z^{-1}$. Inverted delta is then an integrator.

The polynomial $C\left(z^{-1}\right)$ will be further considered as $C\left(z^{-1}\right)=1$. The CARIMA description of the system is then in the form

$$\Delta A\left(z^{-1}\right)y(k)=B\left(z^{-1}\right)\Delta u(k-1)+n_c(k) \tag{34}$$

The difference equation of the CARIMA model without the unknown term $n_c(k)$ can be expressed as:

$$\begin{aligned}y(k)&=\left(1-a_1\right)y(k-1)+\left(a_1-a_2\right)y(k-2)+a_2y(k-3)+\\&+b_1\Delta u(k-1)+b_2\Delta u(k-2)\end{aligned} \tag{35}$$

It was necessary to compute three step ahead predictions in straightforward way by establishing of lower predictions to higher predictions. The model order defines that computation of one step ahead prediction is based on three past values of the system output. The three step ahead predictions are as follows

$$\begin{aligned}\hat{y}(k+1)&=\left(1-a_1\right)y(k)+\left(a_1-a_2\right)y(k-1)+a_2y(k-2)+\\&+b_1\Delta u(k)+b_2\Delta u(k-1)\\\hat{y}(k+2)&=\left(1-a_1\right)y(k+1)+\left(a_1-a_2\right)y(k)+a_2y(k-1)+\\&+b_1\Delta u(k+1)+b_2\Delta u(k)\\\hat{y}(k+3)&=\left(1-a_1\right)y(k+2)+\left(a_1-a_2\right)y(k+1)+a_2y(k)+\\&+b_1\Delta u(k+2)+b_2\Delta u(k+1)\end{aligned} \tag{36}$$

The predictions after modification can be written in a matrix form

$$\begin{aligned}&\begin{bmatrix}\hat{y}(k+1)\\\hat{y}(k+2)\\\hat{y}(k+3)\end{bmatrix}=\begin{bmatrix}g_1&0\\g_2&g_1\\g_3&g_2\end{bmatrix}\begin{bmatrix}\Delta u(k)\\\Delta u(k+1)\end{bmatrix}+\begin{bmatrix}p_{11}&p_{12}&p_{13}&p_{14}\\p_{21}&p_{22}&p_{23}&p_{24}\\p_{31}&p_{32}&p_{33}&p_{34}\end{bmatrix}\begin{bmatrix}y(k)\\y(k-1)\\y(k-2)\\\Delta u(k-1)\end{bmatrix}=\\&=\begin{bmatrix}b_1&0\\b_1\left(1-a_1\right)+b_2&b_1\\\left(a_1-a_2\right)b_1+\left(1-a_1\right)^2b_1+\left(1-a_1\right)b_2&b_1\left(1-a_1\right)+b_2\end{bmatrix}\begin{bmatrix}\Delta u(k)\\\Delta u(k+1)\end{bmatrix}+\\&+\begin{bmatrix}\left(1-a_1\right)&\left(a_1-a_2\right)&a_2&b_2\\\left(1-a_1\right)^2+\left(a_1-a_2\right)&\left(1-a_1\right)\left(a_1-a_2\right)+a_2&a_2\left(1-a_1\right)&b_2\left(1-a_1\right)\\\left(1-a_1\right)^3+2\left(1-a_1\right)\left(a_1-a_2\right)+a_2&\left(1-a_1\right)^2\left(a_1-a_2\right)+a_2\left(1-a_1\right)+\left(a_1-a_2\right)^2&a_2\left(1-a_1\right)^2+\left(a_1-a_2\right)a_2&b_2\left(1-a_1\right)^2+\left(a_1-a_2\right)b_2\end{bmatrix}\begin{bmatrix}y(k)\\y(k-1)\\y(k-2)\\\Delta u(k-1)\end{bmatrix}\end{aligned} \tag{37}$$

It is possible to divide computation of the predictions to recursion of the free response and recursion of the matrix of the dynamics. Based on the three previous predictions it is repeatedly computed the next row of the free response matrix in the following way:

$$\begin{aligned} p_{41} &= (1-a_1)p_{31} + (a_1 - a_2)p_{21} + a_2 p_{11} \\ p_{42} &= (1-a_1)p_{32} + (a_1 - a_2)p_{22} + a_2 p_{12} \\ p_{43} &= (1-a_1)p_{33} + (a_1 - a_2)p_{23} + a_2 p_{13} \\ p_{44} &= (1-a_1)p_{34} + (a_1 - a_2)p_{24} + a_2 p_{14} \end{aligned} \tag{38}$$

The first row of the matrix is omitted in the next step and further prediction is computed based on the three last rows including the one computed in the previous step. This procedure is cyclically repeated. It is possible to compute an arbitrary number of rows of the matrix.

The recursion of the dynamics matrix is similar. The next element of the first column is repeatedly computed in the same way as in the previous case and the remaining columns are shifted to form a lower triangular matrix in the way which is obvious from the equation (37). This procedure is performed repeatedly until the prediction horizon is achieved. If the control horizon is lower than the prediction horizon a number of columns in the matrix is reduced. Computation of the new element is performed as follows:

$$g_4 = (1-a_1)g_3 + (a_1 - a_2)g_2 + a_2 g_1 \tag{39}$$

4.2 Second Order System with Time-Delay

The nominal model with two steps time-delay is considered as

$$G(z^{-1}) = \frac{B(z^{-1})}{A(z^{-1})} z^{-2} = \frac{b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}} z^{-2} \tag{40}$$

The CARIMA model for time-delay system takes the form

$$\Delta A(z^{-1}) y(k) = z^{-d} B(z^{-1}) \Delta u(k-1) + n_c(k) \tag{41}$$

where d is the dead time. In our case d is equal to 2. In order to compute the control action it is necessary to determine the predictions from d+1 (2+1 in our case) to $d+N_2$ ($2+N_2$).

The predictor (37) is then modified to

$$\begin{aligned} \begin{bmatrix} \hat{y}(k+3) \\ \hat{y}(k+4) \\ \hat{y}(k+5) \end{bmatrix} &= \begin{bmatrix} p_{31} & p_{32} & p_{33} \\ p_{41} & p_{42} & p_{43} \\ p_{51} & p_{52} & p_{53} \end{bmatrix} \begin{bmatrix} y(k) \\ y(k-1) \\ y(k-2) \end{bmatrix} + \\ &+ \begin{bmatrix} g_1 & 0 \\ g_2 & g_1 \\ g_3 & g_2 \end{bmatrix} \begin{bmatrix} \Delta u(k) \\ \Delta u(k+1) \end{bmatrix} + \\ &+ \begin{bmatrix} g_2 & g_3 & p_{34} \\ g_3 & g_4 & p_{44} \\ g_4 & g_5 & p_{54} \end{bmatrix} \begin{bmatrix} \Delta u(k-1) \\ \Delta u(k-2) \\ \Delta u(k-3) \end{bmatrix} \end{aligned} \tag{42}$$

Recursive computation of the matrices is analogical to the recursive computation described in the previous section.

The predictor can be also modified for arbitrary number of steps of time delay

$$\begin{aligned} \begin{bmatrix} \hat{y}(k+1+d) \\ \hat{y}(k+2+d) \\ \hat{y}(k+3+d) \end{bmatrix} &= \begin{bmatrix} p_{(1+d)1} & p_{(1+d)2} & p_{(1+d)3} \\ p_{(2+d)1} & p_{(2+d)2} & p_{(2+d)3} \\ p_{(3+d)1} & p_{(3+d)2} & p_{(3+d)3} \end{bmatrix} \begin{bmatrix} y(k) \\ y(k-1) \\ y(k-2) \end{bmatrix} + \\ &+ \begin{bmatrix} g_1 & 0 \\ g_2 & g_1 \\ g_3 & g_2 \end{bmatrix} \begin{bmatrix} \Delta u(k) \\ \Delta u(k+1) \end{bmatrix} + \\ &+ \begin{bmatrix} g_{1+d-1} & g_{2+d-1} & p_{(1+d)4} \\ g_{2+d-1} & g_{3+d-1} & p_{(2+d)4} \\ g_{3+d-1} & g_{4+d-1} & p_{(3+d)4} \end{bmatrix} \begin{bmatrix} \Delta u(k-1) \\ \Delta u(k-2) \\ \Delta u(k-3) \end{bmatrix} \end{aligned} \tag{43}$$

5 Simulation Examples

For simulation examples were chosen the systems introduced in the sections 2.1 and 2.2. Control responses are in the Figs. 7-10.

The tuning parameters that are lengths of the prediction and control horizons and the weighting coefficient λ were tuned experimentally. There is a lack of clear theory relating to the closed loop behavior to design parameters. The length of the prediction horizon, which should cover the important part of the step response, was in both cases set to N = 40. The length of the control horizon was also set to N_u = 40. The coefficient λ was taken as equal to 0,5.

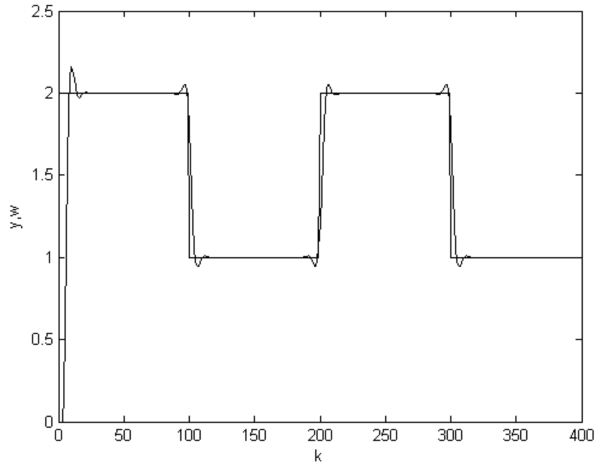

Fig. 7. Control of the model (16)

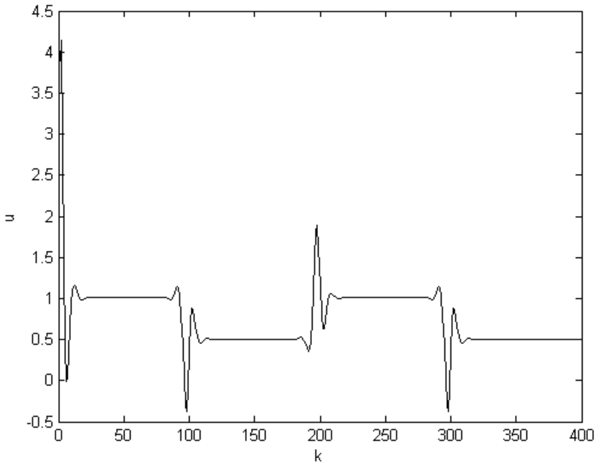

Fig. 8. Control of the model (16) – manipulated variable

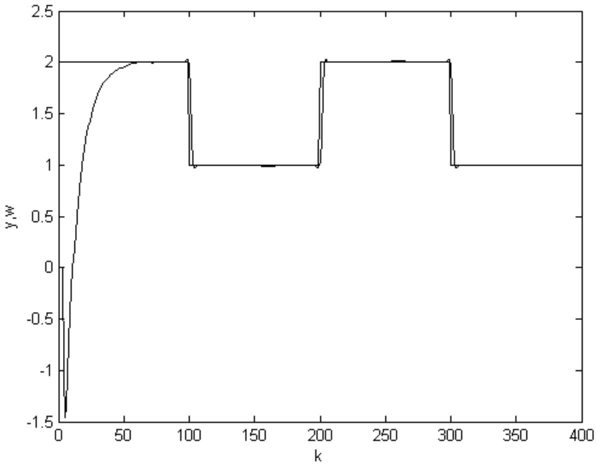

Fig. 9. Control of the model (18)

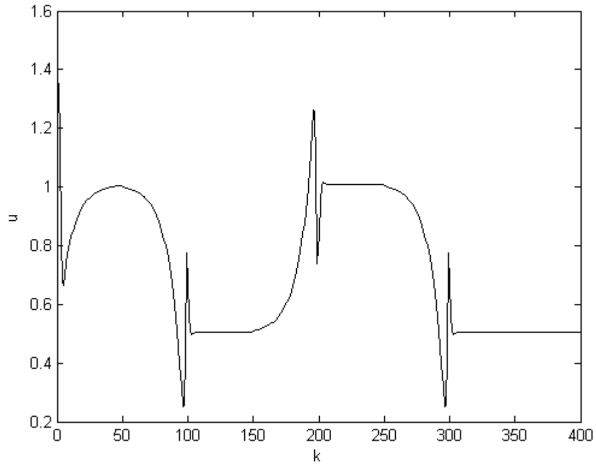

Fig. 10. Control of the model (18) – manipulated variable

Asymptotic tracking of the reference signal was achieved in all cases. The control of non-minimum phase system was rather sensitive to tuning parameters. Experimental tuning of the controller was more complicated in this case.

6 Experimental Example

The use of the predictive control algorithm is also demonstrated on a control of laboratory heat exchanger in simulation conditions. The laboratory heat exchanger [16] is based on the principle of transferring heat from a source through a piping system using a heat transferring media to a heat-consuming appliance.

6.1 Laboratory Heat Exchanger Description

A scheme of the laboratory heat exchanger is depicted in Fig. 11 The heat transferring fluid (e. g. water) is transported using a continuously controllable DC pump (6) into a flow heater (1) with max. power of 750 W. The temperature of a fluid at the heater output T_1 is measured by a platinum thermometer. Warmed liquid then goes through a 15 meters long insulated coiled pipeline (2) which causes the significant delay in the system. The air-water heat exchanger (3) with two cooling fans (4, 5) represents a heat-consuming appliance. The speed of the first fan can be continuously adjusted, whereas the second one is of on/off type. Input and output temperatures of the cooler are measured again by platinum thermometers as T_2, resp. T_3. The laboratory heat exchanger is connected to a standard PC via technological multifunction I/O card. For all monitoring and control functions the MATLAB/SIMULINK environment with Real Time Toolbox was used.

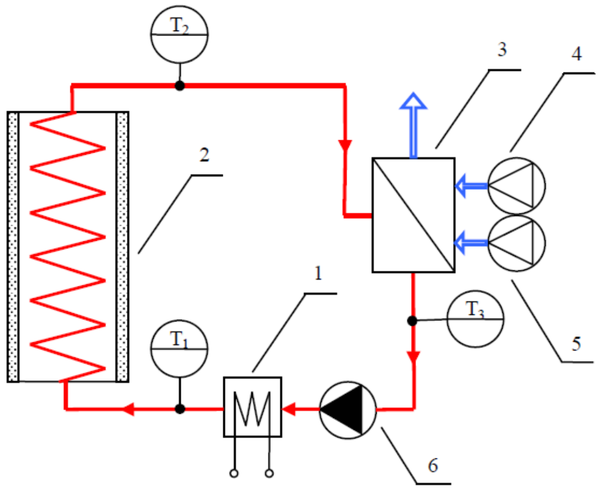

Fig. 11. Scheme of laboratory heat exchanger

6.2 Identification of Laboratory Heat Exchanger

The dynamic model of the laboratory heat exchanger was obtained from processed input (the power of a flow heater P [W]) and output (the temperature of a T_2 [deg] of the cooler) data. As the input signal was again used the Random Gaussian Signal. Following discrete transfer function for sampling period $T_0 = 100$ s was identified

$$G(z^{-1}) = \frac{0.1494z^{-1} + 0.028z^{-2}}{1 - 0.6376z^{-1} - 0.1407z^{-2}} z^{-2} \tag{44}$$

Control responses are in Figs. 12-13.

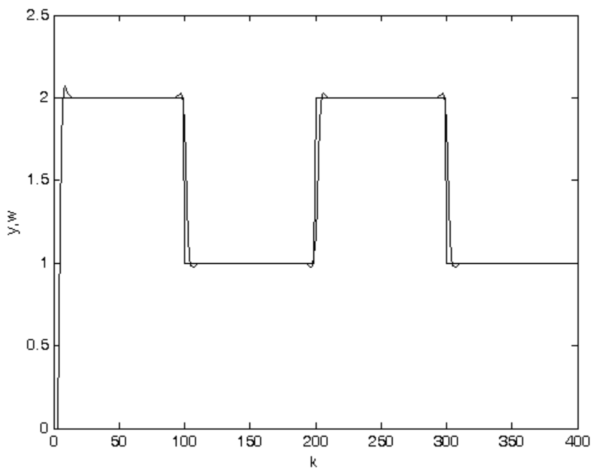

Fig. 12. Control of the model of the laboratory heat exchanger

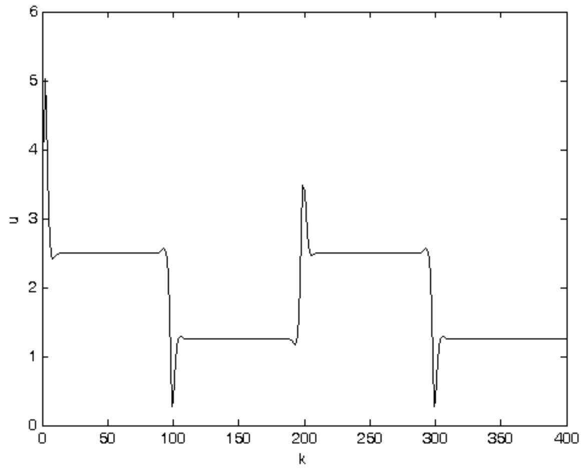

Fig. 13. Control of the model of the laboratory heat exchanger-manipulated variable

7 Conclusion

The algorithm for control of the higher-order processes based on model predictive control was designed. The higher-order process was approximated by the second-order model with time delay. The predictive controller is based on the recursive computation of predictions by direct use of the CARIMA model. The computation of predictions was extended for the time-delay system. The control of two modifications of the higher-order processes (stable and non-minimum phase) were verified by simulation. The laboratory heat exchanger system was identified by an experimental on-line method and its discrete model was also used for verification of the proposed predictive controller. The simulation verification provided good control results. The simulation experiments confirmed that predictive approach is able to cope with the given control problem.

Acknowledgements:
This article was created with support of Operational Programme Research and Development for Innovations co-funded by the European Regional Development Fund (ERDF), national budget of Czech Republic within the framework of the Centre of Polymer Systems project (reg. number: CZ.1.05/2.1.00/03.0111).

References:

[1] T. Hashimoto, T. Amemiya, Stabilization of Linear Time-Varying Uncertain Delay Systems with Double Triangular Configuration, *WSEAS Transactions on Systems and Control*, Vol. 4, Issue 9, September 2009, pp. 465-475, ISSN:1991-8763.

[2] Bahador Makki, Babarak Makki, Control Design for Uncertain Singularly Perturbed Systems with Discrete Time-Delay, *WSEAS Transactions on Systems and Control*, Vol. 6, Issue 12, December 2011, pp. 456-465, ISSN:1991-8763.

[3] F. Neri, Agent Based Modeling under Partial and Full Knowledge Learning Settings to Simulate Financial Markets, AI Communications, IOS Press, 2012.

[4] E. F. Camacho, C. Bordons, *Model Predictive Control,* Springer-Verlag, London, 2004.

[5] J. Mikleš & M. Fikar, *Process Modelling, Optimisation and Control.* (Berlin: Springer-Verlag, 2008).

[6] A. M. Yousef, Model Predictive Control Approach Based Load Frequency Controller, *WSEAS Transactions on Systems and Control*, Vol. 6, Issue 7, July 2011, pp. 265-275, ISSN:1991-8763.

[7] O. J. Smith, Closed control of loops. *Chem. Eng. Progress*, vol. 53, 1957, pp. 217-219.

[8] R. R. Bitmead, M. Gevers, V. Hertz, *Adaptive Optimal Control. The Thinking Man's GPC,* Prentice Hall, Englewood Cliffs, New Jersey, 1990.

[9] Jesus, S. Isabel, Machado, J.A. Tenreiro, Fractional Control of Heat Diffusion Systems, Nonlinear Dynamics, Springer, Vol. 54, Issue 3, pp. 263-282, 2008.

[10] J. A. Rossiter, *Model Based Predictive Control: a Practical Approach* (CRC Press, 2003).

[11] M. Morari, J. H. Lee, Model predictive control: past, present and future. *Computers and Chemical Engineering,* 23, 1999, 667-682.

[12] S. J. Quin & T. A. Bandgwell, An overview of nonlinear model predictive control applications. *Nonlinear Model Predictive Control* (F. Allgöwer & A. Zheng, Ed.), (Basel – Boston – Berlin: Birkhäuser Verlag, 2000), 369-392.

[13] S. J. Quin & T. A. Bandgwell, A survey of industrial model predictive control technology. *Control Engineering Practice*, *11*(7), 2003, 733-764.

[14] W. H. Kwon, H. Choj, D.G. Byun, S. Noh, Recursive solution of generalized predictive control and its equivalence to receding horizon tracking control. *Automatica*, 28(6), 1992, 1235–1238.

[15] M. Kubalčík, V. Bobál, Techniques for Predictor Design in Multivariable Predictive Control, *WSEAS Transactions on Systems and Control*, Vol. 6, Issue 9, September 2011, pp.349-360, ISSN:1991-8763.

[16] L. Pekař, R. Prokop, P. Dostálek, Circuit heating Plant Model with Internal Delays, *WSEAS Transactions on Systems*, Vol. 8, Issue 9, September 2009, pp. 1093-1104, ISSN:1109-2777.