ROBUST CONSTRAINED PID CONTROL

Mikuláš Huba

Slovak University of Technology, Faculty of Electrical Engineering and Information Technology
Ilkovičova 3, 812 19 Bratislava, Slovak Republic
Tel.: +421 2 60291111 Fax: +421 2 60291111
e-mail: mikulas.huba@stuba.sk
MI/PRT, FernUniversität in Hagen, Universitätsstr. 27, D-58084 Hagen, Germany
e-mail: mikulas.huba@fernuni-hagen.de

Abstract: The paper describes some new aspects of the course on Robust Constrained PID Control that is being developed and tought at the Faculty of Electrical Engineering and Information Technology of the STU in Bratislava for more than two decades. New important notions enabling more effective treatment of the carried out analysis and design are introduced and computer tools enabling to achieve homogenous dynamics in the case of parameter uncertainty and control constraints are discussed.

Keywords: PID control, robust control, constrained control, performance portrait method

1 INTRODUCTION

Already several times we have informed participants of the conferences organized by Slovak Society of Cybernetics and Informatics about progress achieved in developing course on Constrained PID Control (Huba & Bistak, 2001; Huba, 2002; 2005). This development that started sometimes around 1980 with the PhD. thesis (Huba 1981a,b) was step by step documented by several textbookes and chapters in textbooks (Huba, 1984a,b; 1992a; 2003a,b; 2006a,b; 2007; Huba and Zakova, 1999) and by many journal and conference papers.

Why do we need such a course? The answer is quit simple: the existing books on PID control bring a lot of interesting content, but still they do not guarantee a consistent and reliable design of simple control tasks. Already when constraining such tasks to Single Input Single Output (SISO) systems, you can meet many problems related e.g. to the control constraints, long time delays, robustness aspects, etc. that are not yet sufficiently covered by a satisfactory and easy to use theory. The aim of our endeavour is to integrate the well known and approved control principles, methods and techniques, if necessary to develope new ones and to interprete all within one conssistent approach that would respond to all main challanges of the practically oriented control design. Despite to the tremendous extent of available literature (and possibly especially due to this tremendous extent), the up to now available textbooks and courses do not offer such integrated approach.

2 PROBLEM STATEMENT

Firstly, we will note that in the robust constrained PID control we are going to deal with plants approximated by the 2nd order dominant dynamics that may include also long time delays (dead time), i.e. with plants that may be approximated by the transfer function

$$S(s)=\frac{b_1 s+b_0}{s^2+a_1 s+a_0}e^{-T_d s} \qquad (1)$$

We might discuss about the possibility to extend this class by giving higher attention to influence of potentially unstable zero dynamics, by considering the 2nd order systems with zero relative degree, etc. But, such cases represent just special not so frequently met situations.

Starting with examples of the well known books on PID control and dead time systems as e.g. Åström &.Hägglund (1995), Åström &.Hägglund (2005), Datta et al. (2000), Glattfelder & Schaufelberger (2003), Johnson & Moradi (2005), Keel et al. (2008), O'Dwyer (2006), Normey-Rico and Camacho (2007) we may mention one important shortness of the traditional PID control. The simplest situation for control design obviously corresponds to the two-parameter plant approximations with the plant transfer functions given as $K/(1+Ts)$ or as $Ke^{-T_d s}$. Both these approximations you can get e.g. by evaluating the average residence time (Åström &.Hägglund, 1995) using the moment method. Both they may be used for tuning simple I-controller (Skogestad, 2003). From the approximation with time constant it is then possible to derive the well known PI controller that represents huge majority of all control applications. However, about equivalent solutions corresponding to the approximation with dead time you find in the literature practically nothing – with exception of some authors proposing also for this situation the PI controller (Åström et al., 1998). What is reason for this assymetry? Are the solutions fitted to the pure dead time compensation not usefull for practice? Going back to the history you may find that solution for it called as the *Disturbance Response Feedback* was proposed by Reswick (Reswick, 1956) yet before the well known Smith predictor (Smith, 1957) – the best known control structure typically used for dead time systems. From that time the Smith predictor underwent a strong development (see e.g. Guzman et al. 2009; Normey-Rico and Camacho, 2007; Normey-Rico et al., 2009). But still there are not available appraoches that might be considered as the definitive ones – still they are in the position of a „trial and error“ approaches and therefore they need interactive tools to be tuned for specific circumstances of a given application. Within this context, due to the mathematical difficulties related to dealing with the dead time systems it is then no wonder that the twin-controller to the PI controller disapeared from the literature. Just few authors (as e.g. Zhong and Li, 2002) report advantages of such controllers. New tools developed for our course enable now to get easily an optimal robust tuning also for this controller, to explain problems of the Reswick solution and to derive basic robustness characteristics of this controller that will show in which application it might be usefull (Huba, 2010).

Yet before comming to these new tools, we should specify condition that must be put on an *admissible control* in order to guarantee under specified constraints transfer of the representative point from one steady state to another one and to inroduce several new notions simplifying formulation of the control problems for constrained systems. We will simply require such constraints that enable to maintain both the original as well as the target steady states what for a linear plant means possibility to keep also all steady states among them.

For an admissible control it has then sense to introduce notion of a *fundamental controller*. Under this denotation we will understand parameterized solution enabling to modify the speed of transient responses from the fully linear (pole assignment, PAC) control up to the relay minimum time control (MTC). The controllers may be parametrized by the closed loop poles, time constants, or the corresponding bandwidths. On the other hand, under „ad hoc“ controllers we will understand solutions offering single (not adjustable) closed loop dynamics, or dynamics adjustable just within a limited range. E.g. the relay MTC may be denoted as „ad hoc“ solution, since it offers unique closed loop dynamics that cannot be simply slowed down.

What do have setpoint step responses corresponding to all three types of control (linear PAC, relay MTC, constrained PAC) similar? As the first dominant property we could identify monotonicity (MO) of the plant output. For the controller output (plant input) it is typical that in controlling a second order plant the transients running from an initial to a final steady state will have two extreme points (intervals) with respect to the steady state control signal value. Whereas in the relay MTC the transitions between the first and the second extreme points and between the second extreme point and the steady state value are ideally instantaneous, in the PAC and the constrained PAC they are continuous, ideally MO. When increasing the closed loop dynamics of the constrained PAC (by shifting the closed loop poles to minus infinity), the control signal transient may approach with any required precission the relay MTC, but it everytime remains MO. With respect to the two extreme points we will denote such a control as a two-pulse control (2P) and a function having two extreme points and MO transients on any interval not including one of these extreme points as the 2P function. 2P functions represent a special case of nP functions, *n* being a non-negative whole number. MO functions may then be denoted as 0P functions.

With *r* being nonnegative whole number denoting the plant relative degree, under *Dynamical Class of control r* (shortly DC*r*) we understand all control tasks and their solutions with the MO step responses at the plant output and with the nP responses at the plant input (Huba, 1999; Huba and Bistak, 1999).

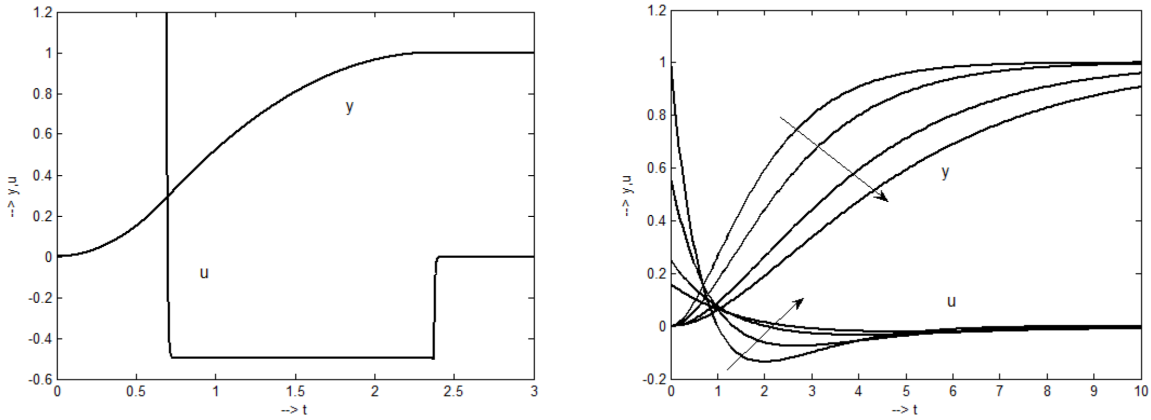

Figure 1 Output and control signal transients of the double integrator plant: left - for the relay MTC with control signal constrained by U_{max}=1.2 and U_{min}=-0.5; right – for the linear PAC with the double real pole values -1, -0.75, -0.5 and -0.4 - arrows indicate pole shifting to zero (i.e. slowing down transients)

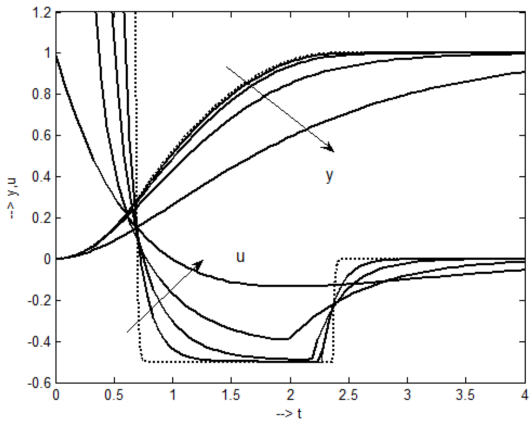

Figure 2 Output and control signal transients of double integrator plant by constrained PAC Umax=1.2 and Umin=-0.5 for the double real pole values -8, -4, -2 and -1 – arrows indicate pole shifting to zero (i.e. slowing down transients); dotted – MTC transients

In difference to the traditional PID control that was build of three basic elements (the P, I and D action), the new approach requires dealing with much higher number of basic fundamental solutions. Example of the not yet definitive table of the fundamental controllers covered up to now by our course (Huba and Simunek, 2007) is shown in Tab. 1. To each entry of the table is is possible to find at least one fundamental controller. Thereby, it is possible to distinguish solutions allowing step control change (discontinuity) at the origin and solutions with reference signal filter guaranteeing a continuous control setpoint step respons for any times.

Similarly as the Mendelejev Periodic Table of Elements, also the Table of Fundamental Controllers organizes controllers into groups having some similarities and contributes to easier and more transparent understanding of their functions, of the dynamics they can offer and of their optimal tuning.

Considering definition of the PID control and by having in mind that when increasing requirements on the dynamics of a fundamental solution the transients must converge to the MTC ones, it is then clear that it is dealing with solutions of the DC0, DC1 and DC2. Since these aspects fully disappeared from the traditional theory, one may also understand reasons for the tremendous inflation of different „optimal“ solutions reported e.g. by O'Dwyer (2006). Many of the known approaches to the PID control design do not fulfill the requirements on the fundamental solutions, since they:

- are linear and so, under consideration of constraints, they do not enable arbitrarily speeding up transients to approach in the limit the MTC transient responses, or
- do not involve free design parameters at all.

Table 1: Table of fundamental controllers (FF = static feedforward control)

		Dominant dynamics							
Dynamic class	I-action	K	Ke^{-T_d}	$\frac{K_s}{s+a}$	$\frac{K_s e^{-T_d s}}{s+a}$	$\frac{K_{s1}}{s+a_1}+\frac{K_{s2}}{s+a_2}$	$\left[\frac{K_{s1}}{s+a_1}+\frac{K_{s2}}{s+a_2}\right]e^{-T_d s}$	$\frac{K_s}{s^2+a_1 s+a_0}$	$\frac{K_s e^{-T_d s}}{s^2+a_1 s+a_0}$
0	N	FF	FF	FF	FF	FF	FF	FF	FF
	Y	I	PrI	PI	PrPI	PID	PrPID	PID	PrPID
1	N	-	-	P	PrP	P-P	PrP-P	PD	PrPD
	Y	-	-	PI	PrPI	P-PI	PrP-PI	PID	PrPID
2	N	-	-	-	-	-	-	PD	PrPD
	Y	-	-	-	-	-	-	PID	PrPID

E.g. Klán and Gorez (2000), similarly as many others, tried to find optimal PI controller tuning for a stable 1st order systems with a relatively long dead time T_d. The problem, however, is that the structure of the PI controller was generically derived for the 1st order plant dynamics without a dead time. Short dead time values used for approximating the nonmodelled dynamics can be allowed by limiting choice of the applicable controller parameters (closed loop poles). This, firstly, violates requirements put on the fundamental solutions. The control task with a nonmodelled dynamics does not allow reaching a dynamics of the MTC. Secondly, for the 1st order plant with a long dead time the fundamental controller will already have more complicated structure than the simple PI controller – it will already involve some features of the Smith predictor (Smith, 1957). So, the traditional PI controller cannot be treated within the group of fundamental solutions for the first order plants with long dead-time, just as a special (ad hoc) suboptimal solution.

The above example represents typical feature of majority of existing solutions. In many situations, the PI controller represents a seamingly easy to use solution, but not the fundamental solution that is able to guarantee high quality control and its validity over broader range of loop parameters and reliable tuning. In the time of analogue pneumatic and electronic controllers the main reason for the rare use of the optimal dead time structures was given by the problems of the dead time implementation. So, for many decades the traditional PID controllers without dead time compensation substituted them. These approaches, howevere, do not guarantee strictly fundamental properties and so they have reasonably contributed to the inflation of different "optimal" controller tuning. They further survive due to the conservativeness of practice despite the fact that the new digital controllers enable an easy dead time modeling and compensation. Of course, it has no sense to fight against their use, but it should be shown what and when they are able to offer. In such a way, all the ambiguity of solutions reported e.g. by O'Dwyer (2006) can be reasonably reduced.

In contrast to Klan and Gorez, Glattfelder and Schaufelberger (2003) tried to design for the first order plant with possible dead-time the pole assignment PI controller in such a way that its control signal step reaction would converge to one pulse of the MTC. This point shows other important discrepancy of the PI control theory. When comparing their design criterion with that one introduced by Klán and Gorez who required the optimal PI control signal step reaction to have a monotonic character approaching a step function (see e.g. the paper by Klan and Gorez , 2000 and the discussion by Strmčnik and Vrančič, 2000) we see a clear contradiction. Who is right in this conflict? The response may surprise many people – both requirements are right.

Simply, there exist two dynamical classes of PI control. Whereas the traditional linear PI control having the control signal response required by Klán and Gorez corresponds to the DC0, on the other hand controllers using anti-windup circuitry and trying to approach the MTC response characteristic by one saturated pulse of control represent already solution of the DC1. It has no sense to ask, which one is better – each has its unique properties that cover requirements of a specific group of applications (Huba et al. 1997).

Generalizing this way of arguing, it is then possible to define three different dynamical classes of the PID control (Tab.1). In this way, introduction of the DCs together with introduction of fundamental solutions enable transparent practically motivated classification of the existing controller structures and tuning rules. Here we should deal with the question under which conditions does the controller tuning yield a homogenous dynamics not depending on fluctuations of parameters and on the degree of the constraint, or on the magnitude of the input changes.

3 NEW DEVELOPMENT – THE PERFORMANCE PORTRAIT METHOD

By systematically mapping and storing information about the plant performance achieved by analysing the loop step responses computed by simulation over a grid of normalized loop parameters we get the so called closed loop performance portrait. Such performance portrait may then be used for optimally localizing a nominal operating point, or an uncertainty set of all possible operating points corresponding to specified intervals of loop parameters, or for deriving different robustness characteristics of the loop. Although the method is connected with numerical problems (related to the nature of computations over grid of points) and with difficulties related to the work in higher dimensional spaces, it gives very promising results especially when dealing with dead time systems.

Which control properties may be evaluated and stored? The key role in different control scenarios is usually played by the closed loop stability (ST). For stable transients, various quality measures may be defined including the *time related measures* (characterizing how fast the system reaches the required state after an initial deviation from it) and the *shapes related measures*, among them we may mention properties like nonovershooting (NO), monotonicity (MO), damping ratio, maximal overshooting and undershooting, etc.

Numerically, the Bounded-Input-Bounded-Output (BIBO) and Internal Model Control (IMC, see e.g. Morari and Zafiriou, 1989) stability require both bounded plant output $y(t)$ and bounded plant input $u(t)$

$$0 \le |y(t)| < Y_{\max} < \infty \ ; \ t \in \langle 0, \infty \ ) \\ -\infty < U_{\min} \le u(t) \le U_{\max} < \infty \qquad (2)$$

For a setpoint reference signal $w(t)$ the mostly used time related measures are the IAE (Integral of Absolute Error) and the ISE (Integral of Squared Error) performance indices defined as

$$IAE = \int_0^\infty |e(t) - e(\infty)| dt \ ; \ ISE = \int_0^\infty [e(t) - e(\infty)]^2 dt \ ; \ ITAE = \int_0^\infty t|e(t) - e(\infty)| dt \ ; \ e(t) = w(t) - y(t) \qquad (3)$$

Shinskey (1990) argues that IAE is a good performance measure because the size and length of error is proportional to lost revenue, so it is mostly enough to use it instead of ISE that was preferred in analytical design because of mathematical convenience. For dealing with real noisy systems the ITAE criterion is not appropriate because of increasing weighting by time that is multiplying effect of the measurement noise. In optimizing controllers the minimal IAE and ISE values usually correspond to transients with some overshooting. So, whereas aiming at transients without overshooting one has to define additional design constraints.

By considering initial and final values $y_0 = y(0)$, $y_\infty = y(\infty)$, $u_0 = u(0)$, $u_\infty = u(\infty)$ and by fulfilling

$$[y(t) - y_\infty] sign(y_0 - y_\infty) \ge -\varepsilon_y \ , \ \forall t > 0 \ ; \\ [u(t) - u_\infty] sign(u_0 - u_\infty) \ge -\varepsilon_u \ , \ \forall t > 0 \qquad (4)$$

the output and control responses will be denoted as nearly nonovershooting (NO). By introducing several measurement precision levels $\varepsilon_y, \varepsilon_u$ it is possible to replace the true/false information by more detailed quantitative one.

Final evaluation precision may also be introduced into the MO tests by requiring besides of (2) and (4)

$$[y(t) - y(t - T)] sign(y_\infty - y_0) \ge -\varepsilon_y \ , \ T \le t < \infty, \ T \in (0, T_h) \qquad (5)$$

T_h has to be chosen long enough to capture amplitudes of higher-harmonics pulses superimposed on the basic MO signal.

By generalizing properties of MO transients to more complicated shapes we may come to notion of the n-pulse (nP) function. This enables establishing bridge from the smooth shapes of the PAC to the rectangular pulses of the relay MTC. From the beginning of the MTC, with the integer n denoting the systems full degree, the necessity of n periods of energy accumulation/dissipation during the setpoint step responses was well known as the Felbaum's theorem (Feldbaum, 1965; Föllinger, 1993). In the constrained PAC the only exception of the control continuity may be allowed by considering transition from the initial actuator initial value to the next extreme point (saturation limit).

Function of time $f(t)$ that is continuous for $t > 0$ (with possible discontinuity at the origin) with the initial value $f(0^-) = \lim_{t \to 0^-} f(t)$, having with respect to the finite final value $f(\infty) = \lim_{t \to \infty} f(t)$ n-extreme points

$$f_{mi} = f(t_{mi}), \ i = 1, \dots, n \ \text{at} \ 0 < t_{m1} < \dots < t_{mn} \qquad (6)$$

and is MO on each interval not including some of these extreme points will be denoted as the n-Pulse (nP) function. By allowing discontinuity of $f(t)$ at the origin, the first extreme point may also move to zero from the right, when $t_{m1} = 0^+$, whereby the first MO interval before t_{m1} shrinks to zero.

By introducing notion of nP function it is then possible to denote the MO transients as 0P one. Since by limiting values of nP function to any interval containing $f(\infty)$ one does not change number of extreme points, the shape of the nP function may be considered to be invariant against admissible constraints. By introducing saturation limits and by decreasing duration of MO intervals among particular extreme points, nP functions may approach rectangular (relay) *n*-pulses of discontinuous MTC, but for $t > 0$ they always remain continuous in time.

Due to the limited space, in this paper we will be able to show just applications of the 0P and 1P functions. Obviously, 1P function may be defined as function with one extreme point that is MO before and beyond this extreme point. Examples of the 1P functions may include single exponential $f(t) = e^{-t}\mathbf{1}(t)$ (with $\mathbf{1}(t)$ denoting the unit step) that has single extreme point $f(0^+) = 1$ and discontinuity at the origin, or the difference of two exponentials $f(t) = \left(e^{-t} - e^{-2t}\right)\mathbf{1}(t)$ having the extreme $f_m = 1/4$ at $t_m = \ln 2$.

Similarly, as the 2P function we denote function with two extreme points that is MO on each interval not including one of them. Example of the 2P function with a discontinuity at the origin is given by $f(t) = \left(2e^{-2t} - e^{-t}\right)\mathbf{1}(t)$ with $t_{m1} = 0^+$; $t_{m2} = 2\ln(2) = 1.3863$. An example of 2P function that is continuous at the origin is given by $f(t) = 4\left((1+2t)e^{-2t} - e^{-t}\right)\mathbf{1}(t)$ having extreme points $t_{m1} = -LambertW(-1/4) = 0.3574$; $t_{m2} = -LambertW(-1,-1/4) = 2.1533$

To cover whole spectrum of transients typical for PID control we should yet complete the above list by definition of periodic functions interpreted as the nP function with $n \to \infty$. Then, after specifying e.g. the damping ratio (as done by Ziegler and Nichols, 1942) we could treat also the oscillatory loop behavior.

By considering the definition of nearly MO transients (5), it is possible to introduce nearly nP transients, for which each interval among the extreme points (6) satisfies to (5). Whereas (5) enables to evaluate the maximal deviation from the strict monotonicity, by modifications of the Total Variance (TV) proposed by Skogestad (2003)

$$TV = \int_0^\infty \left|\frac{du}{dt}\right| dt \approx \sum_i \left|u_{i+1} - u_i\right| \tag{7}$$

it is possible to introduce for such deviations an integral measure denoted as TV_0 criterion

$$TV_0 = \sum_i \left|u_{i+1} - u_i\right| - \left|u(\infty) - u(0)\right| \tag{8}$$

TV_0 characterizes total contribution of superimposed high-frequency signals to the overall MO control effort that is proportional not just to the amplitude, but also to the number of peaks. It takes zero values just for strictly MO control signal transients. Thereby, it may be applied to characterize both the plant output and the plant input deviations from ideal MO shapes.

To stress contribution of superimposed oscillation in systems with 1P dominant control signal it is appropriate to work with the TV_1 criterion introduced as

$$TV_1 = \sum_i \left|u_{i+1} - u_i\right| - \left|2u_m - u(\infty) - u(0)\right| \tag{9}$$

This again gives zero values just for strictly 1P transients, whereby it does not depend on the limit values of the control signal, just on its deviations from the ideal 1P shape. For nearly 1P control signals with superimposed higher harmonics it takes positive values.

After having prepared different performance measures enabling to specify our judgements about achieved control responses, we may now proceed by showing examples of their application in different control tasks.

4 EXAMPLE 1: COMPARING ROBUSTNESS OF P AND PI$_1$ CONTROLLERS

The task is to tune the loop with the P controller and IPDT plant

$$S(s)=\frac{Y_1(s)}{U(s)}=\frac{K_s}{s}e^{-T_d s} \tag{10}$$

to achieve the fastest possible y_1-MO and u-1P transients for interval loop parameters

$$K_s \in \langle 1,2\rangle \; ; \; T_d \in \langle 1,1.5\rangle ; \; c_K = K_{\max}/K_{\min} = 2 ; \; c_T = T_{d\max}/T_{d\min} = 1.5 \tag{11}$$

The closed loop setpoint response is given as

$$F_w(s)=\frac{Y_1(s)}{W(s)}=\frac{K_P K_s e^{-T_d s}}{s+K_P K_s e^{-T_d s}}=\frac{B(s)}{A(s)} \tag{12}$$

After introducing the normalized variable

$$p = T_d s \tag{13}$$

and the normalized loop parameters

$$\kappa = K_{s0}/K_s \; ; \; \Omega = K_p K_{s0} T_d \tag{14}$$

(12) can be transformed to

$$F_w(p)=\frac{\Omega}{\kappa e^p p+\Omega}=\frac{B(p)}{A(p)} \tag{15}$$

It means that the performance portrait (Fig. 3) can be mapped in the 2D space of the loop parameters (κ,Ω).

In the linear case (without control constraints) the minimal IAE value with MO plant output and 1P plant input corresponds to the line

$$\kappa/\Omega = 2.703 \; ; \; \kappa = K_{s0}/K_s \; ; \; \Omega = K_p K_{s0} T_d \tag{16}$$

that is close to the value corresponding to the double real dominant closed loop pole of the characteristic polynomial $A(p)$ (Huba et al. 1998)

$$\kappa/\Omega = \exp(1) = 2.718 \; ; \; \kappa = K_{s0}/K_s \; ; \; \Omega = K_p K_{s0} T_d \tag{17}$$

Robust tuning that would guarantee y_1-MO and u-1P transients means to locate the uncertainty box (UB) containing all possible operating points (Fig. 3) with vertices

$$UB=\begin{bmatrix}(\kappa_{\min},\Omega_{\max}) & (\kappa_{\max},\Omega_{\max})\\ (\kappa_{\min},\Omega_{\min}) & (\kappa_{\max},\Omega_{\min})\end{bmatrix}; \quad \begin{matrix}\kappa_{\max} = K_{s0}/K_{s\min} \; ; \; \Omega_{\max} = K_p K_{s0} T_{d\max} \\ \kappa_{\min} = K_{s0}/K_{s\max} \; ; \; \Omega_{\min} = K_p K_{s0} T_{d\min}\end{matrix} \tag{18}$$

below the critical line (16) given e.g. by the pair $K_{s0}=1; \;\; \Omega = 1/2.703 = 0.369$. From the radial shape of the border of MO&1P control in Fig. 2 it is obvious that the critical role is played by the upper left vertex

$$B(1,1)=(\kappa_{\min},\Omega_{\max})=(K_{s0}/K_{s\max}, K_P K_{s0} T_{d\max}) \tag{19}$$

In the analytical design this UB was traditionally placed at the line (17). The difference between the analytically (17) and the experimentally determined values (16) seems to be relatively small and it should yet be verified, if it is not casued by the finite precission of the numerical integration used in simulation.

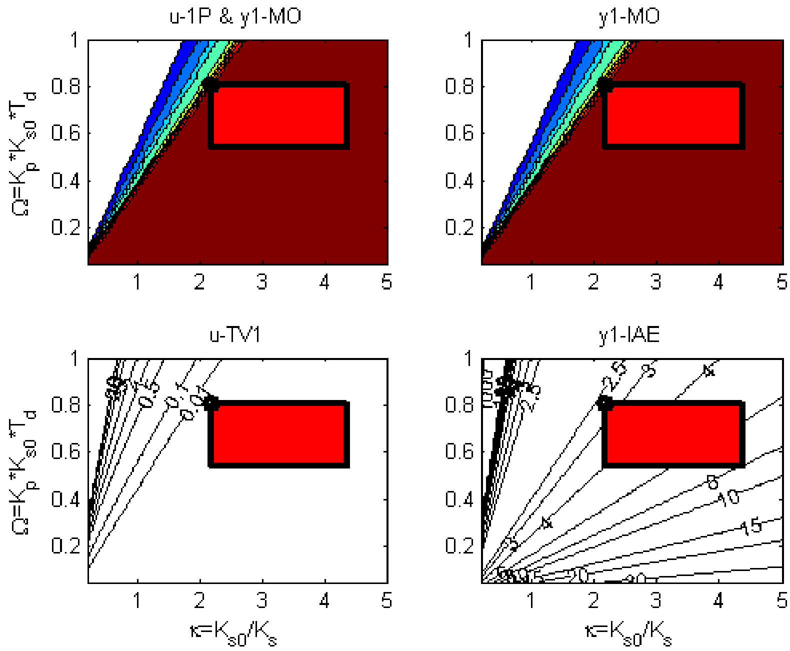

Figure 3 Performance portrait of the linear loop with the P-controller and IPTD plant; optimal point corresponds to the whole line (16); Uncertainty Box (17) corresponding to plan parameters (11) (red)

By considering just the uncertainty in T_d, when the UB (18) reduces to an Uncertainty Line Segment (ULS) with vertices

$$ULS = [\kappa, \Omega_{\min} \quad \kappa, \Omega_{\max}]; \; \Omega_{\min} = K_p K_{s0} T_{d\,\min}; \; \Omega_{\max} = K_p K_{s0} T_{d\,\max}; \; \kappa = const \tag{20}$$

it is then possible to derive robustness characteristic expressing dependance of the mean IAE value over this ULS. This will linearly increase with increasing uncertainty in T_d (Fig. 8). One of the most interesting questions is, if more complex controllers proposed for the plant model (10) will show stronger dependance on this uncertainty than the simplest P controller, or conversely.

As the first more complex controller we will consider the disturbance observer (DOB) based PI_1 controller achieved by extending the P controller by the DOB based I action (Fig. 4).

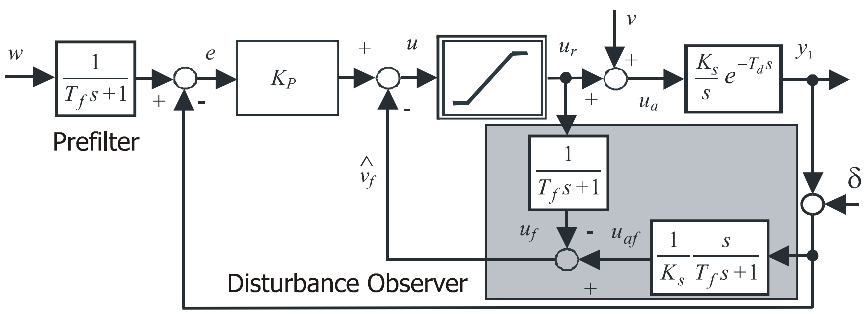

Figure 4 Disturbance observer based PI_1 controller with inverse dominant plant dynamics and possible prefilter

The closed loop setpoint response is now given as

$$F_w(s) = \frac{K_P K_s K_{s0} (1 + T_f s) e^{-T_d s}}{T_f K_{s0} s^2 + K_s e^{-T_d s} \left[s (1 + K_{s0} K_P T_f) + K_{s0} K_P \right]} \tag{21}$$

After introducing normalized variable (13) and the normalized loop parameters

$$\kappa = K_{s0} / K_s \; ; \; \Theta = T_d / T_f \; ; \; \Omega = K_p K_{s0} T_d \qquad (22)$$

it can be transformed to

$$F_w(p) = \frac{\Omega(\Theta + p)}{e^p \kappa p^2 + p(\Theta + \Omega) + \Omega\Theta} \qquad (23a)$$

Frequently, the loop is yet extended by a prefilter with the time constant T_f used to achieve smooth step responses of the control signal and to eliminate potential overshooting. Then, the setpoint transfer function becomes

$$F_w(p) = \frac{\Omega\Theta}{e^p \kappa p^2 + p(\Theta + \Omega) + \Omega\Theta} \qquad (23b)$$

In both cases the performance portrait can be mapped in the 3D space of the loop parameters (κ, Θ, Ω). So, using experimental approach based on identification of NO, MO, or 1P areas and integral performance measures, it will be necessary:

- to choose grid of parameters κ, Θ, Ω (22) with respect to the loop parameters (11), whereby the filter time constant T_f has to be chosen with respect to the required speed of disturbance response (larger filter time constants T_f improve noise filtration and the setpoint response, but prolong the disturbance reconstruction),
- for each point of this grid to run simulations,
- to check fulfillment of conditions for NO, MO and 1P performance and
- to evaluate performance indices (2-9).

Doing so over whole grid of parameters κ, Θ, Ω, it is possible to get information about corresponding control properties that can later be used in the robust controller tuning. This will be based on information about possible plant gains and possible dead time values (11).

After choosing some controller tuning K_P, T_f, one gets in general a 3D uncertainty set (US). It is expected that the robust design with given performance specification can be fulfilled just if the whole US can be located within the specified performance area.

In order to avoid problems of 3D illustrations, the loop performance portrait calculated in a slightly simplified setting in the 2D space (κ, Θ) with chosen value $\Omega = 3$ is in Fig. 5. As it is obvious from this picture, the best strictly y_1-MO and u-1P ($\varepsilon = 10^{-6}$) setpoint step response corresponds to the largest considered DOB filter time constant $T_f = 8.9$. Since this has to be limited with respect to the required speed of disturbance response, after determining its value we may continue with choosing Ω and K_{s0}. Once we specify e.g. $\Omega = 3$, optimal nominal tuning guaranteeing fastest possible MO transients with 1P control signal (with the minimal y_1-IAE=11.7 and with zero u-TV$_1$ value) will be achieved by choosing $\kappa = K_{s0} / K_s = 7.15 \Rightarrow K_{s0} = 7.15$ (this value of $K_{s0} \neq K_s$ compensates the not optimally chosen value of Ω).

When wishing to characterize robustness of this controller, e.g. in respect to the uncertainty in T_d, one should work in the 2D space with coordinates (Θ, Ω), when the US is represented by a skew line segement (ULS, Fig. 6, subplot u-TV$_1$) with vertices

$$ULS = [(\Theta_{\min}, \Omega_{\min}) \quad (\Theta_{\max}, \Omega_{\max})];$$
$$\Theta_{\min} = T_{d\min} / T_f ;\ \Theta_{\max} = T_{d\max} / T_f ;\ \Omega_{\min} = K_p K_{s0} T_{d\min} ;\ \Omega_{\max} = K_p K_{s0} T_{d\max} \qquad (24)$$

It is possible to calculate characteristics showing dependence of the mean IAE value over the ULS with increasing length corresponding to the uncertainty defined by the ratio of the maximal and minimal value $c_d = T_{d\max} / T_{d\min}$ (Fig. 8).

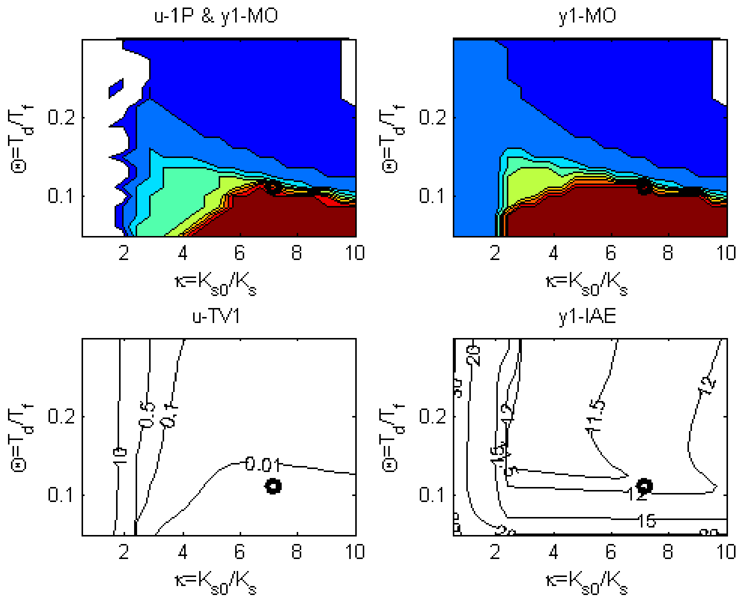

Figure 5 Example of the 2D performance portrait with the point corresponding to y_1-MO & u-1P step response with the lowest IAE value;
$T_f \in \langle 2,10 \rangle; T_d = 1; \Omega = 3; K_s = 1; \varepsilon_y = \{0.1, 0.05, 0.02, 0.01, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-16}\}$ (from blue to brown)

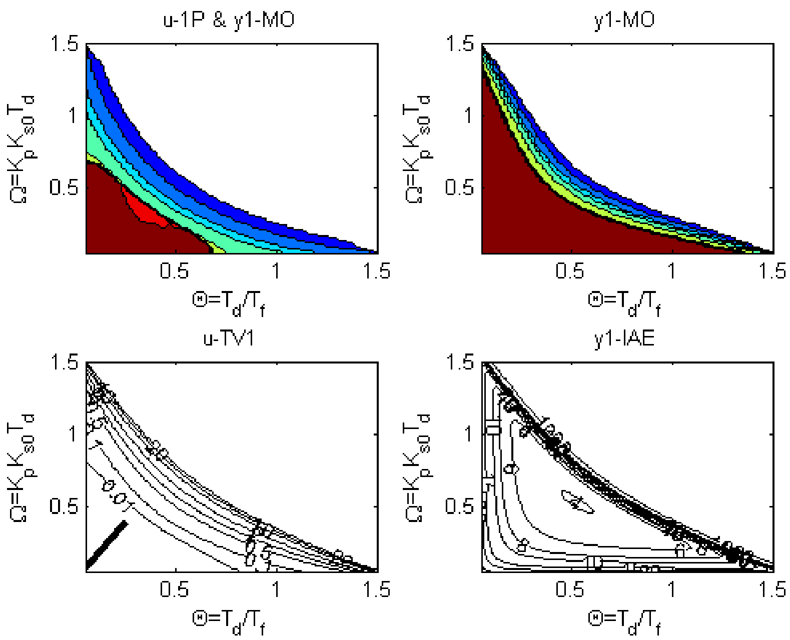

Figure 6 Example of the 2D performance portrait with the ULS corresponding to y_1-MO & u-1P step response ($\varepsilon = 10^{-6}$) with the lowest mean IAE value; $T_f = 20; T_d \in \langle 1,5 \rangle;\ K_s = 1;$
$\varepsilon = \{0.1, 0.05, 0.02, 0.01, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-16}\}$ (from blue to brown)

Thereby, for each possible value $T_d \in \langle T_{d\min}, T_{d\max} \rangle$ the transients (Fig. 7) remain y_1-MO & u-1P.

By comparing characteristics of the PI_1-DOB controller corresponding to different tolerances on the MO & 1P behaviour with those of the P-controller (Fig. 8), we may conclude that at least for smaller uncertainties the more complex PI_1-DOB does not show more radical IAE increase than the P one. Furthermore, in all cases transients with larger tolerated deviations from the strictly MO & 1P behaviour are less dependant on the uncertainty c_d than those requiring higher precission. However for larger uncertainties the IAE values of the PI_1 controller increase more rapidly than for the P controller and it may happen that it is not possible to position whole ULS into MO & 1P area with acceptable deviation from ideal shapes, as it happened for $T_f = 10; \varepsilon = 10^{-16}$, Fig. 8. This is due to the shift of the ULS to the righ by decreasing the filter time constant T_f, whereby also the upper border of the area with $TV_1 \leq 0.01$ decreases.

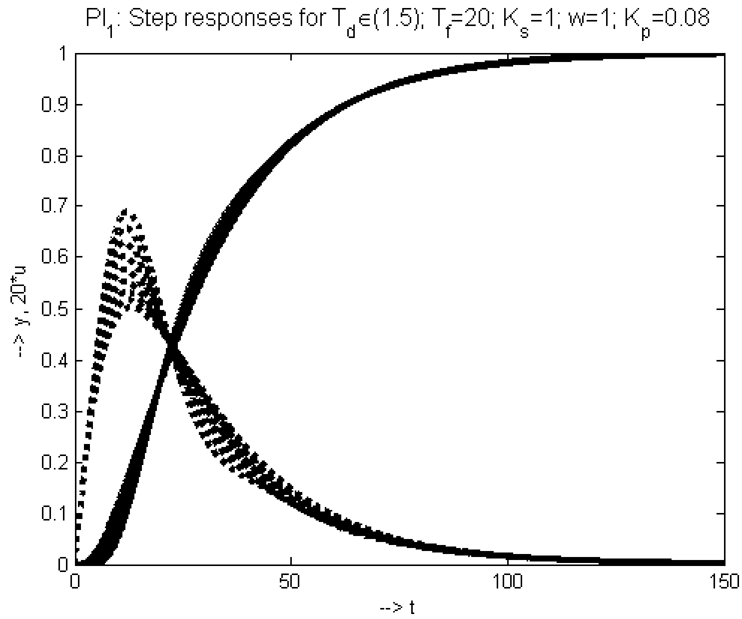

Figure 7 Example of y_1-MO & u-1P step responses corresponding to the ULS from Fig. 6 with the lowest mean IAE value; $T_f = 20; T_d \in \langle 1,5 \rangle; K_s = 1;$

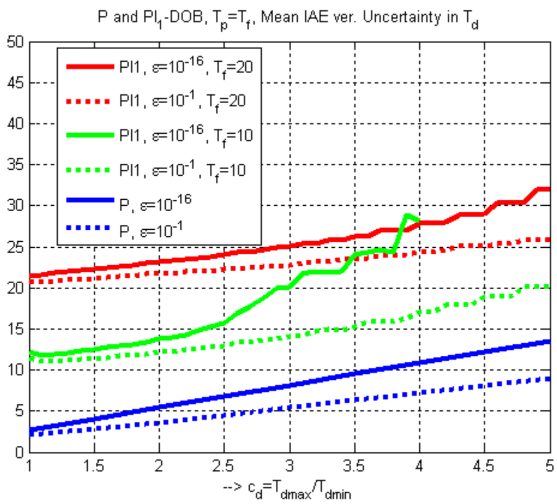

Figure 8 Influence of the uncertainty in T_d characterized by $c_d = T_{d\,\max} / T_{d\,\min}$ on the mean IAE values of the P and of the PI_1-DOB with $T_f = 20$ and $T_f = 10$

5 EXAMPLE 2 EVALUATING EFFECT OF CONTROL SIGNAL CONSTRAINTS ON P AND PI$_1$ CONTROLLERS

Traditionally (see e.g. Glattfelder and Schaufelberger, 2003; Föllinger, 1993; Hsu and Mayer, 1968) stability of constrained control was tested by using the Popov and the circle criterion, or by the hyperstability tests. However, fulfilment of stability conditions does not yet guarantee the required control performance. The performance portrait may be used also for evaluating and designing quality of constrained control.

Next we are going to analyze how the control constraints influence performance and optimal tuning of the constrained PID control. Ideally, we would like having such a control that has for any admissible constraints no deviation of the output and input signals from their ideal shapes defined in the linear case without control saturation. Of course, in defining ideal shapes of the linear case we will surelly be restricted – with respect to this we have already introduced performance measures that seem to might be invariant against the control constraints: we will require monotonicity for the plant output and depending on the dynamical class of control shapes of the nP functions for the plant input. For evaluating deviations from ideal shapes it is possible to use formulas (5) and the TV$_n$ criteria, $n = 0,1,\dots$.

When constraining plant input in the loop with simple P controller, by the performance portrait (Fig. 9a) it is possible to identify that just the optimal tuning corresponding to strictly y_1-MO & u-1P behaviour remains invariant against saturation. All other performance measures enabling some deviations from ideal shapes do already depend on the saturation. From this point of view, giving controller tunig with some tolerated output overshooting (as e.g. do many methods reported by O'Dwyer, 2006) has sense just for unconstrained control and so it is frequently not appropriate for practical use. It is also clear that when dealing with static plants with output depending on the steady state control signal value, all measures with possible exception of the strict nP shapes will depend also on the reference signal values. So, these shapes seem to play a special role in the control design.

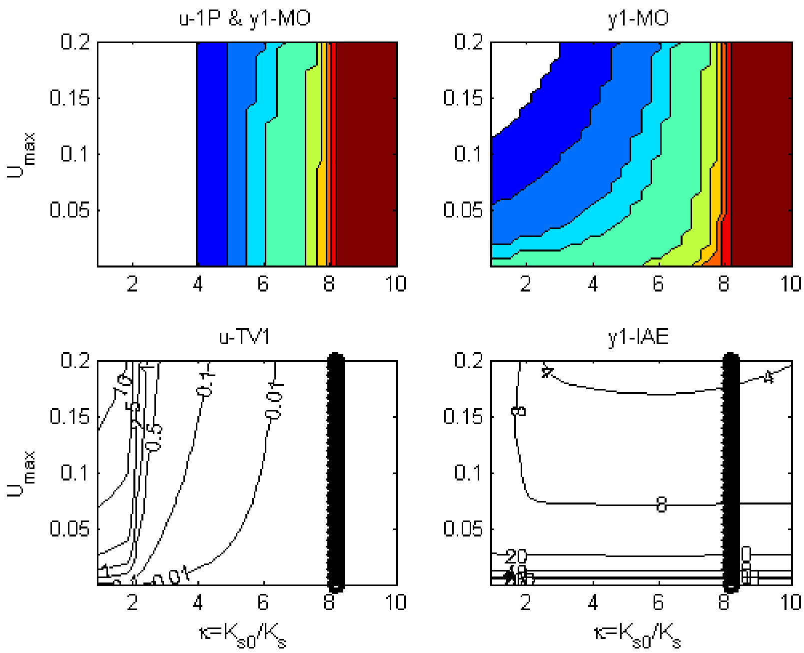

Figure 9a Influence of the saturation level U_m on the optimal working point κ of the P controller; $\varepsilon = \varepsilon_y = \varepsilon_u = \{0.1, 0.05, 0.02, 0.01, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-16}\}$ (from blue to brown)

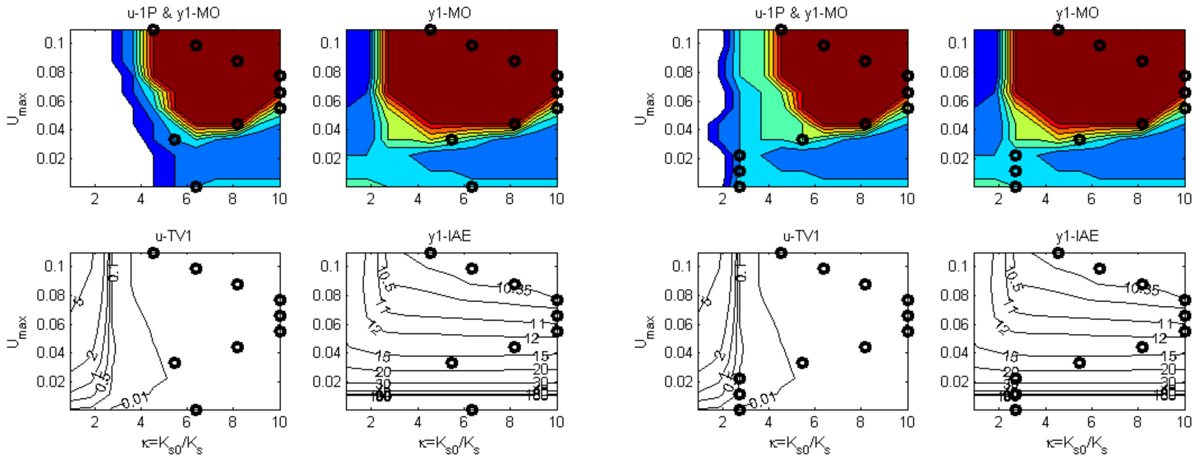

Figure 9b Influence of the saturation level U_m on the optimal working point κ of the PI$_1$-DOB controller for $\varepsilon = \varepsilon_y = \varepsilon_u = 0.02$ for the DOB filter time constant $T_f = 10;\ \Omega = 3;\ T_d = 1;\ K_s = 1;$ $\varepsilon = \{0.1, 0.05, 0.02, 0.01, 10^{-3}, 10^{-4}, 10^{-5}, 10^{-6}, 10^{-16}\}$ (from blue to brown, left relative precission for *u*-1P, right absolute precission)

Because we already know that by using the PI$_1$-DOB controller with prefilter the speed of transient responses is always limited by the time constant T_f, in evaluating influence of the saturation level U_m here it is enough to tune single parameter κ that can compensate also not optimaly chosen values of Ω. So, in the case of nominal tuning the loop analysis may be visualized in the plane of parameters $(\kappa, U_{\max})$. In the corresponding performance portrait in Fig. 9b it is to remember that the deviation of the controller output from the ideal *u*-1P shape are different if evaluated with the relative precission related to the control amplitude U_m, or if evaluated with an absolute precission.

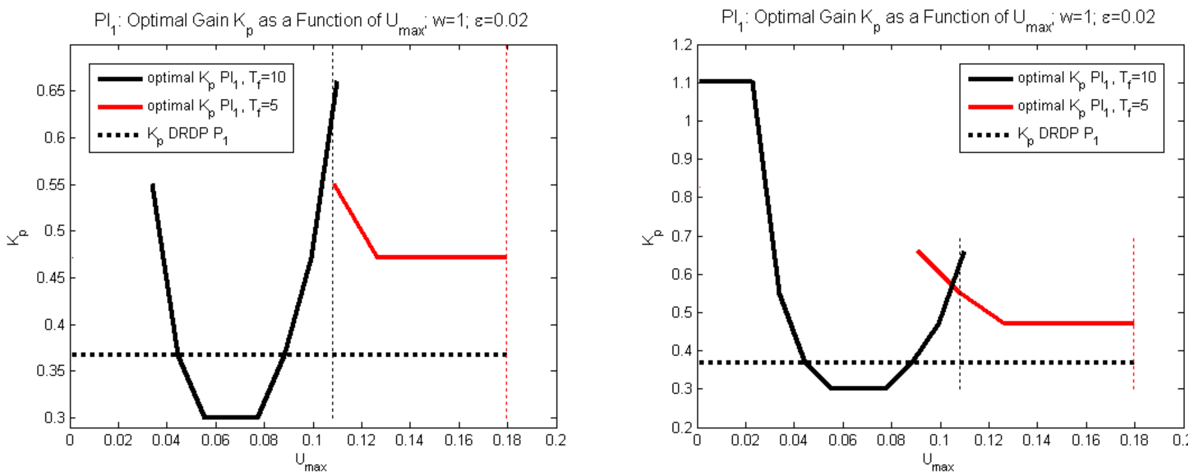

Figure 10 Influence of the saturation level U_m on the optimal gain of the PI$_1$-DOB controller for two different values of the DOB filter time constant T_f and $\varepsilon = \varepsilon_y = \varepsilon_u = 0.02$ (left relative precission for *u*-1P, right absolute precission)

By expressing changes of the optimal working point by the proportional gain K_P (Fig. 10) we may see that the optimal controller gain changes in a relatively broad range. However, from the subplot y_1-IAE in Fig. 9b it is to see that the dependence of the IAE contours on the ratio κ is relatively flat. From the subplot *u*-TV$_1$ in Fig. 9b it is to see that by using tuning optimal for the unconstrained controller in the worst constrained case the maximal TV$_1$ values would be about 0.01 and from the subplot y_1-MO it follows that the output deviations from monotonicity would be about 1% of the reference value. For many applications it would represent a sufficient quality, so we might conclude that the resulting performance is just weekly dependant on the control constraints and the changes in the position of the optimal working point could be neglected.

The second possible conclusion of this example is that it is possible to find such a controller gain that the shape of transients remains unchanged in a broad range of saturation limits. However, they may slightly deviate from the optimal ones giving minimal IAE values.

The third possible conclusion is that whereas the differences in optimal constrained and unconstrained gains seem relatively large, in the time scale of the constrained case are the differences in transients corresponding to both these gains negligible (Fig. 11).

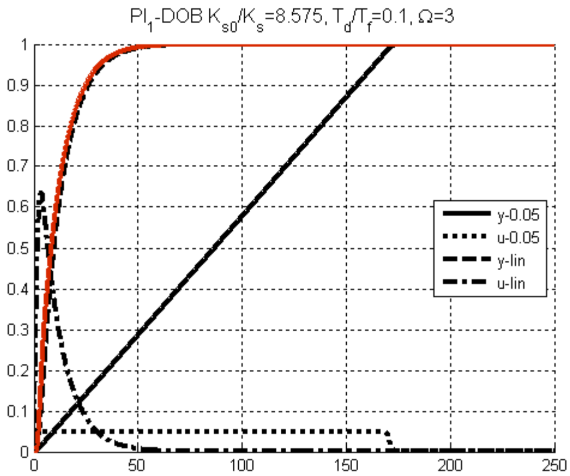

Figure 11 Step responses corresponding to the optimal tuning identified for the saturation level $U_m = 0.05$ and used then also for the unconstrained case (black) and the output step response corresponding to the optimal gain for the unconstrained case (red); $T_f = 10; T_d = 1; K_s = 1; \Omega = 3$

Similarly it may be shown that e.g. the IMC like controllers based on the parallel plant model (including also solutions with modified Smith predictor) that enable just nearly MO & 1P transient responses will not be invariant against control signal constraints, but a reasonable robustness increase (Huba, 2009) may be achieved by compensating dead time in the DOB of the PI_1 controller from Fig. 4.

6 CONCLUSION

This paper has shown that the newly developed approaches to the robust constrained PID control analysis and design using performance portrait enable to solve with a reasonably increased efficiency and elegancy also the traditionally hard control problems.

The method enables determining optimal controller tuning for both the nominal plant parameters and for its parameters known over some uncertainty intervals. Furthermore, it also enables to derive characteristics describing general robustness properties of the particular control loop and to decide, which controller is better appropriate for particular problem. By systematically deriving such characteristics for all possible controllers one gets an efficient support in choosing the optimal solution. Results of such analysis for several other controllers will be presented at the conference.

However, as every time in the research history, each newly answered question opens several new ones. Now, one could ask about possible precission of the computations over grid of points, about guaratnees that no unwanted phenomena occur, about rules for optimally choosing the number of calculated points and intervals for their calculations, etc.

Yet more questions appear when starting to work with different weakened versions of the shape-related properties like the monotonicity and the nP shapes, when the achieved transient may simultaneously fall into several performance classes.

Despite it seems at this moment that the method is fully numerical one, we expect that it will generate several impulses also for the development of analytical methods.

7 ACKNOWLEDGEMENTS

This work was partially supported by the Project VEGA 1/0656/09: Integration and development of nonlinear and robust control methods and their application in controlling flying vehicles, by the project KEGA 3/7245/09 Building virtual and remote experiments for network of online laboratories. It was also supported by the grant (No. NIL-I-007-d) from Iceland, Liechtenstein and Norway through the EEA Financial Mechanism and the Norwegian Financial Mechanism. This project is also co-financed from the state budget of the Slovak Republic

REFERENCES

ÅSTRÖM, K. J., T.HÄGGLUND (1995). *PID controllers: Theory, design, and tuning* – 2nd ed., Instrument Society of America, Research Triangle Park, NC.

ÅSTRÖM, K. J., PANAGOPOULOS, H. AND HÄGGLUND, T. (1998): Design of PI Controllers based on Non-Convex Optimization. *Automatica*, 34, 585–601

DATTA, A., HO, M. T. AND S.P. BHATTACHARYYA (2000). *Structure and synthesis of PID controllers*. Springer London.

FÖLLINGER, O. (1993). *Nichtlineare Regelungen.* 7. Auflage, R. Oldenbourg Verlag München.

GLATTFELDER,A.H. UND SCHAUFELBERGER,W. (2003). *Control Systems with Input and Output Constraints*. Springer, London.

GUZMÁN,J.L ET ALL. (2008). Interactive tool for analysis of time - delay systems with dead - time compensators. *Control Eng. Practice*, 16, 7, 824-835.

HSU,J.C. AND MAYER, A.U. (1968). *Modern Control Principles*. McGraw Hill N.York.

HUBA, M. (1981A). *Číslicový časovo optimálny regulátor s premenlivou štruktúrou.* Písomný referát k odbornej ašpirantskej skúške. [Digital Time Optima Controller With Variable Structure. Written Report to the PhD study exam]. EF SVŠT Bratislava.

HUBA, M. (1981B). *Číslicový časovo optimálny regulátor s premenlivou štruktúrou.* [*Digital Time Optima Controller With Variable Structure.* PhD Thesis]. EF SVŠT Bratislava.

HUBA, M. (1984A). *Použitie mikropočítača ako jedného typu adaptívneho číslicového korekčného člena.* Učebné texty PGŠ Bratislava. [*Using microcontroller as one type of adaptive digital compensator.* Study materials for continuing education courses] EF SVŠT-Inštitút Slovchémie Bratislava.

HUBA,M.:(1984B). *Algoritmy identifikácie*. Učebné texty PGŠ. Bratislava. [*Identification algorithms.* Study materials for continuing education courses] ES SVŠT –Inštitút Slovchémie.

HUBA, M. (1992A). Teória automatického riadenia 3 [*Theory of Automatic Control 3*], STU Bratislava.

HUBA,M., SKACHOVÁ,Z., BISTÁK,P. (1997). Minimum Time Pole Assignment Controllers Based on I1 and I2 Models with Dead Time. *Proc. Int. Summer School on Computer Aided Education in Automation and Control,* STU Bratislava, 253-276.

HUBA,M., KUĽHA,P., BISTÁK, P., SKACHOVÁ, Z. (1997). Two Dynamical Classes of PI-Controllers for the 1st Order Loops. *2nd IFAC Workshop New Trends in Design of Control Systems, Smolenice*, 293-298.

HUBA,M., BISTÁK,P., SKACHOVÁ,Z., ŽÁKOVÁ,K. (1998). P- and PD-Controllers for I1 and I2 Models with Dead Time. *In: Proc. 6th IEEE Mediterranean Conference on Control and Automation*, World Scientific, Alghero, 514-519.

HUBA, M. (1999). Dynamical Classes in the Minimum Time Pole Assignment Control. *In: CASYS '98, Computing Anticipatory Systéme, Liege*. Woodbury: American Institute of Physics.

HUBA, M. AND BISTÁK, P.(1999). Dynamic Classes in the PID Control. *In: Proceedings of the 1999 American Control Conference*. San Diego: AACC.

HUBA, M., ŽÁKOVÁ, K. (1999). Analysis of P and PD controller setting. *In: Modern Control Theory.* Huba, M., Menini, L., Tornambe, A., Edts., Progetto Leonardo Societa editrice esculapio, Bologna

HUBA, M. &.BISTAK, P. (2001). Minimum Time Pole Assignment Control – A Unifying Approach to the Control of Constrained Systems. CD: *Medzinarodna konferencia SSKI – Kybernetika a Informatika* [*International Conference "Cybernetics and Informatics"*], Piestany, 50-51.

HUBA, M. (2002). New Control Concept for Constrained Systems. In: *Kybernetika a informatika. Bratislava : SSKI*, 2002, [*International Conference "Cybernetics and Informatics"*] Trebisov. Slovak Republic, 4.-6.9.2002.

HUBA, M. (2003A). *Nelineárne systémy*. [*Nonlinear Systems*] Vydavateľstvo STU Bratislava.

HUBA, M. (2003B). Syntéza systémov s obmedzeniami I. Základné regulátory. II. Základné štruktúry. [Constrained Systéme Design. Vol.I. Basic Controllers; Vol. 2 Basic Structures] STU Bratislava.

HUBA M. (2005). Constrained pole-assignment control for SISO plants. *In: Medzinárodná konf. Kybernatika a informatika*, 9.-11.2.2005, Dolný Kubín, 317-323.

HUBA, M. (2006A). Constrained pole assignment control, *In: Current Trends in Nonlinear Systems and Control,* L. Menini, L. Zaccarian, Ch. T. Abdallah, Edts., Boston: Birkhäuser, 163-183.

HUBA, M. (2006B) *Teória automatického riadenia 3. PID regulátory s obmedzeniami.* [*Theory of Automatic Control 3. Constrained PID Control*] STU Bratislava.

HUBA, M., ŠIMUNEK, M.(2007). Modular Approach to Teaching PID Control. *IEEE Trans. Ind. Electr.*, 54, 6, 3112-3121.

HUBA, M. (2007). Nelineárne systémy. Návrh riadenia s rešpektovaním obmedzení. [Nonlinear systems. Design respecting constraints.] STU Bratislava –Slovenská e-akadémia.

HUBA, M. (2009). Robust Design of Integrating Controllers for IPDT Plant. Edits: Fikar, M., Kvasnica, M., *In Proc. 17th Int. Conf. Process Control '09, Š. Pleso,* 353–357, http://www.kirp.chtf.stuba.sk/pc09/data/papers/109.pdf.

HUBA, M (2010). Predictive I Controllers, their Robust Tuning & Characteristics. *Submitted to IFAC Symposium on Time Delayed Systémes* Prague.

JOHNSON, M. A., MORADI, M. H. (EDITORS) (2005). *PID Control: New Identification and Design Methods*. Springer London.

KEEL, L. H., KIM, Y. C., BHATTACHARYYA, S. P. (2008). Advances in Three Term Control. *Pre-Congress Tutorials & Workshops*. *17th IFAC World Congress* Seoul, Korea.

KLÁN, P. AND R. GOREZ (2000). Balanced Tuning of PI Controllers. *European Journal of Control*, 6, , 541-550.

MORARI,M. AND E.ZAFIRIOU (1989). *Robust Process Control*. Prentice Hall, Englewood Cliffs, NJ.

NORMEY-RICO, J.E., CAMACHO, E.F. (2007). *Control of Dead-time Processes*. Springer, London.

NORMEY-RICO, J.E. ET AL. (2009). An unified approach for DTC design using interactive tools. *Control Eng. Practice* 17, 1234-1244.

O'DWYER, A. (2006). *Handbook of PI and PID controller tuning rules.* 2nd Ed., Springer London.

RESWICK, J.B. (1956). Disturbance-response Feedback – a new control concept. *Trans. ASME* , No.1, 153-162

RIVERA, D. E., M. MORARI AND S. SKOGESTAD (1986). Internal Model Control. 4. PID Controller Design. *Ind. Eng. Chem. Process Des. Dev*., 25, 252-265.

SABERI, A., STOORVOGEL, A.A. AND P.SANNUTI (2000). *Control of Linear Systems with Regulation and Input Constraints*. Springer London.

SHINSKEY, F.G. (1996). *Process Control Systems*, 4th ed., McGraw-Hill, N.York.

SHINSKEY, F.G. (2000). PID-Deadtime Control of Distributed Processes. *Preprints IFAC Workshop on Digital Control: Past, present and future of PID Control PID'2000,* Terassa, Spain, 14-18.

SKOGESTAD, S. (2003). Simple analytic rules for model reduction and PID controller tuning. *Journal of Process Control* 13, 4, 291–309.

SKOGESTAD, S. (2006). Tuning for Smooth PID Control with Acceptable Disturbance Rejection. *Ind. Eng. Chem. Res.*, 45, 7817-7822.

SMITH, O.J.M. (1957). Closser control of loops with dead time. *Chemical Engineering Progress*, Vol.53, No.5, 217-219.

STRMČNIK, S. AND VRANČIČ, D. (2000). Discussion on „Balanced Tuning of PI-Controllers" by P. Klán and R. Gorez. *European J. of Control*, 5, 51-552.

ZHONG, Q. CH., LI, H. X (2002). A delay-type PID controller. Proc. *15th IFAC World Congress*, Vol. 15, Part 1.

ZIEGLER,J.G. A NICHOLS,N.B. (1942) Optimum settings for automatic controllers. *Trans. ASME*, 759-768.