Proceedings of the 17th World Congress
The International Federation of Automatic Control
Seoul, Korea, July 6-11, 2008

Constrained Control for Systems with Relative Degree One

Peter Ťapák*, Mikuláš Huba, Katarína Žáková*****

Slovak University of Technology in Bratislava, Fac. Electr. Eng. & Info. Technology, Ilkovičova 3, 812 19 Bratislava Slovakia (e-mail: peter.tapak @stuba.sk)*
*** e-mail: mikulas.huba@stuba.sk*
**** e-mail: katarina.zakova@stuba.sk*

Abstract: The paper discusses the design of second and third order plants with two/three parallel first order channels i.e. second/third order system with one/two stable zeroes. The designed control respects input constraints and can be easily tuned by one parameter, the closed loop pole. When choosing proper closed loop poles one has to take into account parasitic time constants, measurement noise, plant uncertainty, etc. Nevertheless, in the nominal case the designed controller is able to give the dynamics ranging from the minimum time control to pure linear one - according to the chosen pole. The desired control signal has one interval at the saturation limit, then it converges to a steady value with the dynamics given by the closed loop pole.

1. INTRODUCTION

The control signal saturation can be considered as the elementary nonlinearity present in practically each control loop. In the 50's and 60's its effects have been intensively treated within the scope of the minimum time systems. Simultaneously, the demand on smooth solutions and quiet steady states lead to the development of the linear control technique called pole assignment control. Nowadays, the most popular techniques dealing with the input constraints used in practice are the MPC and anti-windup control. By respecting the control constraints, the new concept of the constrained pole assignment control combines the qualities of both the minimum time and of the pole assignment control. The paper is organized as follows: Chapter 1: introduction, Dynamical classes of control are introduced in chapter 2, control for 2nd/3rd order system with relative degree one I presented in chapter 3/4.

2. DYNAMICAL CLASSES OF CONTROL

By index of the dynamical class (DC) it is understood a non-negative integer denoting number of possible intervals with the limit control signal values that can occur under the limit case of the constrained pole assignment control with poles shifted to minus infinity that is equivalent to the minimum time control. With respect to the Feldbaum's theorem (Feldbaum,A.A., 1965) it is possible to conclude that the PID control corresponds to the transient processes from the dynamical classes 0, 1 and 2.

While in the DC0 the ideal control response following a setpoint step has also step character (Fig.1) and no saturation phase (therefore it can be successfully treated by the linear theory), the dynamical classes 1, 2 or higher (as e.g. in Fig.2) are already typical by a period (periods) with saturated control and so they are already nonlinear.

Processes of the DC0 are typically used in situations, where the dynamics of transients may be neglected, i.e. it is not connected with a reasonable energy accumulation. Such processes can e.g. be met in controlling flows by valves. After constraining the rate of control signal changes after a setpoint of a disturbance step, the transition to a new control signal value can be an exponential one (Fig.1).

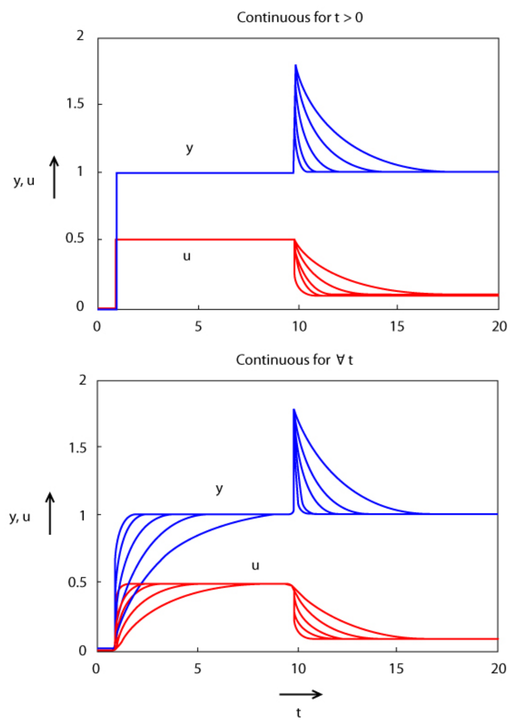

Fig. 1. DC0: Control signal reaction to a setpoint step; up – with the rate constraints just for the disturbance step; down – with a rate constraint also for a setpoint step

Within the DC1 the control signal reaction to a setpoint step change can already involve one control interval with constrained control value (Fig.2) that is later followed by a monotonous transient to the new steady state value. For the initial phase of control it is typical accumulation of energy in the controlled process. This is associated with a gradual increase (decrease) of the controlled output variable that is most rapid under impact of the limit control signal value. E.g. by charging a container with liquid, in the first phase of control the input valve will be fully opened and only in the vicinity of the required level the input flow will decrease to a steady state value keeping the required level. Similar transients can be frequently met in speed control in mechatronic systems, in the temperature, pressure and concentration control, etc.

After limiting the rate of changes during the transients, the span of the period with the limit control action decreases, but the total length of transient to the new steady state increases. When constraining also the control signal rate after a setpoint change, the control signal does not catch to reach the limit value, since the necessary control decrease to the steady state has to start yet before it – the length of transient growths further.

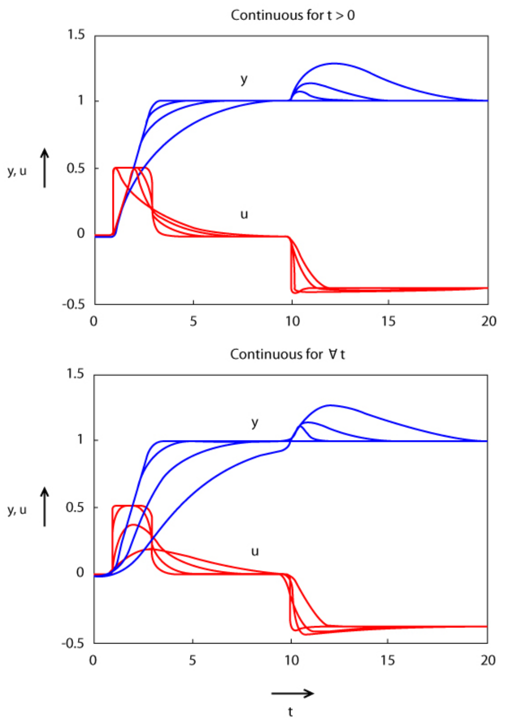

Fig. 2. DC1: Control signal reactions to a setpoint step parametrized by the closed loop poles; left – without rate constraints for t=0; right – with a rate constraint for t=0

With respect to one possible interval with constrained controller output for dealing with this dynamical class it is usually not enough to remain within the linear control. Typical solutions for this dynamical class are frequently achieved with different anti-windup (aw) controllers.

For more information on the Constrained PID control, or constrained pole assignment control see e.g. (Huba, 1990, Huba and Bisták,, 1999, Huba2006, Huba and Simunek, 2007). However the constrained PID control is not the only way how to deal with constraints. Solutions with monotonous output transients parametrized by the closed loop poles that enable to achieve any dynamics ranging from the linear pole assignment control up to the relay minimum time one will be denoted as the fundamental ones. This concept was inspired by the works of Åström et al. (Åström, K. J., 1998),(Åström, K. J., 2003) that tried to develop general parameterized solutions which can be relatively easily adjusted to a particular situation by building on parameterizations as the sensitivity functions, or the complementary sensitivity functions related to the robust control. Having clear-cut physical interpretation of the effect of such tuning parameters and clear picture of its appropriate default values, the tuning should be much simpler and reliable.

However, from more general point of view of the constrained control the sensitivity and complementary sensitivity functions do not represent an optimal solution. They e.g. do not match the natural expectation that by decreasing the range of possible system and state uncertainty (parameter fluctuations), the effect of the non-modelled dynamics (parasitic delays) and the amplitude of the measurement noise - when there are no other specifications on the control quality - the achieved solutions would converge to the results of the minimum time control.

Such a requirement was obviously followed using another way of the closed loop parameterization – the pole assignment method by Glattfelder and Schaufelberger (Glattfelder,A.H. und Schaufelberger,W., 2003). The anti-windup PI controller they have analyzed was very close to the ideal control signal step reaction converging to the one pulse of the minimum time control. But not completely.

In order to introduce an effective controller classification, it is further important to introduce new notion of “fundamental” controllers. Such a controller has to have following properties:

1.For the nominal dynamics S(s) it must yield transient responses reaching from the fully linear up to the time optimal ones that can be simply scalable by the closed loop poles, (or other equivalent parameters as the time constants).

2.For a reliable controller tuning that guarantees monotonous responses the choice of the poles has to be restricted by identifying the perturbation (parasitic) dynamics $\delta S(s)$ chosen usually as a dead-time, or a time constant.

The first point involves the requirement to generalize the two limit solutions – the linear pole assignment control and the relay minimum time control to a compact set of responses that can be simply modified by the closed loop poles by offering properties that combine basic features of both limit solutions.

The second point is related to a reliable controller tuning. It tells that the system has to be approximated in such a way that besides of the nominal dynamics it is also determined the always present parasitic time delay (perturbation dynamics) that determines borders for the closed loop poles choice guaranteeing the expected properties.

Many of the known approaches do not fulfil the requirements on the fundamental solutions, since they do not allow to approach the minimum time transient responses, or they do not involve free design parameters at all. These approaches do not guarantee strictly "global" results 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 modelling and compensation. Of course, it has no sense to fight against their use, but it should be shown that they do not represent globally valid solutions. In such a way, all the ambiguity of solutions reported e.g. by O'Dwyer (O´Dwyer, A., 2006) can be reasonably reduced.

3. SECOND ORDER PLANT

2.1 P-PI Controller design

Let us consider a plant with 2 modes (slow and fast). Its transfer function is

$$F(s)=\frac{K_1}{1+T_1 s}+\frac{K_2}{1+T_2 s}=K\frac{1+T_0 s}{(1+T_1 s)(1+T_2 s)}\,;$$
$$K=K_1+K_2\,;\; T_0=\frac{K_1 T_2+K_2 T_1}{(K_1+K_2)} \tag{1}$$

This transfer function describes e.g. the thermal plant with two ways of heat transmission:

Heat radiation (fast mode)

Heat conduction via body of the plant (slow mode)

The output of both the blocks can be described by the following differential equations

$$\dot y_1=\frac{K_1 u-y_1}{T_1}\,;\; \dot y_2=\frac{K_2 u-y_2}{T_2} \tag{2}$$

For the output of the system

$$y=y_1+y_2 \tag{3}$$

using (1), (2), (3) one can write

$$\dot y=\dot y_1+\dot y_2=\overline{K}u-\frac{1}{T_1 y}-\tau y_2;$$
$$\overline{K}=\left(\frac{K_1}{T_1}+\frac{K_2}{T_2}\right);\; \tau=\left(\frac{1}{T_2}-\frac{1}{T_1}\right) \tag{4}$$

$$e=w-y\,;\; \dot e=-\dot y \tag{5}$$

The value of control signal u_w which maintains the output of the system at the value $w=const$ is

$$u_w=\frac{w}{K_1+K_2} \tag{6}$$

Using the control signal (5) the outputs of the single channels are

$$y_{1\infty}=K_1 u_w=\frac{K_1}{K_1+K_2}w\,;$$
$$y_{2\infty}=K_2 u_w=\frac{K_2}{K_1+K_2}w\overset{!}{=}w_2 \tag{7}$$

The pole assignment control is given by

$$\dot e=\alpha e \tag{8}$$

where α is the chosen closed loop pole. Substituting in (5) one gets

$$-\overline{K}u+\frac{y-w+w}{T_1}+\tau(y_2-w_2+w_2)=\alpha(w-y) \tag{9}$$

which leads to the control design as the parallel structure of the P-P controller

$$u=\frac{1}{K}w+K_R e+K_{R2}e_2\,;\; e_2=w_2-y_2\,;$$
$$K_R=-\frac{\alpha+1/T_1}{\overline{K}}\,;\; K_{R2}=-\frac{T_1-T_2}{K_1 T_2-K_2 T_1} \tag{10}$$

This P-P controller can be expanded to P-PI controller by the I-action as the disturbance reconstruction fig.3 which compensates the steady state error.

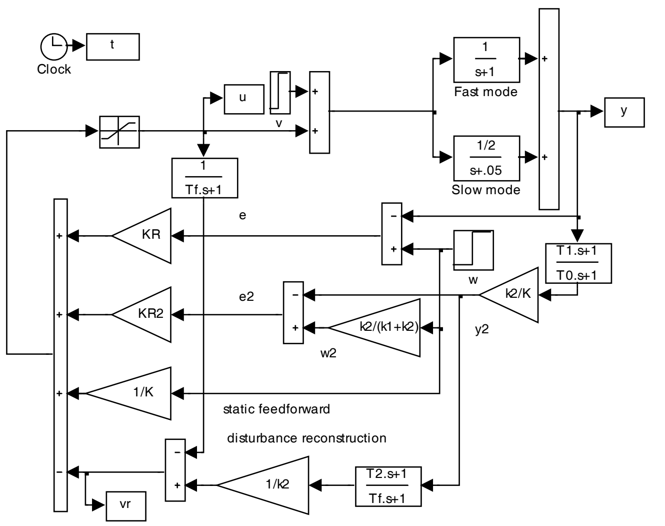

Fig. 3. Simulink model of P-PI controller

2.2 Thermo-optical plant (real-experiment)

The plant used for real experiment was Measurement and Communication System uDAQ28/LT (fig.4.). This product of several years of development offers measurement of 8 process variables (controlled temperature and its filtered value, ambient temperature, controlled light intensity, its filtered value and its derivative, the ventilator speed of rotation and its motor current). The temperature and the light intensity control channels are interconnected by 3 manipulated variables: the bulb voltage (the heat & light source), the light-diode voltage (the light source) and the ventilator voltage (the system cooling). The plant can be easily connected to standard computers via USB, when it enables to work with the sampling periods 40-50 ms and larger.

Within a Matlab/Simulink, or a Scilab/Scicos schemes the plant is represented as a single block, limiting use of costly and complicated software package for the real time control. So, the usual process-computer communication based on standard converter cards (that is also supported) is necessary just for more demanding applications requiring higher sampling frequencies.

Fig. 4. Thermo-optical plant

2.3 *Identification of the plant*

There are several ways how to identify this plant as two parallel first order systems. Despite the problems reported by Åström et al., 1998, we have proposed simple step response method. For the system

$$y(t) = K_1\left(1 - e^{-(t-T_d)/T_1}\right) + K_2\left(1 - e^{-(t-T_d)/T_2}\right) \tag{11}$$

when assuming

$$T_1 >> T_2 >> T_d \tag{12}$$

one gets

$$T_1 = \frac{t_4 - t_3}{\ln \dfrac{y(\infty) - y_3}{y(\infty) - y_4}} \tag{13}$$

$$K_1 = \frac{y_3 - y_4}{e^{-t_4/T_1} - e^{-t_3/T_1}} \tag{14}$$

$$K_2 = y(\infty) - K_1 \tag{15}$$

$$K_2 = T_2 \frac{y_2 - y_1}{t_2 - t_1} \tag{16}$$

The identification depends on the selection of the points, where $t_1, t_2, y_1 = y(t_1), y_2 = y(t_2)$ represent two points from the initial phase of the step response, $t_3, t_4, y_3 = y(t_3), y_4 = y(t_4)$ represent two points from the vicinity of the steady state.

Fig. 5 shows, how the identification fits the measured data.

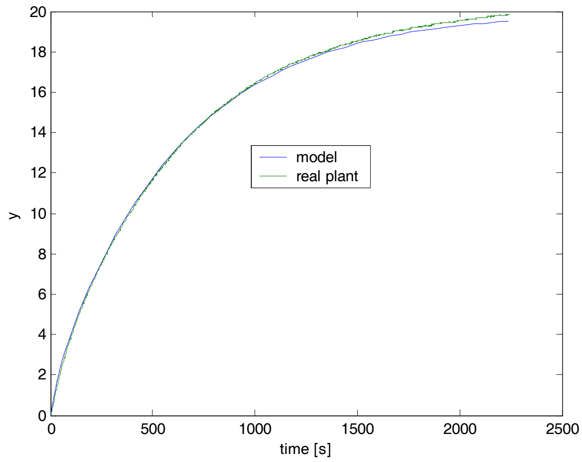

Fig. 5. Thermo-optical plant: Step response comparison of the real plant and the model.

2.4 *Simulation*

Fig. 6 shows the simulation results achieved in Matlab/Simulink. For the relatively "slow" poles the dynamics of the control signal is similar to the first order plants control. When choosing "fast" poles or when the control is close to the time-optimal one (poles are close to minus infinity) the control signal is obviously affected by the zero dynamics of the plant. One interval at the saturation shows that the controller belongs to the DC1.

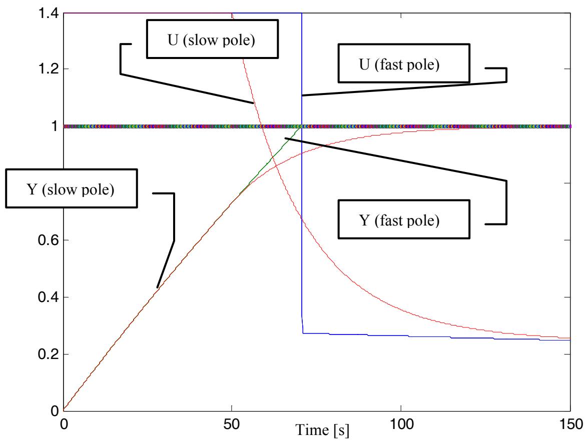

Fig. 6. Control signal and the output for the "fast & slow" poles.

The experiment using the real plant described above has been used to verify the control design. Fig. 7 and 8 show the achieved results, whereby the chosen closed loop pole is $\alpha = -0.03$. The plant has been identified as

$$G(s) = \frac{0.53}{41.9s+1} + \frac{7.5}{610.5s+1} \qquad (17)$$

Several steps of desired value have been used in the experiment. There was a disturbance produced by the 1V step on the fan voltage at the time of 2250s.

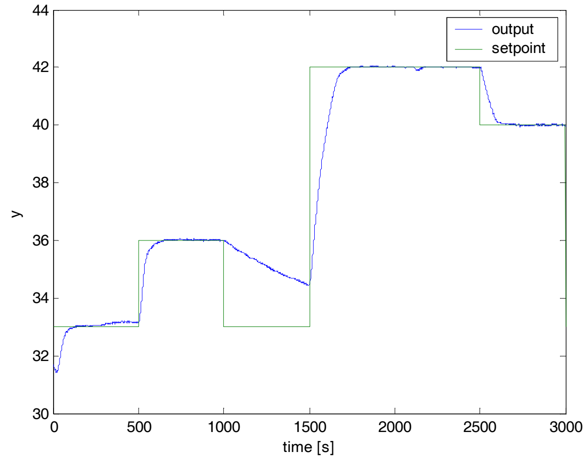

Fig. 7. Real experiment results

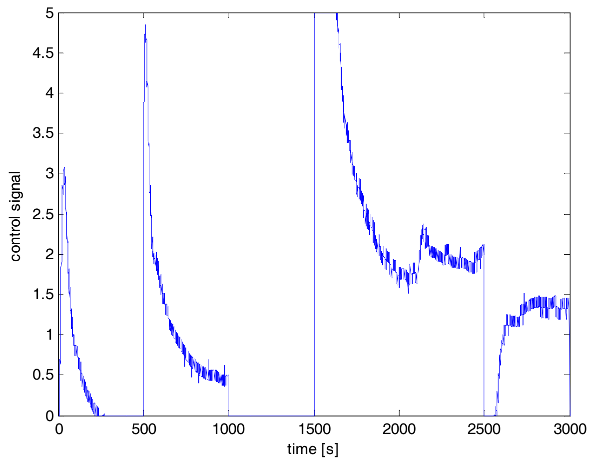

Fig. 8. Real experiment results

The experiment shows that the designed control coincides well with the expected results. There is just a little overshoot during the first setpoint step, caused by the ambient temperature increase. The third setpoint step shows that the plant is non-linear (the dynamics of heating and cooling the plant are not the same). Nevertheless, the zero dynamics is suppressed and the control signal has one interval at the saturation then it converges to desired value with the dynamics given by the closed loop poles. For smaller setpoint steps the control signal does not attack the limit value, however for the larger steps it does.

4. THIRD ORDER PLANT

The control design used for the second order plant (P-P, P-PI controllers) can be easily extended to e.g. third order system. Let us consider the third order plant

$$G(s) = \frac{K_1}{T_1 s+1} + \frac{K_2}{T_2 s+1} + \frac{K_3}{T_3 s+1} \qquad (18)$$

Let y be the output of the system (18), then

$$y = y_1 + y_2 + y_3 = \left(\frac{K_1}{T_1 s+1} + \frac{K_2}{T_2 s+1} + \frac{K_3}{T_3 s+1}\right)u = \\ = K\frac{b_2 s^2 + b_1 s + b_0}{(T_1 s+1)(T_2 s+1)(T_3 s+1)}u \qquad (19)$$

where u is the control signal and K is

$$K = K_1 + K_2 + K_3 \qquad (20)$$

For a piecewise constants setpoint signal, the control error is defined as

$$e = w - y, \dot{e} = -\dot{y} \qquad (21)$$

The pole assignment control is defined by the requirement of a regular control error decrease

$$\dot{e} = \alpha e \qquad (22)$$

In other words

$$-\dot{y}_1 - \dot{y}_2 - \dot{y}_3 = \alpha(w - y) \qquad (23)$$

where α is the chosen closed loop pole. Then one can write

$$-\frac{1}{T_1}(K_1 u - y_1) - \frac{1}{T_2}(K_2 u - y_2) - \frac{1}{T_3}(K_3 u - y_3) = \\ = \alpha(w - y) \qquad (24)$$

$$-\bar{K}u + \frac{1}{T_1}(y - y_2 - y_3) + \frac{1}{T_2}y_2 + \frac{1}{T_3}y_3 = \alpha(w - y) \qquad (25)$$

whereby

$$\bar{K} = \left(\frac{1}{T_1}K_1 + \frac{1}{T_2}K_2 + \frac{1}{T_3}K_3\right) \qquad (26)$$

Then

$$-\alpha(w - y) + \frac{1}{T_1}(y - w + w) - \frac{1}{T_1}(y_2 + y_3) + \frac{1}{T_2}y_2 + \\ + \frac{1}{T_3}y_3 = \bar{K}u \qquad (27)$$

and one can write

$$\bar{K}u = -\left(\alpha + \frac{1}{T_1}\right)(w - y) + \left(\frac{1}{T_1} - \frac{1}{T2}\right)e_2 + \left(\frac{1}{T_1} + \frac{1}{T_3}\right)e_3 + \\ + [K_1 T_2 T_3 + K_2(T_2 T_3 + T_1 - T_2) + K_3(T_2 T_3 + T_1 - T_3)]w \qquad (28)$$

So the control is

$$u = w[K_1 T_2 T_3 + K_2(T_2 T_3 + T_1 - T_2) + K_3(T_2 T_3 + T_1 - T_3)]/\overline{K} + K_R e + K_{R2} e_2 + K_{R3} e_3 \quad (29)$$

which is the parallel P-P-P controller with parameters

$$K_R = -\left(\alpha + \frac{1}{T_1}\right)/\overline{K}$$
$$K_{R2} = \left(\frac{1}{T_1} - \frac{1}{T_2}\right)/\overline{K} \quad (30)$$
$$K_{R3} = \left(\frac{1}{T_1} - \frac{1}{T_3}\right)/\overline{K}$$

In practice there will be mostly measured just the system output. Therefore, the reconstruction of the auxiliary outputs y_2, y_3 is necessary. There are two easy ways to obtain these outputs. They can be reconstructed from the output of the system, or from the control signal. In this case the reconstruction from the system output will be used again.

The simulation results show that by choosing appropriate closed loop poles the control gives the results ranging from the pure linear control to the time optimal one. It could be shown that the achieved controller represents a fundamental solution. The control required by given application is usually specified somewhere between the two limit cases. The desired control signal has one phase at the saturation and then it converges to the steady value with the dynamics given by the closed loop pole. In all cases the output of the system is monotonous and the controller does not need an additional anti-windup circuitry.

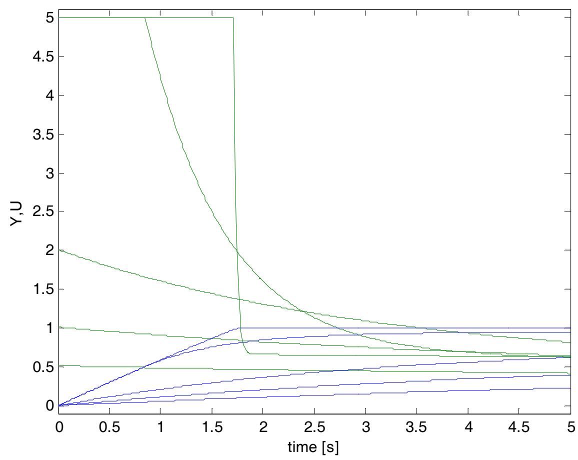

Fig. 9. The simulation results for the »fast« & »slow« poles. The blue curves represent the output and the green ones represent the control signal.

Nevertheless the P-P-P controller can be expanded to P-P-PI controller by adding the observer based disturbance rejection.

ACKNOWLEDGEMENTS

This work was partially supported by:

ESF Project 13120200015 Human Resource Development in Automation and Control

ESF project JPD 3 2005/NP1-047 „PhD students for Modern Industrial Automation in SR", code No. 13120200115

VEGA Project 1/3089/06 Development and Integration of Methods of the Nonlinear System Theory

REFERENCES

Åström, K. J., Panagopoulos, H. and Hägglund, T., 1998: Design of PI Controllers based on Non-Convex Optimization. Automatica, 34, 585–601.

Åström, K. J., Hägglund, T. and Ingimundarson, A, 2003: Criteria for Design of PID Controllers. 2nd IFAC Symposium CSD'03 Bratislava

Feldbaum,A.A., 1965: Optimal control systems. Academic Press, N.York,

Glattfelder,A.H. und Schaufelberger,W., 2003: Control Systems with Input and Output Constraints. Springer, London,

Huba, M., 1990: Saturating Pole Assignement Control-Construction and geometrical interpre-tation on the phase plane, In: 1. IFAC-Workshop „NTDCS", Smolenice, Slovakia.,

Huba, M. and P. Bisták,, 1999: "Dynamic Classes in the PID Control." In: Proceedings of the 1999 American Control Conference. AACC, San Diego

Huba, M, 2006: Theory of Automatic Control 3. STU Bratislava (in Slovak).

Huba M., Šimunek M. (2007). Modular Approach to Teaching PID Control. In: IEEE Transactions on Industrial Electronics, VPL.54 No.6 December pp.3112-3121.

O´Dwyer, A., 2006: Handbook of PI and PID Controllers Tuning Rules. 2nd Edition. Imperial College Press. London.,