15th Triennial World Congress, Barcelona, Spain

ON THE REACHABLE PERFORMANCE/ROBUSTNESS LIMITS FOR LINEAR CONTROL SYSTEMS

L. Keviczky and **Cs. Bányász**

Computer and Automation Research Institute
Hungarian Academy of Sciences
H-1111 Budapest, Kende u 13-17, HUNGARY
Phone: +361-466-5435; Fax: +361-466-7503
e-mail: keviczky@sztaki.hu ; banyasz@sztaki.hu

Abstract: A new decomposition method is presented to handle optimal control design for two-degree of freedom time delay control systems. In this approach exact relationships between the actuator, process and design parameters furthermore the Nyquist stability margin are developed for a first order time delay process. The ultimate robustness limit of any control can be explicitly calculated using this approach. *Copyright©2002 IFAC*

Keywords: two-degree of freedom controller, robustness measure, power surplus

1. INTRODUCTION

A *generic two-degree of freedom* (*G2DF*) system (Keviczky, 1995) was introduced by the authors, which was successfully applied for several linear and *nonlinear* control (*NG2DF*) problems (Haber and Keviczky, 1999). The *G2DF* system is based on the *Youla-parametrization* providing all realizable stabilizing regulators (*ARS*) for open-loop stable plants and on a special structure, which is a certain extension of the well known *IMC* approach.

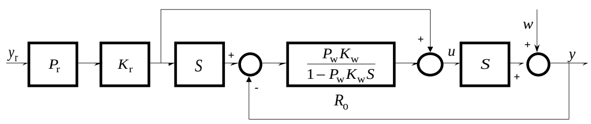

Figure 1 The *generic 2DF* (*G2DF*) control system

A *G2DF* control system is shown in Fig. 1, where y_r, u, y and w are the reference, process input, output and disturbance signals, respectively. The optimal *ARS* regulator (Maciejowski, 1989) of the *G2DF* scheme (Keviczky and Bányász, 1999) is given by

$$R_o = \frac{P_w K_w}{1 - P_w K_w S} = \frac{Q_o}{1 - Q_o S} = \frac{P_w G_w S_+^{-1}}{1 - P_w G_w S_- z^{-d}} \quad (1)$$

where

$$Q_o = Q_w = P_w K_w = P_w G_w S_+^{-1} \quad (2)$$

is the associated *Y-parameter* (Maciejowski, 1989) furthermore

$$Q_r = P_r K_r = P_r G_r S_+^{-1} \ ; \ K_w = G_w S_+^{-1} \ ; \ K_r = G_r S_+^{-1} \quad (3)$$

assuming that the process is factorable as

$$S = S_+ \overline{S}_- = S_+ S_- z^{-d} \quad (4)$$

where S_+ means the inverse stable (*IS*) and S_- the inverse unstable (*IU*) factors, respectively. z^{-d} corresponds to the discrete time delay, which is the integer multiple of the sampling time.

It is interesting to see how the transfer characteristics of this system looks like:

$$\begin{aligned} y &= P_r K_r S y_r - (1 - P_w K_w S) w = \\ &= P_r G_r S_- z^{-d} y_r - \left(1 - P_w G_w S_- z^{-d}\right) w = y_t + y_d \end{aligned} \quad (5)$$

where y_t is the tracking (servo) and y_d is the regulating (or disturbance rejection) independent behaviors of the closed-loop response, respectively. Here P_r and P_w are assumed stable and proper transfer functions, that are partly capable to place desired poles in the servo and the regulatory transfer functions, furthermore they are usually referred as reference signal and output disturbance predictors. They can even be called as reference models, so reasonably $P_r(\omega = 0) = 1$ and $P_w(\omega = 0) = 1$ are selected.

The ultimate optimal goal of any control system could be to exactly follow a prescribed external (usually a unit step) excitation by the (step) response of the closed-loop system. Using the *G2DF* system we required to follow the transients prescribed by P_r and P_w (more exactly $(1-P_w)$), i.e. the ideal overall transfer function of the *G2DF* control system would be

$$y^o = P_r y_r - (1-P_w) w = y_t^o + y_d^o \tag{6}$$

Equation (5) shows that we can not reach these ideal tracking $y_t^o = P_r y_r$ and regulatory behaviors $y_d^o = (1-P_w) w$, because of the uncompensable time-delay and the so-called invariant zeros in the *IU* factor S_-. The reachable best transient is given by $P_r G_r S_- z^{-d}$ and $(1-P_w G_w S_- z^{-d})$ respectively, where G_r and G_w can optimally attenuate the influence of S_-. (Unfortunately S_- and z^{-d} do not depend on the control design and only slightly can be influenced via the proper selection of the sampling time. These factors are basic behaviors of the process, so they can be considerably changed only via certain technological changes.) Express the deviation between the ideal y^o and the best reachable (realizable and optimal) closed-loop (given by (5)) as

$$\begin{aligned}\Delta y &= y^o - y = P_r(1-K_r S) y_r + P_w(1-K_w S) w = \\ &= P_r\left(1-G_r \bar{S}_- z^{-d}\right) y_r + P_w\left(1-G_w \bar{S}_- z^{-d}\right) y_r = \\ &= \mathcal{E}_w y_r + \mathcal{E}_w w = \Delta y_t^o + \Delta y_d^o\end{aligned} \tag{7}$$

So the deviation transfer functions from the ideal ones for both the tracking Δy_t^o and the regulatory properties Δy_d^o have the same structure: $P_x\left(1-G_x \bar{S}_- z^{-d}\right)\Big|_{x=r,w}$. This deviation form is excellent for (sometimes called model matching) optimization of the *generic scheme* as it was shown in (Keviczky and Bányász, 1999). An interesting result was (Keviczky and Bányász, 1999) that the optimization of the *G2DF* scheme can be performed in $\mathcal{H}_2$ and $\mathcal{H}_\infty$ norm spaces by the proper selection of the serial K_r and embedded K_w filters (compensators). Observe that in these optimizations both K_r and K_w use full cancellation of the *IS* factor of the process and the originally quite sophisticated optimization could be reduced to the optimal computation of the G_r and G_w filters (Keviczky and Bányász, 1999). If G_r and G_w are optimally selected, then R_o in (1) denotes the optimal *ARS* regulator. (The reasonable factorization of S means that the *IU* factor S_- is monic. So if the optimal G_w is also monic, then a unity gain selection for P_w - which was assumed above - provides that R_o is integrating, so has a pole at 1. The same considerations can be derived for G_r and P_r.)

Do not forget that the optimal *ARS* regulator R_o is introduced for discrete-time systems, when the above cancellation process does not result in a nonrealizable order condition, so it is not so restrictive, than in the equivalent continuous case. However, in some special cases even the continuous version is applicable, as it will be shown later.

2. A DECOMPOSITION APPROACH FOR CONTROLLER DESIGN

The control error transfer functions of the *G2DF* system are given by

$$\begin{aligned}e &= \left(1-P_r G_r \bar{S}_- z^{-d}\right) y_r - \left(1-P_w G_w \bar{S}_- z^{-d}\right) w = \\ &= E_r y_r + E_w w = E_r y_r + E w = e_r + e_w\end{aligned} \tag{8}$$

where the control tracking performance can be best evaluated from the deviation Δy_r^o

$$\begin{aligned}E_r &= 1-P_r G_r \bar{S}_- z^{-d} = 1-P_r+P_r-P_r G_r S_- z^{-d} = \\ &= (1-P_r) + P_r\left(1-G_r S_- z^{-d}\right) = \mathcal{E}_r^o + \mathcal{E}_r\end{aligned} \tag{9}$$

Here the first term characterizes the design performance: how close P_r is to the ideal unity and the second term characterizes the performance degradation caused by the invariant factors of the process, i.e., how close $P_r G_r S_- z^{-d}$ is to P_r. The form of (9) is a reasonable decomposition of the general controller design paradigm. A corresponding cost function can be constructed by using the triangle inequality and applying an appropriate norm

$$\begin{aligned}J_{\text{control}}^r &\le J_{\text{design}}^r + J_{\text{degradation}}^r = \\ &= \|1-P_r\| + \left\|P_r - P_r G_r S_- z^{-d}\right\|\end{aligned} \tag{10}$$

The same observations can be made for the tracking errors, too. Formally both terms in (8) are the same, however the control sensitivity function is the second one

$$E = E_w = 1-P_w G_w \bar{S}_- z^{-d} \tag{11}$$

which does not equal to the sensitivity function

$$\mathcal{E}_w = P_w\left(1-G_w \bar{S}_- z^{-d}\right) \tag{12}$$

in (7) generating the regulatory model matching error Δy_t^o. A short analysis shows that

$$\begin{aligned}E &= E_w = 1-P_w G_w \bar{S}_- z^{-d} = \\ &= 1-P_w+P_w-P_r G_r S_- z^{-d} = \\ &= (1-P_w) + P_w\left(1-G_w S_- z^{-d}\right) = \mathcal{E}_w^o + \mathcal{E}_w\end{aligned} \tag{13}$$

Similar cost function can be constructed for the regulatory performance, too

$$J_{\text{control}}^{\text{w}} \le J_{\text{design}}^{\text{w}} + J_{\text{degradation}}^{\text{w}} = \\ = \left\| 1 - P_{\text{w}} \right\| + \left\| P_{\text{w}} - P_{\text{w}} G_{\text{w}} S_{-} \, z^{-d} \right\| \tag{14}$$

Summarizing the above results: the control error (resulting either from tracking or from disturbance rejection) is the sum of a design error and a reference model performance degradation error, so the overall control performance is the sum of the design and degradation performances. The authors believe that the relatively easy and reasonably optimal solution of a generally very sophisticated control problem strongly depends on the proper decomposition of the original paradigm. These decompositions would correspond to a natural control engineering practice, too, where the best reachable design goal and the way how to obtain it appear in a generally iterative sequential procedure.

A large percentage of papers suggesting optimal controller design do not follow the above decomposition possibility. Most of them introduce an optimization technique only and stops there. Some, who are familiar with the practical needs, apply further special detuning methods to increase the robustness of the solution or reduce the control action. The basic theoretical optimal design methods usually result in too sensitive controllers and assume no amplitude constraints for the control action signal. However, in the control engineering practice one should always assume a nonlinear limiter, corresponding to a real actuator. In many cases the optimal regulator obtained by sophisticated theoretical methods generates too big control actions (amplitude changes at the output of the regulator). These big changes mostly can not, of course, "go through" the amplitude (sometimes rate) constrained real actuator. Therefore industrial control experts used to laugh at the optimal regulators of theorists because they state that the resulting transient rather depends on the practical limits than the optimality of the design algorithm.

This is the case for almost all dead-beat, pole-cancellation and $\mathcal{H}_2$ optimal regulators, except if this input action is not penalized in the control criterion, which reduces this effect. Therefore the energy of the plant input is generally included in the control cost function at the *LQG* and model-predictive controls. In this way it is generally possible to reduce the variation of the regulator output considerably. Input penalization is always a possible way of detuning. Optimization and detuning is also a certain decomposition approach. However, it is not a simple procedure to find the proper weighting (penalizing) factors and filters in the criterion and there is no easy way to calculate the obtained bandwidth for the closed-loop system. The practice is usually based on a "trial and check" method. The recent advanced methodology tries to fulfill both performance and robustness requirements via special compromising loop-shaping techniques. These techniques can also be considered certain decomposition methodology.

3. NEW RELATIONSHIPS FOR ROBUSTNESS MEASURES

In our recent research, application projects and studies we stick on the above decomposition, which considers the optimal design and optimal performance degradation the two major steps (14). For the minimization of the second terms wide class of solutions exists depending on the applied norm, process and some existing constraints if the process parameters are known. If these parameters are not available, then an iterative combined ID and control technique or its adaptive version can be used. The recent advanced methods try to use the available or assumed plant uncertainties in the optimization.

Relatively much less papers deal with the optimization of the first term. The simple formal description of this paradigm is

$$P_{\text{r}}^{\text{opt}} = \arg\left\{ \min_{P_{\text{r}}} \left(J_{\text{design}}^{\text{r}} \right) \Big|_{u \in \mathcal{U}} \right\} = \arg\left\{ \min_{P_{\text{r}}} \left\| 1 - P_{\text{r}} \right\| \Big|_{u \in \mathcal{U}} \right\} \\ P_{\text{w}}^{\text{opt}} = \arg\left\{ \min_{P_{\text{w}}} \left(J_{\text{design}}^{\text{w}} \right) \Big|_{u \in \mathcal{U}} \right\} = \arg\left\{ \min_{P_{\text{w}}} \left\| 1 - P_{\text{w}} \right\| \Big|_{u \in \mathcal{U}} \right\} \tag{15}$$

where $\mathcal{U}$ is the (mostly amplitude: $\mathcal{U}: |u| \le 1$) constrained input signal domain. A nonlinear limiter representing a real actuator is always an important source of considerable performance degradation comparing to the original linear optimal system designed. The general solution of this paradigm is a certain rescaling of the nonlinear practical system to achieve a linear operational domain. The performance degradation caused by the limiting actuator and by the necessary rescaling is calculable and known, if we use a reference model redesign technique to fulfill the amplitude constraints requirements and to solve the paradigm (Keviczky and Bányász, 1997).

Using the iterative reference model redesign technique the fastest reference model (so the highest closed-loop bandwidth) can be found reachable within the linear operational range. The situation is more complex, because changing the reference model the robustness margins of the closed-loop also change. It would be desirable to know how the limiting robustness measure depend on the limiting reference model. In the sequel this relationship will be investigated.

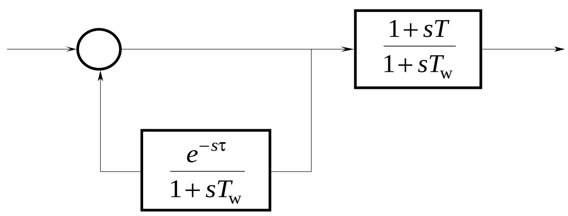

Figure 2 The simple realization of the optimal continuous-time *ARS* regulator

In our analysis the continuous time equivalent of the *G2DF* will be used, where the following simple assumptions are used:

$$P_w = 1/(1+sT_w) \quad ; \quad S = e^{-s\tau}/(1+sT) \tag{16}$$

so the *IS* process is a first order time delay lag and the reference model is a first order lag. The continuous-time optimal *ARS* regulator based on (1) is now

$$R_o = \frac{1}{1-P_w e^{-s\tau}}\left(P_w S_+^{-1}\right) = \frac{1}{1-e^{-s\tau}/(1+sT_w)}\frac{1+sT}{1+T_w} \tag{17}$$

which can be easily realized, e.g., by a simple closed-loop according to Fig. 2. Note that R_o has a pole at $s=0$, so it is an integrating regulator.

It is easy to compute that the open-loop transfer function for the *G2DF* system is

$$Y = R_o S = \frac{P_w}{1-P_w e^{-s\tau}} \tag{18}$$

so the crossover frequency ω_c can be obtained from the condition $|Y|=1$, i.e., when

$$\left|P_w e^{-s\tau}\right| = \left|1-P_w e^{-s\tau}\right| \tag{19}$$

For the ideal no delay case this condition means that

$$\left|P_w\right| = \left|1-P_w\right| \tag{20}$$

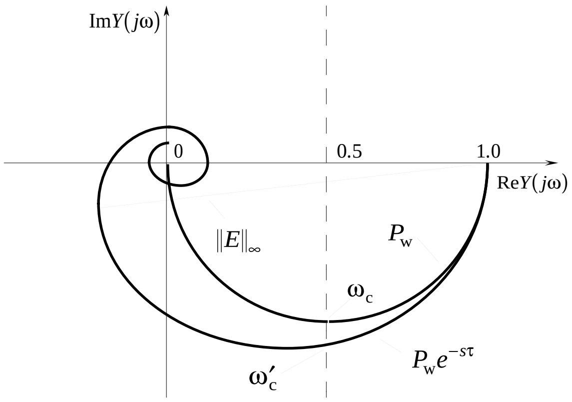

Figure 3 The demonstration of the change of the crossover frequency ω_c to ω_c' for the uncompensable process time delay $e^{-s\tau}$

This equation gives a very simple geometrical tool to determine the crossover frequency ω_c, demonstrated on Fig. 3. This method gives ω_c by the intersection of $P_w(j\omega)$ and the vertical line drawn at the point $(0.5+0j)$, where $\left|P_w(j\omega_c)\right|$ equals to the distance from the point $(1+0j)$, i.e., $\left|1-P_w(j\omega_c)\right|$. This method can also be applied to determine the distorted ω_c' crossover frequency for the uncompensable process time $e^{-s\tau}$ using $P_w(j\omega)e^{-j\omega\tau}$.

The corresponding sensitivity function (11) of the *G2DF* system is

$$E = 1-P_w e^{-s\tau} = 1-\frac{1}{1+sT_w}e^{-s\tau} = \frac{1+sT_w-e^{-s\tau}}{1+sT_w} \tag{21}$$

The $\|E\|_\infty$ of the sensitivity function can also be determined graphically on Fig. 3, which is the farthest distance of $P_w(j\omega)e^{-j\omega\tau}$ from the point $(1+0j)$.

Let us denote the well-known robustness measure: the distance between the point $(-1+0j)$ and any $Y(j\omega)$ point of the Nyquist curve of the open-loop frequency characteristics by ρ

$$\rho = \rho(\omega,R) = |1+RS| = \frac{1}{|E(\omega,R)|} \tag{22}$$

Because ρ changes by ω and the shape of $\rho(\omega)$ is difficult to characterize by one scalar indicator, therefore the real stability/robustness measure is

$$\rho_m = \rho_{min}(R) = \min_\omega|\rho(\omega,R)| = \min_\omega|1+RS| = \frac{1}{\|E\|_\infty} \tag{23}$$

which is the distance between the point $(-1+0j)$ and the closest point of $Y(j\omega)$ and the reciprocal value of the norm $\|E\|_\infty$. (Note that in the most general case of $|Y(\omega=\infty)|=0$ this measure falls into the range $0 \le \rho_m \le 1$ for stable closed-loops.) It can be well seen in Fig. 3 that the real part of $P_w(j\omega)e^{-j\omega\tau}$ at first the intersection at $\bar{\omega}$ with the real axis can be used as a lower limit for $\|E\|_\infty$, so its reciprocal value is an appropriate upper limit for $\rho_m = \rho_{min}$

$$\bar{\rho}_m \le \frac{1}{1-\mathrm{Re}\left\{P_w(j\bar{\omega})e^{-j\bar{\omega}\tau}\right\}} \tag{24}$$

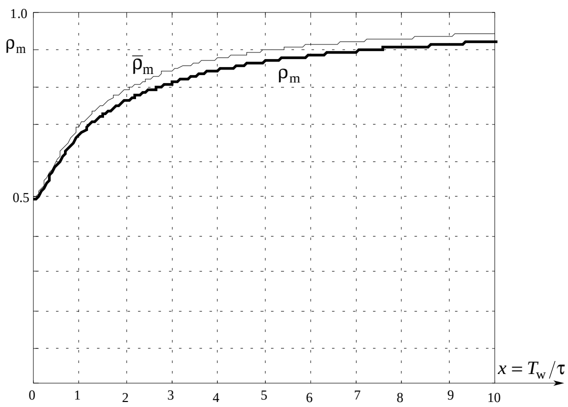

Figure 4 The reachable robustness measure ρ_m and a practical upper bound

One can see from Fig. 3 that ρ_m depends only on our design goal (T_w) and on the process behavior time delay (τ), more exactly on their relative value $x = T_w/\tau$. Unfortunately there is no simple analytical solution to obtain the relationship $\rho_m(x)$, (and $\bar{\rho}_m(x)$) only a numerical procedure can be applied following the graphical interpretation for $\|E\|_\infty$ in Fig. 3. The $\rho_m(x)$ and $\bar{\rho}_m(x)$ curves obtained by MATLAB numerical calculations are plotted on Fig. 4.

The interpretation of $\rho_m(x)$ is very important, because this curve gives the theoretically best reachable robustness measure with any controller for an arbitrary *IS* time-delay plant. This measure is $\rho_m(0) = 0.5$ for cases when the P_w reference model requires a very fast transient response from the time-delay process and the measure is $\rho_m(\infty) = 1$, if the time-delay is negligible comparing to the time-delay of P_w.

A typical time-response of a *PID* regulator for a square-wave input excitation is shown in Fig. 5 for continuous-time case. Here the steady-state value $\Delta u^+(\infty) = -\Delta u^-(\infty)$ is the virtual gain of the regulation between the excitation and the control action. The ratio of the initial peak $\Delta u^+(0)$ to the change $2\Delta u^+(\infty)$ in the steady-state value

$$p_s = \frac{\Delta u^+(0)}{\Delta u^+(\infty) - \Delta u^-(\infty)} = \frac{\Delta u^+(0)}{2\Delta u^+(\infty)} = \\ = \frac{u^+_{max}(0) + \Delta u^+(\infty)}{2\Delta u^+(\infty)} = \frac{1}{\alpha} \tag{25}$$

is sometimes called the power-surplus or the virtual differential effect of the regulator. Its reverse α gives the ratio of the time lag and the differential time of a D-effect $\left(sT_D / (1 + s\alpha T_D)\right)$.

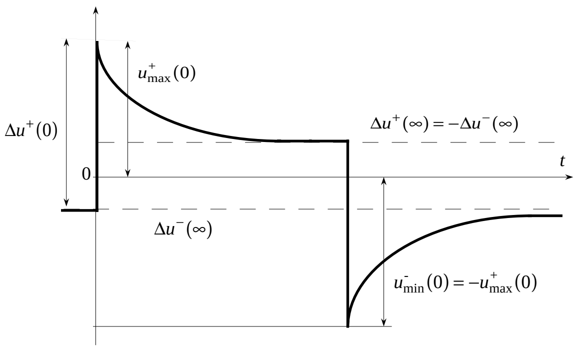

Figure 5. A typical time-response of a *PID* regulator for a square-wave reference signal excitation in a closed-loop

It is easy to see from the form of R_o in our first order example that the initial peak for a unit step excitation is $R_o(\omega = \infty) = T/T_w$, so the power surplus is

$$p_s = T/T_w \tag{26}$$

which comes from a simple physical interpretation: we should like to speed up the closed-loop from the original open-loop bandwidth $\omega_{b,o} = 1/T$ to the desired closed-loop reference bandwidth $\omega_{b,c} = 1/T_w$, therefore p_s is their ratio

$$p_s = \omega_{b,c}/\omega_{b,o} = T/T_w \tag{27}$$

Introducing an auxiliary variable $y = T_w/T$ it is possible to draw a complex four quadrant figure representing the relationships between ρ_m, x, y and p_s parametrized by the ratio T/τ as the Fig. 6 shows.

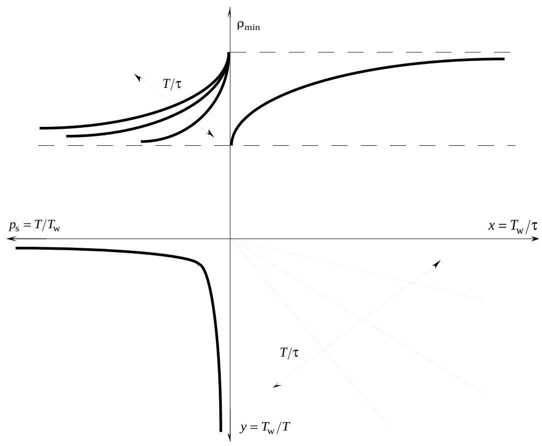

Figure 6. Complex relationships between ρ_m, x, y and p_s parametrized by the ratio T/τ

In a practical application the bottom two quadrants are not necessary, these stand here only for explanation. The numerically computed exact values are plotted in Fig. 7.

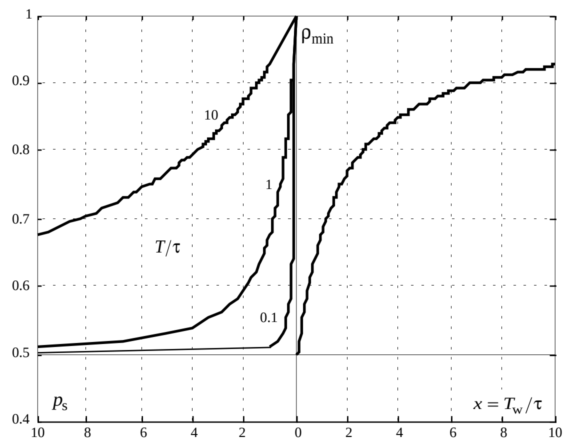

Figure 7. Dependence of ρ_m from x, p_s and T/τ

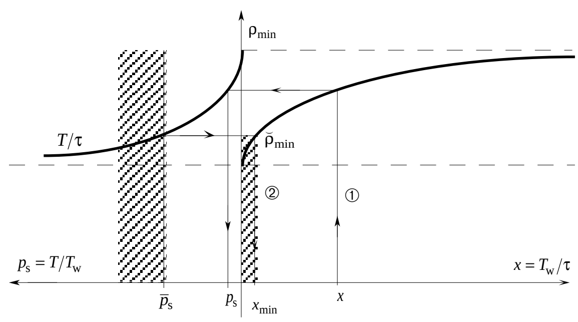

Figure 8. Two possible design problems

The first possibility ① to use Fig. 7 is how one can determine the necessary power surplus p_s to a required design goal $x = T_w/\tau$ for a given process characterized by T/τ. The another possibility ② is when the technically allowable maximal power surplus $\bar{p}_s$ is given and one can determine the fastest possible reference model $x_{\min}$ and the corresponding worst (smallest) minimal robustness measure $x_{\min}$. These two design problems are represented in Fig. 8. It is also possible that $\breve{\rho}_{\min}$ is given and we need the corresponding $x_{\min}$ and $\bar{p}_{\mathrm{s}}$.

4. EXAMPLE

Assume a first order time delay plant as

$$S = e^{-5s}/(1+5s) \qquad (28)$$

where $T/\tau = 1$ and use the design goals

$$P_{\mathrm{r}} = 1/(1+s) \qquad \text{and} \qquad P_{\mathrm{w}} = 1/(1+2.5s) \qquad (29)$$

so in this case $x = T_{\mathrm{w}}/\tau = 0.5$. Using Fig. 7 $p_{\mathrm{s}} = 2.5$ and $\rho_{\mathrm{m}} = 0.56$ correspond to these process and design parameters. Therefore if we need higher robustness value ρ_{m} it can only be reached by applying a slower reference model P_{w}. Assume a unit-step reference signal excitation $y_{\mathrm{r}} = 1(t-5)$ furthermore a step output disturbance $w = 0.5 \times 1(t-25)$ at the *G2DF* closed-loop. (Here $1(t)$ is the classical unit-step signal, i.e. $1(t) = 1$ for $t \geq 0$ and $1(t) = 0$ for $t < 0$.)

Fig. 9 shows the output response $y(t)$ of the closed-loop for the $y_{\mathrm{r}}(t)$ and $w(t)$ excitations. The output of the regulator is shown in Fig 10, where it is easy to see that $p_{\mathrm{s}} = 2$ ($0.5p_{\mathrm{s}}$ is shown in the figure, because the amplitude of $w(t)$ was 0.5 !!!). Do not miss p_{s} with the power surplus $p'_{\mathrm{s}} = 5$ necessary to P_{r}, which does not depend on the closed-loop properties directly and independent of p_{s}.

5. CONCLUSIONS

The full cross relationships of the most important actuator, process parameters and robustness measures are presented, according to the decomposition approach discussed previously. The developed plots are very important, because they give the ultimate control limits reachable by any regulator. The assumed plant was a simple first order time delay lag and the question arises what can we say for higher order and nonminimum phase plants. In case of a higher order plant one can always use the dominant time constant as T in these investigations. All further lag term (higher order denominator in S) makes the situation worst lowering the $\rho_{\mathrm{m}}(x)$ curve (decreasing the robustness of the closed-loop). The influence of minimum phase lead terms (higher order numerator in S) improves the situation by increasing the $\rho_{\mathrm{m}}(x)$ curve. The influence of *IU* non-minimum phase lead terms (higher order unstable numerator in S) have the same effects as further lags in the denominator of S. This is how the limiting character of the above results should be interpreted.

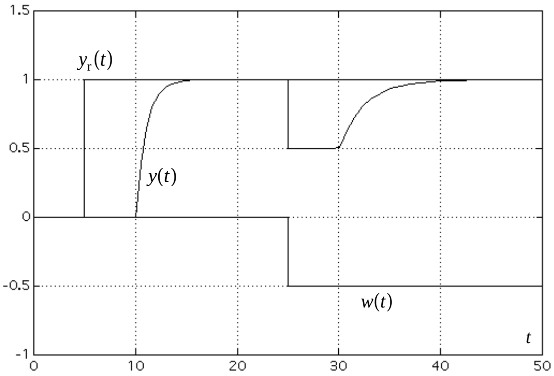

Figure 9. Model tracking (P_{r},) and disturbance rejection (P_{w}) properties shown by $y(t)$

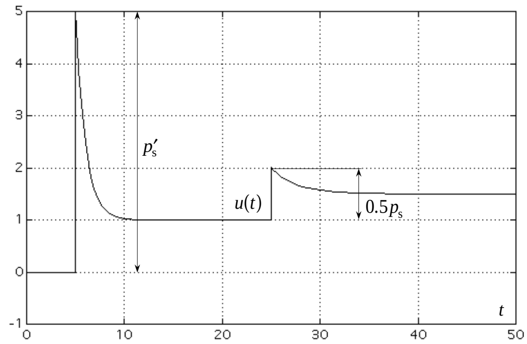

Figure 10. The output of the regulator $u(t)$ with p_{s} and p'_{s}

REFERENCES

Haber, R. and L. Keviczky (1999). *Nonlinear System Identification — Input-Output Modeling Approach*. Kluwer Academic Publishers.

Keviczky, L. (1995). Combined identification and control: another way. (Invited plenary paper.) *5th IFAC Symp. on Adaptive Control and Signal Processing, ACASP'95*, 13-30, Budapest, H.

Keviczky, L. and Cs. Bányász (1997). An iterative redesign technique of reference models: How to reach the maximal bandwidth? *11th IFAC Symposium on System Identification SYSID'97*, Fukuoka, Japan, 619-624.

Keviczky, L. and Cs. Bányász (1999). Optimality of two-degree of freedom controllers in $\mathcal{H}_2$- and $\mathcal{H}_\infty$-norm space, their robustness and minimal sensitivity. *14th IFAC World Congress*, **F**, 331-336, Beijing, PRC.

Maciejowski, J.M. (1989). *Multivariable Feedback Design*, Addison Wesley.

This work was supported by the Hungarian NSF (OTKA).