Appl. Math. and Comp. Sci., 1996, vol.6, No.4, 707–732

DISCRETE-TIME AND CONTINUOUS-TIME GENERALISED PREDICTIVE CONTROLLERS WITH ANTICIPATED FILTRATION: TUNING RULES

ZDZISŁAW KOWALCZUK*, PIOTR SUCHOMSKI*
ANDRZEJ MARCIŃCZYK*

A general class of control algorithms based on the generalised predictive control (GPC) strategy with anticipated filtering (AF) of the control error is considered in both the discrete- and continuous-time domains. It is shown that in the discrete-time settings, under certain conditions, a solution of the AF-GPC design always exists and the design leads to stable control systems with definite closed-loop characteristics. The plant cancellation issue is taken into account. Conditions for the existence of the solution of the GPC design and the corresponding rules of tuning of the resulting controller are given. A suitable iterative procedure for a simultaneous determination of the AF-GPC design parameters (the control horizon and the order of plant cancellation, as well as the controller gain) and a root locus interpretation of the design are also supplied. The continuous-time predictive control (CGPC) has properties similar to those of the discrete-time GPC strategy. It is shown that the idea of using the anticipated filtering approach to the GPC design can also be effectively applied in the continuous-time restatement. Rules for tuning the CGPC controller, which are based on parameters of a system rate of reaction, identified by a starting phase of system step response, are given and shown to be practically effective. With the anticipated filtering, applied in both the discrete- and continuous-time frameworks, the excitation of the closed-loop system is suitably abated by performing a moderating filtration in the anticipated-time domain. The pertinence of the anticipated filtering lies in shaping the closed-loop characteristics of the control system, reducing the disagreeable control effort and, consequently, based on a certain balance obtained in the cost function, in making the λ-tuning more practicable. The proposed tuning rules are validated via simulation experiments.

1. Introduction

The development and research of identification and control procedures, including those for adaptive control systems, in the discrete-time domain took considerable attention in the literature (see e.g. Clark *et al.*, 1987a; 1987b; Clark and Mohtadi, 1987; Favier, 1987; Gorez *et al.*, 1987; Kowalczuk, 1992a). The general predictive

* Automatic Control Department, Faculty of Electronics, Technical University of Gdańsk, ul. Narutowicza 11/12, P.O.Box 612, 80–952 Gdańsk, Poland, e-mail: kova@sunrise.pg.gda.pl

control (GPC) strategy, discussed in this paper, that uses a multi-step cost function has superior robustness in comparison with some other control methods (such as the GMV, the pole-placement, etc.). Therefore, the GPC strategy is a good candidate for modern control system design procedures. The purpose of additional filtration of the control error, which was originally proposed for the GPC design in (Clark *et al.*, 1987a; 1987b; Clark and Mohtadi, 1987; Demircioglu and Gawthrop, 1991), is to abate the excitation of the closed-loop system. The filtration is performed in an anticipation-time domain and is referred to as the anticipated filtering (AF).

This approach has recently been exercised in (Kowalczuk and Suchomski, 1995a; 1995b), where it has been shown that with the AF approach, under certain conditions including the cancellation issue, the solution to the GPC design always exists and the design leads to stable control systems with a definite closed-loop characteristics. Moreover, it has been demonstrated that some bounds on the GPC design parameters have to be taken into account. The effectiveness of this approach stems from the possibility of using an iterative procedure, which has been proposed for solution to the problem of determining some design parameters (the control horizons and the order of plant cancellation) and the controller gain itself.

There are, however, a number of hindrances connected to the discrete-time approach to design. The loss of information on the relative order of the process, the residual delay, the choice of the sampling time and the non-minimum phase property, as well as the effects of roots clustering and the resulting system parameter sensitivity can be enumerated here.

The above factors have had a considerable impact on the recently observed restoration of interest in the continuous-time approach to the design of digital control systems, where the fundamental design is carried out in the continuous-time domain prior to digital mechanisation (Kowalczuk, 1991, 1993, 1994).

A continuous-time version of the general predictive control, referred to as the CGPC (Demircioglu and Gawthrop, 1991), seems to be worth of consideration for adaptive purposes. It has properties similar to those of the discrete-time GPC strategy and can be suitable for stable, unstable, minimum and non-minimum phase systems.

An important feature of that approach both in the discrete- and continuous-time domains is that the anticipated filtering makes it possible to reduce the disagreeable control effort associated with GPC and to make the λ-tuning more practicable.

2. The Discrete-Time AF-GPC Design Principles

We consider the following CARIMA model of a linear system:

$$A(q^{-1})y(n) = B(q^{-1})u(n) + \Delta^{-1}C(q^{-1})e(n) \tag{1}$$

where n is the discrete-time index, $\{u(n)\}$ and $\{y(n)\}$ are the input and output of the controlled system, $\{e(n)\}$ is a zero-mean white-noise disturbance, q^{-1} is the backward shift operator, $\Delta = 1 - q^{-1}$ is the difference operator, and

$$A(q^{-1}) = 1 + a_1 q^{-1} + \cdots + a_{N_A} q^{-N_A} \tag{2}$$

$$C(q^{-1}) = b_1 q^{-1} + \cdots + b_{N_B} q^{-N_B} \tag{3}$$

$$A(q^{-1}) = 1 + c_1 q^{-1} + \cdots + c_{N_C} q^{-N_C} \tag{4}$$

and a cost function of the form

$$J = E\left\{ \sum_{i=N_1}^{N_2} \Big[y(n+i) - w(n+i)\Big]^2 + \lambda \sum_{i=1}^{N_U} \Delta u(n+i-1)^2 \right\} \tag{5}$$

where $\{w(n)\}$ denotes the reference sequence, N_1 and N_2 are the bottom and top of the observation horizon, respectively, N_U is the control horizon, $\lambda \geq 0$ is a control weighting factor, and E is the expectation operator conditioned on data up to time n. To facilitate further discussion, let us also introduce an auxiliary notion of the effective observation horizon $N_0 = N_2 - N_1 + 1$. The optimal, in the minimum variance sense, i-step ahead predictor of y is given by (Clark and Mohtadi, 1987; Favier, 1987)

$$\widehat{y}(n+i) = H_i(q^{-1})\Delta u(n+i-1) + \widehat{y}(n+i \,|\, n) \tag{6}$$

where $\widehat{y}(n+i \,|\, n)$ satisfies the following equation:

$$C(q^{-1})\widehat{y}(n+i \,|\, n) = G_i(q^{-1})y(n) + E_i(q^{-1})\Delta u(n-1) \tag{7}$$

The polynomials H_i, G_i and E_i can be obtained from the Diophantine equations (Clark and Mohtadi, 1987; Favier, 1987; Gorez *et al.*, 1987)

$$C(q^{-1}) = \widehat{A}(q^{-1})F_i(q^{-1}) + q^{-i}G_i(q^{-1}) \tag{8}$$

$$F_i(q^{-1})\overline{B}(q^{-1}) = C(q^{-1})H_i(q^{-1}) + q^{-i}E_i(q^{-1}) \tag{9}$$

where

$$\widehat{A}(q^{-1}) = \Delta A(q^{-1}), \quad \overline{B}(q^{-1}) = qB(q^{-1})$$

$$\deg E_i(q^{-1}) = \max(N_B - 2, N_C - 1), \quad \deg F_i(q^{-1}) = i - 1$$

$$\deg G_i(q^{-1}) = N_A, \quad \deg H_i(q^{-1}) = i - 1$$

Note that for discrete-time models of continuous-time systems without a transportation delay we have $\deg E_i(q^{-1}) = N_A - 2$.

Assuming that $\Delta u(n+i-1) = 0$ for $i > N_U$, where N_U denotes a control horizon, one obtains (Clark *et al.*, 1987a; 1987b; Clark and Mohtadi, 1987) the optimal incremental control $\Delta \boldsymbol{u}(n)$ that minimises (5)

$$\Delta \boldsymbol{u}(n) = \boldsymbol{K}\Big(\boldsymbol{w}(n) - \widehat{\boldsymbol{y}}(n)\Big) \tag{10}$$

$$\boldsymbol{K} = (\boldsymbol{H}^T\boldsymbol{H} + \lambda \boldsymbol{I})^{-1}\boldsymbol{H}^T \tag{11}$$

where $\Delta \boldsymbol{u}(n) \in \mathbb{R}^{N_U}$, $\boldsymbol{K} \in \mathbb{R}^{N_U \times N_0}$, $\boldsymbol{w}(n) \in \mathbb{R}^{N_0}$, $\widehat{\boldsymbol{y}}(n) \in \mathbb{R}^{N_0}$, $\boldsymbol{H} \in \mathbb{R}^{N_0 \times N_U}$ and

$$\begin{aligned} \Delta \boldsymbol{u}(n) &= \Big[\Delta u(n) \ \cdots \ \Delta u(n+N_U-1)\Big]^T \\ \boldsymbol{w}(n) &= \Big[w(n+N_1) \ \cdots \ w(n+N_2)\Big]^T \\ \widehat{\boldsymbol{y}}(n) &= \Big[\widehat{y}(n+N_1 \,|\, n) \ \cdots \ \widehat{y}(n+N_2 \,|\, n)\Big]^T \end{aligned} \tag{12}$$

$$\boldsymbol{H} = \begin{bmatrix} h_{N_1-1} & h_{N_1-2} & \cdots & h_{N_1-N_U} \\ h_{N_1} & h_{N_1-1} & \cdots & h_{N_1-N_U+1} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_2-1} & h_{N_2-2} & \cdots & h_{N_2-N_U} \end{bmatrix} \tag{13}$$

with $h_j = 0$ for $j < 0$.

Let $\boldsymbol{k} \in \mathbb{R}^{N_0}$ be a column vector formed of the first row of $\boldsymbol{K}$

$$\boldsymbol{k}^T = \Big[k_1 \ \cdots \ k_{N_2-N_1+1}\Big] \tag{14}$$

Then the GPC control law is given by the formula

$$\Delta u(n) = \boldsymbol{k}^T\Big(\boldsymbol{w}(n) - \widehat{\boldsymbol{y}}(n)\Big) \tag{15}$$

Note that only the first element of all N_U elements of $\Delta \boldsymbol{u}(n)$ is used as the control input. That means that a *relative range of realization* of the control sequence is $RRR = 1/N_U$.

Assuming that the future set point is known, i.e. $w(n+1) = w(n)$ for $i = 1, 2, \ldots$, we define a reference signal, resulting from anticipated filtering (Kowalczuk and Suchomski, 1995a; 1995b)

$$w^*(n+i) = r_i\Big(w(n) - y(n)\Big) \tag{16}$$

where r_i, $i = 1, 2, \ldots$ are the coefficients of the step response of the filter used in anticipation. The above reference signal goes from the current output $y(n)$ to $w(n)$ as illustrated in Fig. 1.

Now the objective of the design is to drive the predicted output to the reference signal $w^*(n+i)$, $i = 1, 2, \ldots$. Consider therefore the modified cost function

$$J = E\left\{ \sum_{i=N_1}^{N_2} \Big[y^*(n+i) - w^*(n+i)\Big]^2 + \lambda \sum_{i=1}^{N_U} \Delta u(n+i-1)^2 \right\} \tag{17}$$

where $y^*(n+i) = y(n+i) - y(n)$. Minimisation of the above criterion yields the following control action:

$$\Delta u(n) = \boldsymbol{k}^T\Big(\boldsymbol{w}^*(n) - \widehat{\boldsymbol{y}}^*(n)\Big) \tag{18}$$

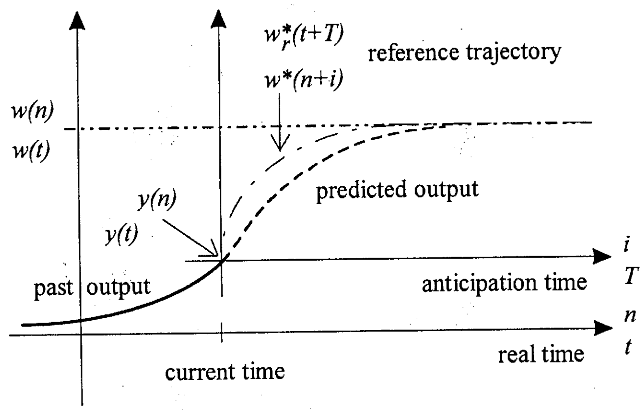


Fig. 1. Anticipation of the set point signal in discrete- and continuous-time.

where the vectors

$$\begin{cases} \boldsymbol{w}^*(n) = \Big(w(n) - y(n)\Big)\boldsymbol{r} = \Big(w(n) - y(n)\Big)\Big[r_{N_1} \cdots r_{N_2}\Big]^T \\ \widehat{\boldsymbol{y}}^*(n) = \Big[\widehat{y}(n + N_1 \,|\, n) - y(n) \;\cdots\; \widehat{y}(n + N_2 \,|\, n) - y(n)\Big]^T \end{cases} \tag{19}$$

are from $\mathbb{R}^{N_0}$, $N_0 = N_2 - N_1 + 1$.

By virtue of (7) it can be shown that the incremental control law (18) can be expressed by

$$C(q^{-1})\Delta u(n) = gC(q^{-1})\Big(w(n) - y(n)\Big) - L(q^{-1})\Delta u(n) - M(q^{-1})y(n) \tag{20}$$

where

$$g = \sum_{i=1}^{N_0} k_i r_{N_1+i-1} \tag{21}$$

$$L(q^{-1}) = q^{-1}\sum_{i=1}^{N_0} k_i E_{N_1+i-1}(q^{-1}) \tag{22}$$

$$M(q^{-1}) = \sum_{i=1}^{N_0} k_i\Big(G_{N_1+i-1}(q^{-1}) - C(q^{-1})\Big) \tag{23}$$

and $\deg L(q^{-1}) = \max(N_B - 1,\, N_C)$, $\deg M(q^{-1}) = \max(N_A,\, N_C)$.

From (20) it follows that the closed-loop characteristic polynomial $D(q^{-1})$ of the resulting GPC control system shown in Fig. 2 is given by

$$\begin{aligned} D(q^{-1}) &= \widehat{A}(q^{-1})C(q^{-1}) + gB(q^{-1})C(q^{-1}) \\ &\quad + \widehat{A}(q^{-1})L(q^{-1}) + B(q^{-1})M(q^{-1}) \end{aligned} \tag{24}$$

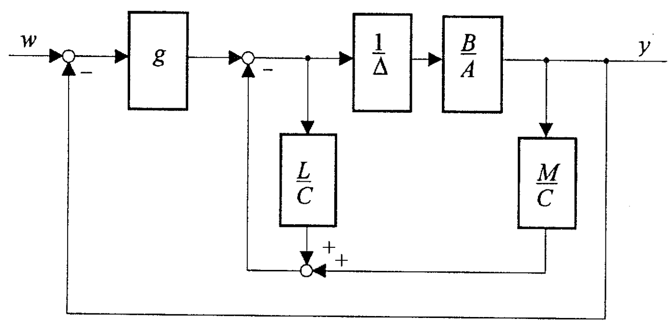


Fig. 2. The GPC system structure.

Finally, by virtue of (8) and (9)

$$D(q^{-1}) = C(q^{-1})\widetilde{D}(q^{-1}) \tag{25}$$

where

$$\widetilde{D}(q^{-1}) = \widehat{A}(q^{-1}) - q^{-1}\widetilde{A}(q^{-1}) + g^*B(q^{-1})$$

$$g^* = g - \sum_{i=1}^{N_0} k_i = \sum_{i=1}^{N_2-N_1+1} k_i(r_{N_1+i-1} - 1) \tag{26}$$

$$\widetilde{A}(q^{-1}) = \sum_{i=1}^{N_0} k_i\Big(\widehat{A}(q^{-1})H_{N_1+i-1}(q^{-1}) - \overline{B}(q^{-1})\Big)q^{N_1+i-1}$$

Note that because $C(q^{-1})$ is assumed to be stable, the closed loop system is stable if and only if $\widetilde{D}(q^{-1})$ is stable.

3. Parameter Properties and Tuning of the AF-GPC Controller

For $\lambda = 0$ two cases, distinguished in the following two Theorems (1 and 2), will be considered. The first case concerns relatively prime polynomials $\widehat{A}(q^{-1})$ and $B(q^{-1})$ and the other treats those polynomials, which have a common factor.

Note that an interesting relationship between $\widehat{A}(q^{-1})$ and $\overline{B}(q^{-1})$ of (9) and $\{h_k\}$, which is the sequence of the Markov parameters of the open-loop system $\overline{B}(q^{-1})/\widehat{A}(q^{-1})$, can be demonstrated by introducing a double infinite lower-triangular Toeplitz matrix, in which $\boldsymbol{H}$ of (13) is included as a submatrix

$$\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_{N_B} \\ 0 \\ \vdots \end{bmatrix} = \begin{bmatrix} h_0 & 0 & 0 & \cdots \\ h_1 & h_0 & 0 & \cdots \\ h_2 & h_1 & h_0 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} 1 \\ \widehat{a}_1 \\ \vdots \\ \widehat{a}_{N_A+1} \\ 0 \\ \vdots \end{bmatrix} \tag{27}$$

It has been shown (Kowalczuk and Suchomski, 1995a; 1995b) that the anticipated filtering approach to the GPC design leads to stable control systems with a desired corresponding closed-loop pole placement (for $\lambda = 0$) and that under certain conditions the solution (11) always exists. The main results, summarised in the form of two theorems, are given and discussed below (see Appendix A for the derivation).

Theorem 1. *If $\widehat{A}(q^{-1})$ and $B(q^{-1})$ are relatively prime and one of the following set of conditions*

(C1′)	$N_2 \geq N_1 + N_U - 1$		(C1″)	$N_2 \geq N_1 + N_U - 1$
(C2′)	$N_1 \geq N_B$	*or*	(C2″)	$N_1 = N_B$
(C3′)	$N_U = N_A + 1$		(C3″)	$N_U \geq N_A + 1$

is satisfied, then $\boldsymbol{H}$ has full column rank ($\operatorname{rank} \boldsymbol{H} = N_U$)*, the solution (11) exists, and the closed-loop characteristic polynomial $D(q^{-1})$ is determined by*

$$D(q^{-1}) = C(q^{-1})\widetilde{D}(q^{-1}) \tag{28}$$

$$\widetilde{D}(q^{-1}) = 1 + g^* B(q^{-1}) \tag{29}$$

Remarks:

1. Note that the stability of the closed-loop control system can be regulated via the anticipated filter and that in the special case of $r_i = 1,\ i = N_1, \ldots, N_2$ in (26) the characteristic polynomial $D(q^{-1})$ is completely determined by the observer polynomial $C(q^{-1})$

$$D(q^{-1}) = C(q^{-1}) \tag{30}$$

2. Setting N_1 to the value of the plant delay (κ) does not guarantee solvability of the design problem (11) and (28)–(29) because Theorem 1 states that for this purpose (with $\lambda = 0$) it is sufficient that N_1 is not less than the order of the numerator N_B. Thus it can be interesting to notice that—by using this suggestion—we simply reject the first $N_B - \kappa$ nonzero samples. Consequently, we can interpret

$$N_1 = N_B$$

the lowest number of the samples in the output sequence that carry the most "essential information" necessary for the design solution, as an information boundary.

3. With reference to the first set of triple condition C1′ it is now evident that in order to assure the main result of Theorem 1 one has to set

$$N_U = N_A + 1$$

Note that in the case of a parsimonious choice of $N_1 = N_B$ (see the second triple — C1″) the upper bound does not exist.

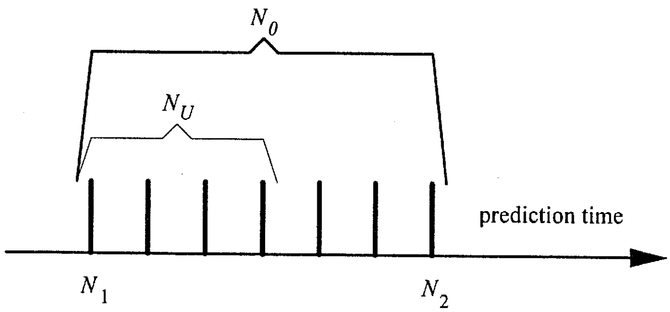


Fig. 3. The control horizon (N_U) and the observation (N_0, N_1, N_2) horizons.

4. Conditions C1′ and C1″ are identical and equivalent to the relation

$$N_U \leq N_0$$

according to which that the effective observation horizon $N_0 = N_2 - N_1 + 1$ should not be shorter than the assumed length N_U (on the prediction time axis) of the designed control sequence $\Delta \boldsymbol{u}(t)$ (see eqns. (10)–(12) and Fig. 3). It is clear from the above relation that conditions C3′ and C3″ lead to a bilateral restriction on N_U

$$N_A + 1 \leq N_U \leq N_0 = N_2 - N_1 + 1$$

Since $\operatorname{rank} \boldsymbol{H} = N_U$, matrix $\boldsymbol{H}$ is $N_0 \times N_U$ and N_U is the critical parameter sought after. From the above restriction it also results that the GPC observation horizon is limited as well

$$N_0 \geq N_A + 1$$

5. Note that $N_U = N_A + 1$ fulfils both the C3 conditions at the same time. Hence $N_1 \geq N_B$ can be accepted as a general restriction C2.

6. Having in mind both the design parsimony (with respect to N_2 and N_U) and the maintenance of the "essential information" (N_1), one can propose two parsimonious (P and S) ways of selecting the observation and control horizons based on the orders of the plant transfer function N_A and N_B:

(P1)	$N_1 = N_B$		(S1)	$N_1 = N_B + 1$
(P2)	$N_2 = N_A + N_B$	or	(S2)	$N_2 = N_A + N_B + 1$
(P3)	$N_U = N_A + 1$		(S3)	$N_U = N_A + 1$

With both the tuning settings given above the effective observation horizon is

$$N_0 = N_2 - N_1 + 1 = N_A + 1 = N_U$$

Conditions C′ and the suboptimal tuning set (S) play a key role in the derivation of a numerical design support (the CD-HAG algorithm presented in the sequel).

Let us assume now that

$$\widehat{A}(q^{-1}) = \widehat{A}_0(q^{-1})\Lambda(q^{-1}) \tag{31}$$

$$B(q^{-1}) = B_0(q^{-1})\Lambda(q^{-1}) \tag{32}$$

where $\widehat{A}_0(q^{-1})$ and $B_0(q^{-1})$ are relatively prime, and $\Lambda(q^{-1})$ with $\deg\Lambda(q^{-1}) = N_\Lambda > 0$, denotes the greatest common divisor of $\widehat{A}(q^{-1})$ and $B(q^{-1})$, which will be referred to as the cancellation order. Note that the Markov parameters of $\overline{B}(q^{-1})/\widehat{A}(q^{-1})$ are identical to the Markov parameters of $\overline{B}_0(q^{-1})/\widehat{A}_0(q^{-1})$, where $\overline{B}_0(q^{-1}) = qB_0(q^{-1})$. Therefore from (25)–(26) it follows that the closed-loop characteristic polynomial $D(q^{-1})$ has now the form

$$D(q^{-1}) = C(q^{-1})\Lambda(q^{-1})\Big(\widehat{A}_0(q^{-1}) - q^{-1}\widetilde{A}_0(q^{-1}) + g^* B_0(q^{-1})\Big) \tag{33}$$

where

$$\widetilde{A}_0(q^{-1}) = \sum_{i=1}^{N_0} k_i\Big(\widehat{A}_0(q^{-1})H_{N_1+i-1}(q^{-1}) - \overline{B}_0(q^{-1})\Big)q^{N_1+i-1}$$

Theorem 2. *If $\widehat{A}(q^{-1})$ and $B(q^{-1})$ are not relatively prime and one of the two following triple conditions:*

(C1′)	$N_2 \geq N_1 + N_U - 1$		(C1″)	$N_2 \geq N_1 + N_U - 1$
(C2′)	$N_1 \geq N_B - N_\Lambda$	*or*	(C2″)	$N_1 = N_B - N_\Lambda$
(C3′)	$N_U = N_A - N_\Lambda + 1$		(C3″)	$N_U \geq N_A - N_\Lambda + 1$

is satisfied, then $\boldsymbol{H}$ has full column rank ($\operatorname{rank}\boldsymbol{H} = N_U$), *the solution (11) exists, and the closed-loop characteristic polynomial is*

$$D(q^{-1}) = C(q^{-1})\Lambda(q^{-1})\widetilde{D}_\Lambda(q^{-1}) \tag{34}$$

$$\widetilde{D}_\Lambda(q^{-1}) = 1 + g^* B_0(q^{-1}) \tag{35}$$

Remarks:

1. Note that this time for $r_i = 1$, $i = N_1, \dots, N_2$ in (26) the characteristic polynomial $D(q^{-1})$ is partly determined by the observer polynomial $C(q^{-1})$

$$D(q^{-1}) = C(q^{-1})\Lambda(q^{-1}) \tag{36}$$

This means that the closed-loop system will be stable for all stabilisable systems of (1).

2. While designing with another choice of $\boldsymbol{r}$ from (19), in fundamental stability considerations, the locations of the zeros of the factor $\widetilde{D}_\Lambda(q^{-1})$ have to be analysed. To this end the idea of root loci can be applied to show that for sufficiently small absolute values of g^* the polynomial $\widetilde{D}_\Lambda(q^{-1})$ will always be stable. In Fig. 4 the root locus plot is given for a second-order system (Kowalczuk and Suchomski, 1995a; 1995b) controlled by GPC defined by $N_1 = 3$, $N_2 = 8$ and $N_U = 3$, and

$$\boldsymbol{k} = \begin{bmatrix} 1.4288 & -0.1471 & -0.7960 & -0.6664 & -0.0175 & 0.8649 \end{bmatrix}^T$$

$$L(q^{-1}) = 0.4754q^{-1}, \quad M(q^{-1}) = 1.3578 - 2.0235q^{-1} + 0.6657q^{-2}$$

$$\boldsymbol{r} = [r \ 1 \ 1 \ 1 \ 1 \ 1]^T$$

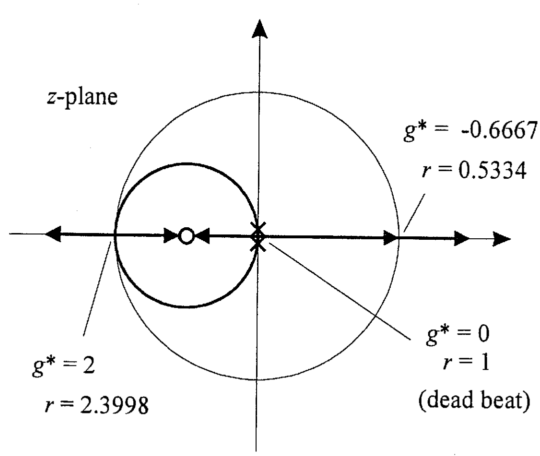


Fig. 4. Root loci of a control system ($g^* = k_1(r-1)$).

3. A *relative stability index* (cf. (26)) can be evaluated as

$$\zeta = \frac{g^*}{g} = 1 - \frac{g_{DB}}{g} \tag{37}$$

where

$$g_{DB} = \sum_{i=1}^{N_0} k_i$$

4. The conditions of Theorem 2 are analogous to those of Theorem 1, provided that the orders of the numerator (N_B) and denominator (N_A) of the plant transfer function are reduced by the cancellation order. In other words, the parameters (i.e. the "effective orders") of a minimal realisation of the plant should be used. In Theorem 2 the equivalent restriction concerning the knowledge of the cancellation order (N_Λ) is used explicitly. This is especially inconvenient in the case of

condition C2″, which imposes the necessity of precise knowledge of the effective order of the numerator

$$N_B^0 = N_B - N_\Lambda$$

5. Since condition C3′ has been chosen (Kowalczuk and Suchomski, 1995a) as the upper bound on N_U, the first set of conditions C′ seems to be a convenient basis for choosing the design parameters. Consequently, with reference to Remark 6 given after Theorem 1, it is the suboptimal tuning procedure (S), resulting in the same observation horizon N_0 and the same order of the controller, that should be preferred.

6. There is a practical problem with the determination of the cancellation order N_Λ. Note that cancellation can take place in the controlled plant or be induced by an overparametrised model used in identification of the plant. Because the cancellation order diminishes the bound on N_U it can be evaluated via detection of the upper bound of N_U that guarantees nonsingularity on $\boldsymbol{H}^T\boldsymbol{H}$.

An iterative algorithm for concurrent determination of the control horizon N_U and gain $\boldsymbol{K}$ (CD-HAG), derived in Appendix B for $\lambda = 0$, is described in the following procedure.

Procedure (CD-HAG)

Starting with $i = 0$, $\boldsymbol{P}_0 = \boldsymbol{I}$, and having $\boldsymbol{h}_i = [h_{N_1-i} \cdots h_{N_2-i}]^T$ as the i-th column of $\boldsymbol{H}$, the upper bound on N_U can be obtained from the conditions of termination of the procedure that follows:

$$\begin{aligned}
\boldsymbol{n}_{i+1} &= \boldsymbol{P}_i \boldsymbol{h}_{i+1} \\
\|\boldsymbol{n}_{i+1}\|_2^2 &= \boldsymbol{n}_{i+1}^T \boldsymbol{n}_{i+1} \qquad \text{(terminate if } \|\boldsymbol{n}_{i+1}\|_2^2 = 0) \\
\boldsymbol{n}_{i+1}^+ &= \|\boldsymbol{n}_{i+1}\|_2^{-2} \boldsymbol{n}_{i+1}^T \\
\boldsymbol{p}_{i+1} &= \boldsymbol{H}_i^+ \boldsymbol{h}_{i+1} \\
\boldsymbol{H}_{i+1}^+ &= \begin{bmatrix} \boldsymbol{H}_i^+ - \boldsymbol{p}_{i+1}\boldsymbol{n}_{i+1}^+ \\ \boldsymbol{n}_{i+1}^+ \end{bmatrix}, \qquad \text{for the first run } (i=0) \text{ use } \boldsymbol{H}_1^+ = \boldsymbol{n}_1^+ \\
\boldsymbol{P}_{i+1} &= \boldsymbol{P}_i - \boldsymbol{n}_{i+1}\boldsymbol{n}_{i+1}^+
\end{aligned} \tag{38}$$

Remarks:

1. For practical reasons, the condition $\|n_{i+1}\|_2^2 < \varepsilon$ should be monitored, where ε is a small computer-dependent value used for detecting zero (and linear dependence).

2. Observe that if for some i rank$\cdot\boldsymbol{H}_i = i$, then $\boldsymbol{H}_i^+ = (\boldsymbol{H}_i^T\boldsymbol{H}_i)^{-1}\boldsymbol{H}_i^T$ (Boullion and Odell, 1971). Thus, setting $N_U = i$ yields the corresponding solution $\boldsymbol{K}$ of (11) for $\lambda = 0$.

3. After termination of the procedure CD-HAG, index i has a maximum allowable value $i_{\max}$ and $\operatorname{rank} \boldsymbol{H}_i = i_{\max}$. Hence, putting $\boldsymbol{H} = \boldsymbol{H}_i$ and $N_U = n_{U,\max} = i_{\max}$, we have

$$\boldsymbol{H}^+ = (\boldsymbol{H}^T \boldsymbol{H})^{-1} \boldsymbol{H}^T$$

and the sought solution for $\boldsymbol{K}$.

4. Consequently, for an arbitrarily assumed value of N_A the cancellation order can simply be calculated based on the detected maximum value of N_U (see also C3′ in Theorem 2):

$$N_\Lambda = N_A - N_U + 1 \tag{39}$$

5. It is interesting to note that by using this procedure we have found that the parameter N_U is an equivalent of the effective system order

$$N_U = N_A^0 + 1, \quad \text{where} \quad N_A^0 = N_A - N_\Lambda \tag{40}$$

6. It is clear from Theorem 2 that the second set of conditions (Cl″) does not lead to analogous results, since the parameter N_U can be of arbitrarily large value.

Simulated Performance. The basic form of the plant under examination is given by the following minimal model with relatively prime polynomials $\widehat{A}(q^{-1})$ and $B(q^{-1})$

$$A(q^{-1}) = (1 - 0.67032q^{-1})(1 - 0.76593q^{-1})(1 - 0.81873q^{-1})$$

$$B(q^{-1}) = 0.0028689(1 + 0.21523q^{-1})(1 + 3.01224q^{-1})q^{-1}$$

$$C(q^{-1}) = (1 + 0.9q^{-1})\Big[(1 + 0.63639q^{-1})^2 + (0.63639q^{-1})^2\Big]$$

According to the first parsimonious rule of selecting the observation and control horizons (P1–P3), the design settings are as follows:

$$\lambda = 0, \quad N_1 = N_B = 3, \quad N_2 = N_A + N_B = 6, \quad N_U = N_A + 1 = 4$$

and

$$r_{N_1} = r, \quad r_{N_1+1} = \cdots = r_{N_2} = 1, \quad r = r_{N_1} \in [0.8,\ 1]$$

The resulting discrete-time control and output signals corresponding to a unit step excitation $w(n)$ are demonstrated in Figs. 5(a) and 5(b), respectively. It is clear that the desired effect of reduction of control effort is obtained at the cost of a slight deterioration of transient of the controlled process.

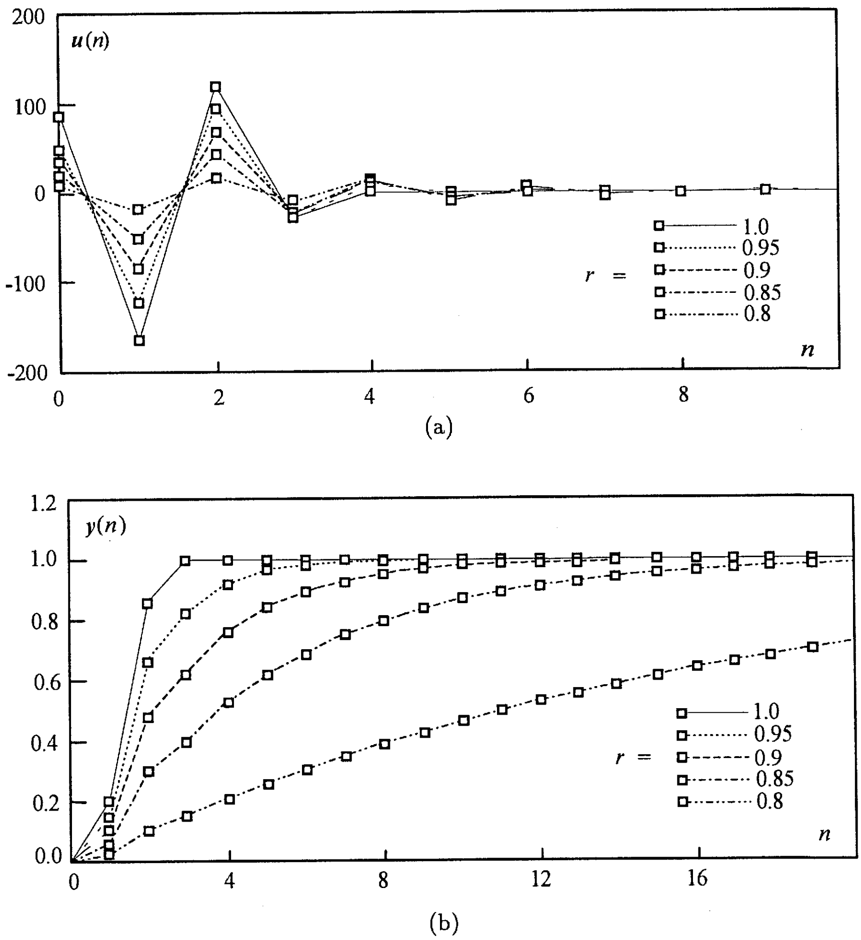


Fig. 5. The GPC system performance for different AF: (a) control signals, (b) step responses.

4. The Continuous-Time GPC Design Principles

In the derivation of the continuous-time generalised predictive controller CGPC (Demircioglu and Gawthrop, 1991) the following aggregated plant model is considered:

$$Y(s) = \frac{B(s)}{A(s)}U(s) + \frac{C(s)}{A(s)}V(s) \tag{41}$$

This model covers non-minimum phase objects and unstable objects without disturbances ($V(s) = 0$). For a disturbed unstable plant model another structural output-error (OE) plant model is necessary

$$Y(s) = \frac{D_0(s)}{D_1(s)}\left(\frac{B_0(s)}{B_1(s)}U(s) + \frac{C_0(s)}{C_1(s)}V(s)\right) \tag{42}$$

from which we have

$$A(s) = B_1(s)C_1(s)D_1(s)$$

$$B(s) = C_1(s)B_0(s)D_0(s)$$

$$C(s) = B_1(s)C_0(s)D_0(s)$$

In the aggregated plant models, $C(s)$ must be stable, while in the OE plant models all the polynomials, except for $B_0(s)$ and $D_1(s)$, must be stable to assure the stability of the regulator (Kowalczuk and Marcińczyk, 1995a; 1995b).

The CGPC regulator design (Demircioglu and Gawthrop, 1991) is based on the model (41), which can be characterised by the model order N_A and the relative model order ρ

$$N_A = \deg A(s) \tag{43}$$

$$\rho = N_A - N_B, \quad \text{where} \quad N_B = \deg B(s) \tag{44}$$

and on calculation of the first k derivatives of the model output signal function

$$Y_k(s) = \frac{s^k B(s)}{A(s)}U(s) + \frac{s^k C(s)}{A(s)}V(s) \tag{45}$$

The output prediction can then be calculated from

$$y^*(t+T) = \sum_{k=0}^{N_Y} y_k^*(t)\frac{T^k}{k!} \tag{46}$$

which can be expressed in the matrix form

$$y^*(t+T) = \boldsymbol{T}_{N_Y}\boldsymbol{H}\boldsymbol{u} + \boldsymbol{T}_{N_Y}\boldsymbol{Y}^0 \tag{47}$$

$$Y_T^*(s) = \boldsymbol{T}_{N_Y}\boldsymbol{H}\boldsymbol{S}_H U(s) + \boldsymbol{T}_{N_Y}\boldsymbol{Y}^0(s) \tag{48}$$

where

$$\boldsymbol{Y}^0(s) = \frac{\boldsymbol{G}\boldsymbol{S}_G}{C(s)}U(s) + \frac{\boldsymbol{F}\boldsymbol{S}_F}{C(s)}Y(s) \tag{49}$$

$$\boldsymbol{Y}^0 = \begin{bmatrix} y_0^0(t) & y_1^0(t) & \cdots & y_{N_Y}^0(t) \end{bmatrix}^T \tag{50}$$

$$\boldsymbol{u} = \begin{bmatrix} u(t) & u_1(t) & \cdots & u_{N_U}(t) \end{bmatrix}^T \tag{51}$$

$$\boldsymbol{T}_{N_Y} = \begin{bmatrix} 1 & T & \frac{T^2}{2!} & \cdots & \frac{T^{N_Y}}{N_Y!} \end{bmatrix} \tag{52}$$

$$\boldsymbol{S}_H = \begin{bmatrix} 1 & s & s^2 & \cdots & s^{N_U} \end{bmatrix}^T \tag{53}$$

$$\boldsymbol{S}_G = \begin{bmatrix} 1 & s & s^2 & \cdots & s^{N_A-2} \end{bmatrix}^T \tag{54}$$

$$\boldsymbol{S}_F = \begin{bmatrix} 1 & s & s^2 & \cdots & s^{N_A-1} \end{bmatrix}^T \tag{55}$$

$$\boldsymbol{H} = \begin{bmatrix} h_0^0 & 0 & 0 & \cdots & 0 \\ h_0^1 & h_1^1 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \ddots & \cdots \\ h_0^{N_Y-\rho} & h_1^{N_Y-\rho} & h_2^{N_Y-\rho} & \cdots & h_{N_U}^{N_Y-\rho} \\ \cdots & \cdots & \cdots & \ddots & \cdots \\ h_0^{N_Y} & h_1^{N_Y} & h_2^{N_Y} & \cdots & h_{N_U}^{N_Y} \end{bmatrix} \tag{56}$$

$$\boldsymbol{G} = \begin{bmatrix} g_0^0 & g_1^0 & \cdots & g_{N_A-2}^0 \\ g_0^1 & g_1^1 & \cdots & g_{N_A-2}^1 \\ \cdots & \cdots & \cdots & \cdots \\ g_0^{N_Y} & g_1^{N_Y} & \cdots & g_{N_A-2}^{N_Y} \end{bmatrix} \tag{57}$$

$$\boldsymbol{F} = \begin{bmatrix} f_0^0 & f_1^0 & \cdots & f_{N_A-1}^0 \\ f_0^1 & f_1^1 & \cdots & f_{N_A-1}^1 \\ \cdots & \cdots & \cdots & \cdots \\ f_0^{N_Y} & f_1^{N_Y} & \cdots & f_{N_A-1}^{N_Y} \end{bmatrix} \tag{58}$$

where $N_U \leq N_Y - \rho$, and f_i^k, g_i^k and h_i^k are the coefficients of polynomials $F_k(s)$, $G_k(s)$ and $H_k(s)$, respectively, which results from polynomial division

$$\frac{s^k C(s)}{A(s)} = E_k(s) + \frac{F_k(s)}{A(s)} \tag{59}$$

$$\frac{E_k(s)B(s)}{C(s)} = H_k(s) + \frac{G_k(s)}{C(s)} \tag{60}$$

Consider now the following cost function:

$$J(t) = \int_{T_1}^{T_2} \Big[y_r^*(t+T) - w_r^*(t+T)\Big]^2 \mathrm{d}T + \lambda \int_0^{T_2-T_1} \Big[u^*(t+T)\Big]^2 \mathrm{d}T \tag{61}$$

where $w_r^*(t+T)$ is a reference trajectory obtained by the anticipated filtration whose purpose is to protect against excessive control actions and overshoots. This idea is identical to the one used in the discrete-time domain, which has been explained in Fig. 1. Note that $y_r^*(t+T)$ and $w_r^*(t+T)$ are now respectively the output and reference trajectory related to $y(t)$

$$w_r^*(t+T) = w_r^*(t,T) = F_r\Big[w(t+T) - y(t)\Big] = F_{r,T}\Big[w(t) - y(t)\Big]$$

or

$$w_r^*(t,T) = \mathcal{L}_T^{-1}\left\{\frac{R_n(s)}{R_d(s)}\,\frac{w(t)-y(t)}{s}\right\} \tag{62}$$

and

$$y_r^*(t+T) = y^*(t+T) - y(t) \tag{63}$$

Note that the inverse Laplace transform results in the T-time domain. Consequently, (49) can be replaced by

$$\boldsymbol{Y}'^0(s) = \frac{\boldsymbol{G}'\boldsymbol{S}_G}{C(s)}U(s) + \frac{\boldsymbol{F}'\boldsymbol{S}_F}{C(s)}Y(s) \tag{64}$$

where

$$\boldsymbol{Y}'^0 = \begin{bmatrix} 0 & y_1^0(t) & \cdots & y_{N_Y}^0(t) \end{bmatrix}^T \tag{65}$$

$$\boldsymbol{G}' = \operatorname{diag}\begin{bmatrix} 0 & 1 & 1 & \cdots \end{bmatrix}\boldsymbol{G} \tag{66}$$

$$\boldsymbol{F}' = \operatorname{diag}\begin{bmatrix} 0 & 1 & 1 & \cdots \end{bmatrix}\boldsymbol{F} \tag{67}$$

Since the reference trajectory can be approximated by the Maclaurin series expansion

$$w_r^*(t+T) = F_{r,T}\Big[w(t) - y(t)\Big] \approx \sum_{i=0}^{N_Y} \frac{T^i}{i!}\,\frac{d^i w_r^*(t,T)}{dT^i}\bigg|_{T=0} = \boldsymbol{T}_{N_Y}\boldsymbol{w} \tag{68}$$

where $\boldsymbol{T}_{N_Y}$ is defined in (52) and

$$\boldsymbol{w} = \boldsymbol{r}\Big[w(t) - y(t)\Big] \tag{69}$$

$$\boldsymbol{r} = \begin{bmatrix} r_0 & r_1 & \cdots & r_{N_Y} \end{bmatrix}^T \tag{70}$$

$$\frac{R_n(s)}{R_d(s)} \approx \sum_{i=0}^{N_Y} r_i s^{-i} \tag{71}$$

the cost function becomes (see(61))

$$J(t) = \int_{T_1}^{T_2} (\boldsymbol{T}_{N_Y}\boldsymbol{H}\boldsymbol{u} + \boldsymbol{T}_{N_Y}\boldsymbol{Y}'^0 - \boldsymbol{T}_{N_Y}\boldsymbol{w})^2 \,\mathrm{d}T + \lambda \int_0^{T_2-T_1} \boldsymbol{u}^T\boldsymbol{T}_{N_U}^T\boldsymbol{T}_{N_U}\boldsymbol{u}\,\mathrm{d}T \tag{72}$$

and its minimisation results in (51) given by

$$\boldsymbol{u} = \boldsymbol{K}(\boldsymbol{w} - \boldsymbol{Y}'^0) \tag{73}$$

$$\boldsymbol{K} = (\boldsymbol{H}^T\boldsymbol{T}_Y\boldsymbol{H} + \lambda\boldsymbol{T}_U)^{-1}\boldsymbol{H}^T\boldsymbol{T}_Y \tag{74}$$

$$\boldsymbol{Y}'^0 = \begin{bmatrix} 0 & y_1^0(t) & \cdots & y_{N_Y}^0(t) \end{bmatrix}^T \tag{75}$$

$$\boldsymbol{T}_Y = \int_{T_1}^{T_2} \boldsymbol{T}_{N_Y}^T\boldsymbol{T}_{N_Y}\,\mathrm{d}T \tag{76}$$

$$\boldsymbol{T}_U = \int_0^{T_2-T_1} \boldsymbol{T}_{N_U}^T\boldsymbol{T}_{N_U}\,\mathrm{d}T \tag{77}$$

$$\boldsymbol{T}_{N_U} = \begin{bmatrix} 1 & T & \frac{T^2}{2!} & \cdots & \frac{T^{N_U}}{N_U!} \end{bmatrix} \tag{78}$$

Note that simply implementable is the control signal defined as the first element of $\boldsymbol{u}$ in (73)

$$u(t) = \begin{bmatrix} 1 & 0 & 0 & \cdots \end{bmatrix}\boldsymbol{K}(\boldsymbol{w} - \boldsymbol{Y}'^0) \tag{79}$$

Hence the CGPC formula (similar to Fig. 2 if one sets $L := \overline{G}_0(s)$ and $M := \overline{F}_0(s)$) becomes

$$U(s) = g\Big[W(s) - Y(s)\Big] - \frac{\overline{G}_0(s)}{C(s)}U(s) - \frac{\overline{F}_0(s)}{C(s)}Y(s) \tag{80}$$

where

$$g = \begin{bmatrix} 1 & 0 & 0 & \cdots \end{bmatrix}\boldsymbol{K}\boldsymbol{r} \tag{81}$$

$$\overline{G}_0 = \begin{bmatrix} 1 & 0 & 0 & \cdots \end{bmatrix}\boldsymbol{K}\boldsymbol{G}'\boldsymbol{S}_G \tag{82}$$

$$\overline{F}_0 = \begin{bmatrix} 1 & 0 & 0 & \cdots \end{bmatrix}\boldsymbol{K}\boldsymbol{F}'\boldsymbol{S}_F \tag{83}$$

The effect of the s-transform of the reference trajectory (see (62)) is reduced here to a single scalar gain coefficient g (see (21), (26), and (81)). Therefore, a simple anticipated filter characterised by a time constant r will be sufficient for generation of the reference trajectory:

$$\frac{R_n(s)}{R_d(s)} = \frac{1}{rs + 1} \tag{84}$$

5. Tuning of the AF-CGPC Controller

The values of the preliminary design parameters (N_Y, N_U, T_1, T_2, r and λ) that have to be determined for the algorithm are extremely critical for the closed-loop CGPC system behaviour and therefore the way of tuning is of great importance. While yielding unstable control systems, the approximate tuning indications given by Demircioglu and Gawthrop (1991) apparently do not work in many practical cases.

By examining the problem, it can easily be established that the classically recognised step response and its dominant time constant cannot serve as a rational basis for the determination of the design time-parameters, including their time scale. At the same time, these parameters can be effectively derived based on an initial phase of the system step response.

The origins of the proposed tuning rules can be found in a simple claim that from the prediction viewpoint *the most substantial characteristics of the controlled object is its ability to follow the command signal in an open-loop operation.* Namely, it is important to know when and how the value of the controlled variable will acquire the value of the controlling variable. Consequently, the top of the prediction horizon T_2 can be determined by the time of the value replication (approximately "1-to-1") and the bottom of the prediction horizon T_1 can be found as an approximate delay-time.

To illustrate the basic tuning principles, let us consider an exemplary plant model with undamped modes

$$\begin{cases} A(s) = s(s^2+1) \\ B(s) = 1 \\ C(s) = 0.2s^2 + s + 1 \end{cases} \tag{85}$$

The initial phase of a step response of such a plant is given in Fig. 6. Note that this part of the step responses of objects of other types, including non-minimum phase

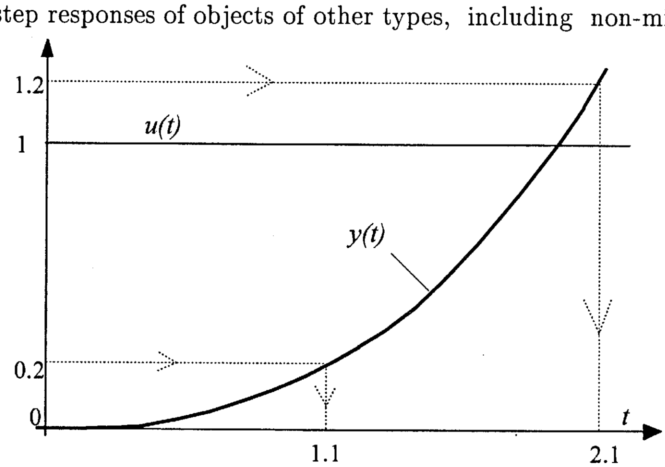


Fig. 6. System step response and parameters PRR.

and unstable ones, can be parametrised in a similar manner, provided that the object has a positive gain, an integral property, and that the non-minimum phase and unstable properties are not overly dominant.

PRR Tuning Principles. In view of observations of many simulation runs for plant models of different types, the tuning principles based on the Parameters of the Reaction Rate (PRR) can be stated as follows:

1. T_1 — corresponding to $\approx 20\%$ of the height of the input step function,
2. T_2 — corresponding to $\approx 120\%$ of the height of the input step function,
3. $r = T_2 - T_1$,
4. $N_Y = N_A + N_B + 1$,
5. $N_U = \min\{N_C, 2N_B + 1\}$, where $N_C = \deg C(s)$, and
6. $\lambda \in [0.0005, 0.05]$ (e.g. $\lambda = 0.01$).

If there is a lack of the integral property in the object model, the above principles cannot be used and in the closed-loop system the steady-state error is bound to appear. To circumvent this problem, it is recommended to attach an integral part at the input of the plant — and to modify accordingly the plant model.

Note that the value of the output order $N_Y = N_A + N_B + 1$ is in accordance with its discrete-time equivalent $N_2 = N_A + N_B + 1$ of the suboptimal tuning set (S) given in Remark 6 after Theorem 1.

Simulated Performance. In the simulation study, a programming package for simulation of analogue and digital control systems has been used (Kowalczuk, 1992b; Kowalczuk and Marcińczyk, 1995a). Different types of aggregated plant models, including unstable and non-minimum-phase ones, with the preliminary design parameters set according to the PRR tuning principles have been tested (Kowalczuk and Marcińczyk, 1995b) successfully. In the study, the weighting coefficient λ has been set to 0.01. For the model (85) represented in Fig. 6 the tuning parameters have been set as follows:

$$T_1 = 1.1, \quad T_2 = 2.1, \quad r = 1, \quad N_Y = 4, \quad N_U = 1 \tag{86}$$

The correspponding simulation run of the CGPC system is shown in Fig. 7.

In order to characterise the quality of the step response of the CGPC systems, the following performance indices have been used:

- ISE – the integral of squared error,
- ISC – the integral of squared control,
- $T_{5\%}$ – 5% regulation time,
- $MACS$ – max absolute value of control signal, and
- $I = \frac{1}{4}\left(\frac{ISE}{ISE_{\mathrm{PRR}}} + \frac{ISC}{ISC_{\mathrm{PRR}}} + \frac{T_{5\%}}{T_{5\%\,\mathrm{PRR}}} + \frac{MACS}{MACS_{\mathrm{PRR}}}\right)$

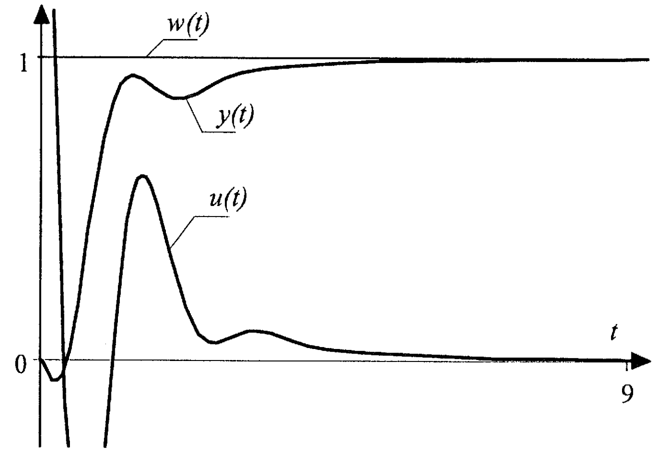


Fig. 7. Simulation run for CGPC control.

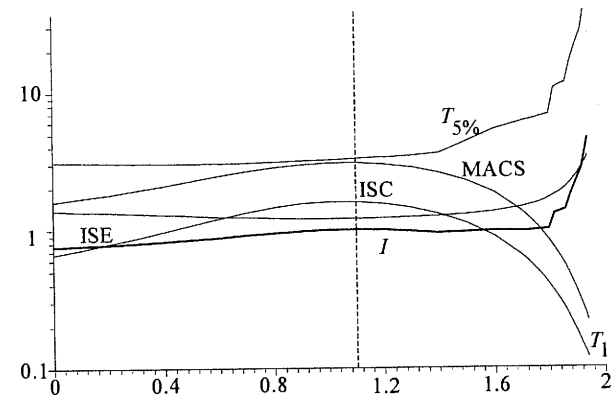
Index I is a global index composed of the four basic indices which are normalised with respect to their 'optimal' values that have been obtained during simulation runs with the CGPC regulators designed using the preliminary design parameters, set according to the PRR tuning principles. The results of a sensitivity study are shown in Fig. 8(a)–8(c), where the vertical dashed lines are used to mark the proposed values of the design parameters under consideration. Note also that the global index I manifests its minimum close to an optimal value that takes into account all of the component indices (ISE, ISC, $T_{5\%}$, $MACS$).

It can easily be seen that the proposed PRR tuning method minimises in practice the global cost function I with a certain stability margin. Note that for the plant model under consideration the chosen value of the parameter T_2 is placed inside a sharply outlined and confined stability region, in the plot using a logarithmic scale (Fig. 8(b)). With a linear scale the I function would have a rectangular shape. Therefore, it can be concluded that the value of T_2 obtained via the PRR method practically minimises the global cost function I. Analogous efects have also been obtained with respect to other parameters and for other types of plant models.

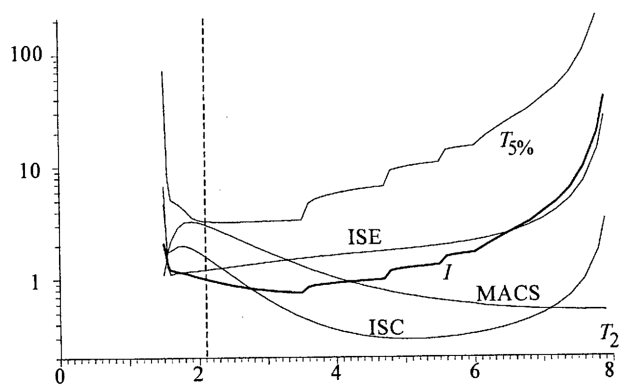
Thus the PRR method has been succesfully verified by simulation of the continuous-time control loops with continuous-time objects of different types with respect to the stability and minimum-phase properties (Kowalczuk and Marcińczyk, 1995b).

6. Conclusions

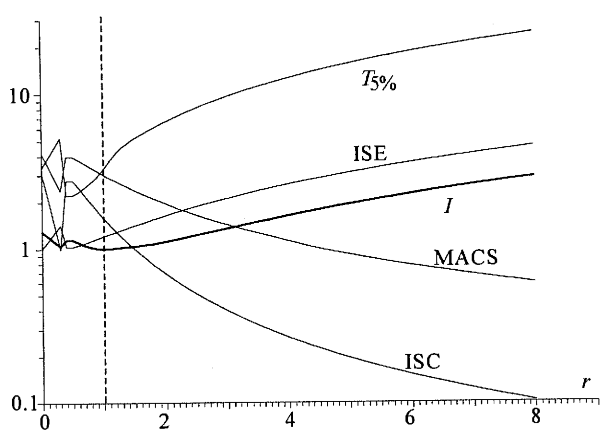
It is known that designs based on the dead-beat approach lead to an excessive control action, which can in many cases result in a very limited practical range of λ and some sensitivity problems. It has been shown that within the discrete-time approach the anticipated filtering can have a desired effect on the closed-loop behaviour of the controlled plant in terms of pole placement and that an appropriate design of



(a)



(b)



(c)

Fig. 8. Performance indices versus the PRR parameters: (a) T_1, (b) T_2, (c) r.

the anticipated filter can reduce the disagreeable control effort leading to a certain balance in the cost function.

Another problem of practical importance is to provide the user with a set of tuning rules. The proposed iterative algorithm for the simultaneous determination of the control horizon N_U, the matrix controller gain $\boldsymbol{K}$, and the cancellation order are in line with this requirement. The tools developed make it possible to apply the algorithm along with an identification procedure using models with an overestimated order and, at the same time, to design a regulator of reduced order.

The results concerning the continuous-time approach show that the proposed tuning principles, based on the controlled system's dynamics, outcome in stable closed loop CGPC systems for all the preliminary design parameters (time constants: T_1, T_2, r). Moreover, the principles lead to approximately optimally tuned regulator. Namely, it is shown that the method minimises in practice the global cost function I with a certain stability margin. More examples of the continuous-time results are given in (Kowalczuk and Marcińczyk, 1995b).

It is thus clear that the CGPC algorithm has properties similar to those of the GPC and can be used for unstable and non-minimum-phase object models. The algorithm can also be equipped with anticipated filtration, which can produce a salutary effect on the control loop in terms of the closed loop system characteristics, the regulation overshoot and the control effort.

The nice discrete-time tuning rules have been analytically obtained due to the assumption that λ is equal to zero. There are, however, no restrictions as to using non-zero λ in (11) and (74) after the determination of the tuning design parameters (see also the continuous-time example). Nevertheless, one has to keep in mind that the effect of the anticipated filter is similar to the effect of $\lambda \neq 0$. On the other hand, since without the AF mechanism the value of λ necessary to practise the cost functions (5) and (17) or (61) can appear to be extremely small, and in effect unimplementable, in fact, it is the anticipated filtration that makes the λ-tuning practicable and thus allows for a further improvement of the GPC control that is originally attributed to λ.

The idea of application of the simplest anticipation filter results from the mathematical characteristics of the design in both the disctete-time and continuous-time domains that consists in reducing the effect of the AF filter to a scalar gain coefficient g. This fact, supported also by experience, leads to a conclusion that it is sufficient to apply a single parameter r tuning the AF filter (or, equivalently, the outer loop gain g).

Within both the discrete- and continuous-time developments, a constant setpoint is considered, but, principally, this implicitly takes place only in the *anticipated* time domains. On the other hand, there are no stringent restrictions imposed on the input signal, and it is apparent that the discussed design solution can also be used for a tracking problem. Note that the unit step function applied in the simulation study is considered in industry as one of the most severe types of excitation signals.

Appendices

A. Derivation of Theorem 1

1. Note that assuming $\lambda = 0$, $N_2 \geq N_1 + N_U - 1$ and $\operatorname{rank} \boldsymbol{H} = N_U$, we have $\boldsymbol{K} = (\boldsymbol{H}^T\boldsymbol{H})^{-1}\boldsymbol{H}^T$ and $\boldsymbol{KH} = \boldsymbol{I}$, where $\boldsymbol{I} \in \mathbb{R}^{N_U \times N_U}$ is the identity matrix. Thus the entries of $\boldsymbol{k}^T\boldsymbol{H}$ can be expressed as

$$\sum_{i=1}^{N_2-N_1+1} k_i h_{N_1+i-k-1} = \begin{cases} 1 & \text{for } k = 1 \\ 0 & \text{for } k = 2, \ldots, N_U \end{cases} \tag{A1}$$

2. By using (26), (27) and $N_1 \geq N_B$, the polynomial $\widetilde{A}(q^{-1})$ can be written down as

$$\widetilde{A}(q^{-1}) = \sum_{j=0}^{N_A} \nu_j q^{-j} \tag{A2}$$

where

$$\nu_j = \sum_{k=1}^{N_A-j+1} \widehat{a}_{k+j}\tau_k, \qquad j = 0, \ldots, N_A$$

and

$$\tau_k = \sum_{i=1}^{N_0} k_i h_{N_1+i-k-1}, \qquad k = 1, \ldots, N_A + 1$$

3. If $N_U \geq N_A + 1$, then from steps 1 and 2 it results that $\nu_j = \widehat{a}_{j+1}$, $j = 0, \ldots, N_A$ and

$$\widetilde{A}(q^{-1}) = \sum_{j=0}^{N_A} \widehat{a}_{j+1} q^{-j} \tag{A3}$$

with

$$\widehat{A}(q^{-1}) - q^{-1}\widetilde{A}(q^{-1}) = 1$$

4. Hence it may be concluded that for relatively prime $\widehat{A}(q^{-1})$ and $B(q^{-1})$, $N_2 \geq N_1 + N_U - 1$, $\operatorname{rank} \boldsymbol{H} = N_U$, $N_1 \geq N_B$, and $N_U \geq N_A + 1$, the closed-loop characteristic polynomial $D(q^{-1})$ is determined by (25) with

$$\widetilde{D}(q^{-1}) = 1 + g^* B(q^{-1}) \tag{A4}$$

5. Now, we shall assume that $N_2 \geq N_1 + N_U - 1$, $N_1 > N_B$, and $N_U = N_A + 1$

6. From (13) and (27) it follows that

$$\boldsymbol{H}\begin{bmatrix} 1 \\ \widehat{a}_1 \\ \vdots \\ \widehat{a}_{N_A} \end{bmatrix} = -\widehat{a}_{N_A+1}\begin{bmatrix} h_{N_1-N_U-1} \\ h_{N_1-N_U} \\ \vdots \\ h_{N_2-N_U-1} \end{bmatrix} \tag{A5}$$

7. If the polynomials $\widehat{A}(q^{-1})$ and $B(q^{-1})$ are not relatively prime, then the right-hand side of (A5) is zero. This means that the matrix $\boldsymbol{H}$ cannot have full rank. Hence, by using contradiction, we can claim that if $\boldsymbol{H}$ has full column rank, then the polynomials $\widehat{A}(q^{-1})$ and $B(q^{-1})$ are relatively prime.

8. It results from (A5) that if $\operatorname{rank}\boldsymbol{H} < N_A + 1$, then there must be a reduced-order system having the same Markov parameters as the original system (1). This means that there exists a polynomial different from $\widehat{A}(q^{-1})$ of a lower order that fulfils the "matching" relation similar to (A5). This means that, by contradiction, we have proved that if the polynomials $\widehat{A}(q^{-1})$ and $B(q^{-1})$ are relatively prime, then the matrix $\boldsymbol{H}$ has full column rank.

9. Accordingly augmenting the considerations of steps 5–8 (to the "restricted" case of $N_U \geq N_A+1$ and to a "complementary" case of $N_1 = N_B$, and $N_U \geq N_A+1$) and using a similar judgement, we can show that if ($N_2 \geq N_1+N_U-1$, $N_1 > N_B$ and $N_U \leq N_A+1$) or ($N_2 \geq N_1+N_U-1$, $N_1 = N_B$ and $N_U \geq N_A+1$), then $\boldsymbol{H}$ has full column rank if and only if $\widehat{A}(q^{-1})$ and $B(q^{-1})$ are relatively prime.

10. Combining steps 4 and 9 together results in Theorem 1, and reconsidering the whole discussion for the cancellation case proves Theorem 2.

B. Derivation of Algorithm CD-HAG

1. Taking into account solvability of the design problem with $\lambda = 0$, as it results from Theorem 2 and its first set of conditions C1′–C3′, where $N_2 \geq N_1 + N_A$ and $N_1 > N_B \geq N_B - N_\Lambda$ (see also Remark 4), there always exists an upper bound on N_U:

$$N_U \leq N_A - N_\Lambda + 1 \tag{B1}$$

Note that this bound establishes the largest value of the controller parameter N_U necessary to assure the claim corresponding to the claim of step 9 from Appendix A. This means that the cancellation order diminishes appropriately the bound on N_U.

2. Let the matrix $\boldsymbol{H}_{i+1}$, $i = 1, 2, \ldots$ be partitioned as follows:

$$\boldsymbol{H}_{i+1} = \begin{bmatrix} \boldsymbol{H}_i & \boldsymbol{h}_{i+1} \end{bmatrix} \tag{B2}$$

where $\boldsymbol{H}_i \in \mathbb{R}^{N_0 \times i}$, $\boldsymbol{h}_i \in \mathbb{R}^{N_0}$, and (with $h_k = 0$ for $k < 0$)

$$\boldsymbol{H}_i = \begin{bmatrix} h_{N_1-1} & h_{N_1-2} & \cdots & h_{N_1-i} \\ h_{N_1} & h_{N_1-1} & \cdots & h_{N_1-i+1} \\ \vdots & \vdots & \ddots & \vdots \\ h_{N_2-1} & h_{N_2-2} & \cdots & h_{N_2-i} \end{bmatrix} \quad \text{and} \quad \boldsymbol{h}_i = \begin{bmatrix} h_{N_1-i} \\ h_{N_1-i+1} \\ \vdots \\ h_{N_2-i} \end{bmatrix}$$

3. First, consider the projection $\boldsymbol{I} - \boldsymbol{H}_i\boldsymbol{H}_i^+$ on $\mathcal{R}(\boldsymbol{H}_i)^\perp$, where $\boldsymbol{H}_i^+ \in \mathbb{R}^{i \times N_0}$ denotes the Moore-Penrose pseudo inverse of $\boldsymbol{H}_i$ and $\mathcal{R}(\boldsymbol{H}_i)^\perp$ denotes the orthogonal complement of the range of $\boldsymbol{H}_i$. Let us assume that, for some i, $\boldsymbol{h}_{i+1}$ belongs to the null space of $\boldsymbol{I} - \boldsymbol{H}_i\boldsymbol{H}_i^+$:

$$\boldsymbol{h}_{i+1} \in \mathcal{N}(\boldsymbol{I} - \boldsymbol{H}_i\boldsymbol{H}_i^+)$$

Hence $\boldsymbol{H}_i\boldsymbol{H}_i^+\boldsymbol{h}_{i+1} = \boldsymbol{h}_{i+1}$ and $\boldsymbol{h}_{i+1} \in \mathcal{R}(\boldsymbol{H}_i)$. Consecutively,

$$\operatorname{rank}\begin{bmatrix} \boldsymbol{H}_i & \boldsymbol{h}_{i+1} \end{bmatrix} = \operatorname{rank} \boldsymbol{H}_i \tag{B3}$$

4. Now we may recall the common rule of computing the pseudo inverse of partitioned matrices (Boullion and Odell, 1971)

$$\begin{bmatrix} \boldsymbol{H}_i & \boldsymbol{h}_{i+1} \end{bmatrix}^+ = \begin{bmatrix} \boldsymbol{H}_i^+ - \boldsymbol{H}_i^+\boldsymbol{h}_{i+1}\boldsymbol{n}_{i+1}^+ \\ \boldsymbol{n}_{i+1}^+ \end{bmatrix} \tag{B4}$$

where $\boldsymbol{n}_{i+1} \in \mathbb{R}^{N_0}$ is defined as follows:

$$\boldsymbol{n}_{i+1} = (\boldsymbol{I} - \boldsymbol{H}_i\boldsymbol{H}_i^+)\boldsymbol{h}_{i+1}$$

5. Let $\boldsymbol{P}_i = \boldsymbol{I} - \boldsymbol{H}_i\boldsymbol{H}_i^+$ such that $\boldsymbol{P}_i \in \mathbb{R}^{N_0 \times N_0}$. Note that assuming that $\boldsymbol{n}_{i+1} \neq \boldsymbol{0}$, we obtain the recursive solution

$$\boldsymbol{P}_{i+1} = \boldsymbol{P}_i - \frac{\boldsymbol{n}_{i+1}\boldsymbol{n}_{i+1}^T}{\boldsymbol{n}_{i+1}^T\boldsymbol{n}_{i+1}} \tag{B5}$$

References

Boullion T.L. and Odell P.L. (1971): *Generalized Inverse Matrices.* — New York: Wiley-Interscience, John Wiley and Sons.

Clark D.W., Mohtadi C. and Tuffs P.S. (1987a): *Generalized predictive control - Part I. The basic algorithm.* — Automatica, v.23, No.2, pp.137–148.

Clark D.W., Mohtadi C. and Tuffs P.S. (1987b): *Generalized predictive control - Part II. Extensions and interpretations.* — Automatica, v.23, No.2, pp.149–160.

Clark D.W. and Mohtadi C. (1987): *Properties of generalised predictive control.* — Proc. 10th IFAC World Congress, Munich, Germany, v.10, pp.63–74.

Demircioḡlu H. and Gawthrop P.J. (1991): *Continuous-time generalised predictive control.* — Automatica, v.27, No.1, pp.55–74.

Favier G. (1987): *Self-tuning long-range predictive controllers.* — Proc. 10th IFAC World Congress, v.10, Munich, Germany, pp.80–87.

Gorez R., Wertz V. and Zhu K.-Y. (1987): *On a generalised predictive control algorithm.* — Systems and Control Letters, v.9, No.5, pp.369–377.

Kowalczuk Z. (1991): *On discretisation of continuous-time state-space models: A stable normal approach.* — IEEE Trans., v.CAS–38, No.12, pp.1460–1477.

Kowalczuk Z. (1992a): *Competitive identification for self-tuning control: Robust estimation design and simulation experiments.* — Automatica, v.28, No.1, pp.193–201.

Kowalczuk Z. (1992b): *DAML — a language for modelling and simulation of digital control systems.* — Proc. IFAC Symp. *Low Cost Automation*, Vienna, pp.385-390.

Kowalczuk Z. (1993): *Closed-form discrete-time modelling of continuous systems.* — Proc. 2nd European Control Conf., v.2, Groningen, the Netherlands, pp.444–449.

Kowalczuk Z. (1994): *Discrete-time realisation of on-line continuous-time process identification.* — Proc. IASTED Int. Conf. *Systems and Control*, Lugano, pp.19–21.

Kowalczuk Z. and Marcińczyk A. (1995a): *Modelling and simulations of continuous-time generalised predictive control systems.* — Proc. Int. Conf. *System Modelling Control*, Zakopane, Poland, v.1, pp.427–432.

Kowalczuk Z. and Marcińczyk A. (1995b): *Tuning of continuous-time generalised predictive control systems with anticipated filtration.* — Proc. 2nd Int. Symp. *Mathematical Models in Automation and Robotics MMAR'95*, Międzyzdroje, v.1, pp.253–258.

Kowalczuk Z. and Suchomski P. (1995a): *Generalised predictive control algorithms with anticipated filtration of control error.* — Proc. Int. Conf. *System, Modelling, Control*, Zakopane, Poland, v.1, pp.433–438.

Kowalczuk Z. and Suchomski P. (1995b): *Anticipated filtering approach to generalised predictive control.* — Proc. 3rd European Control Conf., Rome, Italy, v.4, pp.3591–3596.

Received: November 16, 1995
Revised: January 15, 1996