Predictive PID Control of Non-Minimum Phase Systems

Kenny Uren and George van Schoor
North-West University, Potchefstroom Campus
South Africa

1. Introduction

Control engineers have been aware of non-minimum phase systems showing either undershoot or time-delay characteristics for some considerable time (Linoya & Altpeter, 1962; Mita & Yoshida, 1981; Vidyasagar, 1986; Waller & Nygardas, 1975). A number of researchers that addressed this problem from a predictive control point of view mainly followed one of two approaches: a classical (non-optimal) predictive approach or a modern optimisation based predictive approach (Johnson & Moradi, 2005). The common characteristic of all these approaches is that they are model-based. Predictive control allows the controller to predict future changes in the output signal and to use this prediction to generate a desirable control variable. The classical predictive controllers that are most widely considered include the Smith predictor structure and the internal model control (IMC) structure (Katebi & Moradi, 2001; Morari & Zafiriou, 1989; Tan et al., 2001). Modern predictive controllers consider generalised predictive control (GPC) or model-based predictive control (MPC) structures (Johnson & Moradi, 2005; Miller et al., 1999; Moradi et al., 2001; Sato, 2010).

The performance of a PID controller degrades for plants exhibiting non-minimum phase characteristics. In order for a PID controller to deal with non-minimum phase behaviour, some kind of predictive control is required (Hägglund, 1992). Normally the derivative component of the PID controller can be considered as a predictive mechanism, however this kind of prediction is not appropriate when addressing non-minimum phase systems. In such a case the PI control part is retained and the prediction is performed by an internal simulation of plant inside the controller.

This chapter starts with a quick review of the system-theoretic concept of a pole and zero and then draws the relationship to non-minimum phase behaviour. The relationship between the undershoot response and time-delay response will be discussed using Padé approximations. Classical and modern predictive PID control approaches are considered with accompanying examples. The main contribution of the chapter is to illustrate the context and categories of predictive PID control strategies applied to non-minimum phase systems by:

- Considering the history of predictive PID control;

- The use of models in predictive control design;

- Exploring recent advances in predictive PID control where GPC (Generalised Predictive Control) algorithms play a prominent role;

• Appreciating the control improvements achieved using predictive strategies.

2. The influence of poles and zeros on system dynamics

When considering the compensation of systems it is of great importance to first understand the system-theoretic concept of a system pole and zero in the realm of system dynamics and control theory. Consider a continuous-time single-input, single-output (SISO) system

$$\dot{\mathbf{X}}(t) = \mathbf{A}\mathbf{X}(t) + \mathbf{B}u(t), \tag{1}$$

$$y(t) = \mathbf{C}\mathbf{X}(t) + Du(t), \tag{2}$$

where $u(t)$ and $y(t)$ are the scalar-valued input and output respectively. The column vector $\mathbf{X}(t)$ is called the state of the system and comprises n elements for an nth-order system. The $n \times n$ matrix $\mathbf{A}$ is called the system matrix and represents the dynamics of the system. The $n \times 1$ column vector $\mathbf{B}$ represents the effect of the actuator and the $1 \times n$ row vector $\mathbf{C}$ represents the response of the sensor. D is a scalar value called the direct transmission term. If $D = 0$, it is assumed that the input $u(t)$ cannot affect the output $y(t)$ directly.

If $\mathbf{X}(0) = \mathbf{0}$ and $D = 0$ (in the case where the output is not directly influenced by the input), then the system transfer function $G(s)$ is given by

$$G(s) = \frac{Y(s)}{U(s)} = \mathbf{C}(s\mathbf{I} - \mathbf{A})^{-1}\mathbf{B}. \tag{3}$$

The poles and zeros can be determined by writing $G(s)$ as

$$G(s) = \frac{N(s)}{D(s)}, \tag{4}$$

where the numerator polynomial is

$$N(s) \triangleq \det \begin{bmatrix} s\mathbf{I} - \mathbf{A} & -\mathbf{B} \\ \mathbf{C} & 0 \end{bmatrix}, \tag{5}$$

and the denominator polynomial is

$$D(s) \triangleq \det(s\mathbf{I} - \mathbf{A}). \tag{6}$$

Then the roots of $N(s)$ and $D(s)$ are defined as the *zeros* and *poles* of $G(s)$ respectively (Franklin et al., 2010; Hag & Bernstein, 2007). This holds only in the case where $N(s)$ and $D(s)$ do not have common roots.

The poles of $G(s)$ can be used to determine damping and natural frequencies of the system, as well as determining if the system is stable or unstable. As can be seen from Eq. (6) the poles depend only on the system matrix $\mathbf{A}$, but the zeros depend on matrices $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$. This leads to the question as to how the zeros influence the dynamic response of a system?

Consider a normalised transfer function of a system with two complex poles and one zero (Franklin et al., 2010):

$$T(s) = \frac{(s/a\zeta\omega_n) + 1}{s^2/\omega_n^2 + 2\zeta(s/\omega_n) + 1}. \tag{7}$$

The zero is therefore located at $s = -a\zeta\omega_n$. By replacing the s/ω_n with s results in a frequency normalising effect and also a time normalising effect in the corresponding step response. Therefore the normalised version of Eq.(7) can be rewritten as

$$T_n(s) = \frac{s/a\zeta + 1}{s^2 + 2\zeta s + 1}. \tag{8}$$

The normalised transfer function can be written as the sum of two terms

$$T_n(s) = T_1(s) + T_2(s), \tag{9}$$

$$= \frac{1}{s^2 + 2\zeta s + 1} + \frac{1}{a\zeta}\frac{s}{s^2 + 2\zeta s + 1}, \tag{10}$$

where $T_1(s)$ can be viewed as the original term with no added zeros, and $T_2(s)$ is introduced by the zero. Since the Laplace transform of a derivative dy/dt is $sY(s)$, the step response of $T_n(s)$ can be written as

$$y_n(t) = y_1(t) + y_2(t) = y_1(t) + \frac{1}{a\zeta}\dot{y}_1(t) \tag{11}$$

where y_1 and y_2 are the step responses of $T_1(s)$ and $T_2(s)$ respectively. The step responses for the case when $a > 0$ (introduction of a left half plane zero, $a = 1.1, \zeta = 0.5$) are plotted in Fig. 1(a). The derivative term y_2 introduced by the zero lifts up the total response of $T_n(s)$ to produce increased overshoot. The step responses for the case when $a < 0$ (introduction of a right half plane zero, $a = -1.1, \zeta = 0.5$) are plotted in Fig. 1(b). In this case the right half plane zero, also called a non-minimum phase zero causes the response of $T_n(s)$ to produce an initial undershoot.

In general a substantial amount of literature discusses the dynamic effects of poles, but less is available on the dynamic effects of zeros.

## 3. A closer look at non-minimum phase zeros

Before a formal definition of non-minimum phase zeros can be given, some definitions and assumptions are given. In this chapter only proper transfer functions will be considered. Eq. (4) may be expanded so that

$$G(s) = \frac{N(s)}{D(s)} = \frac{b_m s^m + b_{m-1} s^{m-1} + \cdots + b_1 s + b_0}{s^n + a_{n-1} s^{n-1} + \cdots + a_1 s + a_0}. \tag{12}$$

$G(s)$ is *strictly proper* if the order of the polynomial $D(s)$ is greater than that of $N(s)$ (i.e. $n > m$) and *exactly proper* if $n = m$ (Kuo & Golnaraghi, 2010). If $G(s)$ is *asymptotically stable*, that is, when the roots of $D(s)$ are all in the left half plane, each zero has a specific effect on the system for specific inputs. The roots of $N(s)$ (the zeros) can either be real or complex.

In general, a zero near a pole reduces the effect of that term in the total response. This can be shown by assuming that the poles, p_i, are real or complex but distinct and $G(s)$ can be written as a partial fraction expansion

$$G(s) = \frac{C_1}{s - p_1} + \frac{C_2}{s - p_2} + \cdots + \frac{C_n}{s - p_n}. \tag{13}$$
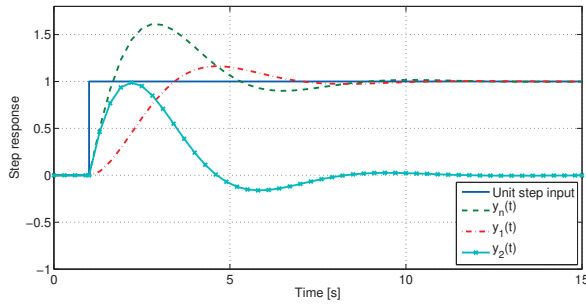
(a) Effect of a left half plane zero
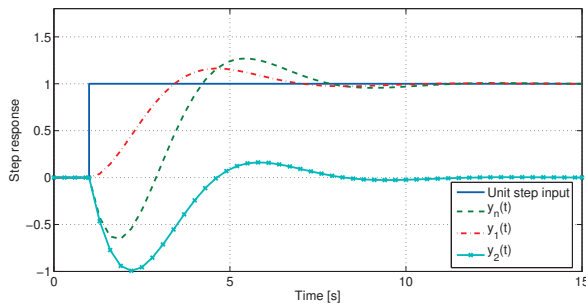


(b) Effect of a right half plane zero

Fig. 1. Step response of $T_n(s)$

When considering Eq. (13), and the equation for the coefficient C_1 given by

$$C_1 = (s - p_1)G(s)|_{s=p_1}, \qquad (14)$$

it can be seen that in the case where $G(s)$ has an left half plane zero near the pole at $s = p_1$, the value of C_1 will decrease. This means that the coefficient C_1, which determines the contribution of the specific term in the response will be small. From this observation it can also be said that in general, each zero in the left half plane blocks a specific input signal (Hag & Bernstein, 2007). The question is what happens in the case of a right half plane zero?

(Hag & Bernstein, 2007) illustrated this by looking at the response of a transfer function to an unbounded input signal such as $u(t) = e^t$. Fig. 2 shows the responses of two transfer functions, $G_1(s) = 2(s+1)/(s+1)(s+2)$ and $G_2(s) = 2(s-1)/(s+1)(s+2)$. It can be seen that what distinguishes a right half plane zero is the fact that it blocked the unbounded signal. With a better understanding of the character of right half plane zeros, a formal definition of a non-minimum phase system will be given. Interesting enough, a non-minimum phase system is defined as a system having either a zero or *a pole* in the right-half s-plane (Kuo & Golnaraghi, 2010). (Morari & Zafiriou, 1989) defined a non-minimum phase system as having a transfer function that contains zeros in the right half plane or time delays or both.
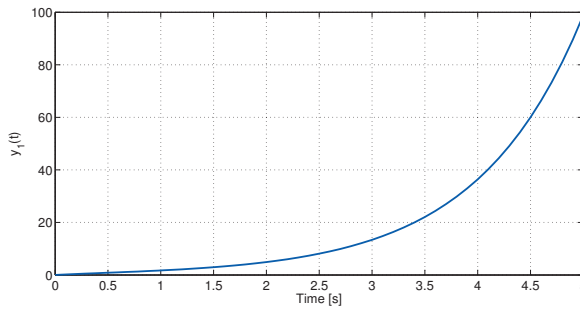
(a) Unbounded response of $G_1(s)$
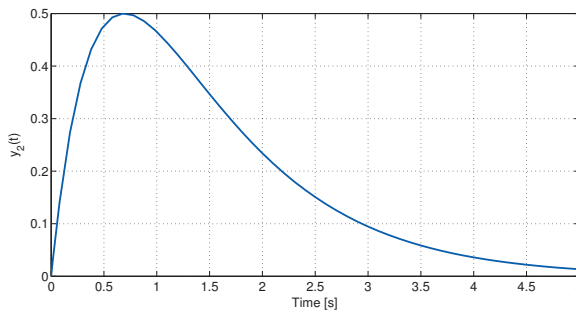
(b) Bounded response of $G_2(s)$

Fig. 2. Responses due to an unbounded input signal $u(t) = e^t$

In this chapter the focus is on non-minimum phase systems showing either *inverse response* (undershoot) or *time-delays*. Undershoot refers to an initial response in the opposite direction from the steady state. According to (Bernardo & Leon de la Barra, 1994) continuous systems having an odd number of real open right half plane zeros are characterised by an initial inverse response to a step input.

Systems having a time-delay characteristic is a special case of non-minimum phase systems (Waller & Nygardas, 1975). The Padé approximation is often used to approximate a time delay by a rational transfer function. Consider a first-order system with time-delay given by

$$G(s) = \frac{K}{1 + s\tau} e^{-sT}. \tag{15}$$

K represents the gain constant, τ the time constant, and T the time-delay of the system. The Padé approximation for the term e^{-sT} is given by

$$e^{-sT} \cong \frac{N_r(sT)}{D_r(sT)} \tag{16}$$

where

$$N_r(sT) = \sum_{k=0}^{r} \frac{(2r-k)!}{k!(r-k)!}(-sT)^k \tag{17}$$

$$D_r(sT) = \sum_{k=0}^{r} \frac{(2r-k)!}{k!(r-k)!}(sT)^k \tag{18}$$

and r is the order of the approximation (Silva et al., 2005).

Consider the function $G(s) = 2e^{-s}/(s+1)(s+2)$. The time-delay term can be approximated by a first order Padé approximation given by

$$e^{-s} \cong \frac{2T-s}{2T+s} = \frac{2-s}{2+s}, \tag{19}$$

and therefore the rational approximated version of $G(s)$ is given by

$$G_r(s) = \frac{2(-s+2)}{(s+1)(s+2)(s+2)}. \tag{20}$$

Fig. 3 plots the step responses of G and G_r and it is interesting to note that the response of G_r exhibits an initial inverse response. This also demonstrates a link between time delay-systems and inverse response systems.
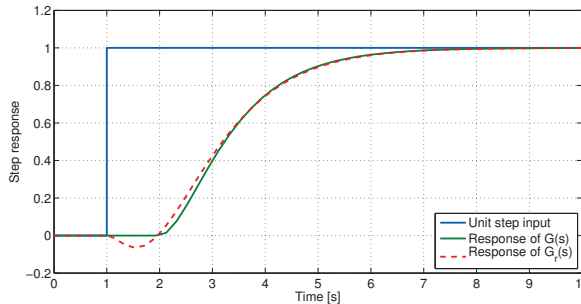

Fig. 3. First order Padé approximation of a time-delay system

4. Practical systems exhibiting non-minimum phase characteristics

Control engineers are often confronted with practical systems exhibiting an inverse response or a time-delay. An inverse response is found in boiler level control systems. An inverse response occurs when the flow rate of the feed water is increased by a step change, and the total volume of boiling water decreases for a short period (Astrom & Hägglund, 1995; Linoya & Altpeter, 1962). The same effect occurs in some aircraft regarding the step response of the elevator deflection to pitch angle (Franklin et al., 2010). In some high temperature gas-cooled reactors (HTGRs) such as the pebble bed modular reactor (PBMR) the power output also shows undershoot phenomena when helium is injected at the low-pressure side of the power conversion unit (PCU) (Uren et al., 2010).

5. Predictive PID controllers

5.1 Classical predictive approaches

Considering industrial applications, the simplicity and effectiveness of a control structure are vital attributes. This consideration can be more important than trying to implement more complex control structures. Therefore the basic structure of PID control is still considered, but with some enhancing adjustments. Like the Zielger-Nichols PID tuning rules, the Smith predictor concept have been around since the late 1950s (Bahill, 1983; Meyer et al., 1976; Smith, 1957; 1958). The internal model control (IMC) method which will also be discussed can be viewed as an extension of the Smith predictor (Astrom & Hägglund, 1995; Rivera et al., 1986)

5.1.1 Smith predictor structure

The undesirable effect of time-delays in feedback control loops are well known. One of the most popular ways to control systems with time delays effectively is by using a Smith predictor. The goal of the Smith predictor is mainly to eliminate the time delay from the characteristic equation and consequently allowing a larger controller gain to be used. The control structure of the Smith predictor is shown in Fig. 4. Let the plant be represented by

$$G_p(s) = G(s)e^{-Ts}, \tag{21}$$

where $T > 0$ is the time-delay. Along with the controller $G_c(s)$ an internal loop is added that simulates the plant dynamics. The plant model is given by

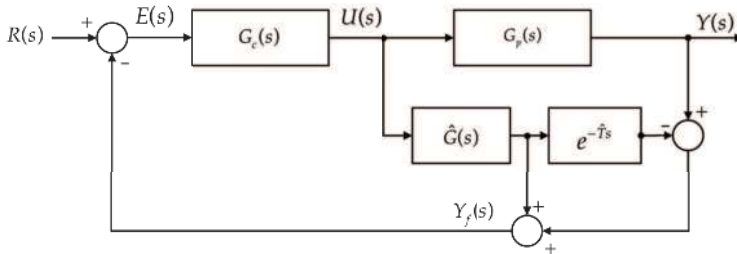$$G_m(s) = \hat{G}(s)e^{-\hat{T}s}. \tag{22}$$

Fig. 4. Smith predictor control structure

Considering Fig. 4 the feedback signal, $Y_f(s)$ is given by

$$Y_f(s) = \hat{G}(s)U(s) + (G(s)e^{-Ts} - \hat{G}(s)e^{-\hat{T}s})U(s) \tag{23}$$

If a "perfect" model of the plant is considered then

$$G(s) = \hat{G}(s), \tag{24}$$
$$T = \hat{T}. \tag{25}$$

This means that the feedback is only dependent on the model of the plant, that is

$$Y_f(s) = \hat{G}(s)U(s). \tag{26}$$

The relationship between the control variable and the system output is

$$U(s) = \frac{1}{G(s)e^{-Ts}}Y(s), \tag{27}$$

and since $G(s) = \hat{G}(s)$, Eq.(26) becomes

$$Y_f(s) = \hat{G}(s)\frac{1}{G(s)e^{-Ts}}Y(s) = e^{Ts}Y(s). \tag{28}$$

This shows that the internal loop containing the plant model feeds back a signal that is a prediction of the output, since e^{Ts} represents a prediction $y(t+T)$ in the time domain. The closed loop transfer function of the system can be determined by using

$$Y(s) = G(s)e^{-Ts}U(s), \tag{29}$$

$$U(s) = G_c(s)(R(s) - Y_f(s)), \tag{30}$$

and Eq. (26) to obtain

$$\frac{Y(s)}{R(s)} = \frac{G(s)e^{-Ts}G_c(s)}{1 + G(s)G_c(s)}. \tag{31}$$

According to (Dorf & Bishop, 2011) the sensitivity expression in this case can be defined as

$$S(s) = \frac{1}{1 + G(s)G_c(s)}. \tag{32}$$

As can be seen, the controller can now be designed without considering the effect of the time delay. (Hägglund, 1992; 1996) combined the properties of the Smith predictor with a PI controller to control a first order plant with a time delay. The transfer function of the plant is given by

$$G_p(s) = \frac{Ke^{-Ts}}{\tau s + 1}, \tag{33}$$

where $K > 0$ is the plant gain, τ the time constant and T the time-delay of the plant. The PI controller is given by

$$G_c(s) = K_p\left(1 + \frac{1}{\tau_i s}\right), \tag{34}$$

where the K_p is the proportional gain, and τ_i is the integral time constant. The control structure is given in Fig. 5

The time delay can be approximated by a first order Padé approximation with the time delay $\hat{T} > 0$. This control structure results in five parameters that need tuning $(K_p, \tau_i, \hat{K}, \hat{\tau}, \hat{T})$.

Example

Consider the following first order plant with a time-delay of two seconds

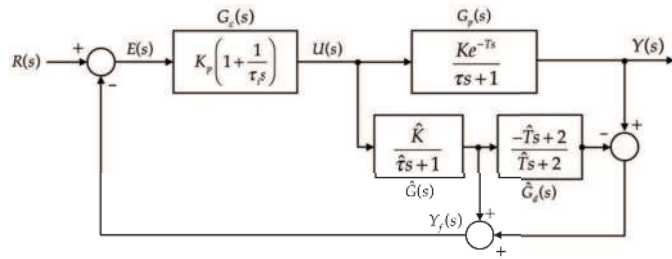$$G_p(s) = G(s)G_d(s) = \frac{2}{2s+1}e^{-2s}, \tag{35}$$

Fig. 5. PI with Smith predictor control structure

where $G_d(s)$ represents the time-delay dynamics. Let the model of the plant be given by

$$G_m(s) = \hat{G}(s)\hat{G}_d(s) = \frac{2}{2s+1}\frac{(-2s+2)}{(2s+2)}, \tag{36}$$

where $\hat{G}_d(s)$ represents the Padé approximation of the time-delay. The PI control constants are set to $K_p = 1$ and $\tau_i = 1.67$, resulting in the following PI controller

$$G_c(s) = (1 + \frac{0.6}{s}). \tag{37}$$

A predictive PID controller $C(s)$ as shown in Fig. 6 needs to be derived based on the predictive properties of the Smith predictor. PID controllers are sometimes augmented with a filter $F(s)$ to improve stability and dynamic response. By comparing the system transfer functions of the
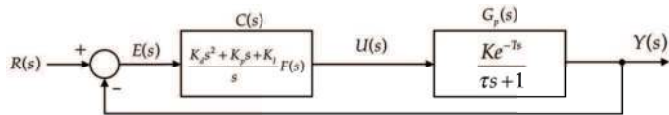

Fig. 6. PID controller based on Smith predictor characteristics

PI with Smith predictor control structure in Fig. 5 and the PID control structure in Fig. 6 a PID controller can be derived based on the Smith predictor qualities:

$$T_{Smith}(s) = T_{PID}(s), \tag{38}$$

$$\frac{\hat{G}(s)\hat{G}_d(s)G_c(s)}{1 + \hat{G}_d(s)G_c(s)} = \frac{C(s)\hat{G}(s)\hat{G}_d}{1 + C(s)\hat{G}(s)\hat{G}_d}, \tag{39}$$

$$C(s) = \frac{G_c(s)}{1 + \hat{G}(s)G_c(s) - \hat{G}(s)G_c(s)\hat{G}_d(s)} \tag{40}$$

$C(s)$ can therefore be considered as a predictive PID controller. Substituting the numerical values leads to

$$C(s) = \frac{4s^4 + 14.4s^3 + 16.2s^2 + 7.4s + 1.2}{4s^4 + 20s^3 + 17.8s^2 + 4.4s}. \tag{41}$$

Applying model reduction techniques $C(s)$ reduces to a PID control structure which is a second order transfer function

$$C(s) = \frac{1.002s^2 + 2.601s + 1.098}{s(s + 4.025)}, \tag{42}$$

where $K_d = 1.002$, $K_p = 2.601$, $K_i = 1.098$ and $F(s) = 1/(s + 4.025)$. Fig. 7 shows the time response of the system output along with the control variable. It can be seen that the control signal acts immediately and not after the occurrence of the time-delay, demonstrating the predictive properties of the PID controller. Fig. 8 shows the time response of the
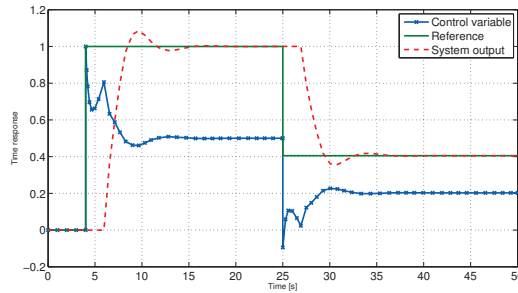

Fig. 7. Time response of system with predictive PID controller $C(s)$ based on Smith predictor

system for larger time-delays. It can be seen that the control performance deteriorates as the time-delay increases. This is due to the limited approximation capabilities of the first order Padé approximation.
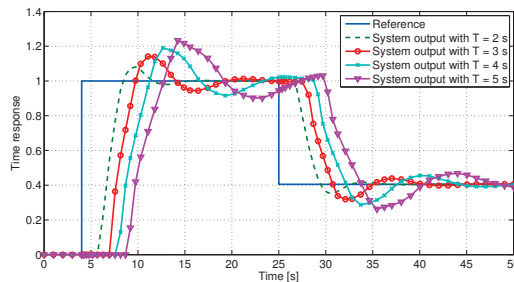

Fig. 8. Time responses of control system based on Smith predictor for different time-delays

5.1.2 Internal model control

The internal model control (IMC) design method starts with the assumption that a model of the system is available that allows the prediction of the system output response due to a output of the controller. In this discussion it is also assumed that the model is a "perfect" representation of the plant. The basic structure of IMC is given in Fig. 9 (Brosilow & Joseph, 2002; Garcia & Morari, 1982). The transfer functions of the plant, the IMC controller and plant model is given by $G_p(s, \varepsilon)$, $G_{IMC}(s)$ and $G_m(s)$ respectively. In the case when the model is not

a perfect representation of the actual plant the tuning parameter ε is used to compensate for modelling errors.
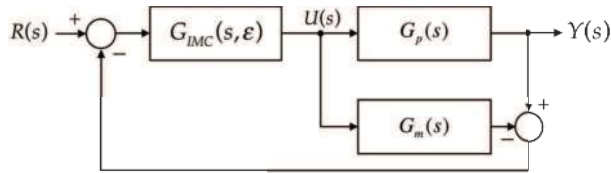

Fig. 9. Internal model control structure

The structure of Fig. 9 can be rearranged into a classical PID structure as shown in Fig. 10. This allows the PID controller to have predictive properties derived from the IMC design.
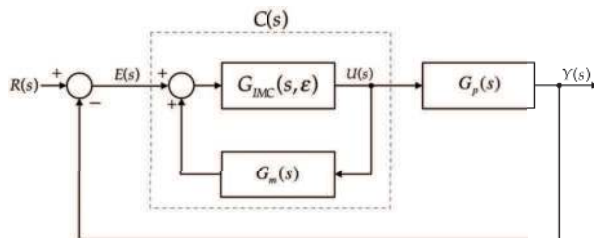

Fig. 10. Classical feedback representation of the IMC structure

The transfer function of the classical controller $C(s)$ is given by

$$C(s) = \frac{U(s)}{E(s)} = \frac{G_{IMC}(s,\varepsilon)}{1 - G_m(s)G_{IMC}(s,\varepsilon)},$$

(43)

and the transfer function of the system is given by

$$T(s) = \frac{Y(s)}{R(s)} = \frac{G_p(s)C(s)}{1 + G_p(s)C(s)}.$$

(44)

A "perfect" controller $C(s)$ would drive the output $Y(s)$ of the system to track the reference input $Y(s)$ instantaneously, that is

$$Y(s) = R(s),$$

(45)

and this requires that

$$G_{IMC}(s,\varepsilon)G_p(s) = 1,$$

(46)

$$G_m(s) = G_p(s).$$

(47)

To have a "perfect" controller, a "perfect" model is needed. Unfortunately it is not possible to model the dynamics of the plant perfectly. However, depending on the controller design method, the controller can come close to show the inverse response of the plant model. Usually the design method incorporates a tuning parameter to accommodate modelling errors.

The plant considered is a non-minimum phase system of the following form

$$G_p(s) = \frac{N(s)}{D(s)} e^{-Ts} = \frac{N_-(s)N_+(s)}{D(s)} e^{-Ts}, \tag{48}$$

where $N_-(s)$ represents a polynomial containing only left half plane zeros, and $N_+(s)$ a polynomial containing only right half plane zeros. The IMC controller of the plant in Eq.(48) is given by

$$G_{IMC}(s, \varepsilon) = \frac{D(s)}{N_-(s)N_+(-s)(\varepsilon s + 1)^r}, \tag{49}$$

where the zeros of $N_+(-s)$ are all in the left half plane and are the mirror images of the zeros of $N_+(s)$. The filter constant ε is a tuning parameter that can be used to avoid noise amplification and to accommodate modelling errors; and r is the relative order of $N(s)/D(s)$ (Brosilow & Joseph, 2002).

Example
Consider the following non-minimum phase system

$$G_p(s) = \frac{2(-2s + 2)}{(2s + 1)(2s + 2)}. \tag{50}$$

The IMC controller can be derived by using Eq.(49), but in order to ensure zero offset for step inputs $G_p(s)$ is adapted as follows

$$G_p(s) = \frac{2(-2s + 2)}{2(2s + 1)(2s + 2)}. \tag{51}$$

Then

$$G_{IMC}(s) = \frac{(2s + 1)(s + 1)}{(s + 1)(\varepsilon s + 1)^r}, \tag{52}$$

and let $\varepsilon = 1$ and $r = 1$ then

$$G_{IMC}(s) = \frac{(2s + 1)(s + 1)}{(s + 1)(s + 1)}. \tag{53}$$

The classical controller for this case is given by

$$C(s) = \frac{G_{IMC}(s)}{1 - G_p(s)G_{IMC}(s)} = \frac{1}{2}\frac{(2s + 1)(s + 1)}{s^2 + 3s} = \frac{s^2 + 1.5s + 0.5}{s(s + 3)}. \tag{54}$$

The form of $C(s)$ corresponds to the form of a PID controller (Dorf & Bishop, 2011):

$$C_{PID}(s) = \frac{K_d(s^2 + as + b)}{s} \tag{55}$$

where $a = K_p/K_d$ and $b = K_i/K_d$. The IMC-based controller, Eq.(54), is therefore a PID controller augmented with a filter $F(s) = 1/(\varepsilon s + 1)^r$ and is called and IMC-PID controller (Lee et al., 2008). Fig.11 shows the time response of the system output along with the control variable.
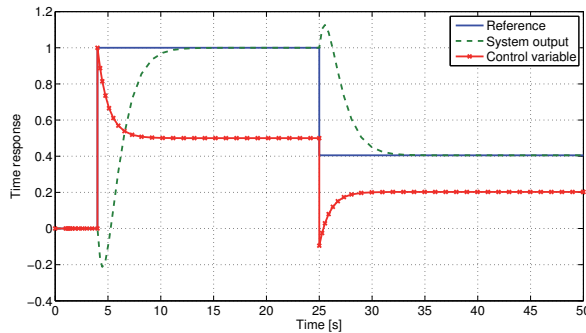
Fig. 11. Time response of control system based on IMC

5.2 Modern predictive approaches

One of the most successful developments in modern control engineering is the area of model predictive control (MPC). It is an optimal control structure utilising a receding horizon principle. This method have found wide-spread application in process industries and research in the field is very active (Wang, 2009). In MPC the control law is computed via optimisation of a quadratic cost function and a plant model is used to predict the future output response to possible future control trajectories. These predictions are computed for a finite time horizons, but only the first value of the optimal control trajectory is used at each sample instant.

Following a model predictive approach for the design of PID controllers is a challenging task. Two routes can be followed namely a *restricted model approach* or a *control signal matching approach* (Johnson & Moradi, 2005; Tan et al., 2000; 2002). In this section the restricted model approach will be considered. This approach formulates the control problem in terms the generalised predictive control (GPC) algorithm. The model used by the controller is restricted to second order such that the predictive control law that emerges has a PID structure. The following control algorithm is discussed in discrete-time since it offers a more natural setting for the derivation of predictive control techniques. It also simplifies the description of the design process and has a strong relevance to industrial applications when presented in discrete-time (Wang, 2009).

5.2.1 The GPC-based algorithm
Augmented state space model

The main idea is to derive an MPC control law equivalent to the second order control law of a PID controller. This can be done by developing an MPC control law, but considering a second-order general plant (Tan et al., 2000; 2002). Consider a single-input, single-output model of a plant described by:

$$\mathbf{X}_m(k+1) = \mathbf{A}_m\mathbf{X}_m(k) + \mathbf{B}_m u(k), \tag{56}$$

$$y(k) = \mathbf{C}_m\mathbf{X}_m(k), \tag{57}$$

where $u(k)$ is the input variable and $y(k)$ is the output variable; and $\mathbf{X}_m$ is the state variable vector of dimension $n = 2$, since a second order plant is considered. Note that the plant model has $u(k)$ as its input. This needs to be altered since a predictive controller needs to be designed. A common first step is to augment the model with an integrator (Wang, 2009). By

taking the difference operation on both sides of Eq.(56) the following is obtained

$$\mathbf{X}_m(k+1) - \mathbf{X}_m(k) = \mathbf{A}_m(\mathbf{X}_m(k) - \mathbf{X}_m(k-1)) + \mathbf{B}(u(k) - u(k-1)). \tag{58}$$

The difference of the state variables and output is given by

$$\Delta\mathbf{X}_m(k+1) = \mathbf{X}_m(k+1) - \mathbf{X}_m(k), \tag{59}$$

$$\Delta\mathbf{X}_m(k) = \mathbf{X}_m(k) - \mathbf{X}_m(k-1), \tag{60}$$

$$\Delta u(k) = u(k) - u(k-1). \tag{61}$$

The integrating effect is obtained by connecting $\Delta\mathbf{X}_m(k)$ to the output $y(k)$. To do so the new augmented state vector is chosen to be

$$\mathbf{X}(k) = \begin{bmatrix} \Delta\mathbf{X}_m(k)^T & y(k) \end{bmatrix}^T. \tag{62}$$

where the superscript ^{T} indicates the matrix transpose. The state equation can then be written as

$$\Delta\mathbf{X}_m(k+1) = \mathbf{A}_m\Delta\mathbf{X}_m(k) + \mathbf{B}_m\Delta u(k), \tag{63}$$

and the output equation becomes

$$y(k+1) - y(k) = \mathbf{C}_m(\mathbf{X}_m(k+1) - \mathbf{X}_m(k)) = \mathbf{C}_m\Delta\mathbf{X}_m(k+1) \tag{64}$$

$$= \mathbf{C}_m\mathbf{A}_m\Delta\mathbf{X}_m(k) + \mathbf{C}_m\mathbf{B}_m\Delta u(k). \tag{65}$$

Eqs. (63) and (64) can be written in state space form where

$$\begin{bmatrix} \Delta\mathbf{X}_m(k+1) \\ y(k+1) \end{bmatrix} = \begin{bmatrix} \mathbf{A}_m & \mathbf{O}_m^T \\ \mathbf{C}_m\mathbf{A}_m & 1 \end{bmatrix} \begin{bmatrix} \Delta\mathbf{X}_m(k) \\ y(k) \end{bmatrix} + \begin{bmatrix} \mathbf{B}_m \\ \mathbf{C}_m\mathbf{B}_m \end{bmatrix} \Delta u(k), \tag{66}$$

$$y(k) = \begin{bmatrix} \mathbf{O}_m & 1 \end{bmatrix} \begin{bmatrix} \Delta\mathbf{X}_m(k) \\ y(k) \end{bmatrix}, \tag{67}$$

where $\mathbf{O}_m = \begin{bmatrix} 0 & 0 & \cdots & 0 \end{bmatrix}$ is a $1 \times n$ vector, and $n = 2$ in the predictive PID case. This augmented model will be used in the GPC-based predictive PID control design.

Prediction

The next step in the predictive PID control design is to predict the second order plant output with the future control variable as the adjustable parameter. This prediction is done within one optimisation window. Let $k > 0$ be the sampling instant. Then the future control trajectory is denoted by

$$\Delta u(k), \Delta u(k+1), \cdots, \Delta u(k+N_c-1), \tag{68}$$

where N_c is called the control horizon. The future state variables are denoted by

$$\mathbf{X}(k+1|k), \mathbf{X}(k+2|k), \cdots, \mathbf{X}(k+m|k), \cdots, \mathbf{X}(k+N_p|k), \tag{69}$$

where N_p is the length of the optimisation window and $\mathbf{X}(k+m|k)$ is the predicted state variables at $k+m$ with given current plant information $\mathbf{X}(k)$ and $N_c \leq N_p$.

The future states of the plant are calculated by using the plant state space model:

$$\mathbf{X}(k+1|k) = \mathbf{A}_m\mathbf{X}(k) + \mathbf{B}_m\Delta u(k),$$
$$\mathbf{X}(k+2|k) = \mathbf{A}_m\mathbf{X}(k+1|k) + \mathbf{B}_m\Delta u(k+1),$$
$$= \mathbf{A}_m^2\mathbf{X}(k) + \mathbf{A}_m\mathbf{B}_m\Delta u(k) + \mathbf{B}_m\Delta u(k+1),$$
$$\vdots$$
$$\mathbf{X}(k+N_p|k) = \mathbf{A}_m^{N_p}\mathbf{X}(k) + \mathbf{A}_m^{N_p-1}\mathbf{B}_m\Delta u(k) + \mathbf{A}_m^{N_p-2}\mathbf{B}_m\Delta u(k+1)$$
$$+ \cdots + \mathbf{A}_m^{N_p-N_c}\mathbf{B}_m\Delta u(k+N_c-1).$$

The predicted output variables are as follows:

$$y(k+1|k) = \mathbf{C}_m\mathbf{A}_m\mathbf{X}(k) + \mathbf{C}_m\mathbf{B}_m\Delta u(k),$$
$$y(k+2|k) = \mathbf{C}_m\mathbf{A}_m^2\mathbf{X}(k) + \mathbf{C}_m\mathbf{A}_m\mathbf{B}_m\Delta u(k) + \mathbf{C}_m\mathbf{B}_m\Delta u(k+1),$$
$$y(k+3|k) = \mathbf{C}_m\mathbf{A}_m^3\mathbf{X}(k) + \mathbf{C}_m\mathbf{A}_m^2\mathbf{B}_m\Delta u(k) + \mathbf{C}_m\mathbf{A}_m\mathbf{B}_m\Delta u(k+1)$$
$$+ \mathbf{C}_m\mathbf{B}_m\Delta u(k+2),$$
$$\vdots$$
$$y(k+N_p|k) = \mathbf{C}_m\mathbf{A}_m^{N_p}\mathbf{X}(k) + \mathbf{C}_m\mathbf{A}_m^{N_p-1}\mathbf{B}_m\Delta u(k) + \mathbf{C}_m\mathbf{A}_m^{N_p-2}\mathbf{B}_m\Delta u(k+1)$$
$$+ \cdots + \mathbf{C}_m\mathbf{A}_m^{N_p-N_c}\mathbf{B}_m\Delta u(k+N_c-1).$$

The equations above can now be ordered in matrix form as

$$\mathbf{Y} = \mathbf{F}\mathbf{X}(k) + \Phi\Delta\mathbf{U}, \tag{70}$$

where

$$\mathbf{Y} = \begin{bmatrix} y(k+1|k) & y(k+2|k) & y(k+3|k) & \ldots & y(k+N_p|k) \end{bmatrix}^T, \tag{71}$$
$$\Delta\mathbf{U} = \begin{bmatrix} \Delta u(k) & \Delta u(k+1) & \Delta u(k+3) & \ldots & \Delta u(k+N_c-1) \end{bmatrix}^T, \tag{72}$$

and

$$\mathbf{F} = \begin{bmatrix} \mathbf{C}_m\mathbf{A}_m \\ \mathbf{C}_m\mathbf{A}_m^2 \\ \mathbf{C}_m\mathbf{A}_m^3 \\ \vdots \\ \mathbf{C}_m\mathbf{A}_m^{N_p} \end{bmatrix}, \tag{73}$$

$$\Phi = \begin{bmatrix} \mathbf{C}_m\mathbf{B}_m & 0 & 0 & \ldots & 0 \\ \mathbf{C}_m\mathbf{A}_m\mathbf{B}_m & \mathbf{C}_m\mathbf{B}_m & 0 & \ldots & 0 \\ \mathbf{C}_m\mathbf{A}_m^2\mathbf{B}_m & \mathbf{C}_m\mathbf{A}_m\mathbf{B}_m & \mathbf{C}_m\mathbf{B}_m & \ldots & 0 \\ \vdots & & & & \\ \mathbf{C}_m\mathbf{A}_m^{N_p-1}\mathbf{B}_m & \mathbf{C}_m\mathbf{A}_m^{N_p-2}\mathbf{B}_m & \mathbf{C}_m\mathbf{A}_m^{N_p-3}\mathbf{B}_m & \ldots & \mathbf{C}_m\mathbf{A}_m^{N_p-N_c}\mathbf{B}_m \end{bmatrix}. \tag{74}$$

Optimisation and control design

Let $r(k)$ be the set-point signal at sample time k. The idea behind the predictive PID control methodology is to drive the predicted output signal as close as possible to the set-point signal. It is assumed that the set-point signal remains constant during the optimisation window, N_p. Consider the following quadratic cost function which is very similar to the one obtained by (Tan et al., 2002)

$$J = (\mathbf{r} - \mathbf{y})^T(\mathbf{r} - \mathbf{y}) + \Delta\mathbf{U}^T\overline{\mathbf{R}}\Delta\mathbf{U}, \tag{75}$$

where the set-point information is given by

$$\mathbf{r}^T = \begin{bmatrix} 1\ 1\ \dots\ 1 \end{bmatrix} \times r(k), \tag{76}$$

and the dimension of $\mathbf{r}$ is $N_p \times 1$. The cost function, Eq.(75) comprises two parts, the first part focus on minimising the errors between the reference and the output; the second part focus on minimising the control effort. $\overline{\mathbf{R}}$ is a diagonal weight matrix given by

$$\overline{\mathbf{R}} = r_w \times \mathbf{I} \tag{77}$$

where $\mathbf{I}$ is an $N_c \times N_c$ identity matrix and the weight $r_w \geq 0$ is used to tune the closed-loop response. The optimisation problem is defined such that an optimal $\Delta\mathbf{U}$ can be found that minimises the cost function J. Substituting Eq.(70) into Eq.(75), J is expressed as

$$J = (\mathbf{r} - \mathbf{FX}(k))^T(\mathbf{r} - \mathbf{FX}(k)) - 2\Delta\mathbf{U}^T\Phi^T(\mathbf{r} - \mathbf{FX}(k)) + \Delta\mathbf{U}^T(\Phi^T\Phi + \overline{\mathbf{R}})\Delta\mathbf{U}. \tag{78}$$

The solution that minimises the cost function J can be obtained by solving

$$\frac{\partial J}{\partial \Delta\mathbf{U}} = 2\Phi^T(\mathbf{r} - \mathbf{FX}(k)) + 2(\Phi^T\Phi + \overline{\mathbf{R}})\Delta\mathbf{U} = \mathbf{0}. \tag{79}$$

Therefore, the optimal control law is given as

$$\Delta\mathbf{U} = (\Phi^T\Phi + \overline{\mathbf{R}})^{-1}\Phi^T(\mathbf{r} - \mathbf{FX}(k)) \tag{80}$$

or

$$\Delta\mathbf{U} = (\Phi^T\Phi + \overline{\mathbf{R}})^{-1}\Phi^T\mathbf{e}(k) \tag{81}$$

where $\mathbf{e}(k)$ represents the errors at sample k.

Emerging predictive control with PID structure

The discrete configuration of a PID controller has the following form (Huang et al., 2002; Phillips & Nagle, 1995):

$$u(k) = K_p e(k) + K_i \sum_{n=1}^{k} e(n) + K_d(e(k) - e(k-1)), \tag{82}$$

or

$$u(z) = \frac{q_0 + q_1 z^{-1} + q_2 z^{-2}}{1 - z^{-1}} e(z), \tag{83}$$

where K_p, K_i and K_d are the proportional, integral and derivative gains, respectively, and

$$q_0 = K_p + K_i + K_d, \tag{84}$$
$$q_1 = -K_p - 2K_d, \tag{85}$$
$$q_2 = K_d. \tag{86}$$

By taking the difference on both sides of Eq.(82), the velocity form of the PID control law is obtained:

$$\Delta u(k) = K_p[e(k) - e(k-1)] + K_i e(k) + K_d[e(k) - 2e(k-1) + e(k-2)]. \tag{87}$$

This equation can be written in matrix form as (Katebi & Moradi, 2001):

$$\Delta \mathbf{U}(k) = \mathbf{K}\mathbf{e}(k) = \mathbf{K}[\mathbf{r}(k) - \mathbf{y}(k)] \tag{88}$$

where

$$\mathbf{K} = \begin{bmatrix} K_p & K_i & K_d \end{bmatrix} \begin{bmatrix} 0 & -1 & 1 \\ 0 & 0 & 1 \\ 1 & -2 & 1 \end{bmatrix}, \tag{89}$$

and

$$\mathbf{y}(k) = \begin{bmatrix} y(k-2) & y(k-1) & y(k) \end{bmatrix}^T \tag{90}$$
$$\mathbf{e}(k) = \begin{bmatrix} e(k-2) & e(k-1) & e(k) \end{bmatrix}^T \tag{91}$$
$$\mathbf{r}(k) = \begin{bmatrix} r(k-2) & r(k-1) & r(k) \end{bmatrix}^T. \tag{92}$$

By equating Eq.(81) to Eq.(88)the following is obtained

$$\Delta \mathbf{U}(k) = (\Phi^T \Phi + \overline{\mathbf{R}})^{-1} \Phi^T \mathbf{e}(k) = \mathbf{K}^T \mathbf{e}(k) \tag{93}$$

and therefore the predictive PID controller constants are given by

$$\mathbf{K}^T = (\Phi^T \Phi + \overline{\mathbf{R}})^{-1} \Phi^T, \tag{94}$$

or

$$\begin{bmatrix} K_d & (-2K_d - Kp) & (K_d + K_i + K_p) \end{bmatrix}^T = (\Phi^T \Phi + \overline{\mathbf{R}})^{-1} \Phi^T. \tag{95}$$

Example
Consider the following discrete-time state space model of a non-minimum phase system

$$\dot{\mathbf{X}}(k) = \begin{bmatrix} -0.0217 & -0.3141 \\ 0.3141 & 0.7636 \end{bmatrix} \mathbf{X}(k) + \begin{bmatrix} 0.3141 \\ 0.2364 \end{bmatrix} u(k), \tag{96}$$
$$y(k) = \begin{bmatrix} -1 & 2 \end{bmatrix} \mathbf{X}(k). \tag{97}$$

The first step is to create the augmented model for the MPC design, and choose the values of the prediction and control horizon. In this example the control horizon is selected to be $N_c = 3$ and the prediction horizon is $N_p = 20$. Also the sampling period in this case is chosen as 1 second and a 100 samples is considered. Then the predicted output is given by Eq. 70 where

$$
\mathbf{F} = \begin{bmatrix} 0.6500 & 1.8413 & 1.0000 \\ 1.2143 & 3.0432 & 1.0000 \\ 1.5796 & 3.7836 & 1.0000 \\ \vdots & \vdots & \vdots \\ 2.1515 & 4.9290 & 1.0000 \\ 2.1516 & 4.9292 & 1.0000 \\ 2.1517 & 4.9294 & 1.0000 \end{bmatrix}, \Phi = \begin{bmatrix} 0.1587 & 0 & 0 \\ 0.7982 & 0.1587 & 0 \\ 1.2595 & 0.7982 & 0.1587 \\ \vdots & \vdots & \vdots \\ 1.9996 & 1.9993 & 1.9989 \\ 1.9998 & 1.9996 & 1.9993 \\ 1.9998 & 1.9998 & 1.9996 \end{bmatrix}, \tag{98}
$$

are matrices having 20 rows and 3 columns. By choosing a weight $r_w = 0.9$ the optimal control law (Eq. (81)) is given by

$$
\Delta \mathbf{U} = \begin{bmatrix} 0.0628 & 0.2602 & 0.2108 & \cdots & -0.0144 & -0.0144 & -0.0145 \\ -0.0554 & -0.1681 & 0.0617 & \cdots & 0.0035 & 0.0035 & 0.0035 \\ -0.0085 & -0.0976 & -0.2766 & \cdots & 0.0452 & 0.0453 & 0.0453 \end{bmatrix} \mathbf{e}(k), \tag{99}
$$

where the matrix multiplied with the error vector has 3 rows and 20 columns.

Fig. 12 shows the closed loop response of the system output along with the control variable. It can be seen that the control variable acts immediately and not after the occurrence of the
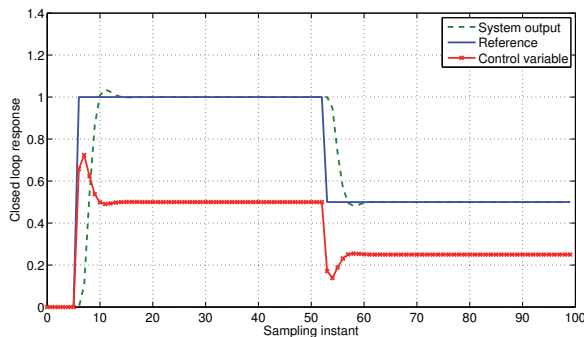

Fig. 12. Closed loop response of a system with an MPC controller having a PID control structure

time-delay. This shows that the MPC controller with a PID structure demonstrates predictive properties. An improvement in the control performance can be seen compared to the previous classical predictive controllers. This is due to the fact that the control law is computed via the optimisation of a quadratic cost function.

6. Conclusions

In this chapter both classical and modern predictive control methods for non-minimum phase systems were considered. Two popular methods considered in the classical approach were the Smith predictor and internal model control (IMC). These two methods utilise a plant model to predict the future output of the plant. This results in a control law that acts immediately on the reference input avoiding instability and sluggish control. In the classical approach the

Smith predictor and IMC structures were used to derive the predictive PID control constants. The predictive PID controller can effectively deal with the non-minimum phase effect.

A modern approach to predictive PID control features a different methodology. A generalised predictive control algorithm was considered. In this approach the model predictive controller is reduced to the same structure as a PID controller for second-order systems (Eq. (87)). In this case the equivalent PID constants changes at every sample since an optimisation routine using a cost function (Eq. (78)) is followed at each sample. The controller structure can further be adapted to be used as a design method to derive optimal values of PID gains (Eq. (95)). The novelty of this method lies in the fact that time-delays are incorporated without any need for approximation.

7. References

Astrom, K. & Hägglund, T. (1995). *PID Controllers:Theory, Design, and Tuning*, 2nd edn, Instrument Society of America.

Bahill, A. (1983). A simple adaptive smith-predictor for controlling time-delay systems: A tutorial, *Control Systems Magazine, IEEE* 3(2): 16 – 22.

Bernardo, A. & Leon de la Barra, S. (1994). On undershoot in siso systems, *Automatic Control, IEEE Transactions on* 39(3): 578 –581.

Brosilow, C. & Joseph, B. (2002). *Techniques of model-based control*, Prentice-Hall international series in the physical and chemical engineering sciences, Prentice Hall.

Dorf, R. & Bishop, R. (2011). *Modern control systems*, 12th edn, Pearson Prentice Hall.

Franklin, G. F., Powell, D. J. & Emami-Naeini, A. (2010). *Feedback Control of Dynamic Systems*, 6th edn, Prentice Hall PTR, Upper Saddle River, NJ, USA.

Garcia, C. E. & Morari, M. (1982). Internal model control. a unifying review and some new results, *Industrial & Engineering Chemistry Process Design and Development* 21(2): 308–323.

Hag, J. & Bernstein, D. (2007). Nonminimum-phase zeros - much to do about nothing - classical control - revisited part ii, *Control Systems Magazine, IEEE* 27(3): 45 –57.

Hägglund, T. (1992). A predictive PI controller for processes with long dead times, *Control Systems Magazine, IEEE* 12(1): 57 –60.

Hägglund, T. (1996). An industrial dead-time compensating PI controller, *Control Engineering Practice* 4(6): 749 – 756.

Huang, S., Tan, K. & Lee, T. (2002). *Applied predictive control*, Advances in industrial control, Springer.

Johnson, M. & Moradi, M. (2005). *PID Control: New Identification and Design Methods*, Springer.

Katebi, M. & Moradi, M. (2001). Predictive pid controllers, *Control Theory and Applications, IEE Proceedings -* 148(6): 478 –487.

Kuo, B. C. & Golnaraghi, F. (2010). *Automatic Control Systems*, 9th edn, John Wiley & Sons, Inc., New York, NY, USA.

Lee, M., Shamsuzzoha, M. & Vu, T. N. L. (2008). Imc-pid approach: An effective way to get an analytical design of robust pid controller, *Control, Automation and Systems, 2008. ICCAS 2008. International Conference on*, pp. 2861 –2866.

Linoya, K. & Altpeter, R. J. (1962). Inverse response in process control, *Industrial & Engineering Chemistry* 54(7): 39–43.

Meyer, C., Seborg, D. E. & Wood, R. K. (1976). A comparison of the smith predictor and conventional feedback control, *Chemical Engineering Science* 31(9): 775 – 778.

Miller, R. M., Shah, S. L., Wood, R. K. & Kwok, E. K. (1999). Predictive pid, *ISA Transactions* 38(1): 11 – 23.

Mita, T. & Yoshida, H. (1981). Undershooting phenomenon and its control in linear multivariable servomechanisms, *Automatic Control, IEEE Transactions on* 26(2): 402 – 407.

Moradi, M., Katebi, M. & Johnson, M. (2001). Predictive pid control: a new algorithm, *Industrial Electronics Society, 2001. IECON '01. The 27th Annual Conference of the IEEE*, Vol. 1, pp. 764 –769 vol.1.

Morari, M. & Zafiriou, E. (1989). *Robust process control*, Prentice Hall.

Phillips, C. & Nagle, H. (1995). *Digital control system analysis and design*, Prentice Hall.

Rivera, D. E., Morari, M. & Skogestad, S. (1986). Internal model control: Pid controller design, *Industrial & Engineering Chemistry Process Design and Development* 25(1): 252–265.

Sato, T. (2010). Design of a gpc-based pid controller for controlling a weigh feeder, *Control Engineering Practice* 18(2): 105 – 113. Special Issue of the 3rd International Symposium on Advanced Control of Industrial Processes.

Silva, G. J., Datta, A. & P., B. S. (2005). *PID controllers for time-delay systems*, Birkhauser Boston.

Smith, O. (1957). Close control of loops with dead time, *Chemical Engineering Progress* 53: 217–219.

Smith, O. (1958). *Feedback Control Systems*, McGraw-Hill, New York.

Tan, K. K., Huang, S. N. & Lee, T. H. (2000). Development of a gpc-based pid controller for unstable systems with deadtime, *ISA Transactions* 39(1): 57 – 70.

Tan, K. K., Lee, T. H., Huang, S. N. & Leu, F. M. (2002). Pid control design based on a gpc approach, *Industrial & Engineering Chemistry Research* 41(8): 2013–2022.

Tan, K. K., Lee, T. H. & Leu, F. M. (2001). Predictive pi versus smith control for dead-time compensation, *ISA Transactions* 40(1): 17 – 29.

Uren, K., Van Schoor, G. & Van Niekerk, C. (2010). Optimal power control of a three-shaft brayton cycle based power conversion unit, *South African institute of electrical engineers* 101: 60–67.

Vidyasagar, M. (1986). On undershoot and nonminimum phase zeros, *Automatic Control, IEEE Transactions on* 31(5): 440 – 440.

Waller, K. V. T. & Nygardas, C. G. (1975). On inverse repsonse in process control, *Industrial & Engineering Chemistry Fundamentals* 14(3): 221–223.

Wang, L. (2009). *Model predictive control system design and implementation using MATLAB*, Advances in Industrial Control, Springer.
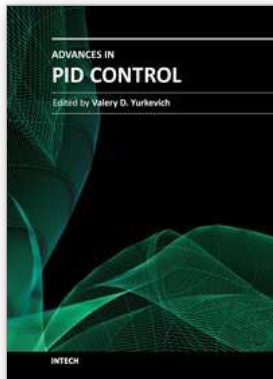
Advances in PID Control

Edited by Dr. Valery D. Yurkevich

Since the foundation and up to the current state-of-the-art in control engineering, the problems of PID control steadily attract great attention of numerous researchers and remain inexhaustible source of new ideas for process of control system design and industrial applications. PID control effectiveness is usually caused by the nature of dynamical processes, conditioned that the majority of the industrial dynamical processes are well described by simple dynamic model of the first or second order. The efficacy of PID controllers vastly falls in case of complicated dynamics, nonlinearities, and varying parameters of the plant. This gives a pulse to further researches in the field of PID control. Consequently, the problems of advanced PID control system design methodologies, rules of adaptive PID control, self-tuning procedures, and particularly robustness and transient performance for nonlinear systems, still remain as the areas of the lively interests for many scientists and researchers at the present time. The recent research results presented in this book provide new ideas for improved performance of PID control applications.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Kenny Uren and George van Schoor (2011). Predictive PID Control of Non-Minimum Phase Systems, Advances in PID Control, Dr. Valery D. Yurkevich (Ed.), ISBN: 978-953-307-267-8, InTech, Available from: http://www.intechopen.com/books/advances-in-pid-control/predictive-pid-control-of-non-minimum-phase-systems

INTECH
open science | open minds