Tuning of PID-type controllers

Document status and date:
Published: 01/01/2004

Document Version:
Publisher's PDF, also known as Version of Record (includes final page, issue and volume numbers)

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

TU/e technische universiteit eindhoven

DAF

TUNING OF PID-TYPE CONTROLLERS: LITERATURE OVERVIEW

G.M. van der Zalm
DCT-report nr. 2004.54
DAF-report nr. 51051/04-050
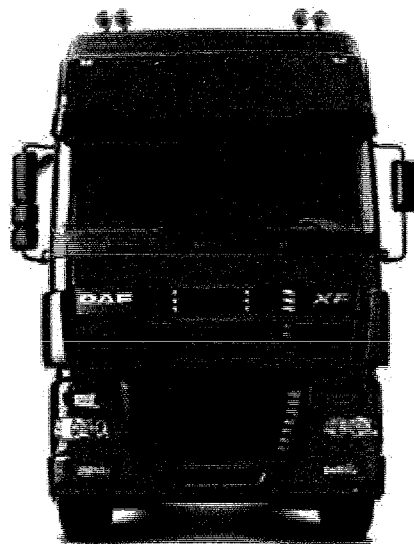
Examination board

Prof. dr. ir. Maarten Steinbuch (TU/e)   - Chairman
Dr. ir. Rudolf Huisman (DAF Trucks N.V.)   - Coach, advisor
Dr. ir. Frans Veldpaus (TU/e)   - Coach, member
Prof. dr. ir. Paul van den Bosch (TU/e)   - Member
Ing. Martin van Ruiswijk (DAF Trucks N.V.)   - Advisor

Contents

Introduction

In the past decades, control theory has gone through major developments. Advanced and intelligent control algorithms have been developed. However, the PID-type controller remains the most popular in industry, studies even indicate that approximately 90% of all industrial controllers is of the PID-type [19]. Reasons for this are the simplicity of this control law and the few tuning parameters. Hundreds of tools, methods and theories are available for this purpose. However, finding appropriate parameters for the PID controller is still a difficult task, so in practice control engineers still often use trial and error for the tuning process. This literature overview gives an impression of a number of the available methods for PID control design, and discusses the (dis)advantages and applicability.

Chapter 1

Simple plant models

Although design methods exist that do not use a process model, it is almost inevitable to derive a model or use some plant information to achieve satisfying results. Often, in the PID-tuning literature very simple models are used. Here, the low-order models most common in the literature are discussed. More sophisticated models can be obtained by either first principle modelling, performing FRF measurements, e.g. discussed in [20] or using approximate realisation techniques [12], [32]. These methods are not discussed here.

1.1 Step response models

The dynamics of a process can also be determined from the response of the process to step inputs. The dynamics of a linear system are in principle uniquely given from such a transient response, provided that the system is at rest initially and the measurement is noise free. In practice, these conditions are not fulfilled, which limits the order of the resulting models. For PID controller tuning, however, simple models are often sufficient. An advantage of step response methods is that they are very simple to use, and are intuitively easy to understand.

1.1.1 Two parameter model

Ziegler and Nichols [42] recognised that in the process industry, step responses of a large number of processes exhibit a step response like that shown in figure 1.1. They called this the process reaction curve, and characterised it with two parameters: a and θ. The model of the process is given by a pure integrator and a delay:

$$H_P(s) = \frac{a}{\theta s} e^{-\theta s} \tag{1.1}$$

The parameters can be determined graphically, by drawing a line tangent to the point where the slope is maximal. The parameter a is determined by the vertical axis intercept of this line, and θ by the horizontal axis intercept. This model has a pure integrator, so it does not catch the steady state value of the process.
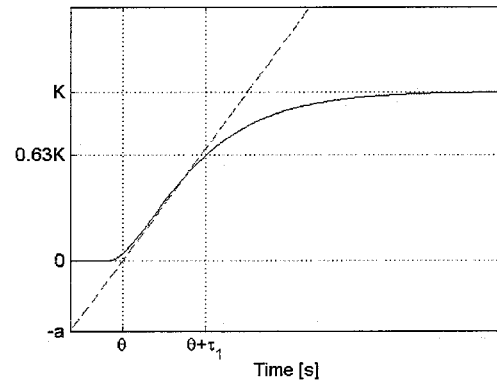
Figure 1.1: Process reaction curve

1.1.2   Three parameter model

A better approximation of the response show in figure 1.1 is obtained by increasing the number of parameters. The model [7]

$$H_P(s) = \frac{Ke^{-\theta s}}{\tau_1 s + 1}$$
(1.2)

has three parameters: the static gain K, the time constant τ_1 and dead time θ. This is the most common model in PID tuning literature, and is often called the FOLPD model (First Order Lag Plus Delay). The parameters of the model can be determined as shown in figure 1.1. More accurate methods to obtain the model parameters are available, for example in Cool [7]. Although the model seems very simple, it is widely used in the process industry. Reasons for this are its simplicity and its applicability to many processes.

1.1.3   Four parameter model

If the model (1.2) is not accurate enough, a second order model can be used. Cool [7] gives a method to the four parameters of the model

$$H_P = \frac{Ke^{-\theta s}}{(\tau_1 s + 1)(\tau_2 s + 1)}.$$
(1.3)

This model is also known as SOLPD (Second Order Lag Plus Delay). The value for τ_1 is calculated as in the FOLPD. The values for θ and τ_2 then are calculated from the relations shown in figure 1.2. The accuracy of the fourth parameter τ_2 is doubtful, so the resulting model should be used with care.
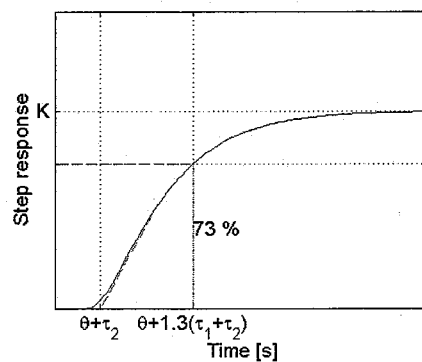
Figure 1.2: Graphical determination SOLPD parameters

1.1.4 Oscillatory processes

It should be noted that although the previously discussed models are widely used in practice, they only describe overcritically damped processes, which are common in the process industry. For mechanical systems, often the step response is undercritically damped, and oscillations occur. Processes with oscillatory step responses as shown in figure 1.3 can be approximated by the two-parameter model (1.1), but this model will not capture the oscillations. None of the three- or four-parameter models described above is suitable either. A three-parameter model that describes the oscillations is given by the transfer function [9]

$$H_P = \frac{K\omega_o^2}{s^2 + 2\zeta\omega_o s + \omega_o^2},$$

(1.4)

with $0<\zeta<1$ and $\omega_0 \geq 0$. K is the static gain, ω_o the natural frequency, and ζ the relative damping. These parameters can be determined approximately from the step response as indicated in figure 1.3. First, the period of oscillation T_p and the decay ratio d are determined. Then, the parameters ω_o and ζ can be calculated with [2]:

$$\zeta = \frac{1}{\sqrt{1+(2\pi/\ln d)^2}}$$

(1.5)

$$\omega_o = \frac{2\pi}{T_p \sqrt{1-\zeta^2}}$$
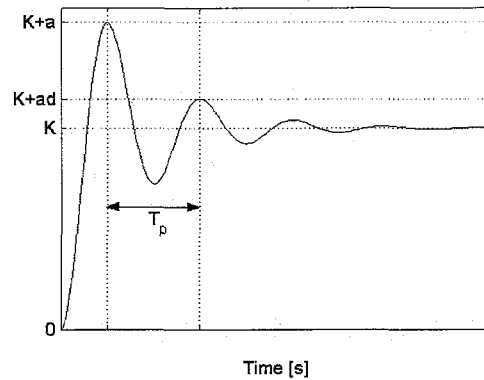
Figure 1.3: Step response of an oscillatory process

1.2 Frequency response measurement

Some methods do not use a complete model, but are based on knowledge of the response to specific frequencies. The idea is that the controller settings can be based on the most critical frequency points for stability.

1.2.1 Ziegler and Nichols [42]

The frequency domain method proposed by Ziegler and Nichols [42] is based on experimentally determining the point of marginal stability. This point can be found by increasing the proportional gain of the controller, until the process becomes marginally stable. The controller gain at this point is called the ultimate gain K_u, and the period of oscillation is called the ultimate period T_u. No explicit model is given, but these two parameters define one point in the Nyquist plot. Based on knowledge of this point, controller settings can be calculated. A disadvantage of the continuous cycling method is that the system is driven towards instability, which can lead to dangerous situations in practice. Further, nothing is known about the course of the Nyquist plot, which may result in a stable, but very oscillatory closed loop system.

1.2.2 Åström and Hägglund [2]

An improvement of the Ziegler-Nichols method is given by Åström and Hägglund [2]. They propose to use a relay feedback, as shown in figure 1.4.
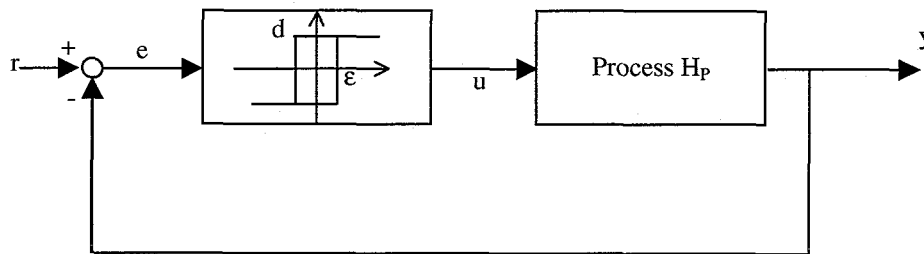
Figure 1.4: Relay feedback

This nonlinear feedback induces a limit cycle oscillation. The period of this oscillation is T_u, and a good estimate for the ultimate gain can be calculated from the oscillation amplitude a with:

$$K_u = \frac{4d}{\pi a} \tag{1.6}$$

The major advantage of using relay feedback is that the system is not driven to instability. Furthermore, including hysteresis in the relay and adding integrators offers the possibility to identify different points on the Nyquist curve, which gives more information about the course of the Nyquist plot.

Chapter 2

Tuning methods

2.1 Controller structure

In the following, the controller is assumed to have the parallel structure, defined as

$$H_C(s) = K_P \left(1 + \frac{1}{\tau_I s} + \tau_D s \right),$$
(2.1)

unless stated otherwise. Further, the control setup of figure 2.1 is assumed, with y the process output, n the measurement noise, r the reference or setpoint, e the tracking error, d the load disturbance, and u the plant input.
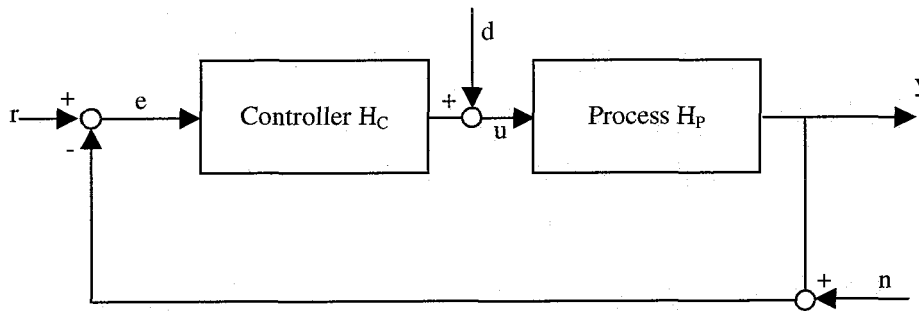
Figure 2.1: Control configuration

2.2 Trial and error

Trial and error is an easy way of tuning a controller, but satisfactory performance is not guaranteed, and it takes a lot of time. Often, the response y to a square wave or step r is taken as the indicator of the quality of the controller. Overshoot, undershoot and rise time can be taken as performance criteria. An example of such a procedure is given in Ellis, [8]. So-called zone-based tuning is used, which means that the low and high frequency parts of the controller can be tuned separately, starting with the high frequency part. For a PID-controller, this means that first the P-and D-action are tuned and then the I-action. An example of a flowchart with steps to follow to tune a PID-controller is given in figure 2.2.

```
┌─────────────────────────────────────┐
│   Zero I- and D-action. Set Kp low.  │
└─────────────────────────────────────┘
                 ↓
┌─────────────────────────────────────┐
│ Apply square wave reference at about 10% of │
│ the desired bandwidth. Use large amplitude,  │
│           but avoid saturation.              │
└─────────────────────────────────────┘
                 ↓
┌─────────────────────────────────────┐
│ Raise Kp for approximately 10 % overshoot. │ ◄────────────────┐
└─────────────────────────────────────┘                        │
                 ↓                              ┌────────────────────────────┐
┌─────────────────────────────────────┐         │   Reduce noise at source   │
│ Raise τD to eliminate most overshoot. │        │            Or              │
└─────────────────────────────────────┘         │    Increase resolution     │
                 ↓                   Yes          │            Or              │
┌─────────────────────────────────────┐ ───────► │         Lower τD           │
│            Too noisy?               │          │            Or              │
└─────────────────────────────────────┘         │         Lower Kp           │
             ↓ No                                └────────────────────────────┘
┌─────────────────────────────────────┐
│   Lower τI for 15% overshoot.       │
└─────────────────────────────────────┘
                 ↓
┌─────────────────────────────────────┐
│              Done                   │
└─────────────────────────────────────┘
```

Figure 2.2: Flowchart for the tuning procedure of a PID-controller,
given in Ellis [8].

Here, the performance criterion is the overshoot of the step response, which is related to the stability margin of the control system. A high stabi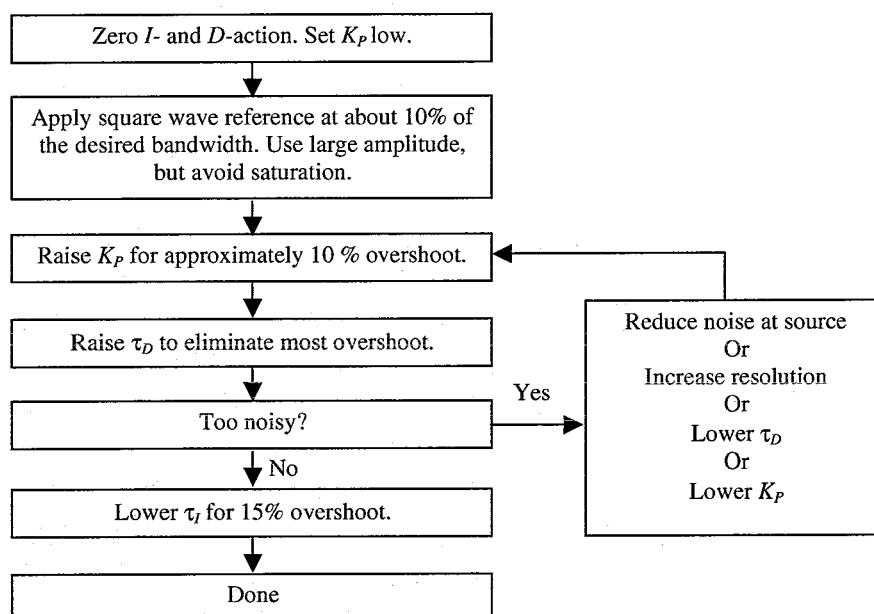lity margin generally leads to low overshoot, and low stability margins to a more oscillatory behaviour. This method is easy to use for process operators, since little or no knowledge of the controlled process is required. The numerical values in the procedure are rules of thumb, based on experience. The disadvantage is that little can be said about the robustness of the system. Further, many iteration steps may be needed before the final settings are obtained.

2.3   Feature based methods

Many tuning methods are presented in literature that are based on a few features of the process dynamics that are easy to obtain experimentally. Typical time-domain features are the parameters of the step response, discussed in section 1.1. Typical frequency-domain features are the ultimate gain and period, discussed in section 1.2.1. An extensive overview of tuning rules for PI or PID controllers is given in O'Dwyer [26].

2.3.1   Ziegler-Nichols step response method (1942)

In 1942, Ziegler and Nichols presented two classical methods to tune a PID-controller [42]. These methods are still widely used, due to their simplicity. In their first method, the controller settings are based on the two parameters θ and a of the process reaction curve of figure 1.1. They based the choice for the controller parameters on simulation results for several processes, by adjusting the controller parameters until the response showed a decay ratio d of 0.25. A quarter decay corresponds to a relative damping $\zeta=0.21$, which they considered a good compromise between quick response and adequate stability margins. The proposed Ziegler-Nichols settings are shown in table 2-1.

Controller type	K_p	τ_I	τ_D
P	$\dfrac{1}{a}$	-	-
PI	$\dfrac{0.9}{a}$	$3.33\,\theta$	-
PID	$\dfrac{1.2}{a}$	$2\,\theta$	$0.5\,\theta$

Table 2-1: Ziegler-Nichols settings for the process reaction curve method

The advantage of the Ziegler-Nichols method is that the tuning rules are very simple to use. Disadvantages are:
- Further fine tuning is needed.
- Controller settings are aggressive, resulting in large overshoot and oscillatory responses.
- Poor performance for processes with a dominant delay.
- Closed loop very sensitive to parameter variations.
- Parameters of the step response may be hard to determine due to measurement noise.

2.3.2 Ziegler-Nichols continuous cycling method

The frequency domain method proposed by Ziegler and Nichols is based on the ultimate gain K_u, and the ultimate period T_u. The controller settings are shown in table 2-2.

Controller type	K_p	τ_I	τ_D
P	0.5 K_u	-	-
PI	0.45 K_u	$T_u/1.2$	-
PID	0.6 K_u	$T_u/2$	$T_u/8$

Table 2-2: Ziegler-Nichols settings for the continuous cycling method

A disadvantage of the continuous cycling method is that the system is driven towards instability, which can lead to dangerous situations in practice. Further, the resulting closed loop behaviour can be very different, depending on the characteristics of the process.

2.3.3 Cohen-Coon (1953)

Cohen and Coon [6] based the controller settings on the three parameters θ, τ_I and K of the open loop step response. The main design criterion is rejection of load disturbances. The method attempts to position closed loop poles such that a quarter decay ratio is achieved. Their suggested controller settings are shown in table 2-3.

Controller type	K_p	τ_I	τ_D
P	$\dfrac{1}{K}\dfrac{\tau_1}{\theta}\left[1+\dfrac{\theta}{3\tau_1}\right]$	-	-
PI	$\dfrac{1}{K}\dfrac{\tau_1}{\theta}\left[0.9+\dfrac{\theta}{12\tau_1}\right]$	$\dfrac{\theta\left[30\tau_1+3\theta\right]}{9\tau_1+20\theta}$	-
PD	$\dfrac{1}{K}\dfrac{\tau_1}{\theta}\left[1.25+\dfrac{\theta}{6\tau_1}\right]$	-	$\dfrac{\theta\left[6\tau_1-2\theta\right]}{22\tau_1+3\theta}$
PID	$\dfrac{1}{K}\dfrac{\tau_1}{\theta}\left[\dfrac{4}{3}+\dfrac{\theta}{4\tau_1}\right]$	$\dfrac{\theta\left[32\tau_1+6\theta\right]}{13\tau_1+8\theta}$	$\dfrac{4\theta\tau_1}{11\tau_1+2\theta}$

Table 2-3: Cohen-Coon settings

Although one more parameter is used in this method, the results are not much better than with the Ziegler-Nichols settings, mainly because of the decay ratio being too small, leading to low damped closed-loop systems.

2.3.4 Åström and Hägglund (1985)

Åström and Hägglund recognised that the Ziegler-Nichols continuous cycling method actually identifies the point $(-1/K_u, 0)$ on the Nyquist curve, and moves it to a predefined point. With PID-control, it is possible to move a given point on the Nyquist curve to an arbitrary position, see figure 2.3. By increasing the gain, the arbitrary point A moves in the direction of $G(j\omega)$. Changing the I- or D-action moves the point in the orthogonal direction.
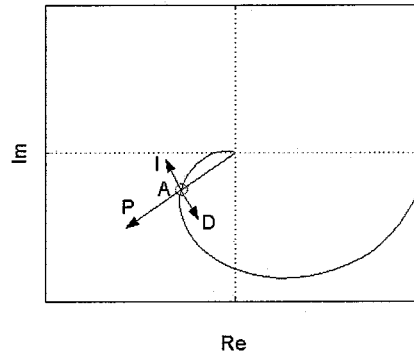
Figure 2.3: Effect of changing the PID parameters shown in the Nyquist plot

Let ω be the frequency that corresponds to A. The frequency response of the controller at ω is:

$$H_C(j\omega) = K_P\left(1 + \frac{1}{\tau_I j\omega} + j\omega\tau_D\right)$$ (2.2)

In the Ziegler-Nichols method, point A is initially the point of marginal stability, located at $(-1/K_u, 0)$ with frequency $\omega_u = 2\pi/T_u$. With the settings from table 2-2, it follows that:

$$H_C = 0.6K_U\left(1 + j\left(\omega_U\tau_D - \frac{1}{\tau_I\omega_U}\right)\right)$$

$$= 0.6K_U\left(1 + j\left(\frac{2\pi}{8} - \frac{1}{\pi}\right)\right) = K_U(0.6 + 0.28j)$$ (2.3)

$$\rightarrow H_P H_C = \frac{-1}{K_U} \cdot K_U(0.6 + 0.28j) = -0.6 - 0.28j$$

The Ziegler-Nichols method can thus be interpreted as finding controller parameters such that the point where the Nyquist curve intersects the negative real axis is moved to $-0.6 - 0.28j$. This corresponds to a phase advance of $25°$ at ω_U. The distance α from this point to the critical point is 0.5, which implies that the sensitivity peak value M_S is always greater than 2.

With this interpretation, it is straightforward to generalise the Ziegler-Nichols method, and move an arbitrary point A to a predefined position B. A convenient choice for point A is the ultimate point. Point B can be determined by a desired gain margin and phase margin, and is written in polar co-ordinates as $B = r_b e^{i(\pi + \varphi_b)}$. The controller settings for these points are given in table 2-4.

Controller type	K_p	τ_I	τ_D
PI	$K_u r_b \cos(\varphi_b)$	$\dfrac{-T_u}{2\pi \tan(\varphi_b)}$	-
PID	$K_u r_b \cos(\varphi_b)$	$\dfrac{-T_u}{\pi}\left(\dfrac{1+\sin(\varphi_b)}{\cos(\varphi_b)}\right)$	$\dfrac{-T_u}{4\pi}\left(\dfrac{1+\sin(\varphi_b)}{\cos(\varphi_b)}\right)$

Table 2-4: Åström and Hägglund stability margin based settings for the frequency response method

This method identifies and moves only one point on the Nyquist curve to a desired position. Since the shape of the Nyquist curve is unknown, good performance or even stability cannot be guaranteed. This can be improved by identifying more points on the Nyquist curve, for example by using the relay based frequency response identification.

2.3.5   Refined Ziegler Nichols (1991)

Hang et al. [11] introduce refined Ziegler-Nichols settings, by adding so called setpoint weighting. This is done by an extra parameter b in the proportional action of the controller:

$$P = K_P(br - y)$$

With this parameter, the overshoot can be reduced to acceptable levels, and thus gives good setpoint response. This modification actually introduces a feedforward action in the controller, since

$$
\begin{aligned}
u &= K_P(br - y) \\
&= \underbrace{K_p(r - y)}_{feedback} + \underbrace{K_P(b-1)r}_{feedforward}
\end{aligned}
\tag{2.4}
$$

Further, they modify the ZN-formulas to account for the normalised dead time, defined as θ/τ_I.

2.3.6 Mann (2001)

Mann [22] classifies the FOLPD in three classes: negligible, middle and long normalised dead time θ/τ_I. Then, tuning rules for each case are derived, where a desired level of overshoot can be chosen. The tuning rules are rather complex, but are reported to give good results.

2.4 Analytical methods

2.4.1 Direct pole placement [20]

The objective of pole placement is to place the closed loop poles at desired locations. Complete knowledge of the transfer function is needed to calculate the appropriate controller settings. Desired closed loop poles are specified, and the controller parameters that move the poles to these positions are analytically calculated. The number of closed loop poles that can be positioned equals the number of controller parameters, e.g. a PI-controller can only position two closed loop poles. For first and second order processes, it is possible to place all closed loop poles with a PID controller. When the process is of higher order, this is not possible anymore, and it is necessary to make approximations to obtain a first or second order model.

2.4.2 Dominant pole design

With direct pole placement it is attempted to place all closed loop poles. One difficulty with this method is that high-order models lead to high-order controllers. Dominant pole design, proposed by Åström and Hägglund [2], is based on placing just a few poles of the closed loop, which are considered dominant for the closed loop response. First, it is noted that the behaviour of many closed loop systems is determined by two dominant poles. A common configuration of the poles and zeros is shown in figure 2.4, where p_1 and p_2 are the dominant poles. If this configuration applies to the controlled system, it is possible to place the dominant poles at desired locations. The influence of the other poles and zeros is assumed to be small, which is only valid if their real part is small enough. The parameters of the controller are calculated, such that the dominant poles are located at desired pole locations, specified by their frequency ω_o and relative damping ζ as:

$$
\begin{aligned}
p_1 &= -\zeta\omega_o + j\omega_o\sqrt{1-\zeta^2} = -\sigma + j\omega \\
p_2 &= -\zeta\omega_o - j\omega_o\sqrt{1-\zeta^2} = -\sigma - j\omega
\end{aligned}
\tag{2.5}
$$

For a PID controller, three closed loop poles can be positioned, so a third pole at $-\alpha_o\omega_o$ is chosen.
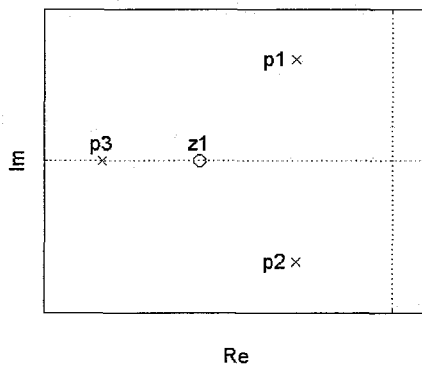

Figure 2.4: Pole-zero configuration of a simple feedback system

If the transfer function of the process is known, the parameters of the controller can be computed analytically. The controller settings for this method are shown in table 2-5.

Controller	K	k_I	k_D
$k + \dfrac{k_I}{s}$	$-\dfrac{\sqrt{1-\zeta^{2}}A(\omega_o)+\zeta B(\omega_o)}{\sqrt{1-\zeta^{2}}\|H_P(p_1)\|}$	$-\dfrac{\omega_o B(\omega_o)}{\sqrt{1-\zeta^{2}}\|H_P(p_1)\|}$	-
$k + k_D s$	$\dfrac{-\sqrt{1-\zeta^{2}}A(\omega_o)+\zeta B(\omega_o)}{\sqrt{1-\zeta^{2}}\|H_P(p_1)\|}$	-	$\dfrac{B(\omega_o)}{\omega_o\sqrt{1-\zeta^{2}}\|H_P(p_1)\|}$
$k + \dfrac{k_I}{s} + k_D s$	$k(\omega_0) + 2\zeta\omega_o k_D$	$k_I(\omega_o) + \omega_o^2 k_D$	$\dfrac{1+\left[k(\omega_o)-\dfrac{k_I(\omega_o)}{\alpha_o\omega_o}\right]H_p(-\alpha_o\omega_o)}{\left(\left(\dfrac{1}{\alpha_o}+\alpha_o\right)-2\zeta\right)\omega_o H_p(-\alpha_o\omega_o)}$

Table 2-5: Controller settings for analytical dominant pole design. $A(\omega_o)$ and $B(\omega_o)$ **stand for the real and imaginary part of** $H_P(p_1)$, **respectively.**

This method is applicable to higher-order processes, with ζ and ω_o as the design parameters. Disadvantage is that these processes should have the pole-zero configuration of figure 2.4. If this is not the case, a pole far away in the left half plane can move towards the right and become a dominating pole, which leads to an unexpected response. This is generally prevented by choosing the desired frequency ω_o not too high.

Åström and Hägglund used this method to design controllers for a test batch of processes, representing time delay, high order and non-minimum phase behaviour. They optimised the controller settings to achieve a minimum sensitivity peak. Then, they plotted normalised controller parameters as a function of the normalised process parameters $\tau=\theta/(\theta+\tau_1)$ or $\kappa=1/(KK_U)$. By fitting an exponential function through the results, they achieved tuning rules valid for processes that fall in the class of test processes. Shen [35] refined these rules by optimisation with a genetic algorithm, and Branica [4] developed a least-squares fitting procedure to connect the PID parameters to parameters of the FOLPD model.

2.4.3 Internal Model Control (IMC)

The IMC procedure [3] uses the process model explicitly in the design of the controller. It is a general design method, but results in PID controllers for low order process models. The controller is formulated as:

$$H_C(s) = \frac{H_F(s)H_{P-}^{-1}(s)}{1 - H_F(s)H_{P-}^{-1}(s)H_P(s)} \qquad (2.6)$$

where $H_P = H_{P+}H_{P-}$, and $H_{P+}(s)$ includes all non-minimum phase dynamics. $H_F(s)$ is a low-pass filter, usually chosen as

$$H_F(s) = \frac{1}{\lambda s + 1} \qquad (2.7)$$

A controller of this type cancels the process poles and zeros, and is usually of high order. However, if the order of the model is low, the procedure results in a PI or PID controller. For example, if the FOLPD model with a first order Padé approximation for the time delay is used, the controller becomes:

$$H_C = \frac{1}{K}\frac{(\tau_1 s + 1)(0.5\theta s + 1)}{(\lambda + 0.5\theta)s},$$

which, after some manipulations, gives for the PID parameters

$$K_P = \frac{1}{K}\frac{(\tau_1 + 0.5\theta)}{(\lambda + 0.5\theta)}; \quad \tau_I = \tau_1 + 0.5\theta; \quad \tau_D = \frac{\tau_1\theta}{2\tau_1 + \theta}$$

With this controller, the closed loop becomes

$$H_{CL}(s) = \frac{1}{\lambda s + 1}$$

An advantage of designing PID controllers via IMC is that the parameter λ represents the trade-off between robustness and performance whereas PID has three parameters that do not provide this clear trade-off. IMC can give very good setpoint responses, but because of the pole cancellation, the load disturbance response can be very sluggish if slow poles are cancelled. Many applications and modifications of the IMC procedure for PID design are reported in literature, such as IMC design for unstable and integrating plants and using a MacLaurin approximation for the time delay [16], [17]. An interesting modification is done by Skogestad [36], who presents simple IMC based tuning rules.

2.5   Optimisation based methods

2.5.1   Minimum optimisation criterion

Many methods are presented in literature that minimise a certain error criterion [26]. To illustrate these methods, the minimum ITAE method is discussed here. ITAE is shorthand for integral of the time weighted absolute error, and is defined as:

$$ITAE = \int_0^\infty t|e(t)|dt \qquad (2.8)$$

The time weighting is used, because the initial error for a step response is always large, and for most setpoint following cases it is reasonable to weigh this error less. Assume that the FOLPD model (1.2) is valid, then it is possible to mathematically determine the controller settings. A distinction is made between responding to a disturbance or responding to a set point. The controller settings that give a minimum ITAE are given in table 2-6.

Type of input	Controller type	K_P	τ_I	τ_D
Disturbance	P	$\dfrac{0.49}{K}\left[\dfrac{\tau_1}{\theta}\right]^{1.084}$	-	-
	PI	$\dfrac{0.859}{K}\left[\dfrac{\tau_1}{\theta}\right]^{0.977}$	$\dfrac{\tau_1}{0.674}\left[\dfrac{\tau_1}{\theta}\right]^{-0.680}$	-
	PID	$\dfrac{1.357}{K}\left[\dfrac{\tau_1}{\theta}\right]^{0.947}$	$\dfrac{\tau_1}{0.842}\left[\dfrac{\tau_1}{\theta}\right]^{-0.738}$	$0.381\tau_1\left[\dfrac{\tau_1}{\theta}\right]^{-0.995}$
Setpoint	PI	$\dfrac{0.586}{K}\left[\dfrac{\tau_1}{\theta}\right]^{0.916}$	$\dfrac{\tau_1^2}{1.03\tau_1 - 0.165\theta}$	-
	PID	$\dfrac{0.965}{K}\left[\dfrac{\tau_1}{\theta}\right]^{0.855}$	$\dfrac{\tau_1^2}{0.796\tau_1 - 0.147\theta}$	$0.308\tau_1\left[\dfrac{\tau_1}{\theta}\right]^{-0.929}$

Table 2-6: Controller settings for minimum ITAE [24], [31]

It should be noted that by optimising with respect to one criterion one should not neglect other constraints.

2.5.2   LQR optimisation of ISE

Argelaguet [1] used LQR optimisation to minimise the ISE-criterion, which is the integral of the squared error. The model used in this optimisation is the FOLPD model with a first-order Padé approximation for the time delay. The tuning rules that result from the minimisation are equal to the IMC tuning rules for this model.

2.5.3   Constrained optimisation

Panagopoulos [27] minimises the integrated error IE by maximising the integral gain K_P/τ_I, since these variables are inversely related for a step load disturbance. This minimisation is done under the constraint that the maximum sensitivity M_S does not exceed a desired value. A three dimensional grid with the values for K_P, K_P/τ_I and $K_P\tau_D$ satisfying this constraint can be drawn, which shows the maximal possible value for K_P/τ_I. Methods to numerically obtain the maximum value are also available. This approach guarantees a desired robustness and gives good load disturbance response. The setpoint response is tuned by setpoint weighting, discussed in section 2.3.5. Ingimundarson [13] uses the same method, but uses as robustness constraint the H_∞-norm of the transfer

from disturbance and measurement noise to process and controller output. Further, a constraint is added such that the Nyquist plot has a smooth course without cusps.

2.5.4 Modulus optimum MO and symmetric optimum SO [20]

MO and SO are methods based on finding a controller that gives the closed loop frequency response a desired shape. The MO criterion requires that the magnitude of the frequency response of the closed loop transfer function should be as flat and as close to one as possible for a large bandwidth. Mathematically, the MO criterion is formulated as

$$T(0) = 1, \quad \left. \frac{d^n |T(j\omega)|}{d\omega^n} \right|_{\omega=0} = 0 , \tag{2.9}$$

for as high n as possible.

If the closed loop transfer is

$$T = \frac{\omega_0^2}{s^2 + \sqrt{2}\omega_0 s + \omega_0^2} , \tag{2.10}$$

n equals three, and for the closed loop transfer

$$T = \frac{\omega_0^3}{(s+\omega_0)(s^2 + \omega_0 s + \omega_0^2)} , \tag{2.11}$$

n equals five. The desired open loop transfer can be calculated with

$$L = \frac{T}{1-T} , \tag{2.12}$$

The controller should be designed such that $H_C H_P$ equals the desired open loop function.

In the SO method, the desired open loop transfer equals

$$H_{SO} = \frac{\omega_0^2 (2s + \omega_0)}{s^2 (2s + \omega_0)} \tag{2.13}$$

The Bode phase diagram of this function is symmetrical around the frequency ω_0. Preitl [30] gives tuning rules to obtain this transfer function for integrating first and second order processes, which results in simple relations between the process and the controller parameters.

2.5.5 MOMI

Vrancic
[38], [39] presents a method to achieve the MO criterion, by using multiple integration of the step response of the process. This is also called the method of moments. This method implicitly uses the process model:

$$H_P(s) = K \frac{b_m s^m + \ldots + b_2 s^2 + b_1 s + 1}{a_n s^n + \ldots + a_2 s^2 + a_1 s + 1} e^{-\theta s} \tag{2.14}$$

By integrating the open loop step response to a step change Δu, the following areas can be calculated

$$A_1 = y_1(\infty)$$
$$\vdots \tag{2.15}$$
$$A_k = y_k(\infty)$$

where

$$y_0(t) = \frac{y(t) - y(0)}{\Delta u}$$

$$y_1(t) = \int_0^t [y_0(\infty) - y_0(\tau)]d\tau$$

$$\vdots$$

$$y_k(t) = \int_0^t [A_{k-1} - y_{k-1}(\tau)]d\tau$$

(2.16)

Then the controller settings that meet the MO criterion (2.9), are calculated with

$$K_P = \frac{A_3}{2(A_1 A_2 - A_3 K - \tau_D A_1^2)}$$

$$\tau_I = \frac{A_3}{A_2 - \tau_D A_1}$$

$$\tau_D = \frac{A_3 A_4 - A_2 A_5}{A_3^2 - A_1 A_5}$$

(2.17)

For a PI controller, τ_D is set to zero.

The advantage of this method is that it is not necessary to calculate model parameters, since the controller parameters are directly calculated from a measured step response. However, the tuning is only applied to overcritically damped processes, which is generally not the case for mechanical systems.

2.5.6 Kristiansson and Lennartson (2002)

Kristiansson and Lennartson [15] perform an optimisation by minimising the H_∞-norm of the disturbance sensitivity function under three constraints with respect to stability robustness, control activity and noise sensitivity. They performed this procedure for a batch of stable non-oscillating plants and derived tuning rules from fitting a curve through the results. This results in simple tuning rules for PI and PID controllers.

2.5.7 Other optimisation methods

Pedret [28] gives a framework for optimisation to refine an initial choice of the controller parameters with guaranteed stability. Zolotas [43] proposes an optimisation algorithm using Quantitative Feedback Theory (QFT), which permits taking into account robustness to model uncertainty.

2.6 Loop shaping

The objective of loop shaping, discussed in e.g. [20], is to shape the open loop L such that both S and T satisfy certain stability and performance criteria. The sensitivity is important for tracking and disturbance rejection, and the complementary sensitivity is important for reducing the influence of measurement noise. Figure 2.5 shows typical bounds on S and T. By shaping the open loop transfer L, these bounds have to be satisfied. The resulting desired open loop is shown in figure 2.5 c).

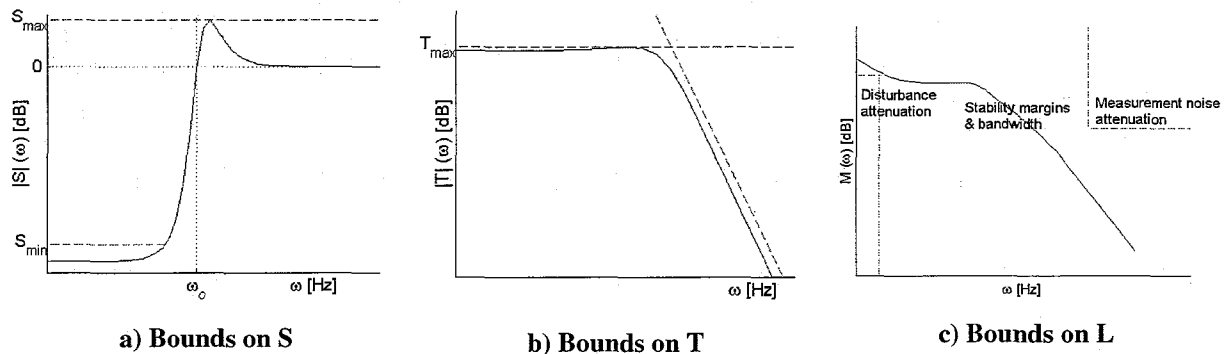a) Bounds on S   b) Bounds on T   c) Bounds on L

Figure 2.5: Bounds on S, T and L

Typically, the disturbance rejection boundary is low-frequent, the mid-frequency range determines stability and performance, and the high-frequency range determines the attenuation of measurement noise. The open loop bandwidth is defined as the frequency where $|L(j\omega)|=1$. The design task is to tune the controller, such that the open loop satisfies the bounds, using the following rules:

- Higher P-action raises the low-frequent part of L, shifts the bandwidth to higher frequencies, but raises the peak value of S and T.

- Higher I-action raises the low-frequent part of L, but raises the peak value of S and T, and reduces the phase until its breakpoint frequency.

- Higher D-action raises the high-frequent part, but reduces the peak value of S and T, and adds extra phase.

These effects are shown in figure 2.6.

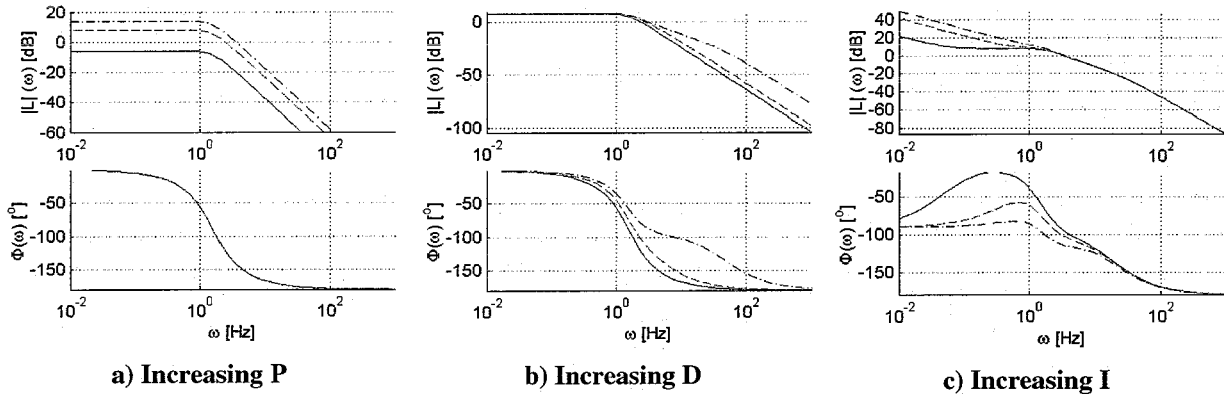| a) Increasing P | b) Increasing D | c) Increasing I |

Figure 2.6: Effect of increasing controller action on the open loop L

With this method, the requirements for the controller are strictly defined, which results in a predictable behaviour.

2.6.1   Fixed structure H∞

An interesting extension of the previous method is the use of H_∞ design methods. Similar to sensitivity loop shaping, boundaries are posed on transfers such as sensitivity, complementary sensitivity etc. This is done by specifying filters that characterise model uncertainties, disturbances, and the reference. Furthermore, filters are specified, which penalise certain signals, such as the control input, the tracking error, etc. The H_∞ algorithm then computes a controller that satisfies the resulting boundaries in an optimal sense. Disadvantage is that the resulting controller often has a high order, which is difficult to implement in practice. Algorithms are developed that restrict the controller to have a fixed structure [25], such as PID. This leads in general to deteriorated performance. Fixed structure H_∞ algorithms may be useful in automating the procedure of sensitivity loop shaping discussed before. The design task in the H_∞ theory is to design the right penalising and characterising filters, which means that sufficient knowledge of the system, including uncertainties, has to be available.

Tan [37] uses H_∞ theory on FOLPD and IFOLPD processes, which results in tuning rules. These rules are almost the same as the IMC rules with Padé approximation discussed earlier, since the algorithm uses pole-cancellation, which is also the case in IMC.

2.7   Other methods

The methods described in the previous sections are the most popular methods, but many other methods are also available in literature. It is not possible to cover them all, but some of them are given below. Chen [5], Lee [18], and Shen [34] use fuzzy neural networks to perform a parameter optimisation. Shen [35] uses a genetic algorithm to arrive at new tuning rules. Schädel [33] designs a PID controller based on the principle of cascaded damping ratios. Wang [40] specifies a desired control signal response to achieve smooth setpoint following. Munro [23] uses the Nyquist crossing count to show the admissible region of PID parameters.

2.8 Autotuning

A logical step is to extend the tuning methods discussed above to auto-tuning, i.e. to automatically tune the controller parameters. Many methods to develop an auto-tuning algorithm are discussed in literature. Most of these methods perform a step response or frequency response measurement and use tuning rules to find the appropriate parameters. Gaikwad [10] presents an algorithm that fits the controller parameters online to achieve a target loopshape, without model identification. Poulin [29] uses excitation by a relay to identify process dynamics, and then minimises the distance between the open loop and a target contour on the Nichols chart. Ingimundarson [14] uses step response experiments to estimate a FOLPD model with the method of moments of section 2.5.5, and automatically tunes the controller. Luo [21] applies the Ziegler-Nichols tuning rules on a process that is estimated by adding a relay parallel to the controller. Woodyatt [41] presents an autotuner based on frequency domain approximation.

Conclusions

This report overviews PID tuning methods found in literature. Only a selection is discussed, since it is impossible to include them all. Many methods have been presented based on the work by Ziegler and Nichols, and use very simple process models to derive tuning rules. The advantage is that the methods are very easy to use, and do in general not require extensive knowledge of the process. However, this leads also to unknown stability robustness and no good control over the resulting performance. The applicability of these methods is generally limited to process industry.

The use of more sophisticated model-based tuning methods, such as loopshaping, H_∞ or IMC allows a better definition of desired closed-loop behaviour and robustness. A disadvantage is that first an accurate model has to be obtained, which is time-demanding and often difficult.

An interesting field is the autotuning of PID-controllers. Advantages are that they are easy to use, and can cope with disturbances and process changes. This area is still subject of research.

Bibliography

[1]   R. Argelaguet, A new tuning of PID controllers based on LQR optimisation. In *Preprints of the IFAC Workshop on Digital Control. Past, Present and Future of PID Control*, pp. 303-308. 1996.

[2]   K.J. Åström and T. Hägglund, *PID Controllers: Theory, Design, and Tuning*. Triangle Park, NC: ISA. 1995.

[3]   B.W. Bequette, *Process Control: Modeling, Design, and Simulation.* Prentice Hall, Inc., New Jersey. 2003.

[4]   Branica, Toolkit for PID Dominant pole design, *Proceedings of 9th IEEE International Conference on Electronics, Circuits and Systems*, pp. 1247-1250. 2002.

[5]   G. R. Chen, Conventional and Fuzzy PID controllers: An Overview, *Int. J. of Intelligent Control Systems*, pp. 235-46. 1996.

[6]   G.H. Cohen, and G.A. Coon, Theoretical considerations of retarded control, In *Transactions of the ASME*, pp 827-834. 1953.

[7]   J.C. Cool, F.J. Schijff and T.J. Viersma, *Regeltechniek*, Delta Press. Amerongen. 1991.

[8]   G. Ellis, *Control systems design guide*. Academic Press, London. 1991.

[9]   G.F. Franklin, J.D. Powell, and A. Emami-Naeini, *Feedback control of dynamic systems*. Addison-Wesley Publishing Company. 1994.

[10]   S.V. Gaikwad, Auto-tuning PID using loop-shaping ideas, *Proceedings of the IEEE International Conference on Control Applications*, p. 589-593. 1999.

[11]   C. C. Hang, K. J. Åström and W. K. Ho, Refinement of the Ziegler-Nichols Tuning Formula, *IEE Proceedings - D, Vol. 138, No. 2*, pp. 111-118. 1991.

[12]   B.L. Ho, *An effective construction of realizations from input/output descriptions*, Ph-D thesis, Stanford University, Stanford, California. 1966.

[13]   A. Ingimundarson, Criteria for design of PID controllers, In *Proceedings of the 2nd IFAC conference Control System Design, CSD'03*, Bratislava, Slovak Republic. 2003.

[14]   A. Ingimundarson, Robust automatic tuning of an industrial PI controller for dead-time systems, In *IFAC Workshop on Digital Control – Past, present, and future of PID Control*, Terrassa, Spain. 2000.

[15]   B. Kristiansson and B. Lennartson, *Evaluation and tuning of robust PID controllers*, Technical Report No R008/2003, Department of Signals and Systems, Chalmers University of Technology. 2002.

[16]   Y. Lee, PID controller tuning for integrating and unstable processes with time delay, In *Chemical Engineering Science, Vol. 55*, pp. 3481 – 3493. 2000.

[17]   Y. Lee, PID controller tuning to obtain desired closed-loop responses for SISO systems, In *AIChE Journal, Vol. 44, No. 1*, pp. 106-115. 1998.

[18]   C.H. Lee and C.C. Teng, Tuning of PID Controllers for Stable and Unstable Processes based on gain and phase margin specifications, *International Journal of Fuzzy Systems, Vol. 3, No. 1*, pp. 346-355. 2001.

[19]   M. Lelic, A Reference Guide to PID Controllers in the Nineties, In *Proceedings of IFAC Workshop: Past, Present and Future of PID Control*. 2000.

[20]   W. Levine, *The Control Handbook*, CRC Press. 1996.

[21]   R. Luo, A new approach to closed loop autotuning for PID controllers, In *Ind. Eng. Chem. Res.*, pp. 2462-2468. 1998.

[22]   G.K.I. Mann, Time domain based design and analysis of new PID tuning rules, In *IEE proceedings-Control Theory and Applications, vol. 148, No. 3*, pp. 251-261. 2001.

[23]   N. Munro, The systematic design of PID controllers, In *Proc.UKACC Control 2000 Conference.* 2000.

[24]   P.W. Murrill, *Automatic control of processes.* International Textbook Co., 1967.

[25]   H.E. Musch, and M. Steiner., *Tuning advanced PID controllers via direct norm minimization,* presented at ECC1997, Brussel, 1997.

[26]   O'Dwyer, *PI and PID controller tuning rules for time delay processes: a summary.* Irish Signals and Systems Conference 1999. NUI Galway. 1999.

[27]   H. Panagopoulos, K.J. Åström and T. Hägglund, Design of PID controllers based on constrained optimisation, In *IEEE Proc. Control Theory Appl. 149:1*, pp.32-40. 2002.

[28]   C. Pedret, A refinement procedure for PID controller tuning, *Computers & Chemical Engineering 26 (6)*, pp. 903 – 908, 2002.

[29]   E. Poulin et al., Development and Evaluation of an Auto-tuning and Adaptive PID Controller, *Automatica, Vol. 32, No. 1*, pp. 71-82. 1996.

[30]   S. Preitl, An extension of tuning relations after symmetrical optimum method for PI and PID controllers, In *Automatica, Elsevier Science, vol. 35, no. 10*, pp.1731-1736. 1999.

[31]   Rovira, A.A., Murrill, P.W. and Smith, C.L., Tuning controllers for setpoint changes. In *Instruments and Control Systems, pp 67-69.* 1969.

[32]   B. de Schutter, Minimal state-space realization in linear system theory: An overview. In *Journal of Computational and Applied Mathematics, Special Issue on Numerical Analysis in the 20th century - Vol. I: Approximation Theory, vol. 121, no.1-2*, pp. 331-354. 2000.

[33]   H.M. Schädel, A new method of direct PID controller design based on the principle of cascaded damping ratios, In *Proceedings European Control Conference*, Brüssel. 1997.

[34]   J.C. Shen, *Fuzzy neural networks for tuning PID controller for plants with under-damped responses*, In *IEEE Trans. On Fuzzy Systems, Vol. 9, No. 2*, pp. 333-342. 2001.

[35]   J.C. Shen, New tuning method for PID controller, In *Proceedings of the 2001 IEEE International Conference on Control Applications*, pp. 459-464. 2001.

[36]   S. Skogestad, Probably the best simple PID tuning rules in the world. In *Journal of process control.* 2001.

[37]   W. Tan, PID tuning based on loop-shaping H_∞ control, In *IEE Proc. Control theory Appl., Vol. 145, No. 6*, pp. 485-490, 1998.

[38]   D. Vrančić, A new PID controller auto-tuning method based on multiple integrations, In *Control Engineering Practice, Vol. 7, No. 5*, pp. 623-633, 1999.

[39]   D. Vrančić, Practical guidelines for tuning PID controllers by using MOMI method, In *Proceedings of the IEEE International Symposium on Industrial Electronics, Vol. 3*, pp. 1130-1134, 1999.

[40]   L. Wang, A new frequency domain design method for PID controllers. In *IEE Proc.-Control Theory and Applications, Vol.142, No.4*, pp.265-271. 1995.

[41]   A. Woodyatt, Auto-tuning PID controller design using frequency domain approximation, In *Proceedings European Control Conference*, Brüssel, 1997.

[42]   Ziegler, J.G. and Nichols, N.B., *Optimum settings for Automatic Controllers.* Trans. ASME, vol. 64, pp.759-768, 1942.

[43]   A.C. Zolotas, Optimal design of PID controllers using the QFT method, In *IEE Proceedings of Control Theory and Applications, Vol. 146, No 6*, pp. 595-599, 1999.