Control Systems

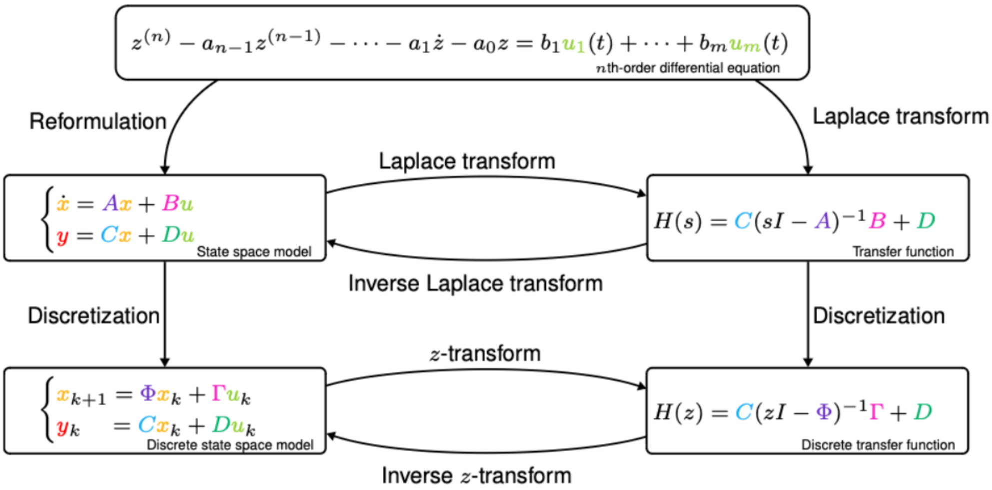
Mathias Balling & Mads Thede

Last updated: March 6, 2024

Contents

1 Linear Time Invariant Systems

Overview



1.1 Time-Domain models

Linear Map

The map $f : \mathbb{R}^n \to \mathbb{R}^m$ is said to be linear if for any $x, y \in \mathbb{R}^n$ and $\alpha \in \mathbb{R}$, the following conditions hold

$$f(x + y) = f(x) + f(y) \qquad \text{Super position}$$
$$f(ax) = \alpha f(x) \qquad \text{Homogeneity}$$

The function has to go through (0,0) in 2D for it to be linear due to homogeneity.

Time-Invariant System

Let $\sigma : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^p$ define the input-output behavior of a system model Σ. The system Σ is time-invariant if for any input signal $u : \mathbb{R} \to \mathbb{R}^m$
and any delay $\tau \in \mathbb{R}$ the following relation holds:

$$y(t - \tau) = \sigma(t, u(t - \tau))$$

for all times $t \in \mathbb{R}$, where y denotes the output signal of the system.

The importance is that the system does not change its behavior due to time. This can be seen as a canon firing at 8am it will not fire different compared to if you do the same at 5pm.

Two types of linear time-domain models.
Continous-time state space models (based on differential equations)

$$\dot{x} = Ax + Bu$$

$$y = Cx + Du$$

- $x \in \mathbb{R}^n$ is state. e.g. position or velocity
- $u \in \mathbb{R}^m$ is input
- $y \in \mathbb{R}^p$ is output
- $A \in \mathbb{R}^{n \times n}$ is system matrix

- $B \in \mathbb{R}^{n \times m}$ is input matrix
- $C \in \mathbb{R}^{p \times n}$ is output matrix
- $D \in \mathbb{R}^{p \times m}$ is the direct feedthrough matrix

Discrete-time state space models (based on difference equations)

$$x_{k+1} = \Phi x_k + \Gamma u_k$$

$$y_k = C x_k + D u_k$$

1.2 Frequency-Domain models

Transfer function:

$$G(s) = \frac{Q(s)}{P(s)}$$

where $Q(s)$ and $P(s)$ are polynomials in s.

- The roots of $P(s)$ are called the **poles** of $G(s)$
- The roots of $Q(s)$ are called the **zeros** of $G(s)$

1.2.1 State space to transfer function

Taking Laplace tranforms of the system and assuming $x_0 = 0$:

$$\dot{x}(t) = Ax(t) + Bu(t)$$

$$y(t) = Cx(t) + Du(t)$$

yields:

$$sX(s) = AX(s) + BU(s)$$
$$Y(s) = CX(s) + DU(s)$$

Which can be rewritten as (I is the identity matrix):

$$X(s) = (sI - A)^{-1} BU(s)$$

$$Y(s) = \left(C (sI - A)^{-1} B + D\right) U(s)$$

where

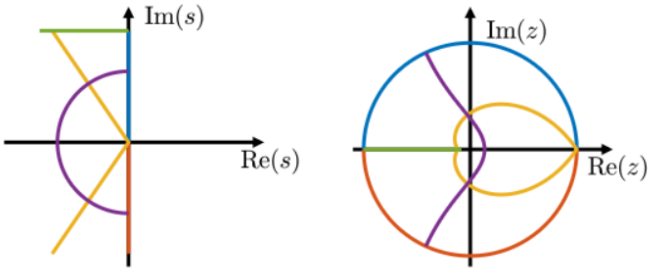
$$Y(s) = G(s)U(s)$$
$$G(s) = C (sI - A)^{-1} B + D$$

1.2.2 Transfer function to state space

1.2.3 Discrete-time transfer function

Discretization from s-domain to z-domain can be done using:

- Matched z-transform
- Bilinear z-transform
- Impulse invariance z-transform



1.3 Examples

2 Stability and Performance Analysis

2.1 Basic System Classes

2.1.1 First Order Systems

State-space representation of first order system:

$$\dot{x} = -\frac{1}{\tau}x + \frac{k}{\tau}u$$
$$y = x$$

A first-order system has one pole and is described by:

$$H(s) = \frac{k}{\tau s + 1}$$

Where k is the DC-gain and τ is the time-constant. The system has a pole in $s = -\frac{1}{\tau}$ i.e., the smaller time-constant, the faster system response.

2.1.2 Second Order Systems

The transfer funciton of a second-order system is given by:

$$H(s) = \frac{k\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}$$

Where $\omega_n > 0$ is the natural frequency and $\zeta > 0$ is the damping ratio and k is the gain.

The system has two poles, which are $s \in \mathbb{C}$ where:

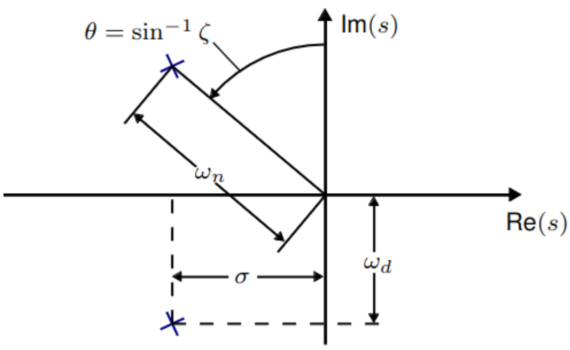
$$s^2 + 2\zeta\omega_n s + \omega_n^2 = 0$$

The values of s is given by:

$$s = -\zeta\omega_n \pm j\omega_n\sqrt{1-\zeta^2}$$

When $\zeta = 1$ the system is critically damped and $H(s)$ has a double pole in $s = -\zeta\omega_n$, when $0 < \zeta < 1$ the system is underdamped and has complex poles. When $\zeta > 1$ the system is overdamped and has real and distinct poles.

$$-\underbrace{\zeta\omega_n}_{\sigma} \pm j\,\underbrace{\omega_n\sqrt{1-\zeta^2}}_{\omega_d}.$$



Impulse response of a underdamped second-order system:

$$h(t) = k\frac{\omega_n}{\sqrt{1-\zeta^2}}e^{-\sigma t}sin(\omega_d t)1(t)$$

The step resonse of a underdamped second-order system:

$$y(t) = k(1 - e^{-\sigma t}(\cos(\omega_d t) + \frac{\sigma}{\omega_d}\sin(\omega_d t)))$$

Impulse response of a critically damped second order system:

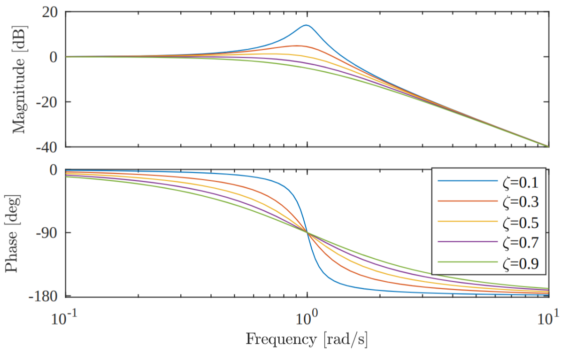
$$h(t) = k\omega_n^2 te^{-\omega_n t}$$

Having a overdamped system with a damping ration that is greater than one leads to a slower impulse response.

The step response of a critically damped second order system:

$$y(t) = k(1 - e^{-\omega_n t} - \omega_n te^{-\omega_n t})$$

Looking at the bode plot of a second-order system depends on the damping ratio. For every pole the magnitude of the bode plot decreases by 20dB/decade. And the phase of the bode plot decreases by 90 degrees for every pole.

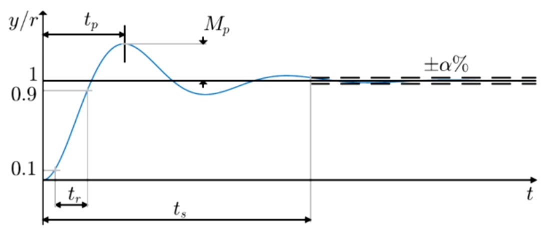


2.2 Performance specifications

Three different performance measurements of dynamical systems are:

1. The rise time t_r (stigetid)
2. The settling time t_s (indsvingingstid)
3. The overshoot M_p (oversving)

The rise time is the time it takes for the system to go from 10% to 90% of the final value. The settling time is the time it takes for the system to reach and stay within a certain % of the final value. The overshoot is the maximum percentage that the system goes over the final value.



The peak time is denoted t_p which is the time where the signal reaches its maximum value.

For a second-order system the rise time can be approximated as:

$$t_r = \frac{1.8}{\omega_n}$$

For a second-order system the settling time can be approximated as:

$$t_s = \frac{-log(\alpha/100)}{\zeta\omega_n}$$

Log is ln in the equation.

The peak time can be found using:

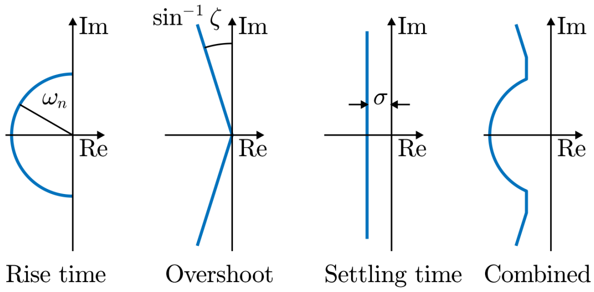
$$t_p = \frac{\pi}{\omega_d}$$

The overshoot is computed from the step response at the peak time. This gives an expression for the overshoot:

$$M_p = e^{-\frac{\zeta\pi}{\sqrt{1-\zeta^2}}}$$

for $0 < \zeta < 1$. The overshoot is given in percentage.

Overview for determining performance variables



To obtain a rise rise shorter than t_r

$$\omega_n \geq \frac{1.8}{t_r}$$

To obtain an overshoot that is smaller than M_p

$$\zeta \geq \sqrt{\frac{(\frac{\log(M_p)}{-\pi})^2}{1+(\frac{\log(M_p)}{\pi})^2}}$$

To obtain an α%-settling time shorter than t_s

$$\sigma \geq \frac{-\log(\alpha/100)}{t_s}$$

2.3 Poles and Zeros of Space-State Models

Using thee state-space matrices:

$$G(s) = C(sI - A)^{-1}B + D$$

Poles can be found using the following equation:

When $G(s) \to \infty$ the system has a pole at $s \to p$.

$$\det(pI - A) = 0$$

Here p is the eigenvalue of A.

Zeros can be found using the following equation:

When $G(s) = 0$ the system has a zero at $s = z$.

$$\begin{vmatrix} A - zI & B \\ C & D \end{vmatrix} = 0$$

2.4 Stability

The stability of the dynamical system can be determined from the eigenvalues of A in the time domain. This is the equivalent of poles in the frequency domain.

$$\dot{x} = Ax + Bu$$

When the eigenvalues of A have a negative real part, the system is stable.

Definition A linear discrete-time system

$$x_{k+1} = \Phi x_k$$

The state of the next time step is given by ϕ mutliplied by x_k.

Where $x_k \in \mathbb{R}^{n \times n}$ is asymptotically (asymptotisk) stable if

$$\lim_{k \to \infty} x_k = 0$$

for any $x_0 \in \mathbb{R}^n$.

A linear discrete-time system is stable if all the eigenvalues are inside the unit circle.

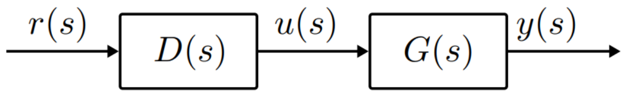
The closer the poles are to origin the more dominating the pole is in the system.

2.5 Examples

3 Introduction to Control

3.1 Open Loop Control

(åbensløjfe regulering)



Steady-State Value of Time Function:

Suppose that Y(s) is the Laplace transform of y(t). Then the final value of y(t) is either:

- Ubounded. If Y(s) has any poles in the open right half-plane (unstable)
- Undefined. If Y(s) has a pole pair on the imaginary axis.
- Constant. If all poles of Y(s) are in the open left half-plane, except for one at s=0

The Final Value Theorem (slutværdi-sætningen)

If all poles of sY(s) are in the open left half-plane, then:

$$\lim_{t \to \infty} y(t) = \lim_{s \to 0} sY(s)$$

The Final Value Theorem determines the constant value that the impulse response of a stable system converges to. The theorem can also be used to determine the DC gain of a system, i.e., the output when a step input is applied to the system.

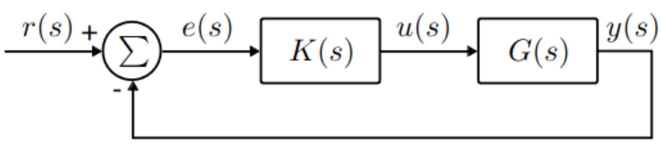
By integrating an impulse response, the step response is obtained. The step response is the integral of the impulse response, and the impulse response is the derivative of the step response.

If the system has a disturbance an open loop control system will not be able to compensate for this.

3.2 Feedback Control

(tilbagekobling). Tilbagekobling anvendes for at elminere forstyrrelsesundertrykkelse.

The connection of a controller K(s) and a system (also called plant) G(s) is called a closed-loop system.



e(s) is the error signal, and is the difference between the reference signal r(t) and the output signal y(t). The error signal is used to adjust the input signal u(t) to the system. The controller K(s) is used to adjust the input signal.

The closed-loop transfer function is given by:

$$\frac{y(s)}{r(s)} = \frac{G(s)K(s)}{1 + G(s)K(s)}$$

K(s) can be designed to meet certain requirements, such as stability, disturbance rejection, and sensitivity analysis.

Stability

The closed-loop system is stable if all the closed-loop poles are in the open left half-plane.

The **loop gain (sløjfe-forstærkning)** is defined as: $L(s) = G(s)K(s)$. The closed-loop poles are given by:

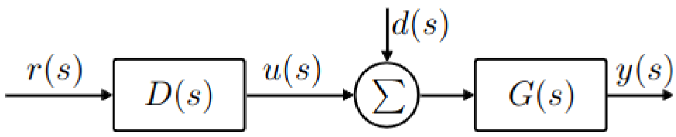
$$1 + L(s) = 0$$

The closed loop poles satisfy:

$$1 + L(s) = 0 \Rightarrow L(s) = -1$$

Disturbance Rejection

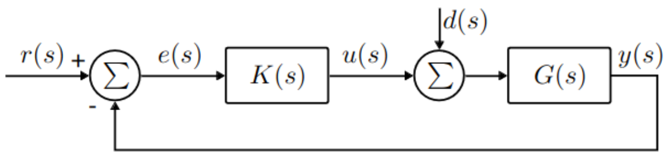
Open loop control system:



The output is given by:

$$y_m(s) = G(s)d(s) + G(s)D(s)r(s)$$

Closed loop control system:



The output is given by (superposition can be applied for deriving the expression):

$$y_m(s) = \frac{G(s)}{1 + G(s)K(s)}d(s) + \frac{G(s)K(s)}{1 + G(s)K(s)}r(s)$$

The larger the amplification of the feedback loop, the smaller the effect of the disturbance on the output.

Sensitivity Analysis

The sensitivity S of the open-loop system is given by the ratio of $\delta T_{ol}/T_o l$ to $\delta G_0/G_0$. For the open loop system S=1. δ is the relative change in the transfer function.

For the closed loop the sensitivity is given by:

$$S = \frac{1}{1 + G_0 K_0}$$

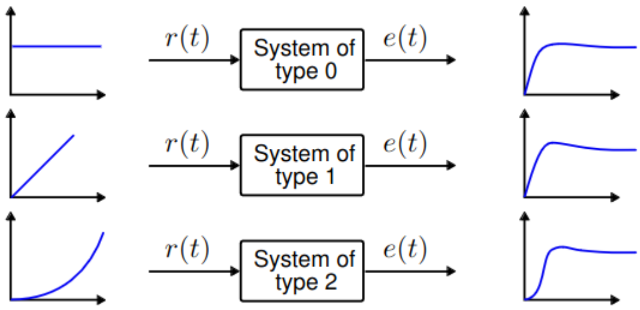
This means that a large amplification leads to a low sensitivity.

Linearization Taylor approximation, YES.

3.3 Steady State Tracking

Definition of system type

Stable systems can be classified according to its system type, defined to be the degree of the polynomial for which the steady-state error in a nonzero finite constant.



An expression for the tracing error is computed to determine the steady-state error. The error is given by:

$$e(s) = \frac{1}{1 + L(s)} r(s)$$

Where r(s) is the reference signal and L(s) is the loop gain.

The amount of poles at s=0 in the denominator of L(s) determines the system type.

Final value theorem is used to determine the steady-state error when the reference signal is a polynomial $r(t) = t^k 1(t)$.

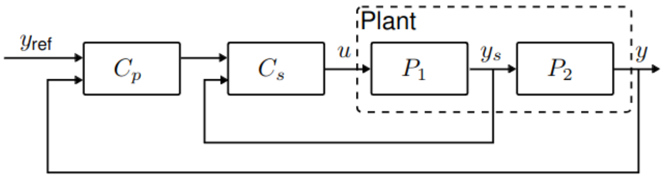
$$e_s s = \lim_{t\to\infty} e(t) = \lim_{s\to 0} e(s)s$$

$$\lim_{s\to 0} e(s)s = \lim_{s\to 0} s \frac{1}{1 + L(s)} r(s)$$

$$\lim_{s\to 0} s \frac{1}{1 + L(s)} \frac{1}{s^{k+1}}$$

3.4 Cascade Control

A cascade control uses the output from one controller as the input to another controller.



The nested control loops are called the inner loop (secondary) and the outer loop (primary). A fundamental reason for applying cascade control is to obtain better disturbance rejection and lower sensitivity to parameter variations.

3.5 DC Motor Dynamics

3.6 Examples

4 Design of PID Controllers

4.1 P-controller

The control law for the proportional controller is given by

$$u(t) = K_p e(t)$$

where K_p is the proportional gain and $e(t)$ is the error signal.

The controller applies a control signal to the systen, which depends linearly on the error signal.

To eliminate the steady state error, feedforward can be added to the P-controller.

$$u(t) = K_p e(t) + u_{ff}$$

Where the term u_{ff} is also called reset. We choose the feedforward according to the DC gain of the system, i.e.

$$u(t) = K_p e(t) + \frac{1}{G(0)} r(t)$$

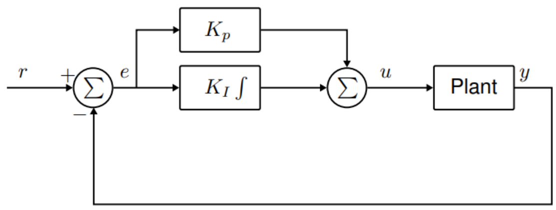
The feedforward does not affect the poles of the closed loop system.

4.2 PI-controller

The control law of a proportional-integral feedback controller is given by:

$$u = K_p e + K_I \int_{t_0}^{t} e(\tau) d\tau$$

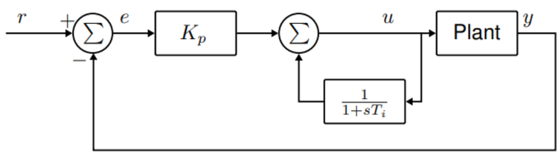
The controller adds a term, which is proportional to the integral of the error e to the P-control term.



The purpose of adding the integral term is to eliminate the steady state error of the system without the need for feed forward, i.e., integral control is less sensitive to modelling errors. Feed forward cannot eliminate modelling errors, but integral control can.

Comparing the PI-controller to the P-controller, the PI-controller has a better response to disturbances, that are unknown or cannot be measured. This is because the feedforward control can only eliminate the steady state error if the model and disturbances are known.

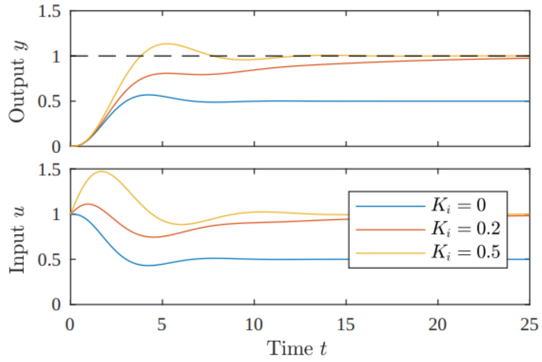
Implementation of a PI-controller is called automatic reset, and is shown in the following figure.



On the figure, T_i is the integration time (integraltid). The transfer function from e to u is given by:

$$T_{ue} = K_p \frac{1 + sT_i}{sT_i} = K_p + \frac{K_p}{sT_i}$$

On the figure below the step respones are shown for $K_i = 0, 0.2, 0.5$ and $K_p = 1$



From the unit step response it can be seen that the integral action removes the steady state error. A large K_i will give a fast response, but also a large overshoot. A small K_i will give a slow response, but also a small overshoot.

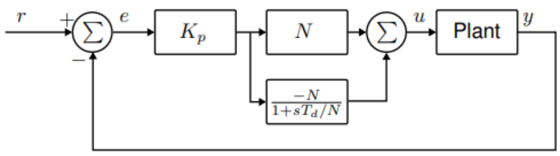
4.3 PD-controller

The derivative term provides an anticipatory action to the control, by doing feedback based on the trend of the error, i.e.

$$u(t) = K_p e(t) + K_d \frac{de(t)}{dt} = K_p \left(e + T_d \frac{de(t)}{dt} \right)$$

Where k_d is the derivative gain and T_d is the derivative time constant and e_p is a prediction of the error (T_d forwards in time). The parenthesis in the equation on the right is a prediction of the error called e_p.

PD-implementation



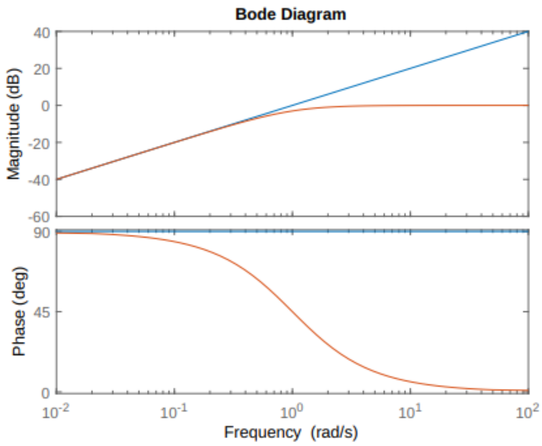
Noise is an issue for controllers that include a derivative term. Therefore, they can be implemented in a low-pass filtered version.

The transfer function of the controller is

$$T_{ue}(s) = K_p \left(N - \frac{N}{1 + sT_d/N} \right) = K_p \frac{sT_d}{1 + sT_d/N}$$

Where N is a filter constant (typical values of N are 2 to 20), and T_d is the derivative time constant.

The bode plot of an ideal PD-controller and a filtered PD controller are similar for low frequencies. Using a low-pass filter, we can at a certain frequency, reduce the gain of the derivative term, and thereby reduce the noise amplification.

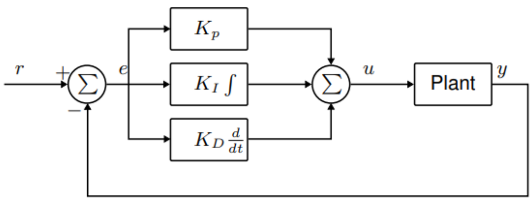


4.4 PID-controller

The control law of a PID feedback controller is

$$u(t) = K_p e(t) + K_i \int_{t_0}^{t} e(\tau) d\tau + K_d \frac{de(t)}{dt}$$

A block diagram of the controller is given below:



Alternatively, the PID-controlelr with filter on the D-term is

$$K(s) = K_p \left(1 + \frac{1}{T_i s} + \frac{sT_d}{1 + sT_d/N}\right)$$

4.5 Tuning a PID Controller

4.5.1 Pole placement

If a model of the system is available, then it is possible to compute an expression for the characteristic polynomial of the closed-loop system. Based on this polynomial, it may be possible to place the poles at desired locations.

4.5.2 Ziegler-Nichols method

Sometimes a model of the plant is not availabe, then the controller should be tuned by only studying the input-output behaviour of the system. Ziegler and Nichols has proposed two methods for tuning PID controllers without explicit use of a plant model.

- Ziegler-Nichols tuning based on step response

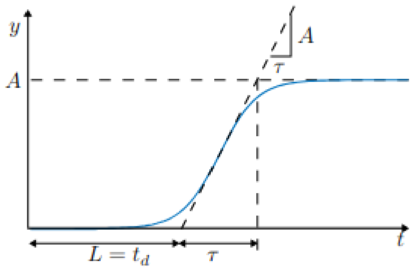
- Ziegler-Nichols tuning - The ultimate sensitivity method

Tuning based on step response

The considered system is assumed to have the transfer function

$$\frac{y(s)}{u(s)} = \frac{A}{\tau s + 1} e^{-s t_d}$$

This is a first-order system with a time delay of t_d.



The parameters for the PID controller can be found using the following table:

Controller type	Gains
P	$K_p = \frac{1}{RL}$
PI	$K_p = \frac{0.9}{RL}$ $T_i = \frac{L}{0.3}$
PID	$K_p = \frac{1.2}{RL}$ $T_i = 2L$ $T_d = 0.5L$

where $R = A/\tau$ is the amplitude of the oscillation, and $L = t_d$.

The ultimate sensitivity method

Start using a P-controller, and increase the gain until the system starts to oscillate. The value of K_p when the output oscillates with a constant amplitude is called the ultimate gain K_u. The period of the oscillation is called the ultimate period P_u.

The parameters for the PID controller can be found using the following table:

Controller type	Gains
P	$K_p = 0.5K_u$
PI	$K_p = 0.45K_u$ $T_i = \frac{P_u}{1.2}$
PID	$K_p = 0.6K_u$ $T_i = 0.5K_u$ $T_d = \frac{1}{8}P_u$

4.6 Examples

5 Root Locus

Rudkurve metoden. Grafisk metode til desgin af regulatorer.

Rule 1

Consider the characteristic equation:

$$1 + KG(s) = 1 + K\frac{Q(s)}{P(s)} = 0$$

This can be written as:

$$P(s) + KQ(s) = 0$$

This is a polynoial of degree $N = max(m, n)$ where m is the number of poles and n is the number of zeros. Lemma. A univariate polynomial of degree d has d roots in $\mathbb{C}$.

There are N lines (loci) where $N = max(m, n)$. Where m is the number of poles and n is the number of zeros.

Rule 2

Similar to previously, the characteristic equation is rewritten as:

$$P(s) + KQ(s) = 0$$

Let $K = 0$, then we observe that the roots of the characteristic equation are the poles of the open-loop system.

Let $K \to \infty$, then we observe that the roots of the characteristic equation are the zeros of the open-loop system.

$$\frac{P(s)}{K} + Q(s) = 0$$

Defintion: As K increases from 0 to ∞, the root move from the poles of $G(s)$ to the zeros of $G(s)$.

Rule 3

When roots are complex they occur in conjugate pairs.

Rule 4

We study the rewritten characteristic equation:

$$\frac{Q(s)}{P(s)} = -\frac{1}{K}$$

And see that the phase of $\frac{Q(s)}{P(s)}$ is 180° to satisfy the equation.

The transfer function can be written as:

$$T(s) = \frac{Q(s)}{P(s)} = \frac{(s - z_1)(s - z_2)\cdots(s - z_m)}{(s - p_1)(s - p_2)\cdots(s - p_n)}$$

where z_i are the zeros and p_i are the poles. Let $z \in \mathbb{C}$ then $zz* = |z|^2$; hence, complex pole pairs and pairs of complex conjugated zeros do not affect the phase of $T(s)$ for $s \in \mathbb{R}$.

The phase of $(s - z_m)$ when $s, z_m \in \mathbb{R}$ is

$$\angle(s - z_m) = \begin{cases} 180° & \text{if } s > z_m \\ 0° & \text{otherwise} \end{cases} \tag{1}$$

The portion of the real axis to the left of an odd number of open loop poles and zeros are part of the loci.

Rule 5

Lines leave and enter the real axis at 90° angles.

Rule 6

For very large values of s the equation:

$$1 + K\frac{s^m + b_1 s^{m-1} + \cdots + b_m}{s^n + a_1 s^{n-1} + \cdots + a_n} = 0$$

can be approximated to:

$$1 + K\frac{1}{(s-\alpha)^{n-m}}$$

The phase of this expression should be 180° for this expression to hold. This implies that:

$$(n-m)\phi_l = 180^\circ + 360^\circ(l-1)$$

where ϕ_l is the phase of the expression.

let $m < n-1$ then:

$$-\sum r_i = -\sum p_i$$

where r_i are the closed-loop poles and p_i are the open-loop poles.

For s going to ∞, it is known that m closed-loop poles go towards the open-loop zeros, and $n-m$ closed-loop poles go towards α, i.e.

$$-\sum r_i = -(n-m)\alpha - \sum z_i = -\sum p_i$$

where z_i are the open-loop zeros.

Thus,

$$\alpha = \frac{\sum z_i - \sum p_i}{n-m}$$

5.1 Examples

6 The Nyquist Stability Criterion

6.1 Frequency response

The frequency response of a system is the steady-state response of a system to a sinusoidal input. The output of a time-invariant system will have the same frequency as the sinusoidal input, but possibly with a different amplitude and phase.

The output of a time-invariant system is given by:

$$y(t) = \int_{-\infty}^{\infty} h(\tau)u(t-\tau)d\tau$$

Where y is the output, u is the input, and h is the impulse response of the system.

To obtain the frequency response, only sinusoidal inputs are considered.

The frequency response of H(s) is given by the magnitude and phase of $H(j\omega)$

$$M = |H(j\omega)|$$
$$\phi = \angle H(j\omega)$$

The M is the amplitude ratio and ϕ is the phase shift.

The bandwidth of a closed-loop system $T(s)$ is defined to be the maximum frequency at which the output y of a system will track a sinusoidal input r in a satisfactory manner. Output attaenuated to $1/\sqrt{2}$ of the input amplitude. Formally the bandwidth ω_{BW} of $T(s)$ is the maximal frequency such that:

$$T(j\omega) >= 1/\sqrt{2}$$

The maximal value of the frequency response is called the resonant peak M_r

6.2 Examples