POLITECNICO DI MILANO

Automatic Control

Frequency domain design

Prof. Luca Bascetta (luca.bascetta@polimi.it)

Politecnico di Milano

Dipartimento di Elettronica, Informazione e Bioingegneria

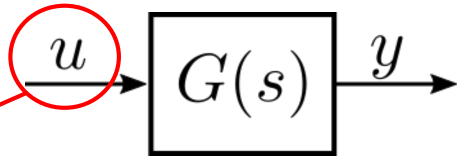
Motivations

This part introduces the fundamentals of classic control theory. We will focus on the design of SISO regulators in the frequency domain.

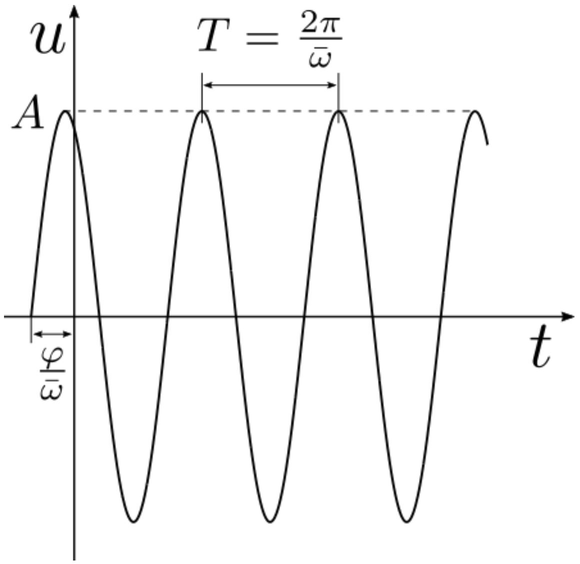
The main topics we will face are:

- frequency response
- introduction to control systems
- loop stability analysis
- loop transient performance analysis
- loop steady-state performance analysis
- control system design
- feedforward compensation
- cascaded control
- PID regulators

Given a general LTI dynamical system, represented by the transfer function $G(s)$, let's consider the sinusoidal response.



$$u(t) = A\sin(\bar{\omega}t + \varphi)$$



If $G(s)$ is <u>asymptotically stable</u>, in steady state sinusoidal inputs generate sinusoidal responses of the same frequency

$$y(t) = B\sin(\bar{\omega}t + \psi)$$

$$B = A\left|G(\mathrm{j}\bar{\omega})\right|$$

$$\psi = \varphi + \angle G(\mathrm{j}\bar{\omega})$$

Frequency response (II)

The frequency response of a system, whose transfer function is $G(s)$, is

$$G(\mathrm{j}\omega) \quad \omega > 0$$

a complex function of the real variable ω.

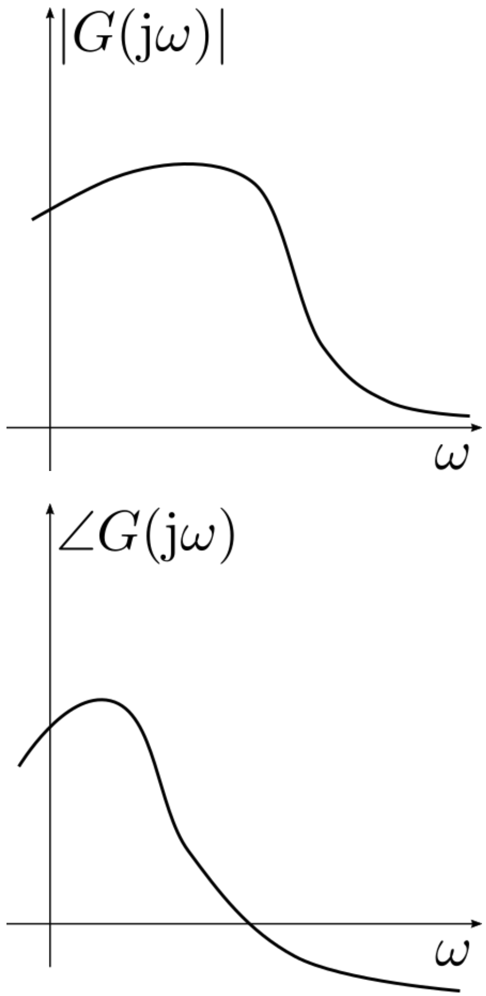
A few remarks:

- the frequency response is the transfer function evaluated along the positive imaginary axis
- ω is called frequency
- the frequency response can be defined for stable and unstable LTI systems
- the sinusoidal response theorem holds only for asymptotically stable LTI systems

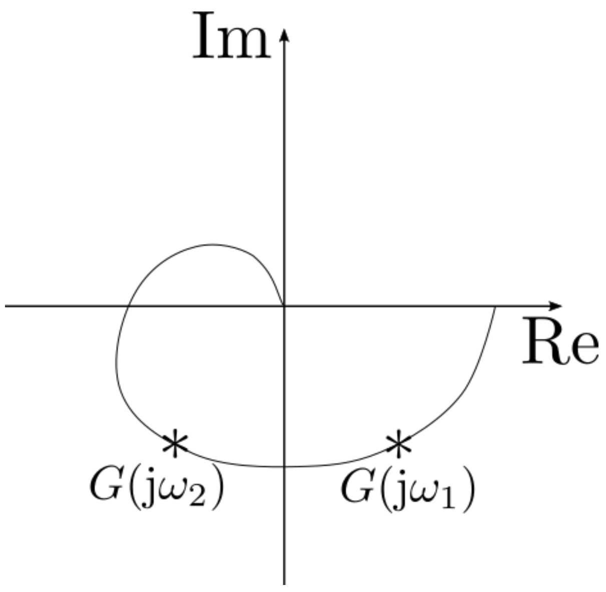
The frequency response is a complex function, how can we plot it?

Frequency response (III)

We can plot separately the absolute value and the argument (Cartesian plots) or we can plot the curve on the complex plane (Polar plots).



Cartesian plot



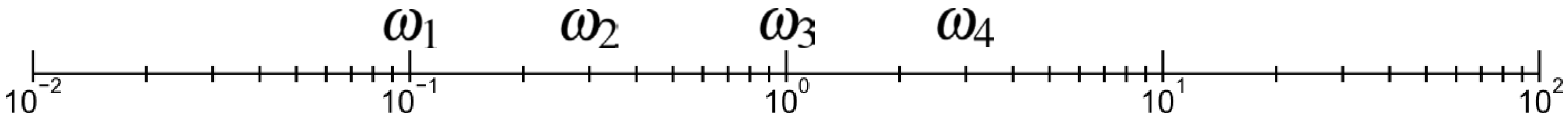
Polar plot

Frequency response (IV)

Let’s start from Cartesian plots and, in particular, from Bode plots.

In Bode plots, the frequency response absolute value and argument are plotted in two separate Cartesian plots.

In both plots the x-axis (frequency axis) is a logarithmic axis.



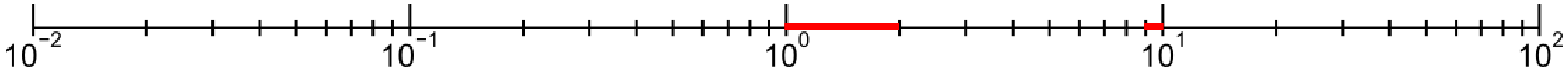
$$\log(\omega_2) - \log(\omega_1) = \log(\omega_4) - \log(\omega_3) \quad \Rightarrow \quad \frac{\omega_2}{\omega_1} = \frac{\omega_4}{\omega_3}$$

Frequency response (IV)

Let's start from Cartesian plots and, in particular, from Bode plots.

In Bode plots, the frequency response absolute value and argument are plotted in two separate Cartesian plots.

In both plots the x-axis (frequency axis) is a logarithmic axis.



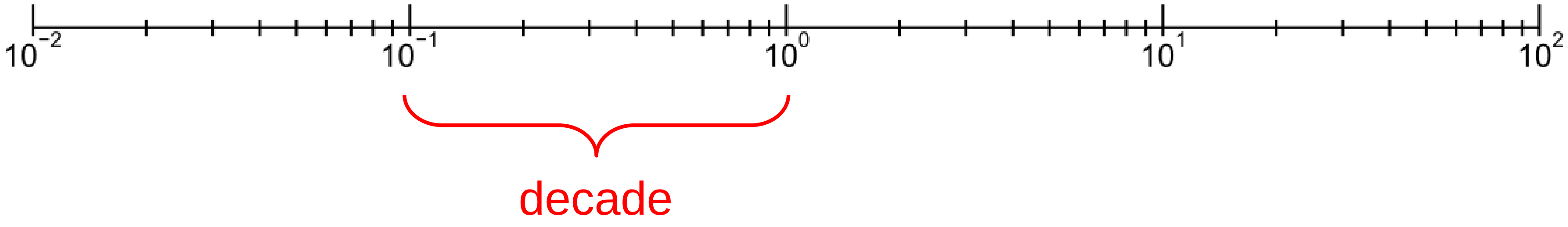
$$\log(2) \gg \log\left(\frac{10}{9}\right)$$

Frequency response (IV)

Let's start from Cartesian plots and, in particular, from Bode plots.

In Bode plots, the frequency response magnitude and argument are plotted in two separate Cartesian plots.

In both plots the x-axis (frequency axis) is a logarithmic axis.



In both plots the y-axis is a linear axis.

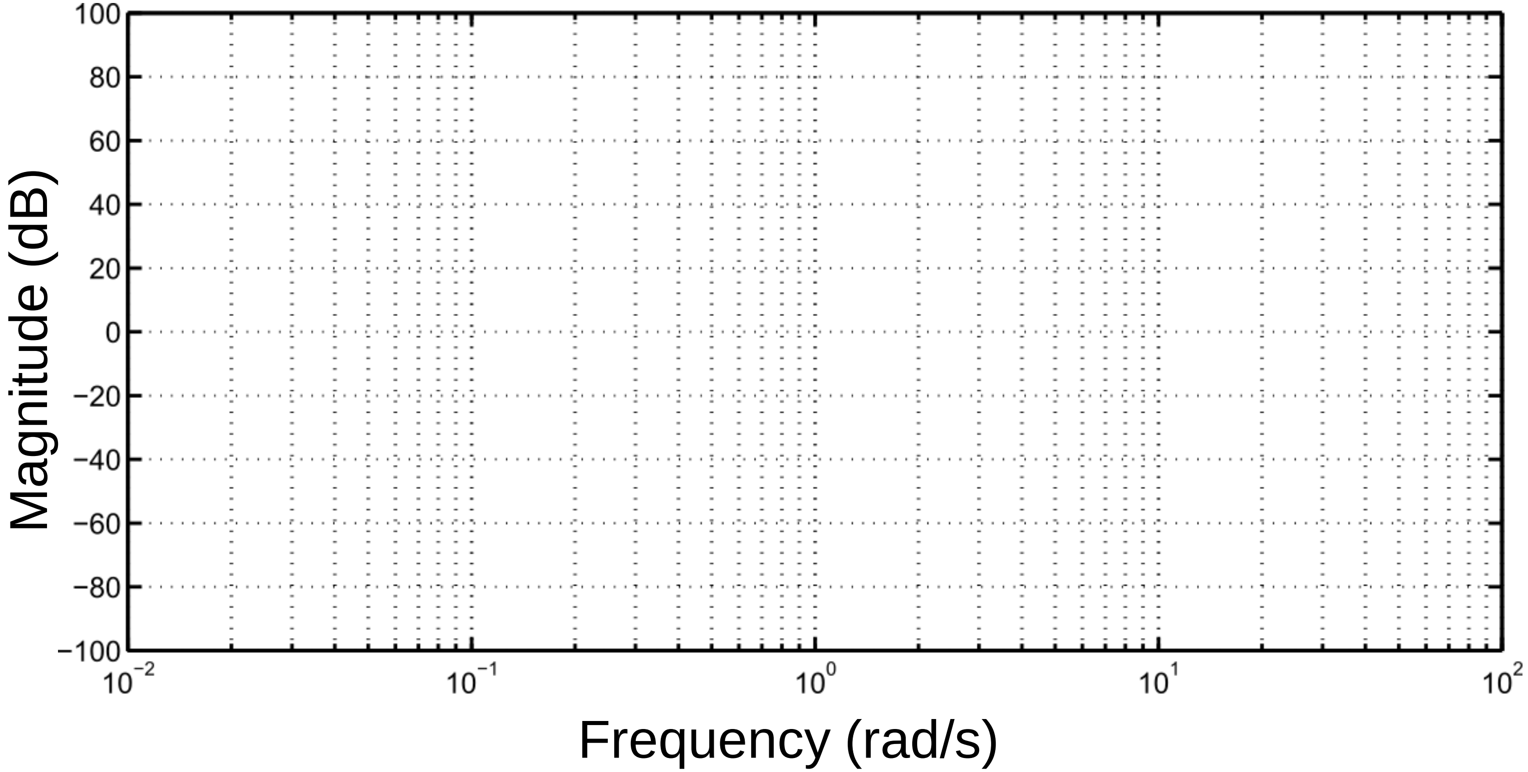
We will call both plots semi logarithmic plots, and draw it on semi logarithmic paper.

Frequency response (V)

In the magnitude plot the y-axis is a linear axis.

The absolute value of the frequency response is plotted in decibel

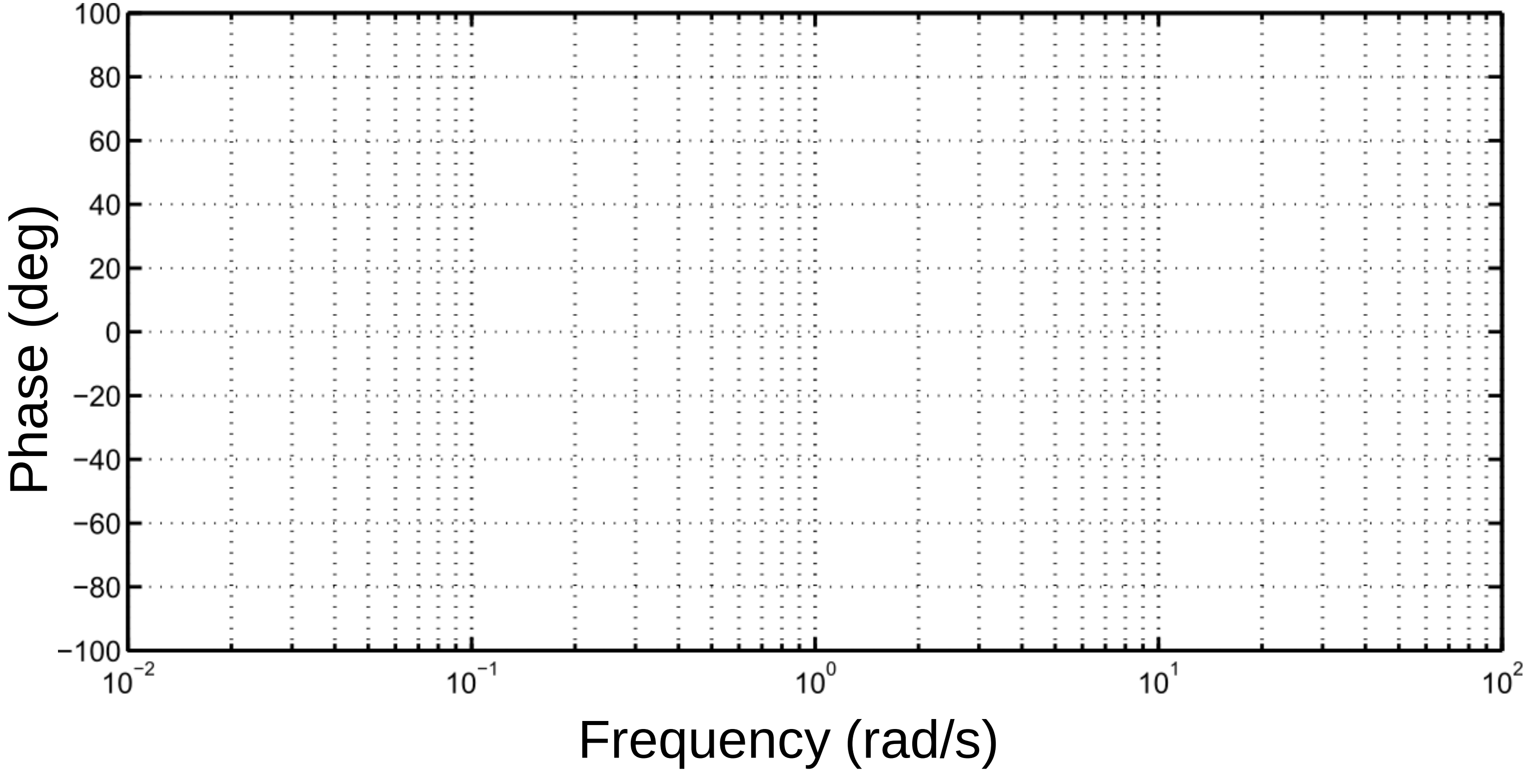
$$\left|G(\mathrm{j}\omega)\right|_{dB} = 20\log_{10}\left|G(\mathrm{j}\omega)\right|$$



Frequency response (VI)

In the phase plot the y-axis is a linear axis.

The argument of the frequency response is plotted in degrees.



For control system analysis and design we will use the asymptotic Bode plots.

Asymptotic Bode plots are an approximation of Bode plots that can be easily manually drawn.

We will now introduce the rules to draw asymptotic Bode plots.

The rules make reference to the gain-time constant form

$$G(s) = \frac{\mu}{s^g} \frac{\prod_i (1 + s\tau_i)}{\prod_j (1 + sT_j)}$$

Magnitude plot

1. For $\omega \to 0$, the plot starts with a line having slope $-g$ and going through point $\omega = 1\ \mathrm{rad/s}$, $|G|_{dB} = |\mu|_{dB}$.

2. At every frequency corresponding to p real poles (zeros), slope decreases (increases) of p units.

3. At every frequency corresponding to the natural frequency of p complex poles (zeros), slope decreases (increases) of $2p$ units.

4. For $\omega \to \infty$, the slope of the plot equals the difference between the number of zeros and the number of poles of the transfer function.

In asymptotic Bode plots slopes are multiples of 20 dB/decade (slope 2 means 40 dB/decade, slope -3 means -60 dB/decade)

Phase plot

1. For $\omega \to 0$, the plot starts with a line parallel to the x-axis and crossing the y-axis at $\angle\mu - g90^{\circ}$.

2. At every frequency corresponding to p real zeros in the left half plane or p real poles in the right half plane, the phase increases of $p90^{\circ}$ degrees.

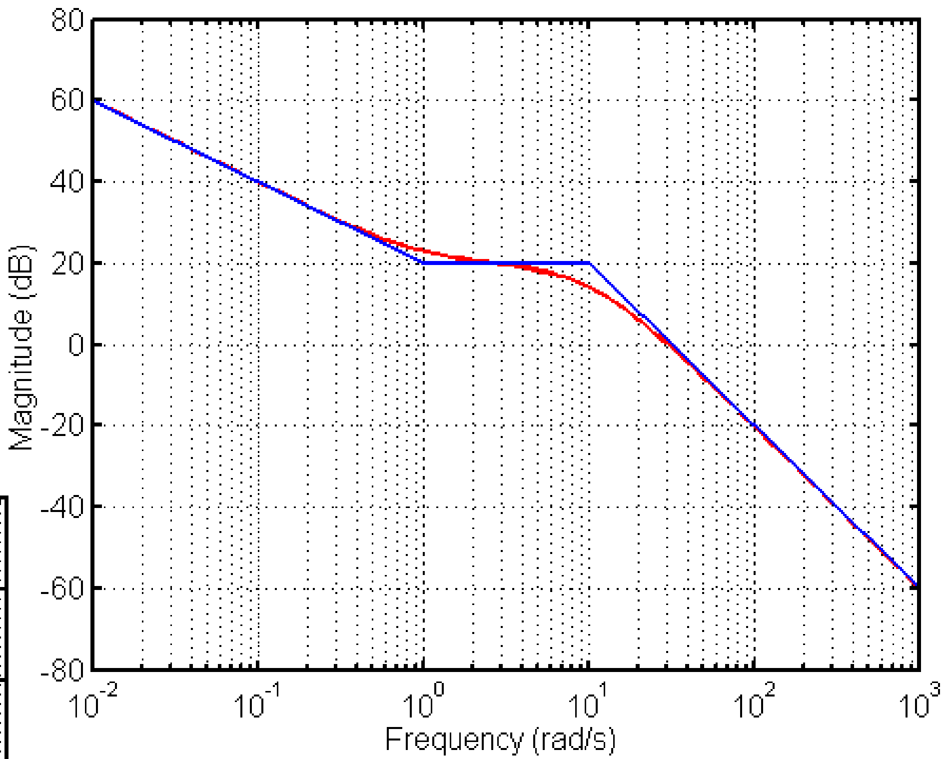
3. At every frequency corresponding to p real zeros in the right half plane or p real poles in the left half plane, the phase decreases of $p90^{\circ}$ degrees.

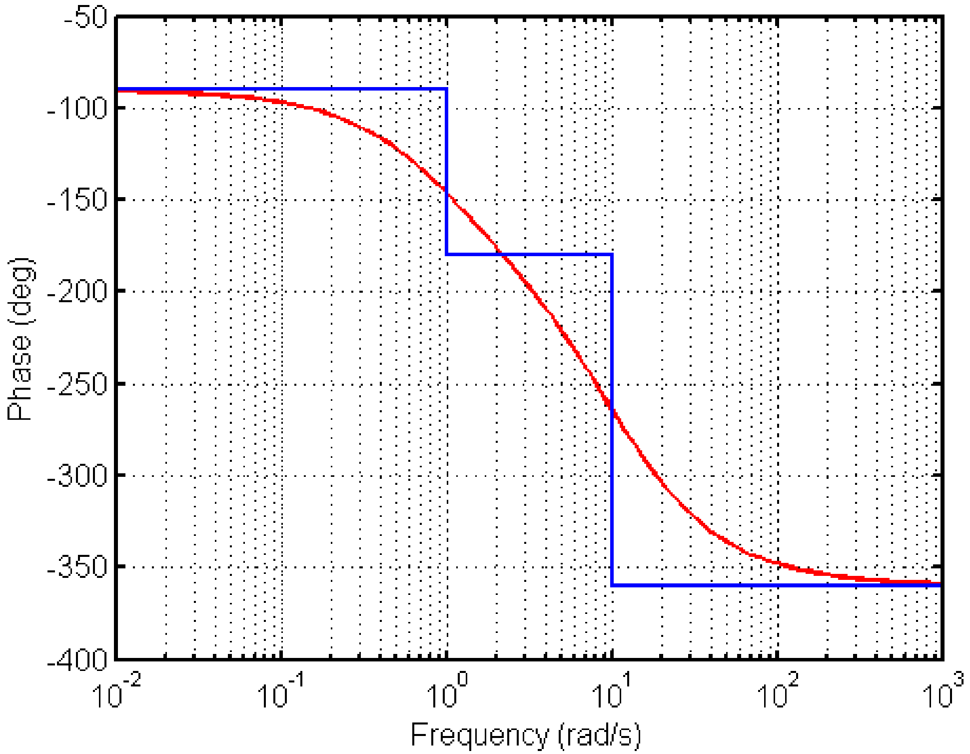
4. At every frequency corresponding to the natural frequency of p complex zeros in the left half plane or p complex poles in the right half plane, the phase increases of $p180^{\circ}$ degrees.

5. At every frequency corresponding to the natural frequency of p complex zeros in the right half plane or p complex poles in the left half plane, the phase decreases of $p180^{\circ}$ degrees.

Frequency response – Example

$$G(s)=\frac{10}{s}\frac{1-s}{(1+0.1s)^2}$$





Frequency response (X)

Polar plots are the other way to plot frequency response.

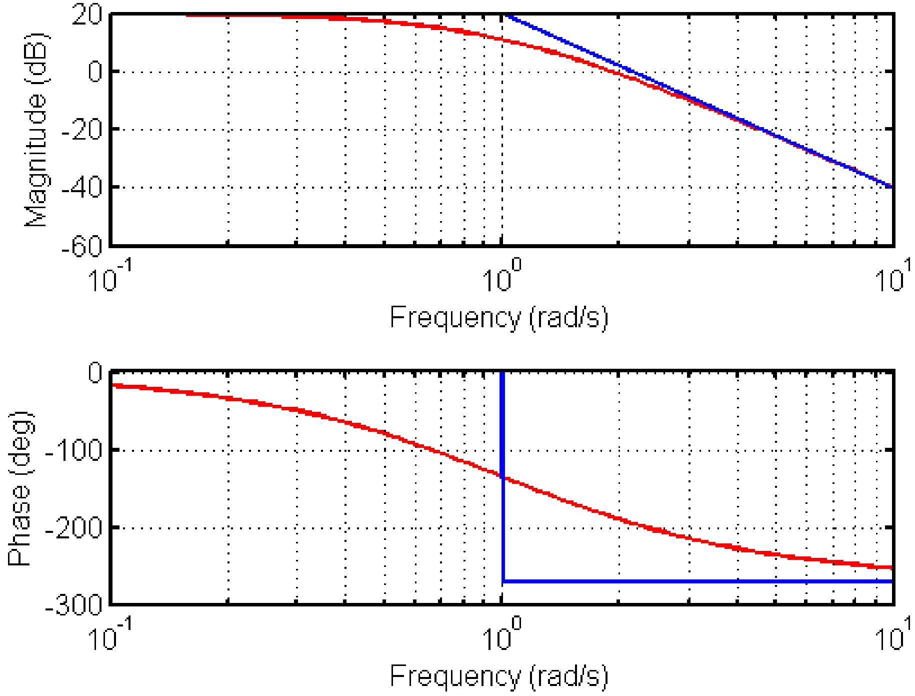
Polar plots are the image of the complex number $G(\mathrm{j}\omega)$ for $\omega \in [0, \infty)$.

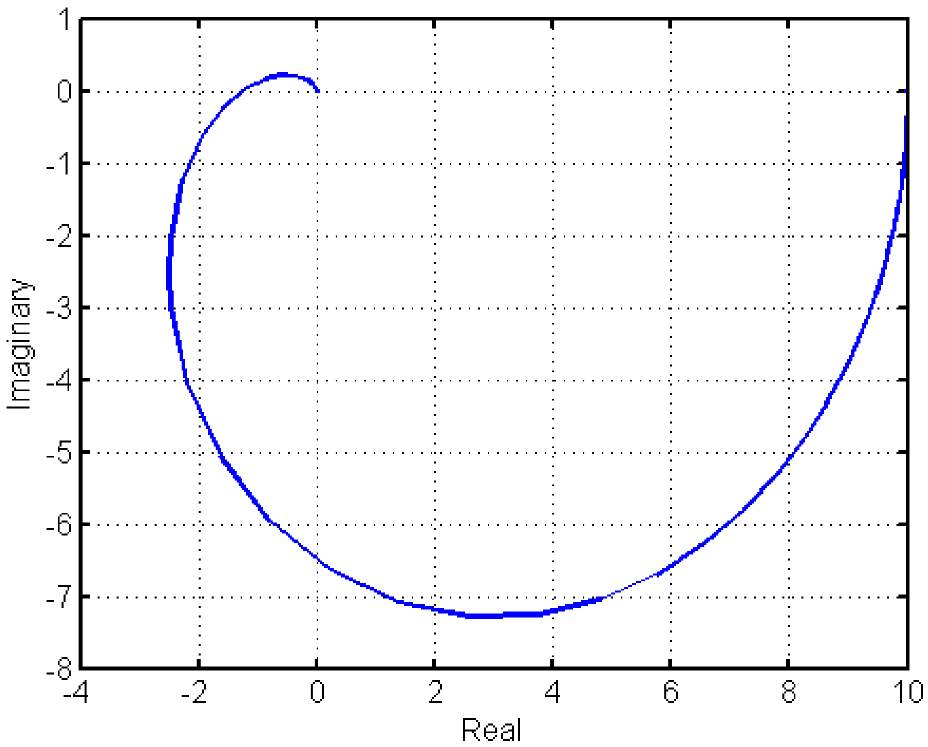
We will see how polar plots can be drawn, deriving from Bode plots how the magnitude and phase of $G(\mathrm{j}\omega)$ change as ω changes.

Caveat: if the transfer function $G(s)$ has poles on the imaginary axis its polar plot tends to infinity along asymptotes whose equations can be analytically computed.

Frequency response – Example

$$G(s) = \frac{10}{(1+s)^3}$$





The sinusoidal response theorem can be easily extended to general periodic or a-periodic input signals using Fourier series and Fourier integral.

Let's start from periodic input signals

$$u(t+T) = u(t) \quad \forall t$$

$$\int_0^T |u(t)| \mathrm{d}t < \infty$$

that can be represented as series of sinusoidal signals as

$$u(t) = U_0 + \sum_{n=1}^{\infty} U_n \cos(n\omega_0 t + \varphi_n) \quad \omega_0 = \frac{2\pi}{T}$$

If $G(s)$ is asymptotically stable, in steady state periodic inputs generate periodic responses of the same fundamental frequency

$$y(t) = Y_0 + \sum_{n=1}^{\infty} Y_n \cos(n\omega_0 t + \psi_n)$$

$$Y_n = |G(\mathrm{j}n\omega_0)| U_n$$

$$\psi_n = \varphi_n + \angle G(\mathrm{j}n\omega_0)$$

Consider an a-periodic input signal

$$\nexists T : u(t+T) = u(t) \quad \forall t$$

$$\int_{-\infty}^{+\infty} |u(t)| \mathrm{d}t < \infty$$

that can be represented as integral of sinusoidal signals as

$$u(t) = \int_0^{\infty} U(\omega) \cos(\omega t + \varphi(\omega)) \, \mathrm{d}\omega$$

If $G(s)$ is asymptotically stable, in steady state a-periodic inputs represented as a Fourier integral generate a-periodic responses that can be represented as a Fourier integral

$$y(t) = \int_0^{\infty} Y(\omega) \cos(\omega t + \psi(\omega)) \, \mathrm{d}\omega$$

$$Y(\omega) = |G(\mathrm{j}\omega)| U(\omega)$$

$$\psi(\omega) = \varphi(\omega) + \angle G(\mathrm{j}\omega)$$

Frequency response (XIII)

Using the sinusoidal response theorem and its extensions based on Fourier series and Fourier integral we can compute the response to any sinusoidal, periodic or a-periodic signal.

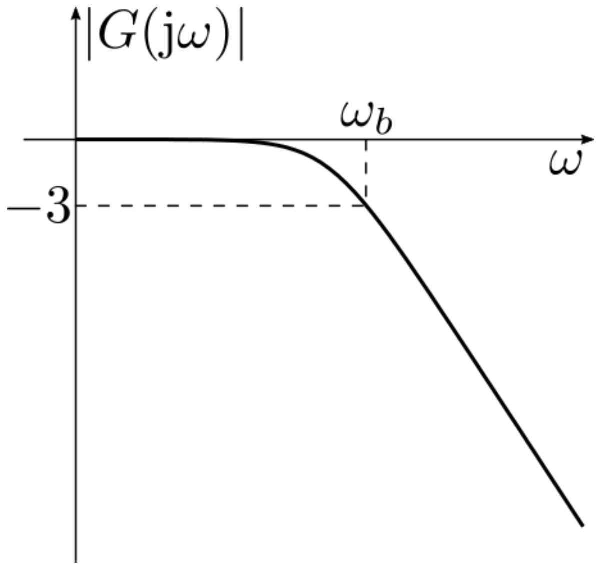
Intuitively, the sinusoidal response theorem states that sinusoidal harmonics are changed in amplitude and phase by the dynamical system, depending on the system frequency response.

In view of this interpretation, an asymptotically stable LTI system acts as a filter on input signals: some harmonics are left unchanged, other are amplified and other are attenuated.

There are many different filters, here we will describe only the low-pass filter that has important applications in control system analysis and design.

Frequency response (XIV)

A <u>low-pass filter</u> is an asymptotically stable LTI system whose frequency response has the following magnitude Bode plot



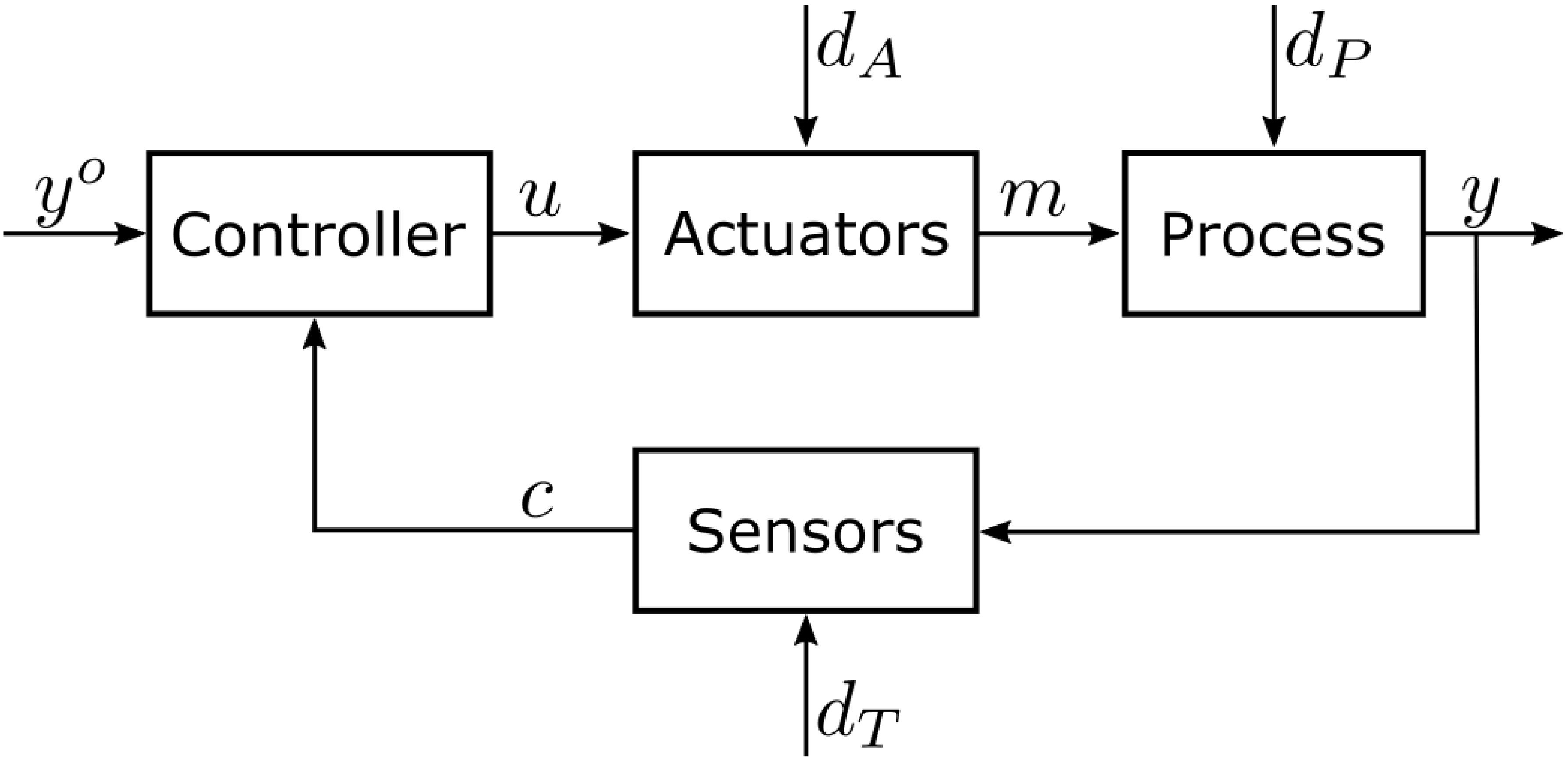
If

$$|G(j\omega)|_{dB} < 3 \quad \forall \omega$$

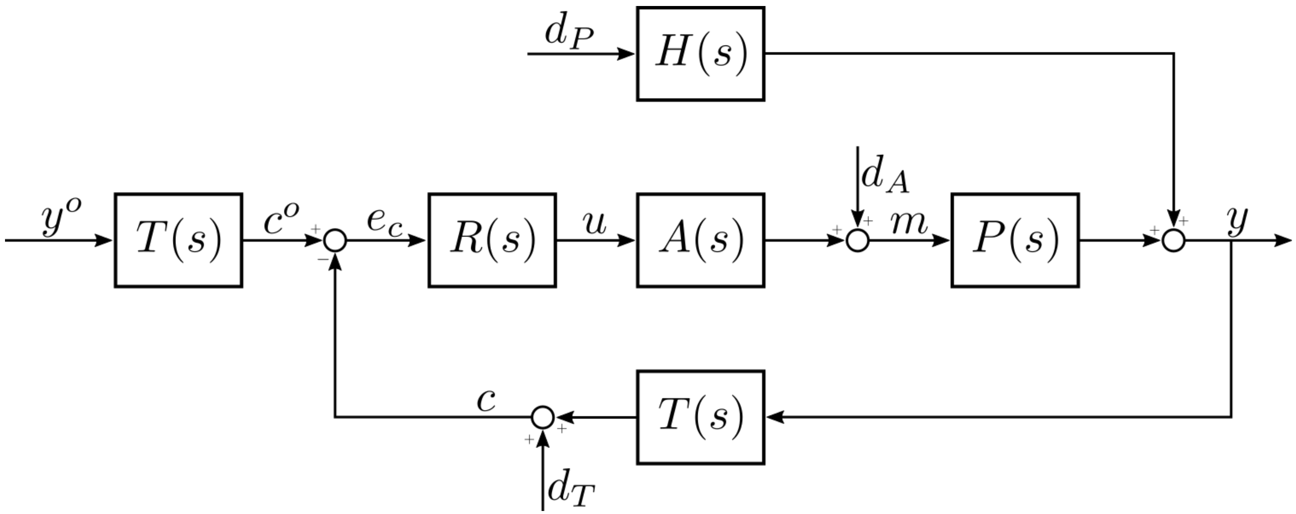
we can define the <u>cutoff frequency</u> ω_b and the <u>passband</u>

$$PB = \{\omega : |G(j\omega)|_{dB} > -3\} = [0, \omega_b]$$

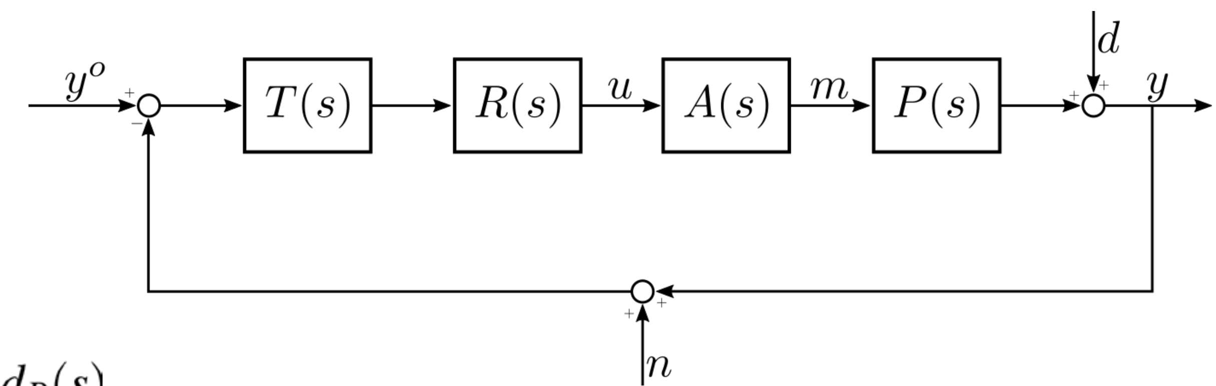
A general feedback control system can be represented by the following block diagram



Assuming that all the components can be described by linear systems, the block diagram can be drawn introducing the transfer functions

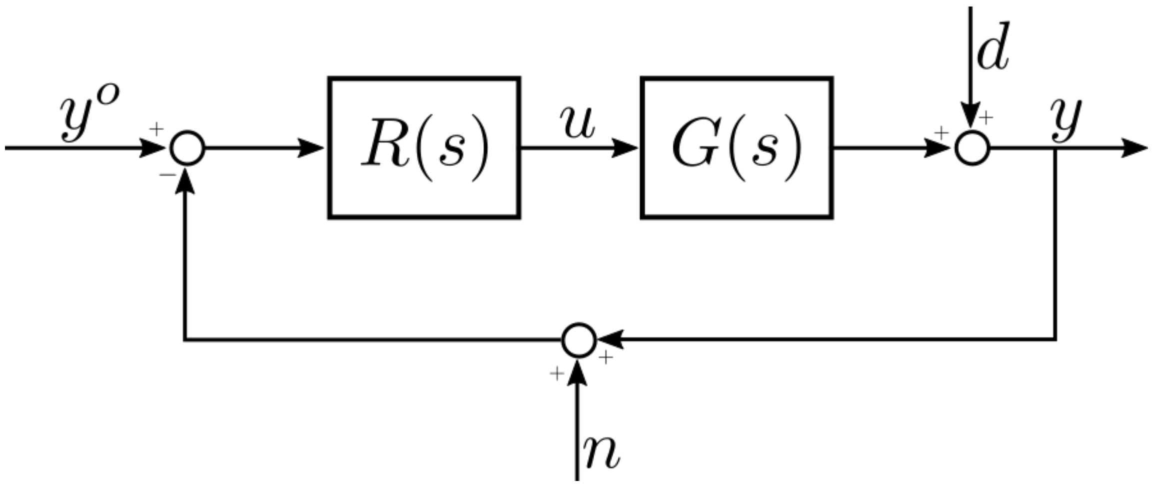


and simplifying the block diagram



$$n(s) = T(s)^{-1} d_T(s)$$
$$d(s) = P(s) d_A(s) + H(s) d_P(s)$$

Finally, the control system is represented by the following block diagram



where

$$G(s) = T(s)P(s)A(s)$$

is the transfer function of the process including actuators and sensors.

In the problem we will study, the goal will be the design of the controller transfer function $R(s)$ given the system transfer function $G(s)$.

In a real problem the control engineer has to model the process and select sensors and actuators as well.

The design of the control system is driven by a set of requirements specification.

The main requirements are:

- asymptotic stability (nominal and robust)
- steady-state performance (steady-state error)
- transient performance (response time, disturbance rejection, oscillation damping)
- control effort mitigation

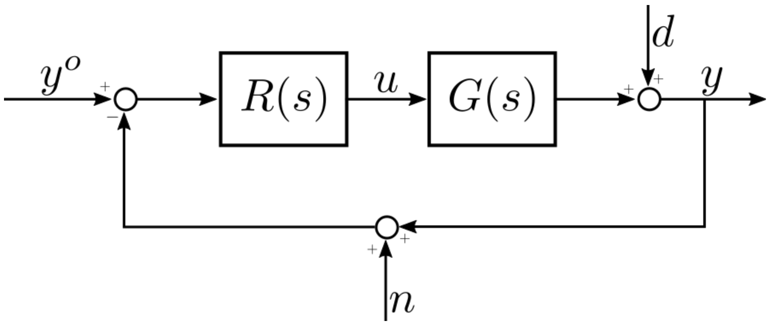
We will first study how to analyze a closed-loop system, given the controller transfer function $R(s)$.

Then control system design techniques based on Bode plots will be discussed.

Loop stability analysis (I)

We start analyzing the stability of the closed-loop system.

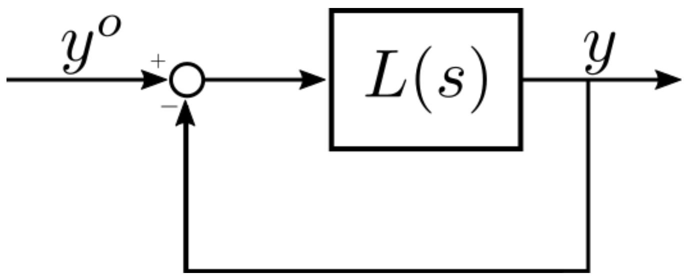
Given the block diagram



to analyze the system stability we can neglect disturbances and introduce the loop transfer function

$$L(s) = R(s)G(s)$$

and the block diagram can be simplified as follows



We will assume $L(s)$ the transfer function of a strictly proper LTI system.

Loop stability analysis (II)

The loop transfer function can be expressed as

$$L(s) = \frac{N(s)}{D(s)}$$

The transfer function from the set point y^o to the controlled variable y is, instead, given by

$$\frac{Y(s)}{Y^o(s)} = \frac{L(s)}{1+L(s)} = \frac{\frac{N(s)}{D(s)}}{1+\frac{N(s)}{D(s)}} = \frac{N(s)}{N(s)+D(s)}$$

The poles of this transfer function are the poles of the closed-loop system, we can thus define a characteristic polynomial of the closed-loop system as

$$\chi(s) = N(s) + D(s)$$

We conclude that the closed-loop system is asymptotically stable if and only if all the roots of the characteristic polynomial lie in the open left half plane.

Loop stability analysis (III)

Caveat: if pole-zero cancellations occur in the product $R(s)G(s)$ and these poles/zeros lie in the closed right half plane, the closed-loop system cannot be asymptotically stable (there is an hidden dynamics that is not asymptotically stable).

The analysis of the characteristic polynomial can be used to assess the stability of the closed-loop system, but it does not represent a viable solution for control design purposes (i.e., to select $R(s)$ in such a way that the closed-loop system is asymptotically stable).

For this reason we will now introduce two graphical methods to assess the stability of a closed-loop system.

Nyquist stability criterion is a graphical technique to determine the stability of a closed-loop system analyzing the frequency response of the loop transfer function.

Let's first introduce the following definitions:

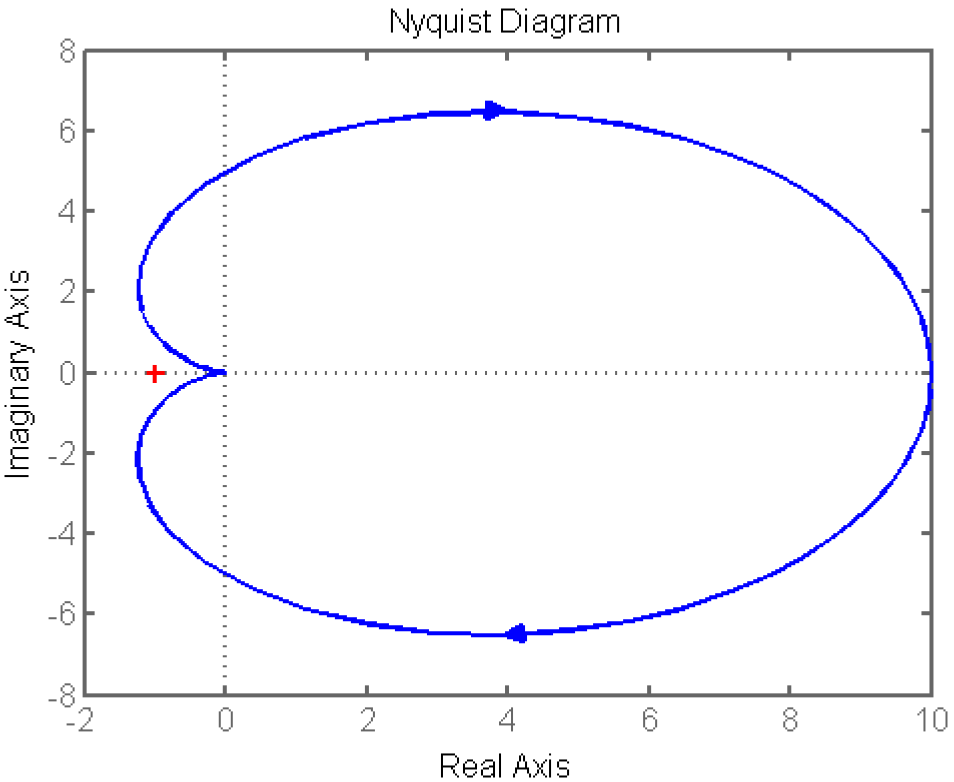
- Nyquist plot, a closed curve constituted by the Polar plot of the loop transfer function $L(s)$ (the curve is oriented in the direction of increasing ω) and its symmetric with respect to the real axis
- P_d, number of poles of $L(s)$ lying in the open right half plane
- N, number of plot's encirclements of -1; we consider clockwise encirclements to be negative and counterclockwise encirclements to be positive (if Nyquist plot crosses the real axis at -1, N is undefined)

Nyquist stability criterion states that the closed-loop system is asymptotically stable if and only if N is not undefined and $N = P_d$.

Caveat: Nyquist stability criterion is a necessary and sufficient condition!

Loop stability analysis – Examples (I)

$$L(s) = \frac{10}{(1+s)^2}$$



We have

$$P_d = 0 \quad N = 0$$

the closed-loop system is thus asymptotically stable.

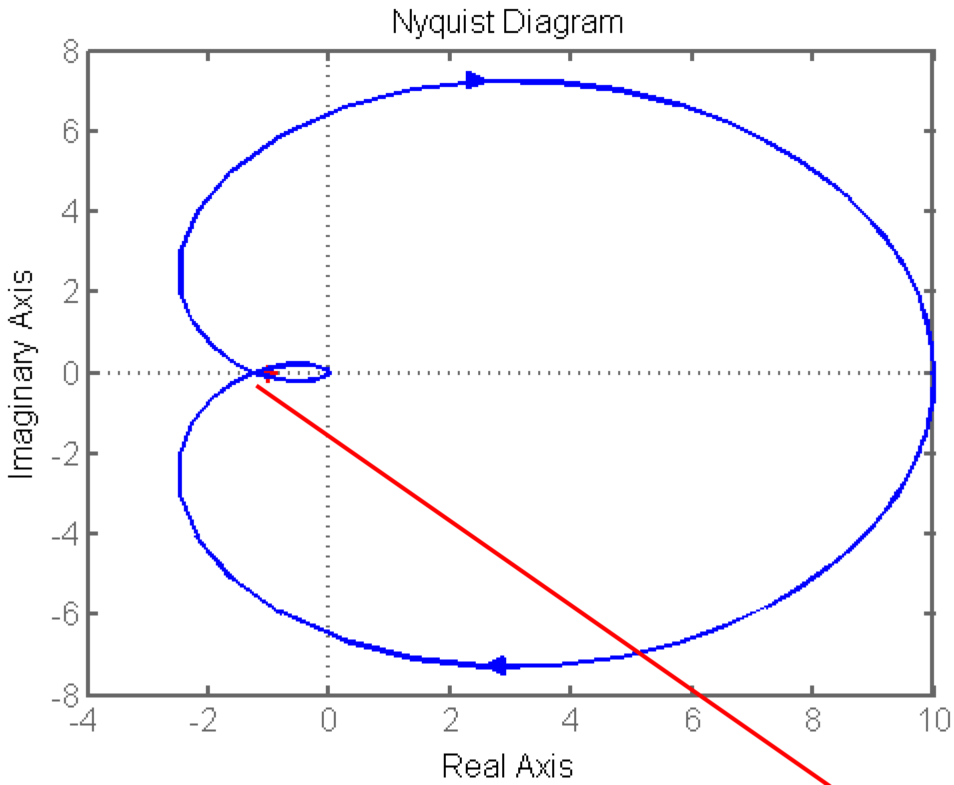
We can verify the result using Routh criterion. The characteristic polynomial is

$$\chi(s) = 10 + (1+s)^2 = s^2 + 2s + 11$$

and has the roots in the open left half plane.

Loop stability analysis – Examples (II)

$$L(s) = \frac{10}{(1+s)^3}$$



We have

$$P_d = 0 \quad N = -2$$

the closed-loop system is thus not asymptotically stable.

We can verify the result using Routh criterion. The characteristic polynomial is

$$\chi(s) = 10 + (1+s)^3 = s^3 + 3s^2 + 3s + 11$$

and has two complex conjugate roots in the right half plane.

the intersection of Nyquist plot with the real axis can be computed determining $\omega_\pi \colon \angle L(\mathrm{j}\omega_\pi) = -180°$ and then computing $|L(\mathrm{j}\omega_\pi)|$

Loop stability analysis (V)

We now introduce the second graphical criterion, the Bode stability criterion, that is based on the Bode plots of the loop transfer function $L(s)$.

The Bode stability criterion holds only if the following conditions are satisfied:

- the loop transfer function $L(s)$ does not have poles lying in the open right half plane
- the magnitude Bode plot of the loop transfer function $L(s)$ crosses the $0\ dB$-axis exactly once

Let's first introduce the following definitions:

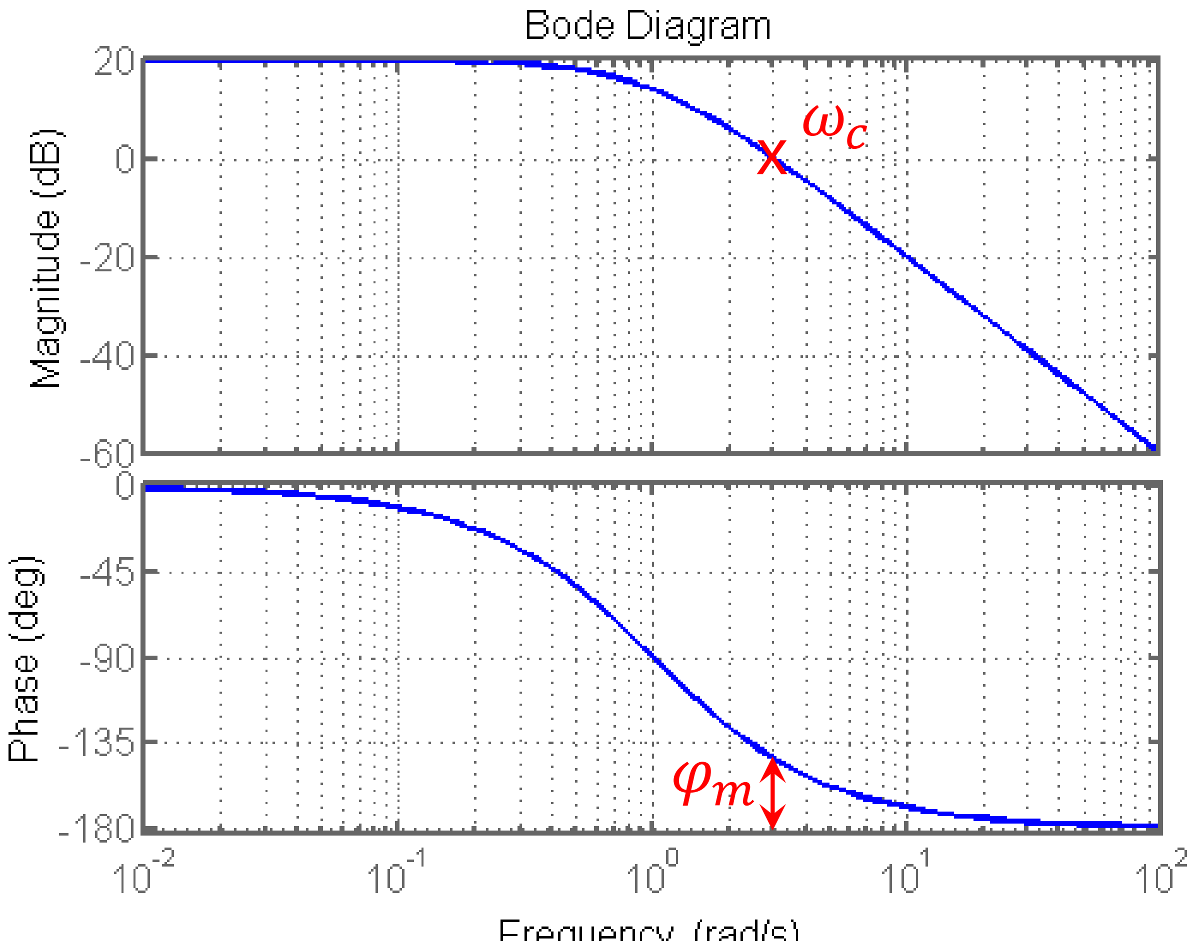
- crossover frequency ω_c, is the frequency ω so that $|L(\mathrm{j}\omega)| = 1$
- phase margin φ_m, is defined as

$$\varphi_m = 180^\circ - |\angle L(\mathrm{j}\omega_c)|$$

- loop gain μ_L, is the gain of the loop transfer function $L(s)$

Loop stability analysis (VI)

The quantities we have just defined can be shown on the magnitude and phase Bode plots of the loop transfer function.



Loop stability analysis (VII)

Bode stability criterion states that the closed-loop system is asymptotically stable if and only if

$$\mu_L > 0 \quad \varphi_m > 0$$

Caveat: Bode stability criterion is a necessary and sufficient condition!

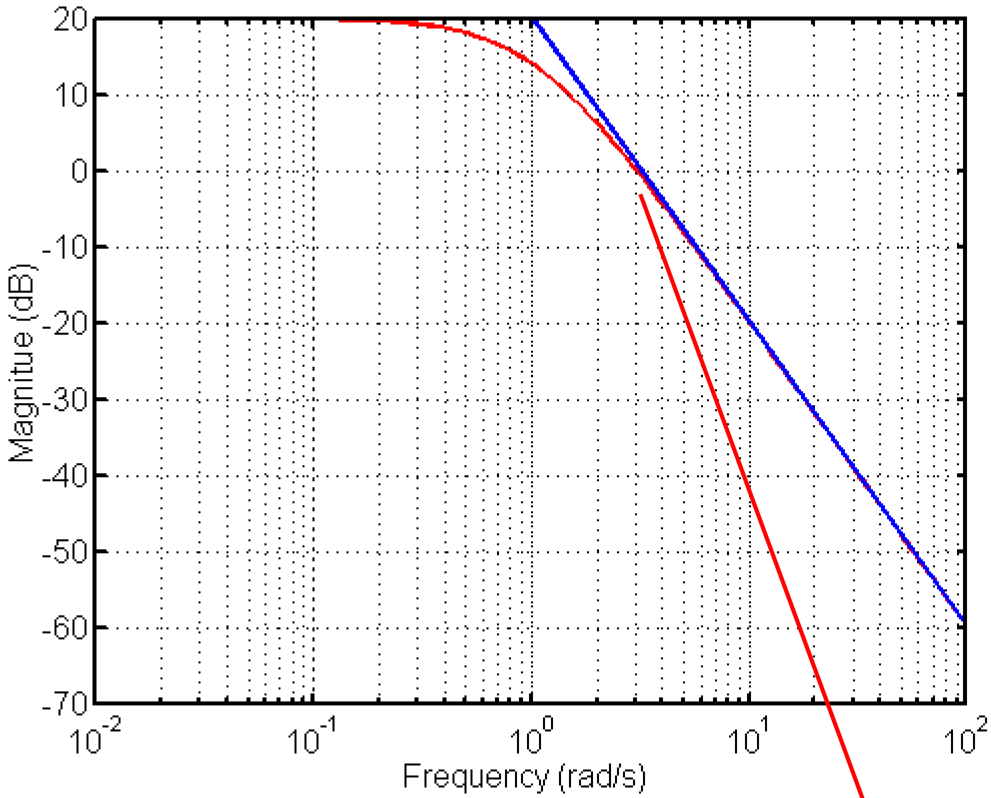
Bode stability criterion can be easily derived from Nyquist stability criterion assuming $P_d = 0$.

Loop stability analysis – Examples (I)

$$L(s) = \frac{10}{(1+s)^2}$$

- Applicability conditions hold
- $\mu_L > 0$
- $\omega_c \approx 3\, rad/s$
- $\angle L(j\omega_c) = -2\arctan(3) = -144°$
- $\varphi_m = 180° - |\angle L(j\omega_c)| = 36° > 0$

We conclude that the closed-loop system is asymptotically stable.



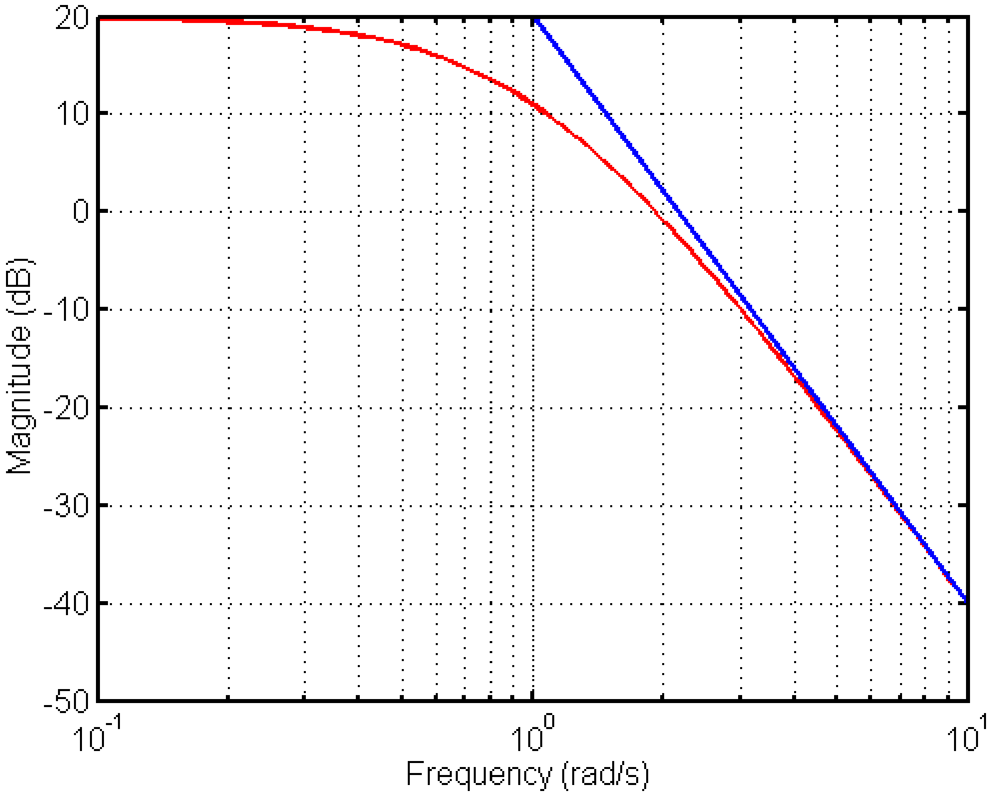
You don't need to analytically compute the crossover frequency, just read it from the asymptotic magnitude Bode plot

Loop stability analysis – Examples (II)

$$L(s) = \frac{10}{(1+s)^3}$$

- Applicability conditions hold
- $\mu_L > 0$
- $\omega_c \approx 2\,rad/s$
- $\angle L(j\omega_c) = -3\arctan(2)$ $= -192°$
- $\varphi_m = 180° - |\angle L(j\omega_c)| = -18° < 0$

We conclude that the closed-loop system is not asymptotically stable.



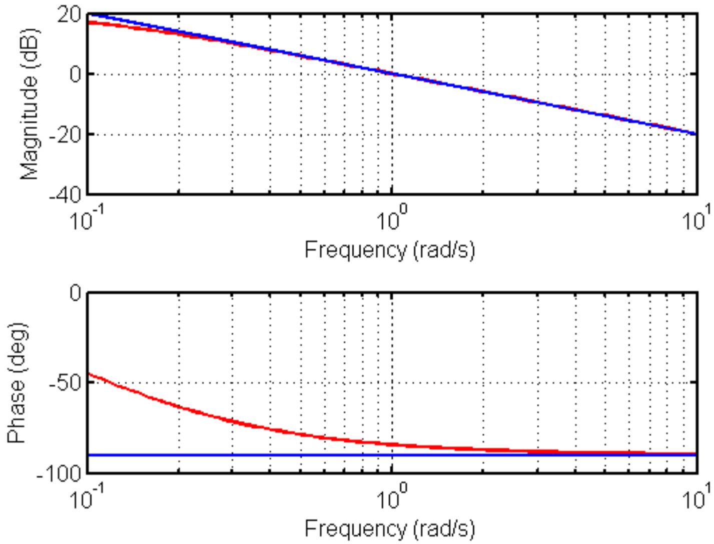
Loop stability analysis (VIII)

We call a LTI system a minimum phase system if:

- the gain is greater than zero
- all the poles lie in the closed left half plane
- all the zeros lie in the closed left half plane

For a minimum phase system the asymptotic phase Bode plot can be drawn from the asymptotic magnitude Bode plot multiplying the slope of each segment by $90°$.

Caveat: if the magnitude Bode plot crosses the $0\ dB$-axis having a slope of -1, and there are no other poles/zeros around the crossover frequency, the phase margin will be close to its asymptotic value ($90°$).
This ensures that the closed-loop system is asymptotically stable.



Loop stability analysis (IX)

Time delay is a system described by the following equation in time domain

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

its transfer function is

$$G(s) = \mathrm{e}^{-s\tau}$$

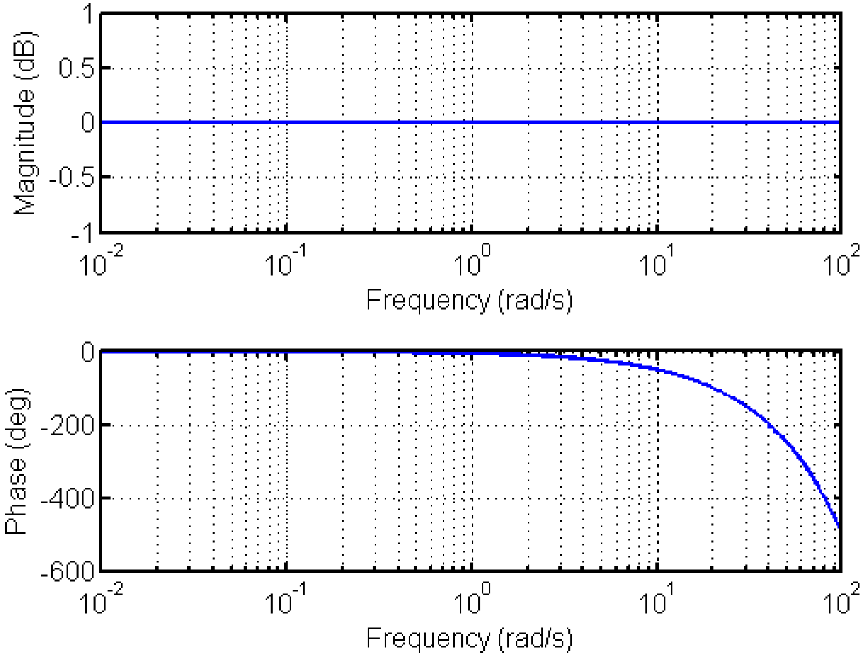
and its frequency response

$$G(\mathrm{j}\omega) = \mathrm{e}^{-\mathrm{j}\omega\tau}$$

analyzing this frequency response we obtain

$$|G(\mathrm{j}\omega)| = |\mathrm{e}^{-\mathrm{j}\omega\tau}| = 1$$

$$\angle G(\mathrm{j}\omega) = |\mathrm{e}^{-\mathrm{j}\omega\tau}| = -\omega\tau$$



Loop stability analysis (X)

Assume that the loop transfer function is composed of a rational transfer function and a delay

$$L(s) = L_r(s)\mathrm{e}^{-s\tau}$$

The magnitude and phase of the frequency response are given by

$$|L(\mathrm{j}\omega)| = |L_r(\mathrm{j}\omega)||\mathrm{e}^{-\mathrm{j}\omega\tau}| = |L_r(\mathrm{j}\omega)|$$

$$\angle L(\mathrm{j}\omega) = \angle L_r(\mathrm{j}\omega) + \angle \mathrm{e}^{-\mathrm{j}\omega\tau} = \angle L_r(\mathrm{j}\omega) - \omega\tau$$

as a consequence

$$\omega_c = \omega_{c_r}$$

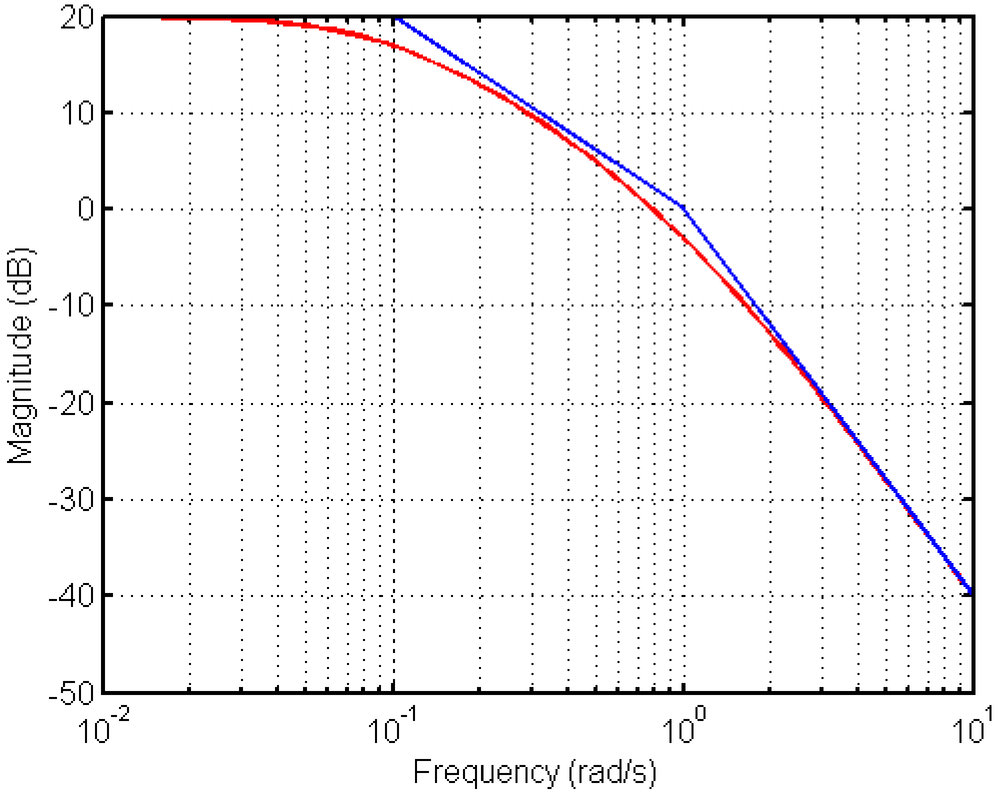
$$\angle L(j\omega_c) = \angle L_r(j\omega_c) \boxed{- \omega_c \tau \frac{180^\circ}{\pi}}$$

Phase decrement due to the delay

Loop stability analysis – Example

$$L(s) = \frac{10}{(1+s)(1+10s)} e^{-s\tau} \qquad \tau > 0$$

- Applicability conditions hold
- $\mu_L > 0$
- $\omega_c \approx 1\, rad/s$
- $\angle L(j\omega_c) = -\arctan\left(\frac{10}{1}\right) - \arctan\left(\frac{1}{1}\right) - \omega_c \tau \frac{180°}{\pi} \approx -129° - 57°\tau$
- $\varphi_m = 180° - |\angle L(j\omega_c)| \approx 51° - 57°\tau$



We conclude that the closed-loop system is asymptotically stable for $\tau < 0.89\ s$

Loop transient performance analysis (I)

We will now analyze the following features that characterize the transient response of a closed-loop system:

- response time, how quickly the closed-loop system responds to a change in the set point

- oscillation damping, the response of the closed-loop system should be overdamped or the oscillations should be characterized by a high damping factor

- disturbance rejection, the control system ensures good set point tracking even in the presence of disturbances

- control effort mitigation, the amplitude of the control signal should not be excessively high

We will try to find a correlation between these features and the parameters that characterize the stability of the closed-loop system.

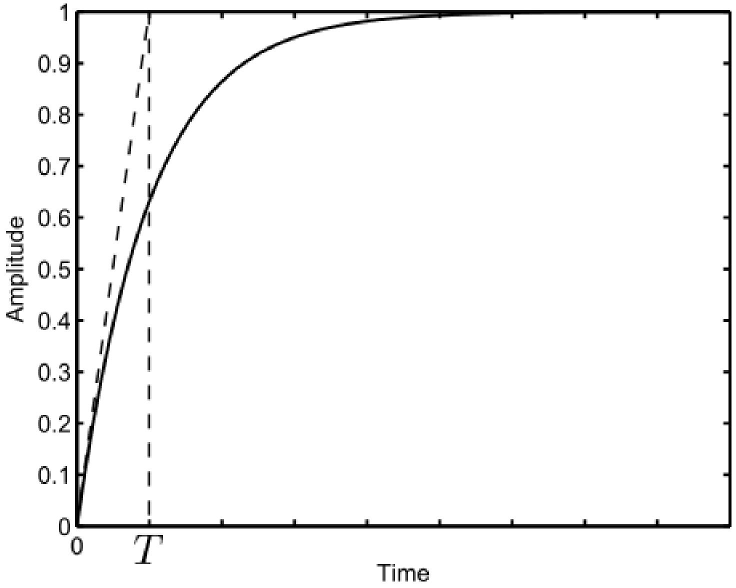
Loop transient performance analysis (II)

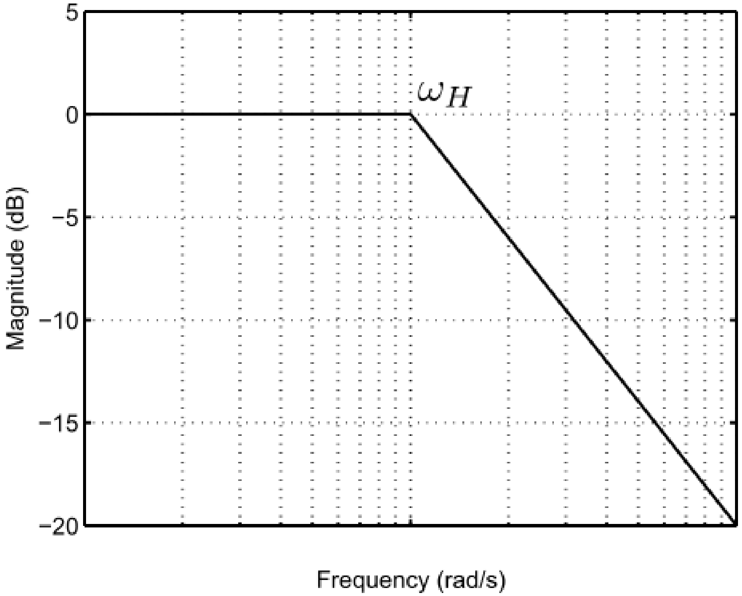
Let's start from response time.

We usually think about response time as a measure to characterize, for example, the step response in time domain.
Can we found an interpretation of this parameter in the frequency domain?

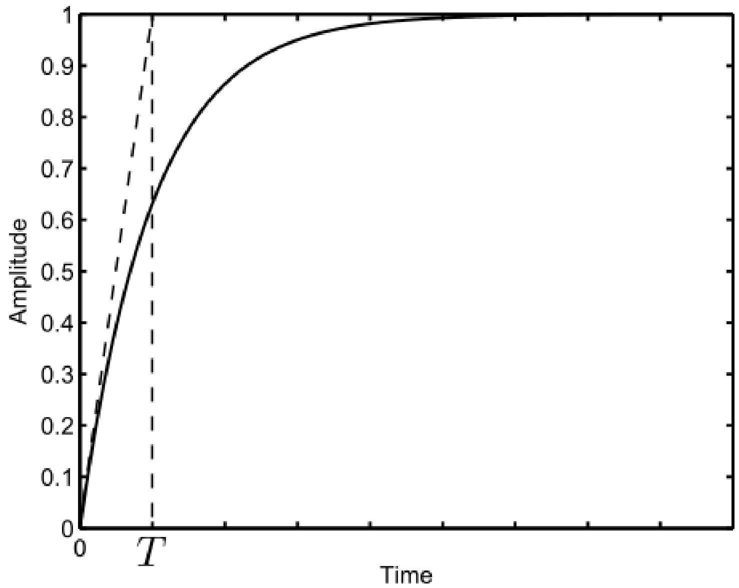
Consider a first order system (low-pass filter)

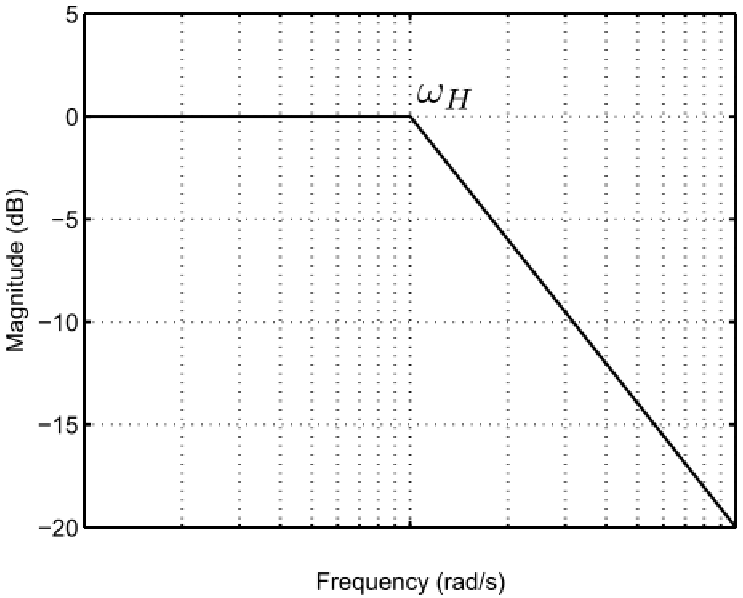
$$H(s) = \frac{1}{1+sT} \quad T > 0$$





Loop transient performance analysis (III)





The speed of the step response increases as the time constant T decreases.

The speed of the step response increases as cutoff frequency $\omega_H = 1/T$ increases.

Consider now the transfer function from the set point to the controlled variable of a closed-loop system

$$\frac{Y(s)}{Y^o(s)} = F(s)$$

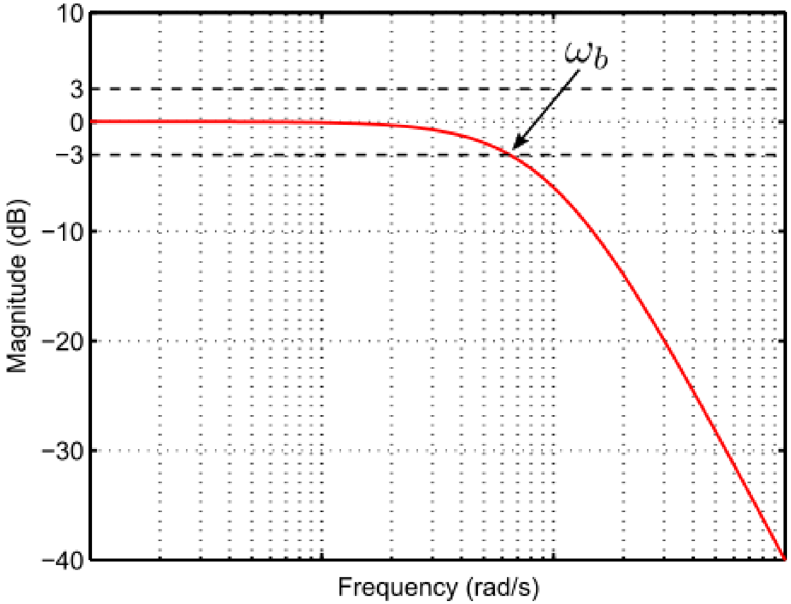
We can assume, for this transfer function, the behavior of a low-pass filter.

As you know, if

$$|F(j\omega)|_{dB} < 3\,dB \quad \forall\omega$$

we can define the cutoff frequency

$$\{\omega : |F(j\omega)|_{dB} > -3\,dB\} = [0, \omega_b]$$



The cutoff frequency ω_b can thus reveal the closed-loop system response time: increasing ω_b the response time decreases.

We discovered a relation between the response time and the cutoff frequency of the closed-loop transfer function $F(s)$.

Can we now relate the response time to a feature of the loop transfer function $L(s)$?

Consider again the closed-loop transfer function

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

and its frequency response

$$F(j\omega) = \frac{L(j\omega)}{1 + L(j\omega)}$$

Let's introduce now the following approximation

$$|F(j\omega)| = \frac{|L(j\omega)|}{|1 + L(j\omega)|} \approx \begin{cases} 1 & \forall \omega : |L(j\omega)| \gg 1 \\ |L(j\omega)| & \forall \omega : |L(j\omega)| \ll 1 \end{cases}$$

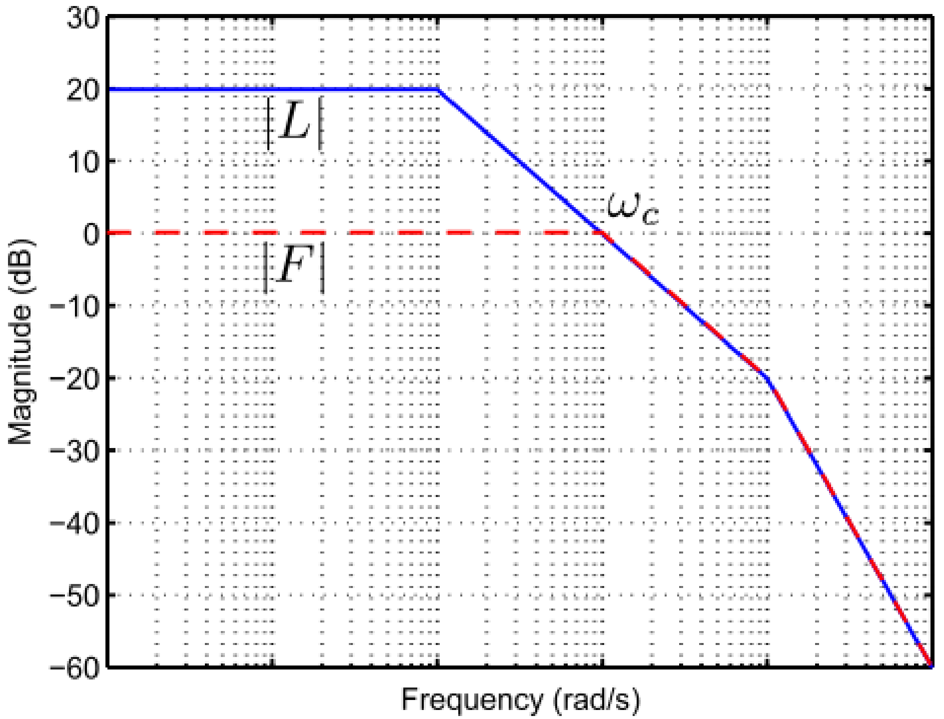
If the assumptions of the Bode criterion hold we have

$$|F(j\omega)| \approx \begin{cases} 1 & \omega \ll \omega_c \\ |L(j\omega)| & \omega \gg \omega_c \end{cases}$$

Is this approximation reliable?

Under witch assumptions?

Is the cutoff frequency well approximated by the crossover frequency?



Loop transient performance analysis (VII)

In order to answer to the previous doubts, we will study the behavior of the closed-loop system transfer function around the crossover frequency

$$|F(j\omega_c)| = \frac{|L(j\omega_c)|}{|1+L(j\omega_c)|} = \frac{1}{|1+e^{j\varphi_c}|} = \frac{1}{|1+\cos(\varphi_c)+j\sin(\varphi_c)|}$$

$$= \frac{1}{\sqrt{1+\cos^2(\varphi_c)+2\cos(\varphi_c)+\sin^2(\varphi_c)}}$$

$$= \frac{1}{\sqrt{2(1+\cos(\varphi_c))}} = \frac{1}{\sqrt{2(1-\cos(\varphi_m))}} = \frac{1}{2\sin\left(\frac{\varphi_m}{2}\right)}$$

If $\varphi_m = 90°$

$$|F(j\omega_c)| = \frac{1}{\sqrt{2}} \quad \Rightarrow \quad |F(j\omega_c)|_{dB} = -3\,dB$$

Cutoff frequency is well-approximated by crossover frequency

and if $\varphi_m > 60°$

$$|F(j\omega_c)| < 1 \quad \Rightarrow \quad |F(j\omega_c)|_{dB} < 0$$

Cutoff frequency and crossover frequency have the same value

We can thus conclude that the closed-loop system can be approximated as:

- if $\varphi_m > 50° \div 60°$

$$F(s) \approx \frac{1}{1 + \frac{s}{\omega_c}} \qquad \Rightarrow \qquad T_{a1} \approx \frac{4.6}{\omega_c}$$

- if $\varphi_m < 30° \div 40°$

$$F(s) \approx \frac{\omega_c^2}{s^2 + 2\xi\omega_c s + \omega_c^2} \qquad \Rightarrow \qquad T_{a1} \approx \frac{4.6}{\xi\omega_c}$$

Loop transient performance analysis (IX)

If the closed-loop system is approximated by a second order transfer function

$$F(s) \approx \frac{\omega_c^2}{s^2 + 2\xi\,\omega_c s + \omega_c^2}$$

we still need to compute the damping factor.

We know that the magnitude of the second order approximation at the crossover frequency is equal to

$$|F(j\omega_c)| = \frac{1}{2\xi}$$

and the magnitude of the closed-loop transfer function at the crossover frequency, instead, is equal to

$$|F(j\omega_c)| = \frac{1}{2\sin\left(\frac{\varphi_m}{2}\right)}$$

merging now these two relations we obtain

Loop transient performance analysis (X)

$$|F(j\omega_c)| = \frac{1}{2\xi} = \frac{1}{2\sin\left(\frac{\varphi_m}{2}\right)} \quad \Rightarrow \quad \xi = \sin\left(\frac{\varphi_m}{2}\right) \approx \frac{\varphi_m}{100}$$

where the last approximation holds only if φ_m is expressed in degrees.

We can conclude that the closed-loop system

- response time
- oscillation damping

are related to the

- crossover frequency ω_c
- phase margin φ_m

respectively (i.e., two features of the loop transfer function).

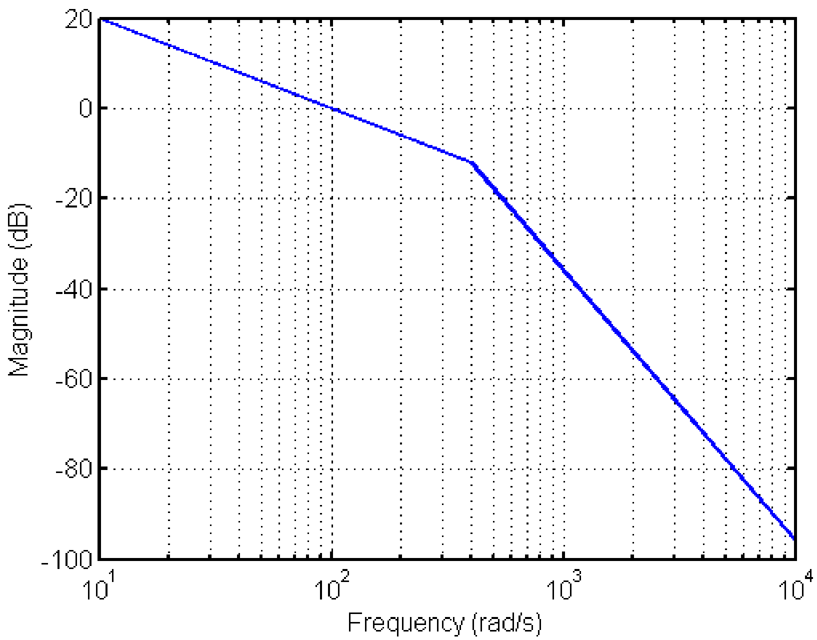
Loop transient performance analysis – Example

$$L(s) = \frac{100}{s(1+0.0025s)^2}$$

We would like to sketch the step response of the closed-loop system

- we first plot the asymptotic magnitude Bode plot of $L(s)$
- we compute $\omega_c \approx 100\ rad/s$
- we compute φ_m

$$\varphi_m = 180^\circ - \left| -90^\circ - 2\arctan\left(\frac{100}{400}\right) \right|$$
$$= 180^\circ - |-90^\circ - 2 \cdot 14^\circ| = 62^\circ$$



As $\varphi_m > 60°$ we can approximate the closed-loop system with a first-order transfer function

$$F(s) \approx \frac{1}{1+0.01s}$$

Loop transient performance analysis – Example

$$L(s) = \frac{100}{s(1+0.0025s)^2}$$

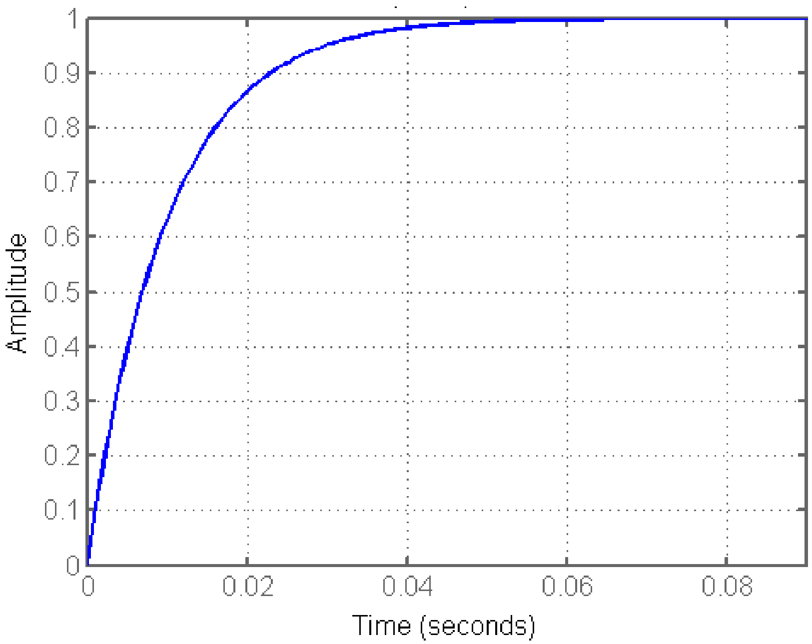
We would like to sketch the step response of the closed-loop system

- we first plot the asymptotic magnitude Bode plot of $L(s)$
- we compute $\omega_c \approx 100\ rad/s$
- we compute φ_m

$$\varphi_m = 180° - \left| -90° - 2\arctan\left(\frac{100}{400}\right) \right|$$
$$= 180° - |-90° - 2 \cdot 14°| = 62°$$

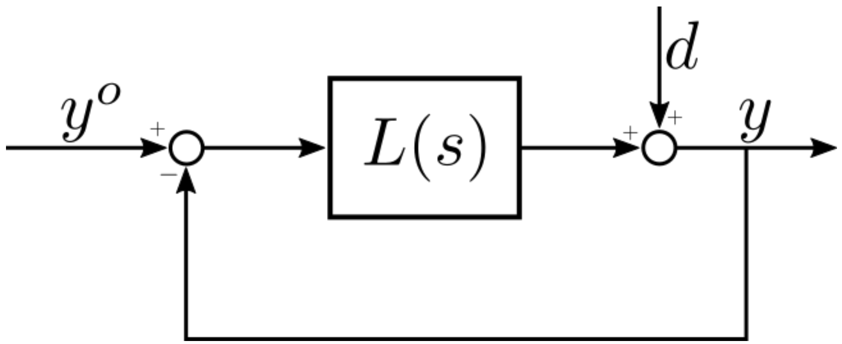
As $\varphi_m > 60°$ we can approximate the closed-loop system with a first-order transfer function

$$F(s) \approx \frac{1}{1+0.01s}$$



We now focus on disturbance rejection and control effort mitigation.

Let's first analyze the effect of a load disturbance d



The transfer function from the disturbance d to the controlled variable y is

$$\frac{Y(s)}{D(s)} = S(s) = \frac{1}{1+L(s)}$$

Sensitivity function

The magnitude of the sensitivity function can be approximated by

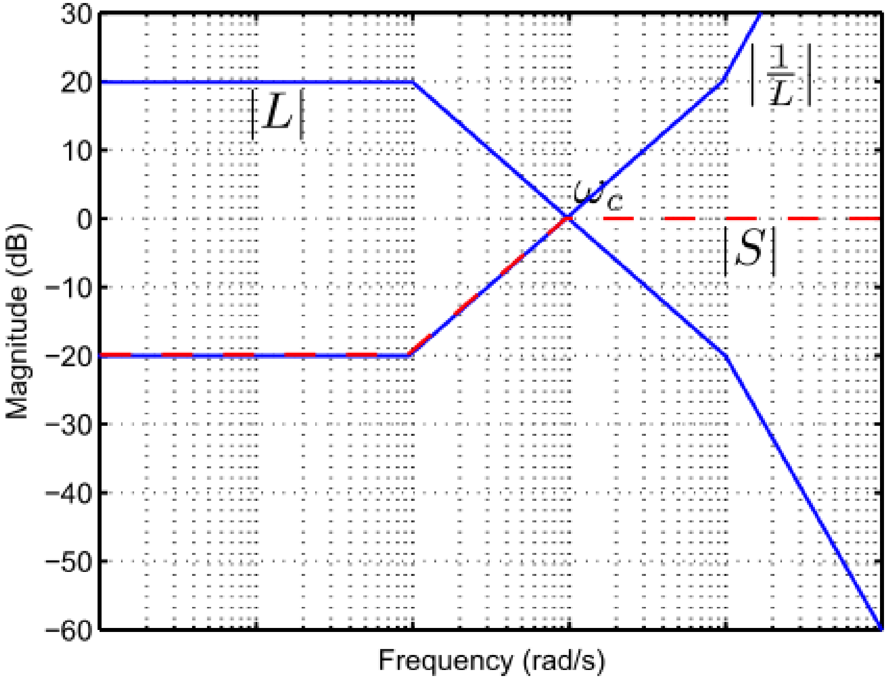
$$|S(j\omega)| = \frac{1}{|1+L(j\omega)|} \approx \begin{cases} \frac{1}{|L(j\omega)|} & \forall\omega : |L(j\omega)| \gg 1 \\ 1 & \forall\omega : |L(j\omega)| \ll 1 \end{cases}$$

Loop transient performance analysis (XII)

If the assumptions of Bode criterion hold we have

$$|S(j\omega)| = \frac{1}{|1+L(j\omega)|} \approx \begin{cases} \frac{1}{|L(j\omega)|} & \omega \ll \omega_c \\ 1 & \omega \gg \omega_c \end{cases}$$

Disturbance harmonics whose frequency is less than the crossover frequency are attenuated

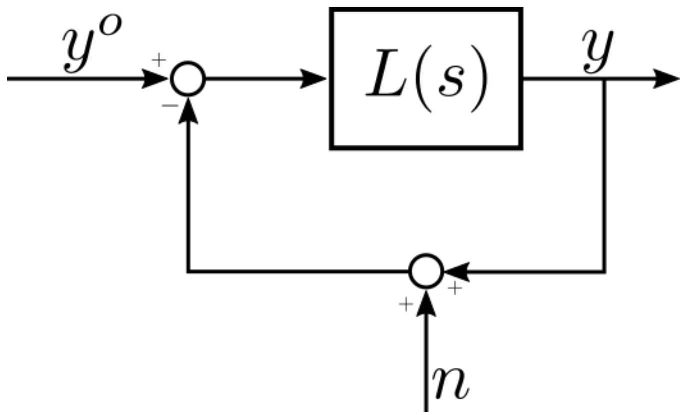


We thus conclude that

- the crossover frequency has to be greater than the highest disturbance harmonic we would like to attenuate
- the higher the magnitude of $L(j\omega)$ for $\omega < \omega_c$, the higher the attenuation

Loop transient performance analysis (XIII)

Consider now the effect of measurement noise n



The transfer function from the noise n to the controlled variable y is

$$\frac{Y(s)}{N(s)} = -F(s) = -\frac{L(s)}{1+L(s)}$$

Complementary sensitivity function

We already now that the magnitude of the complementary sensitivity function can be approximated by

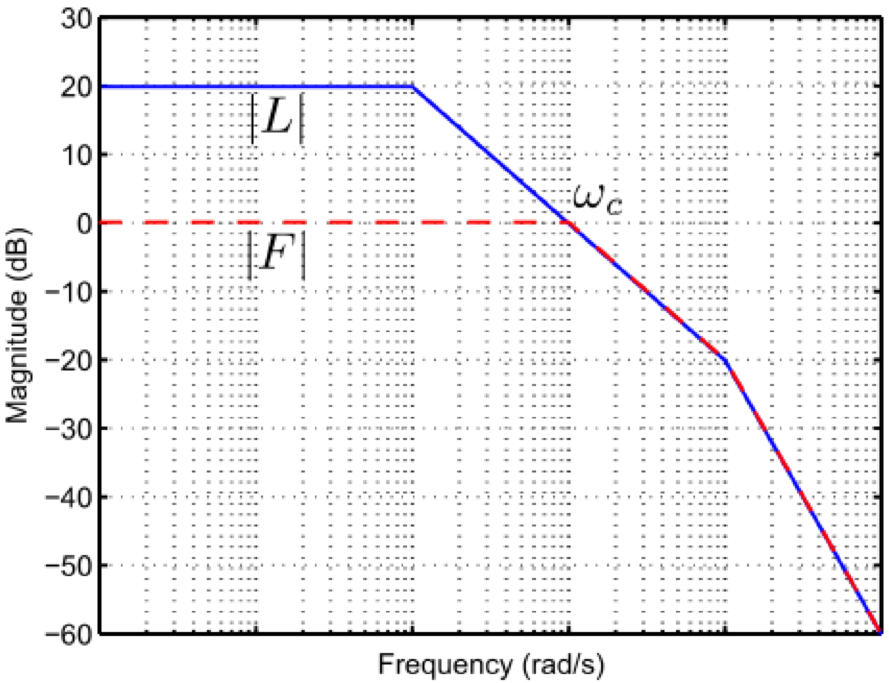
$$|F(j\omega)| = \frac{|L(j\omega)|}{|1+L(j\omega)|} \approx \begin{cases} 1 & \forall\omega : |L(j\omega)| \gg 1 \\ |L(j\omega)| & \forall\omega : |L(j\omega)| \ll 1 \end{cases}$$

Loop transient performance analysis (XIV)

If the assumptions of Bode criterion hold we have

$$|F(j\omega)| \approx \begin{cases} 1 & \omega \ll \omega_c \\ |L(j\omega)| & \omega \gg \omega_c \end{cases}$$

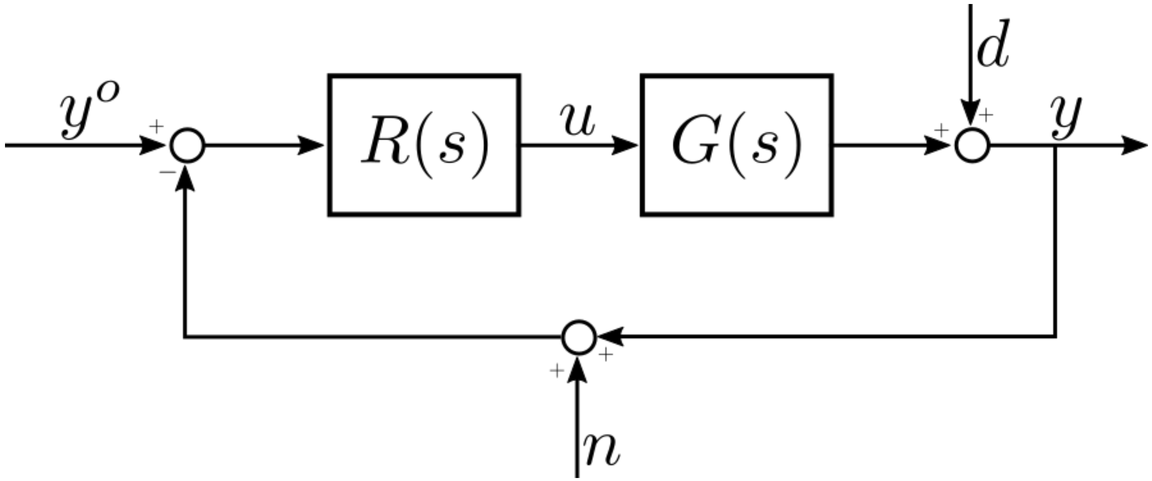
Noise harmonics whose frequency is greater than the crossover frequency are attenuated



We thus conclude that

- the crossover frequency has to be less than the lowest noise harmonic we would like to attenuate
- the lower the magnitude of $L(j\omega)$ for $\omega > \omega_c$, the higher the attenuation

Consider now the control effort mitigation



The transfer function from the set point y^o to the control variable u is

$$\frac{U(s)}{Y^o(s)} = Q(s) = \frac{R(s)}{1+L(s)}$$

Control sensitivity function

The control sensitivity function represents, except for the sign, the transfer function from d to u, and from n to u, as well.

The magnitude of the control sensitivity function should thus be as lower as possible at all frequencies.

Loop transient performance analysis (XVI)

The magnitude of the control sensitivity function can be approximated by

$$|Q(j\omega)| = \frac{|R(j\omega)|}{|1+L(j\omega)|} \approx \begin{cases} \frac{1}{|G(j\omega)|} & \forall\omega : |L(j\omega)| \gg 1 \\ |R(j\omega)| & \forall\omega : |L(j\omega)| \ll 1 \end{cases}$$

If the assumptions of Bode criterion hold we have

$$|Q(j\omega)| = \frac{|R(j\omega)|}{|1+L(j\omega)|} \approx \begin{cases} \frac{1}{|G(j\omega)|} & \omega \ll \omega_c \\ |R(j\omega)| & \omega \gg \omega_c \end{cases}$$

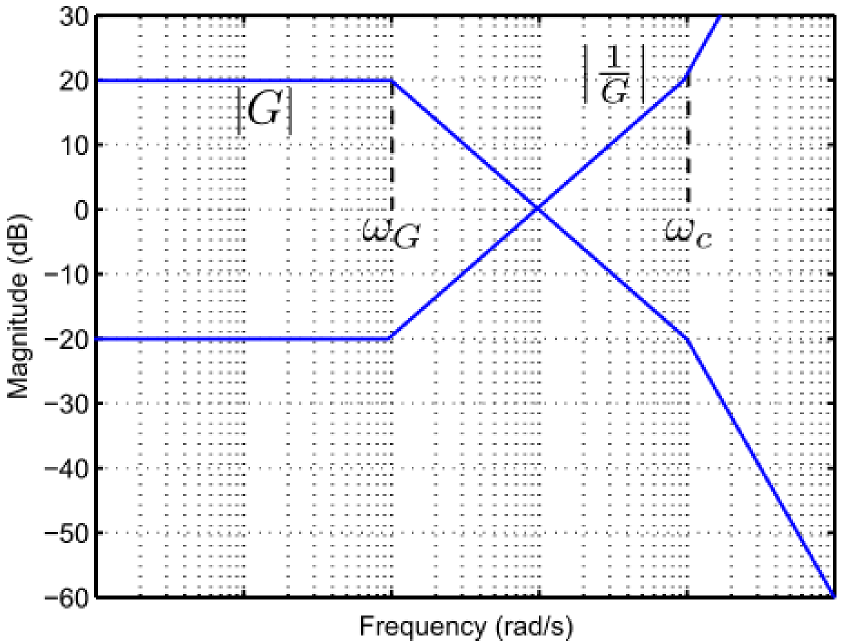
Should be as lower as possible beyond the crossover frequency

Let's assume that $|G|$ can be approximated by a low-pass filter

If $\omega_c \gg \omega_G$ the magnitude of $|1/G|$ can be very high, and the magnitude of $|Q|$ increases.

We thus conclude that

- the crossover frequency should not be too high with respect to the process response time (ω_G)



Loop steady-state performance analysis (I)

We will now study the steady-state behavior of the closed-loop system in terms of the error between the set point and the controlled variable.

We will assume the closed-loop system asymptotically stable (otherwise we cannot have steady-state!).

In the analysis we will exploit the superposition principle (the net response at a given place and time caused by two or more stimuli is the sum of the responses which would have been caused by each stimulus individually), considering the effects of each input individually.

As inputs we will consider steps, ramps and parabolic signals.

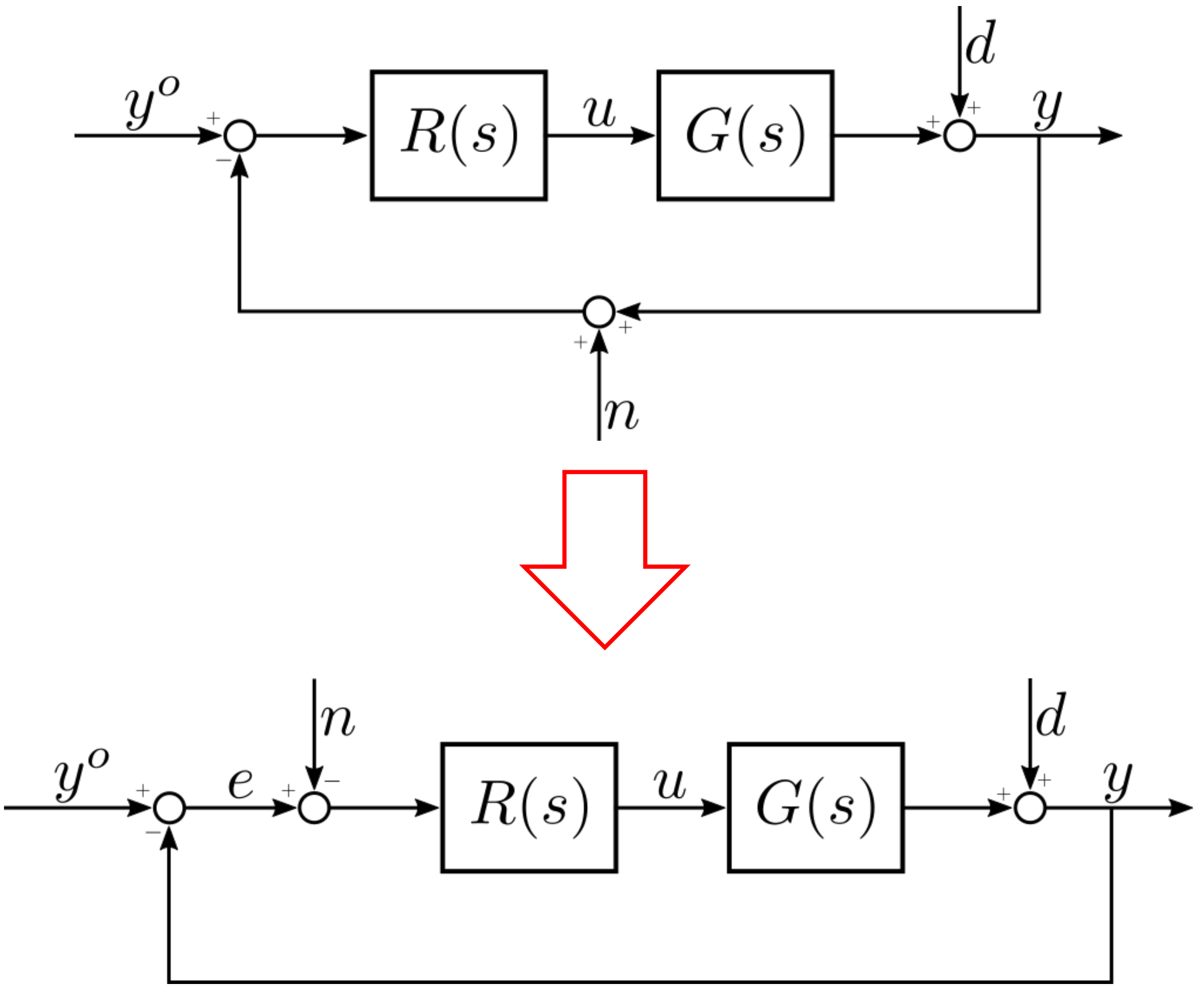
Caveat: you don’t need to memorize any result, you just need to know how to compute transfer functions and how to use the final value theorem!

Loop steady-state performance analysis (II)

Let's first define the error

$$e(t) = y^o(t) - y(t)$$

where is the error in the block diagram?



Loop steady-state performance analysis (III)

Let's start considering the <u>error generated by the set point</u>.

The transfer function from the set point to the error is

$$\frac{E(s)}{Y^o(s)} = \frac{1}{1+L(s)} = S(s)$$

The loop transfer function can be represented as

$$L(s) = \frac{\mu_L}{s^{g_L}} \frac{\prod_i (1+s\tau_i)}{\prod_j (1+sT_j)}$$

and observe that

$$\lim_{s\to 0} L(s) = \lim_{s\to 0} \frac{\mu_L}{s^{g_L}} \frac{\prod_i (1+s\tau_i)}{\prod_j (1+sT_j)} = \lim_{s\to 0} \frac{\mu_L}{s^{g_L}}$$

Assuming that the closed-loop system is asymptotically stable we can now apply the final value theorem to compute the steady-state error e_∞.

Loop steady-state performance analysis (IV)

We have

$$e_\infty = \lim_{t\to\infty} e(t) = \lim_{s\to 0}[sE(s)] = \lim_{s\to 0}\left[s\frac{1}{1+L(s)}Y^o(s)\right]$$

$$= \lim_{s\to 0}\left[s\frac{1}{1+\frac{\mu_L}{s^{g_L}}}Y^o(s)\right] = \lim_{s\to 0}\left[\frac{s^{g_L+1}}{s^{g_L}+\mu_L}Y^o(s)\right]$$

Let's compute now the steady-state error for different set point signals (step, ramp, parabolic signal).

Loop steady-state performance analysis (V)

$y^o(t) = A\,\mathrm{sca}(t)$

$$e_\infty = \lim_{s \to 0}\left[\frac{s^{g_L+1}}{s^{g_L}+\mu_L}\frac{A}{s}\right] = \lim_{s \to 0}\left[A\frac{s^{g_L}}{s^{g_L}+\mu_L}\right] = \begin{cases} A & g_L < 0 \\ \frac{A}{1+\mu_L} & g_L = 0 \\ 0 & g_L \geq 1 \end{cases}$$

$y^o(t) = A\,\mathrm{ram}(t)$

$$e_\infty = \lim_{s \to 0}\left[\frac{s^{g_L+1}}{s^{g_L}+\mu_L}\frac{A}{s^2}\right] = \lim_{s \to 0}\left[A\frac{s^{g_L}}{s^{g_L+1}+s\mu_L}\right] = \begin{cases} \infty & g_L \leq 0 \\ \frac{A}{\mu_L} & g_L = 1 \\ 0 & g_L \geq 2 \end{cases}$$

$y^o(t) = A\,\mathrm{par}(t)$

$$e_\infty = \lim_{s \to 0}\left[\frac{s^{g_L+1}}{s^{g_L}+\mu_L}\frac{A}{s^3}\right] = \lim_{s \to 0}\left[A\frac{s^{g_L}}{s^{g_L+2}+s^2\mu_L}\right] = \begin{cases} \infty & g_L \leq 1 \\ \frac{A}{\mu_L} & g_L = 2 \\ 0 & g_L \geq 3 \end{cases}$$

Worthless!

Loop steady-state performance analysis (V)

$y^o(t) = A\,\mathrm{sca}(t)$

$$e_\infty = \lim_{s\to 0}\left[\frac{s^{g_L+1}}{s^{g_L}+\mu_L}\frac{A}{s}\right] = \lim_{s\to 0}\left[A\frac{s^{g_L}}{s^{g_L}+\mu_L}\right] = \begin{cases} A & g_L < 0 \\ \frac{A}{1+\mu_L} & g_L = 0 \\ 0 & g_L \geq 1 \end{cases}$$

$y^o(t) = A\,\mathrm{ram}(t)$

$$e_\infty = \lim_{s\to 0}\left[\frac{s^{g_L+1}}{s^{g_L}+\mu_L}\frac{A}{s^2}\right] = \lim_{s\to 0}\left[A\frac{s^{g_L}}{s^{g_L+1}+s\mu_L}\right] = \begin{cases} \infty & g_L \leq 0 \\ \frac{A}{\mu_L} & g_L = 1 \\ 0 & g_L \geq 2 \end{cases}$$

$y^o(t) = A\,\mathrm{par}(t)$

$$e_\infty = \lim_{s\to 0}\left[\frac{s^{g_L+1}}{s^{g_L}+\mu_L}\frac{A}{s^3}\right] = \lim_{s\to 0}\left[A\frac{s^{g_L}}{s^{g_L+2}+s^2\mu_L}\right] = \begin{cases} \infty & g_L \leq 1 \\ \frac{A}{\mu_L} & g_L = 2 \\ 0 & g_L \geq 3 \end{cases}$$

Increasing μ_L the error decreases

$y^o(t) = A\,\text{sca}(t)$

$$e_\infty = \lim_{s\to 0}\left[\frac{s^{g_L+1}}{s^{g_L}+\mu_L}\frac{A}{s}\right] = \lim_{s\to 0}\left[A\frac{s^{g_L}}{s^{g_L}+\mu_L}\right] = \begin{cases} A & g_L < 0 \\ \frac{A}{1+\mu_L} & g_L = 0 \\ 0 & g_L \geq 1 \end{cases}$$

$y^o(t) = A\,\text{ram}(t)$

$$e_\infty = \lim_{s\to 0}\left[\frac{s^{g_L+1}}{s^{g_L}+\mu_L}\frac{A}{s^2}\right] = \lim_{s\to 0}\left[A\frac{s^{g_L}}{s^{g_L+1}+s\mu_L}\right] = \begin{cases} \infty & g_L \leq 0 \\ \frac{A}{\mu_L} & g_L = 1 \\ 0 & g_L \geq 2 \end{cases}$$

$y^o(t) = A\,\text{par}(t)$

$$e_\infty = \lim_{s\to 0}\left[\frac{s^{g_L+1}}{s^{g_L}+\mu_L}\frac{A}{s^3}\right] = \lim_{s\to 0}\left[A\frac{s^{g_L}}{s^{g_L+2}+s^2\mu_L}\right] = \begin{cases} \infty & g_L \leq 1 \\ \frac{A}{\mu_L} & g_L = 2 \\ 0 & g_L \geq 3 \end{cases}$$

Zero steady-state error

Consider now the error generated by the disturbance d.

The transfer function from the disturbance to the error is

$$\frac{E(s)}{D(s)} = -\frac{1}{1+L(s)} = -S(s)$$

It's the same transfer function, except for the sign, we have just analyzed, all the previous results hold!

Finally, consider the error generated by the noise n.

The transfer function from noise to error is

$$\frac{E(s)}{N(s)} = \frac{L(s)}{1+L(s)} = F(s)$$

and we have

$$e_\infty = \lim_{t\to\infty} e(t) = \lim_{s\to 0}\left[sE(s)\right] = \lim_{s\to 0}\left[s\frac{L(s)}{1+L(s)}N(s)\right]$$

$$= \lim_{s\to 0}\left[s\frac{\frac{\mu_L}{s^{g_L}}}{1+\frac{\mu_L}{s^{g_L}}}N(s)\right] = \lim_{s\to 0}\left[\frac{\mu_L s}{s^{g_L}+\mu_L}N(s)\right]$$

Loop steady-state performance analysis (VII)

$n(t) = A\ \mathrm{sca}(t)$

$$e_\infty = \lim_{s\to 0}\left[\frac{\mu_L s}{s^{g_L}+\mu_L}\frac{A}{s}\right] = \lim_{s\to 0}\left[A\frac{\mu_L}{s^{g_L}+\mu_L}\right] = \begin{cases} A\frac{\mu_L}{1+\mu_L} & g_L = 0 \\ A & g_L \geq 1 \end{cases}$$

$n(t) = A\ \mathrm{ram}(t)$

$$e_\infty = \lim_{s\to 0}\left[\frac{\mu_L s}{s^{g_L}+\mu_L}\frac{A}{s^2}\right] = \lim_{s\to 0}\left[A\frac{\mu_L}{s^{g_L+1}+s\mu_L}\right] = \infty \quad \forall g_L \geq 0$$

$n(t) = A\ \mathrm{par}(t)$

$$e_\infty = \lim_{s\to 0}\left[\frac{\mu_L s}{s^{g_L}+\mu_L}\frac{A}{s^3}\right] = \lim_{s\to 0}\left[A\frac{\mu_L}{s^{g_L+2}+s^2\mu_L}\right] = \infty \quad \forall g_L \geq 0$$

If the sensor is affected by bias, the steady-state error cannot be less than sensor error

Control system design (I)

We now have all the tools required to address the design of the control system.

The design will be based on the Bode stability criterion, as a consequence the loop transfer function $L(s)$ has to satisfy the assumptions of the Bode criterion.

Caveat: The design method we are introducing cannot be used with processes with poles in the open right half plane.

First of all, we can reformulate the controller requirements as follows:

- asymptotic stability $\varphi_m > 0$
- robust stability and oscillation damping $\varphi_m > \bar{\varphi}_m$
- response time $\omega_c \geq \bar{\omega}_c$
- steady-state performance $|e_\infty| \leq \bar{e}_\infty$
- further requirements …

transient requirements

steady-state requirements

Control system design (II)

The controller design problem is usually divided into two steps:

- steady-state design
    - assumes that the closed-loop system is asymptotically stable
    - considers only the steady-state performance
- transient design
    - considers the remaining specifications (requirements on crossover frequency, phase margin, etc.)

Following this methodology, we can factor the controller transfer function as

$$R(s) = R_1(s)R_2(s) \quad \Rightarrow \quad R_1(s) = \frac{\mu_R}{s^{g_R}} \quad R_2(s) = \frac{\prod_i (1 + s\tau_i)}{\prod_j (1 + sT_j)}$$

We can observe that

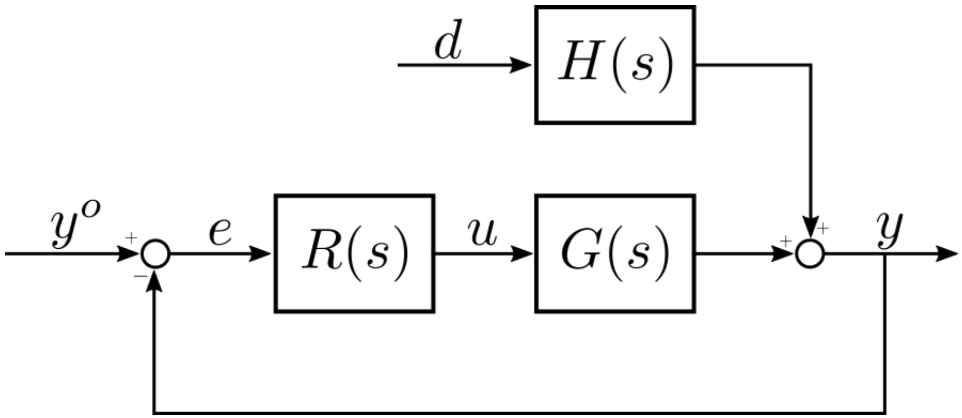
- $R_1(s)$, determines steady-state performance
- $R_2(0) = 1$, and $R_2(s)$ does not contribute to steady-state performance

Design rules:

- steady-state design ($R_1(s)$)
  - we choose the minimum value for g_R that fits with the steady-state requirements
  - once g_R has been selected, we choose the minimum value for μ_R that fits with the steady-state requirements
  - if we can satisfy steady-state requirements without choosing a specific value for μ_R, the selection of this parameter is done during transient design
- transient design ($R_2(s)$)
  - we select zero/pole time constants with a graphical method that aims at shaping the loop transfer function Bode plot so as to satisfy transient requirements

Let's see the controller design procedure with an example.

Control system design – Example (I)



$$G(s) = \frac{50}{(1+0.1s)(1+s)(1+10s)} \qquad H(s) = \frac{5}{1+0.01s}$$

Requirements:

- $|e_\infty| \leq 0.025$ when $y^o(t) = 10\,\mathrm{sca}(t)$ and $d(t) = \pm\mathrm{sca}(t)$
- $\omega_c \geq 1\, rad/s$
- $\varphi_m \geq 60°$

<u>Steady-state design</u>

$$e_{\infty_{y^o}} = \lim_{s\to 0}\left[sE_{y^o}(s)\right] = \lim_{s\to 0}\left[s\frac{1}{1+L(s)}Y^o(s)\right] = \lim_{s\to 0}\left[s\frac{1}{1+\frac{50\mu_R}{s^{g_R}}}\frac{10}{s}\right]$$

$$= \lim_{s\to 0}\left[\frac{10s^{g_R}}{s^{g_R}+50\mu_R}\right] = \begin{cases}\frac{10}{1+50\mu_R} & g_R = 0\\ 0 & g_R \geq 1\end{cases}$$

$$e_{\infty_d} = \lim_{s\to 0}\left[sE_d(s)\right] = \lim_{s\to 0}\left[s\frac{-H(s)}{1+L(s)}D(s)\right] = \lim_{s\to 0}\left[s\frac{-5}{1+\frac{50\mu_R}{s^{g_R}}}\frac{\pm 1}{s}\right]$$

$$= \mp\lim_{s\to 0}\left[\frac{5s^{g_R}}{s^{g_R}+50\mu_R}\right] = \begin{cases}\mp\frac{5}{1+50\mu_R} & g_R = 0\\ 0 & g_R \geq 1\end{cases}$$

Steady-state requirement: $|e_\infty| \leq 0.025$

We can select $g_R = 0$

We have

$$|e_\infty| = \left|e_{\infty_{yo}} + e_{\infty_d}\right| \leq \left|e_{\infty_{yo}}\right| + \left|e_{\infty_d}\right| = \frac{10}{1+50\mu_R} + \frac{5}{1+50\mu_R} = \frac{15}{1+50\mu_R}$$

and considering the requirements

$$\frac{15}{1+50\mu_R} \leq 0.025 \quad \Rightarrow \quad \mu_R \geq \frac{15-0.025}{1.25} \approx 12$$

The steady-state design can be thus concluded with

$$\mu_R = 20 \quad \Rightarrow \quad R_1(s) = \frac{\mu_R}{s^{g_R}} = 20$$

Control system design – Example (IV)

Transient design

The loop transfer function can be rewritten as

$$L(s) = R_1(s)R_2(s)G(s) = R_2(s)L_1(s)$$

where

$$L_1(s) = R_1(s)G(s) = \frac{1000}{(1+0.1s)(1+s)(1+10s)}$$

For the first trial let's assume that

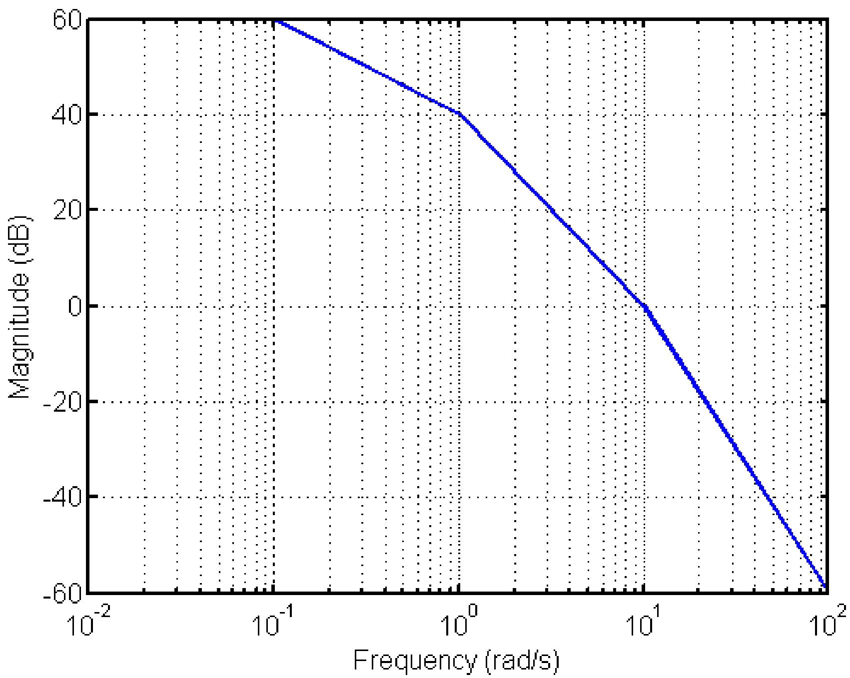
$$R_2(s) = 1 \quad \Rightarrow \quad L(s) = L_1(s)$$

From Bode plot we have

$$\omega_c \gg 1$$

$$\varphi_m < 0$$

We need a second trial with a dynamic regulator.



Remember that, if $L(s)$ is a minimum phase system, crossing the $0\ dB$-axis with slope -1 and having no further zeros/poles around the crossover frequency, the phase margin is close to $90°$.

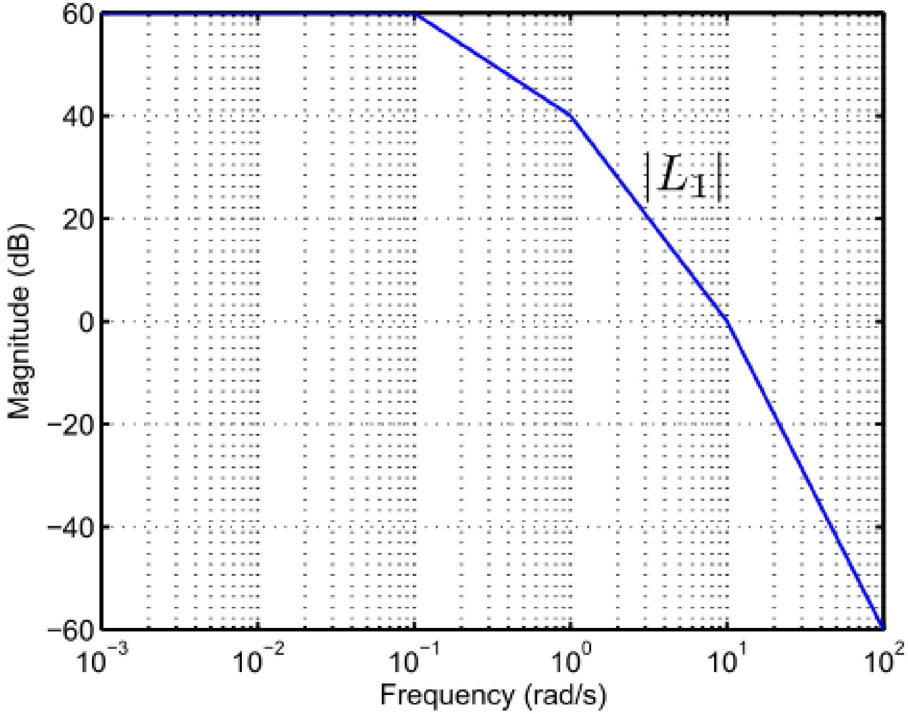
Let's proceed with the design shaping $L(s)$ in such a way that it crosses the $0\ dB$-axis with slope -1 and…

We can proceed as follows:

1. divide the magnitude Bode plot in three parts: low frequency, high frequency, around the crossover frequency
2. start shaping the magnitude Bode plot of $L(s)$ from the part around the crossover frequency

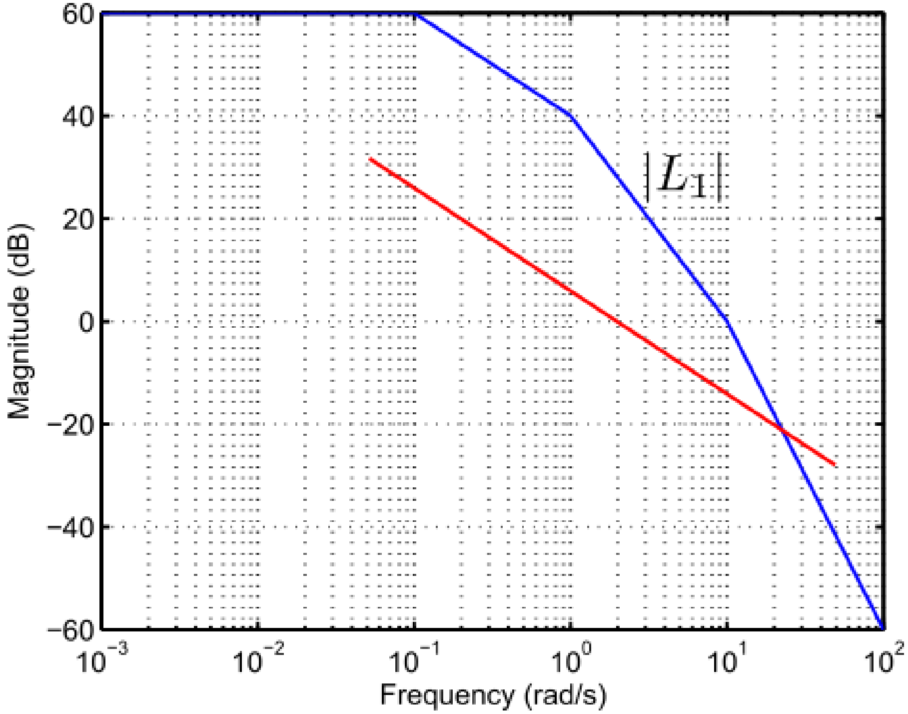
We will now formalize a set of design rules.

Design rules:



Design rules:

1. draw a line, having slope -1, crossing the $0\ dB$-axis at a crossover frequency greater or equal to the minimum value that satisfies the requirements

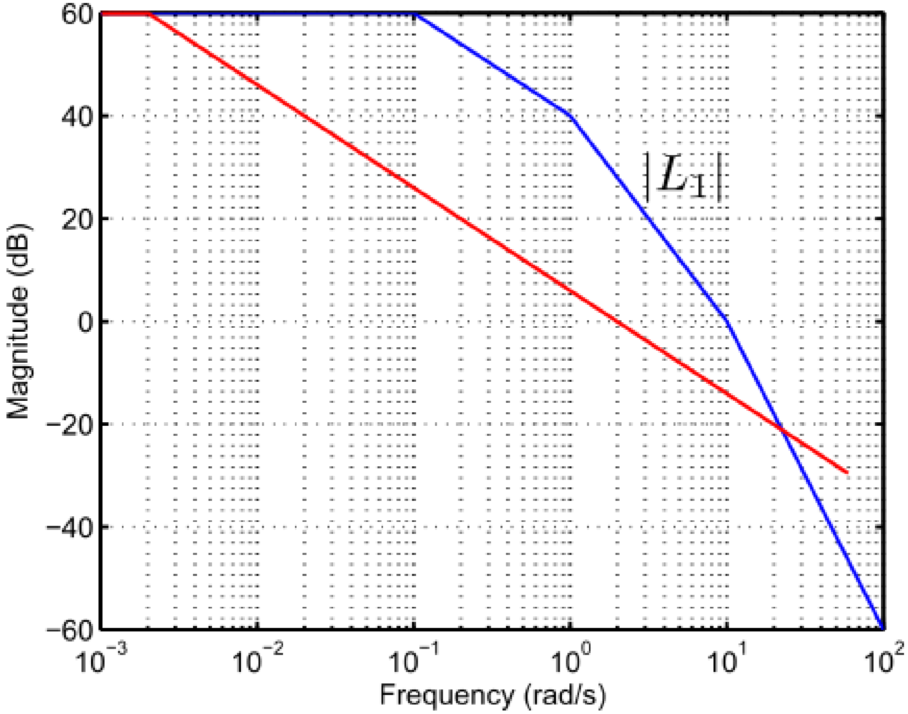


Design rules:

2. low-frequency part
   - $|L|$ and $|L_1|$ must have the same slope

     otherwise you are modifying the controller type selected in steady-state design
   - if the steady-state design set a constraint on the controller gain, then $|L| \geq |L_1|$

     otherwise you are modifying the controller gain selected in steady-state design

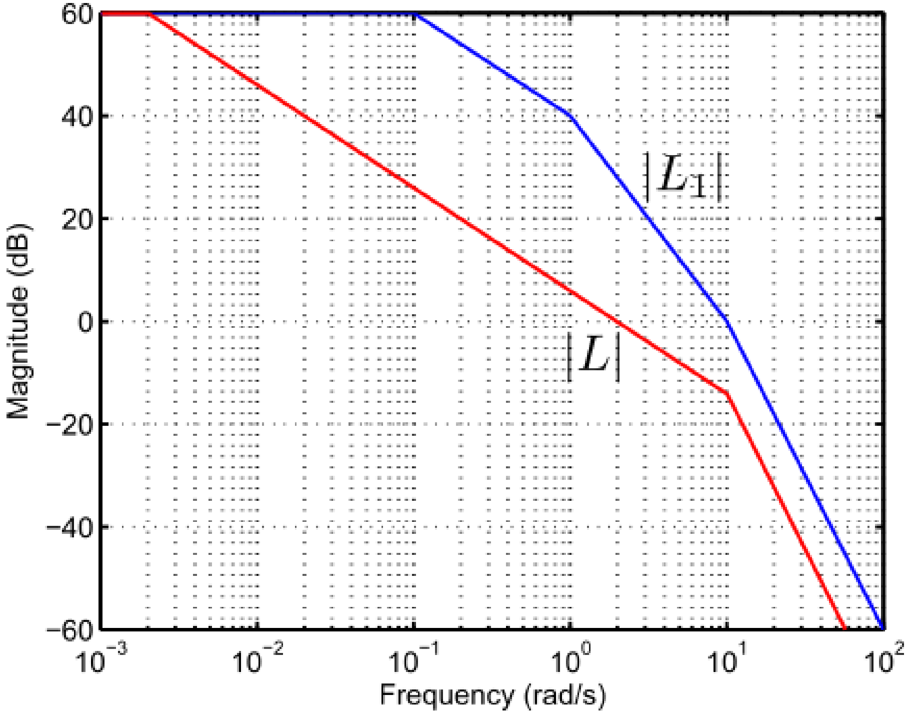


Design rules:

2. high-frequency part
   - the absolute value of the slope of $|L|$ plot must be greater or equal to the one of $|L_1|$ plot

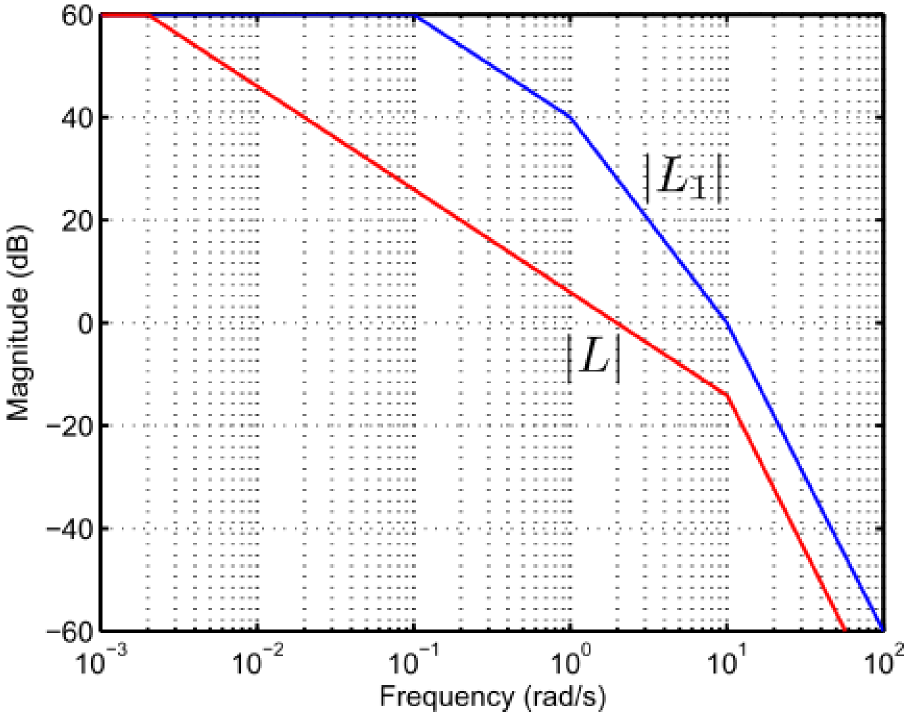
     otherwise the regulator could have more zeros than poles, being an acausal system
   - $|L| \leq |L_1|$

     to guarantee control effort mitigation



From the magnitude Bode plot we can now extract the expression of the loop transfer function

$$L(s) = \frac{1000}{\left(1+\frac{s}{0.002}\right)\left(1+\frac{s}{10}\right)^2} = \frac{1000}{(1+500s)(1+0.1s)^2}$$



and compute the phase margin in order to verify the transient requirements

$$\omega_c \approx 2\, rad/s$$

$$\varphi_m = 180^\circ - \left|-\arctan(2/0.002) - 2\arctan(2/10)\right| = 180^\circ - \left|-90^\circ - 2\cdot 11^\circ\right| = 68^\circ$$

All the requirements have been fulfilled.

We can now compute $R_2(s)$

$$R_2(s) = \frac{L(s)}{L_1(s)} = \frac{1000}{(1+500s)(1+0.1s)^2} \frac{(1+0.1s)(1+s)(1+10s)}{1000}$$
$$= \frac{(1+s)(1+10s)}{(1+500s)(1+0.1s)}$$

and, finally, the expression of the regulator transfer function

$$R(s) = R_1(s)R_2(s) = 20\frac{(1+s)(1+10s)}{(1+500s)(1+0.1s)}$$

Control system design (IV)

Among the further requirements we previously mentioned, there are the disturbance attenuation requirements.

The feedback is required to attenuate load disturbances and measurement noises characterized by sinusoidal signals or any other signal expressed as Fourier series or Fourier integral.

These requirements give rise to further constraints on the frequency response of the loop transfer function we shape in the transient design.

We will now study how to formalize these constraints and how to take them into account during transient design.

Given a load disturbance $d(t)$, whose non negligible harmonics span the range $[0, \omega_{max}]$, the control system should attenuate this disturbance on the controlled variable by a factor A ($A > 1$).

First, the transfer function from d to y is given by

$$\frac{Y(s)}{D(s)} = S(s) = \frac{1}{1+L(s)}$$

and the corresponding frequency response by

$$Y(\mathrm{j}\omega) = \frac{1}{1+L(\mathrm{j}\omega)} D(\mathrm{j}\omega)$$

Assuming the closed-loop system asymptotically stable, we can apply the sinusoidal response theorem. At steady-state, the output generated by the load disturbance has amplitude

$$\left| \frac{1}{1+L(\mathrm{j}\omega)} \right|_{\omega<\omega_{max}} D$$

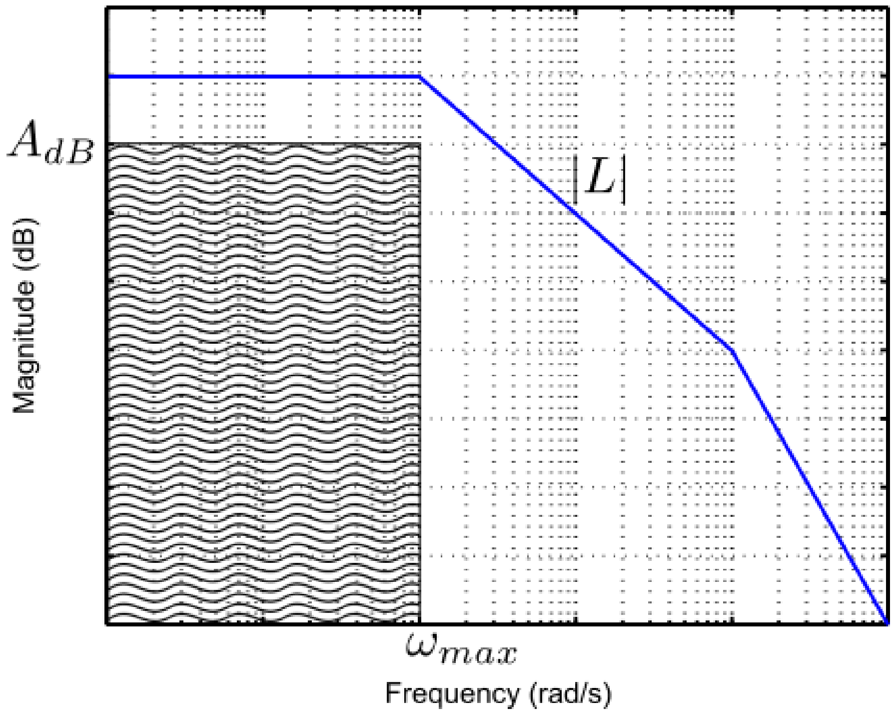
Remembering that the attenuation factor is related to the ratio between the output and input amplitudes, we obtain the following constraint

$$\left|\frac{1}{1+L(j\omega)}\right|_{\omega<\omega_{max}} < \frac{1}{A}$$

assuming that $\omega_{max} < \omega_c$ the constraint can be simplified as

$$|L(j\omega)|_{\omega<\omega_{max}} > A$$

This constraint is equivalent to a forbidden region in the magnitude Bode plot of the loop transfer function.



Given a measurement noise $n(t)$, whose non negligible harmonics span the range $[\omega_{min}, +\infty]$, the control system should attenuate this disturbance on the controlled variable by a factor A ($A > 1$).

First, the transfer function from n to y is given by

$$\frac{Y(s)}{N(s)} = -F(s) = -\frac{L(s)}{1+L(s)}$$

and the corresponding frequency response by

$$Y(j\omega) = -\frac{L(j\omega)}{1+L(j\omega)}N(j\omega)$$

Assuming the closed-loop system asymptotically stable, we can apply the sinusoidal response theorem. At steady-state, the output generated by the measurement noise has amplitude

$$\left|\frac{L(j\omega)}{1+L(j\omega)}\right|_{\omega>\omega_{min}} N$$

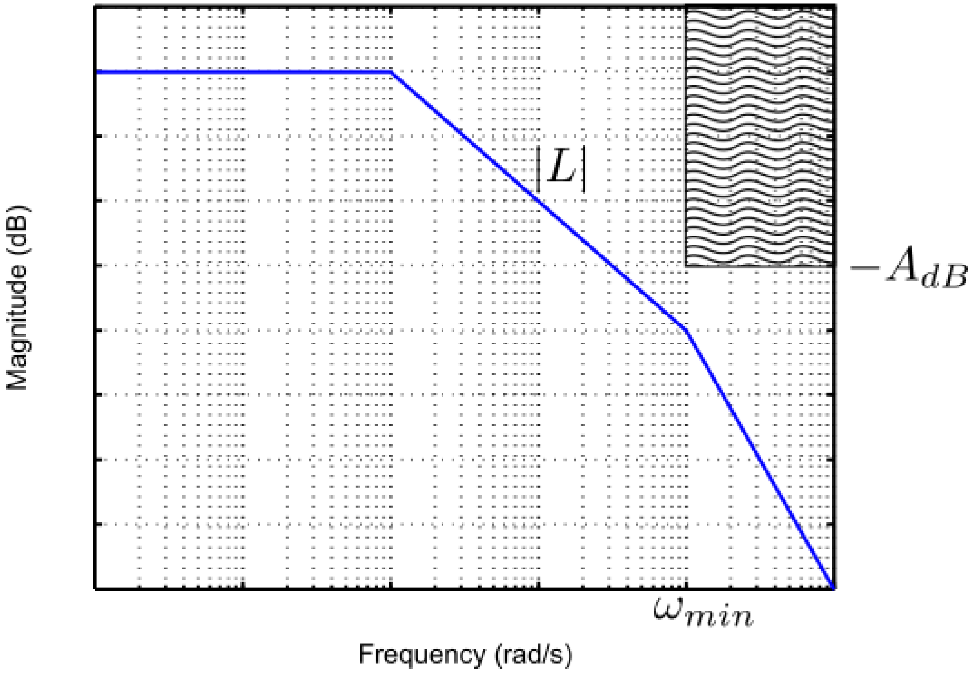
Remembering that the attenuation factor is related to the ratio between the output and input amplitudes, we obtain the following constraint

$$\left| \frac{L(j\omega)}{1+L(j\omega)} \right|_{\omega>\omega_{min}} < \frac{1}{A}$$

assuming that $\omega_{min} > \omega_c$ the constraint can be simplified as

$$|L(j\omega)|_{\omega>\omega_{min}} < \frac{1}{A}$$

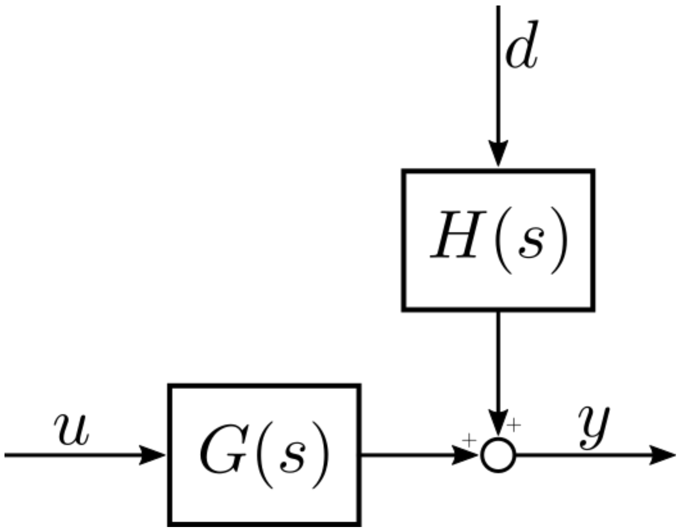
This constraint is equivalent to a forbidden region in the magnitude Bode plot of the loop transfer function.

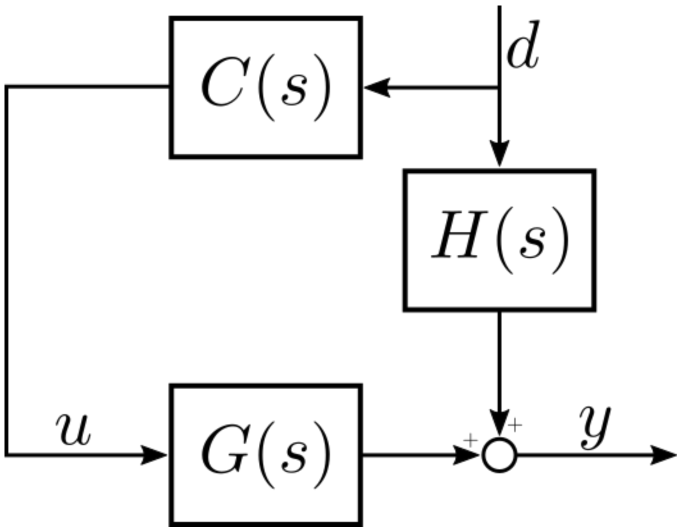


Feedforward compensation (I)

Though the feedback controller is able to attenuate load disturbances and measurement noise, when disturbances are measurable we can try to compensate them.

Disturbance compensation has the advantage, with respect to the attenuation action exerted by feedback, that the compensator exploits the measured disturbance to act directly on the control variable, while feedback has to wait the effect of the disturbance on the controlled variable.





We design the compensator in such a way that it cancels out the effect of the disturbance on the controlled variable. This is equivalent to make the transfer function from d to y identically zero

$$\frac{Y(s)}{D(S)} = H(s) + C(s)G(s) = 0$$

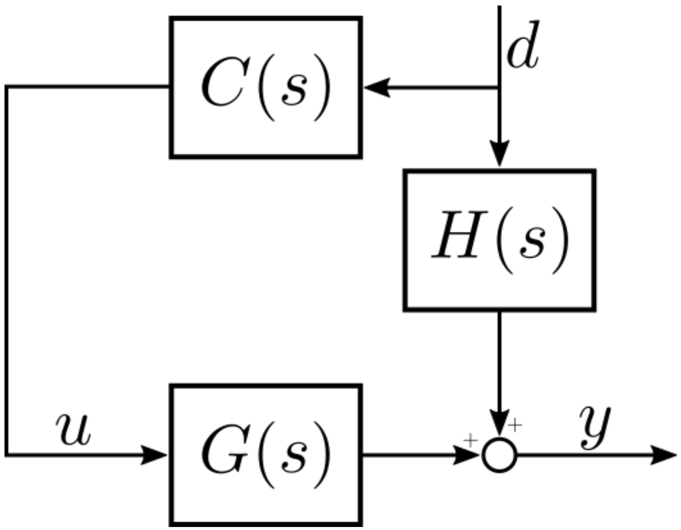
Solving with respect to the compensator transfer function we obtain

$$C(s) = -\frac{H(s)}{G(s)}$$

This relation is a guideline to design the compensator. In fact, in many situations it cannot be directly applied.

For example when:

- $G(s)$ is a non-minimum phase transfer function
- $C(s)$ is an a-causal system



We can consider the following significant situations:

- step disturbance

$$C(s) = \mu_C = -\frac{H(0)}{G(0)}$$

- sinusoidal disturbance at frequency $\bar{\omega}$

$$C(s) \quad : \quad C(j\bar{\omega}) = -\frac{H(j\bar{\omega})}{G(j\bar{\omega})}$$

Design a parametric compensator and select the parameter in such a way that the constraint at $\bar{\omega}$ is satisfied

- disturbance with harmonics in the frequency range $[0, \bar{\omega}]$

$$C(s) \quad : \quad C(j\omega) = -\frac{H(j\omega)}{G(j\omega)} \quad \omega < \bar{\omega}$$

Feedforward compensation (III)

We can consider the following significant situations:

- step disturbance

$$C(s) = \mu_C = -\frac{H(0)}{G(0)}$$

- sinusoidal disturbance at a frequency $\overline{\omega}$

$$C(s) \quad : \quad C(\mathrm{j}\bar{\omega}) = -\frac{H(\mathrm{j}\bar{\omega})}{G(\mathrm{j}\bar{\omega})}$$

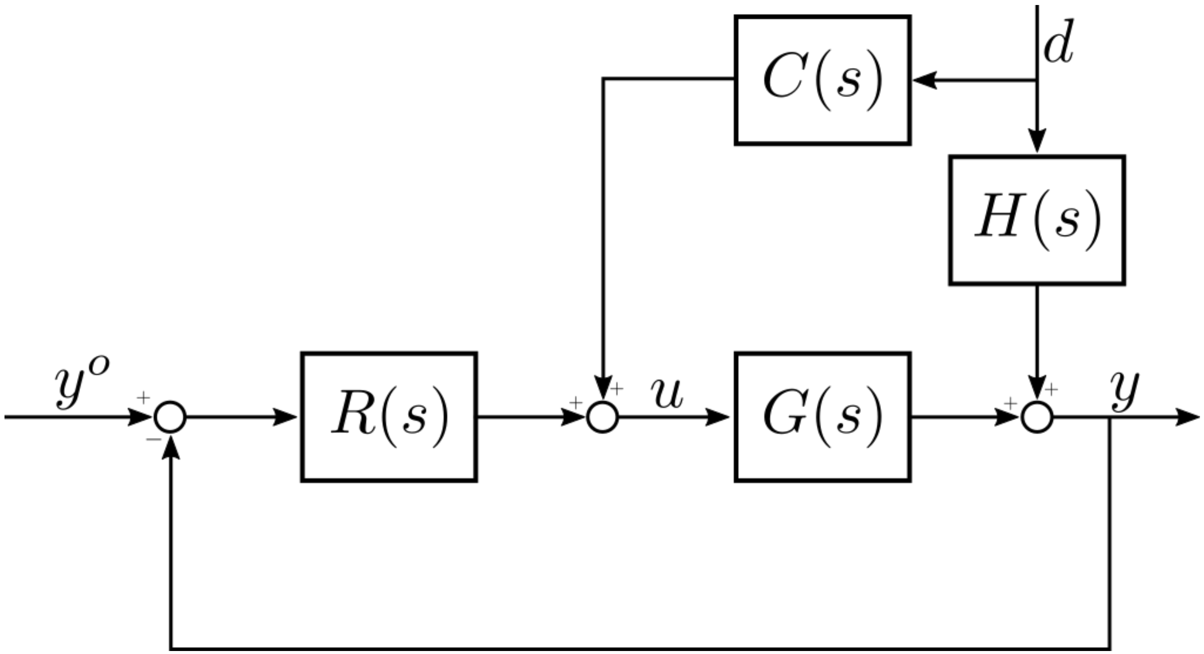
- disturbance with harmonics in the frequency range $[0, \overline{\omega}]$

$$C(s) \quad : \quad C(\mathrm{j}\omega) = -\frac{H(\mathrm{j}\omega)}{G(\mathrm{j}\omega)} \quad \omega < \bar{\omega}$$

Design a compensator that approximates the ideal one until frequency $\overline{\omega}$

Feedforward compensation (IV)

Let's now add the compensator to the standard feedback architecture.



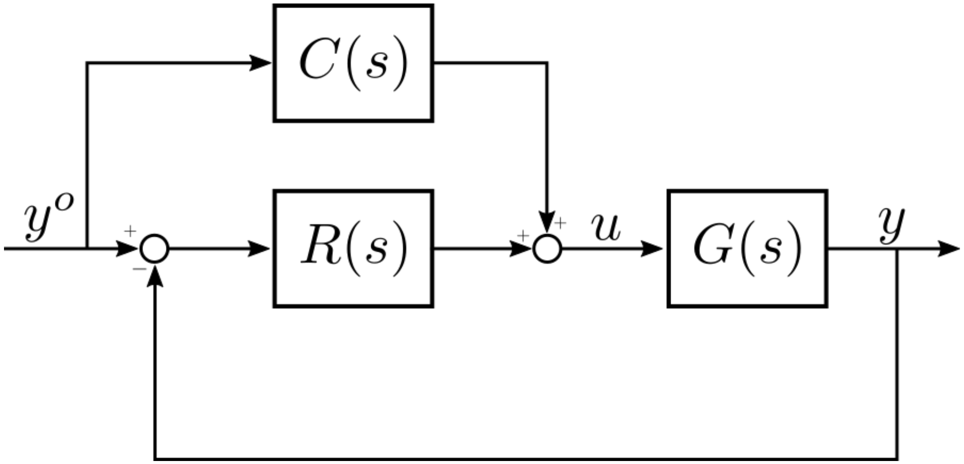
The equation to design the compensator is now

$$\frac{Y(s)}{D(s)} = \frac{H(s)+C(s)G(s)}{1+R(s)G(s)} = 0 \qquad \Rightarrow \qquad C(s) = -\frac{H(s)}{G(s)}$$

Caveat: The feedback controller and the compensator can be designed independently!

Feedforward compensation (V)

The same philosophy can be applied to the set point, in order to improve the tracking performance of the closed-loop system.



The transfer function from y^o to y is

$$\frac{Y(s)}{Y^o(s)} = \frac{R(s)G(s) + C(s)G(s)}{1 + R(s)G(s)}$$

We design the compensator to have a unitary transfer function from y^o to y

$$C(s) = \frac{1}{G(s)}$$

Feedforward compensation (VI)

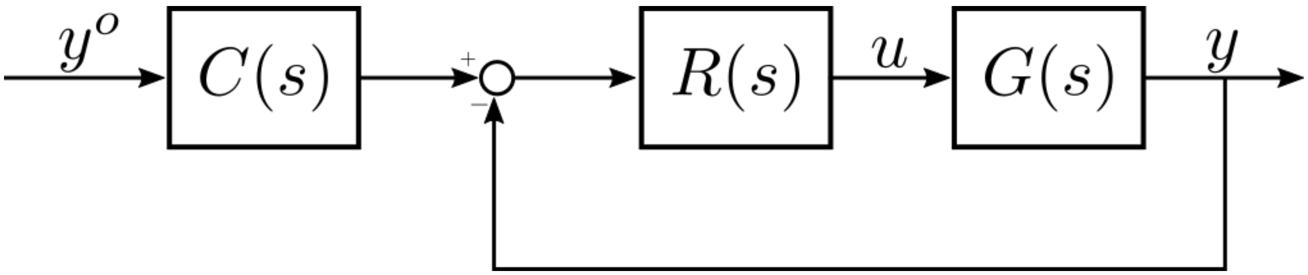
As we have seen for the disturbance compensator, this relation is a guideline for the design.

An example of a more realistic solution is the following

$$C(s) \quad : \quad C(j\omega) = \frac{1}{G(j\omega)} \quad \omega < \omega_{max}$$

where the compensator is designed to replicate the ideal one on a desired range of frequencies.

Another common feedforward action to improve set point tracking is pre-filtering.



Remember that the transfer function from y^o to y is

$$\frac{Y(s)}{Y^o(s)} = C(s)\frac{R(s)G(s)}{1+R(s)G(s)} = C(s)F(s)$$

Two examples of pre-filter are:

- a pre-filter to enforce zero steady-state error

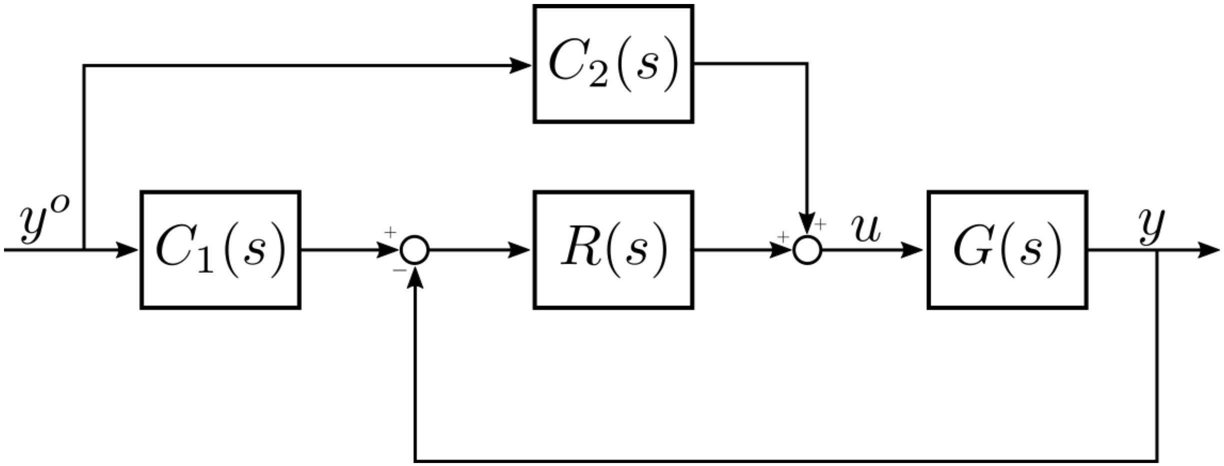
$$C(s) = \mu_C = \frac{1}{F(0)} \quad \Rightarrow \quad \left.\frac{Y(s)}{Y^o(s)}\right|_{s=0} = 1$$

- a pre-filter to increase the crossover frequency

$$C(s) = \frac{1+s/\omega_c}{1+s/\omega_b} \qquad \omega_b > \omega_c \qquad \Rightarrow \qquad \frac{Y(s)}{Y^o(s)} \approx \frac{1}{1+s/\omega_b}$$

Feedforward compensation (IX)

We can now imagine a control architecture where set point pre-filtering and compensation coexist.



This architecture allows to impose to the closed-loop system a desired behavior, imposing that the $y^o \rightarrow y$ relation behaves like a desired transfer function $F^o(s)$ (reference model).

Consider the transfer function from y^o to y

$$\frac{Y(s)}{Y^o(s)} = \frac{C_1(s)R(s) + C_2(s)}{1 + R(s)G(s)} G(s)$$

We can obtain the desire

$$\frac{Y(s)}{Y^o(s)} = F^o(s)$$

selecting

$$C_1(s) = F^o(s) \qquad C_2(s) = F^o(s)G^{-1}(s)$$

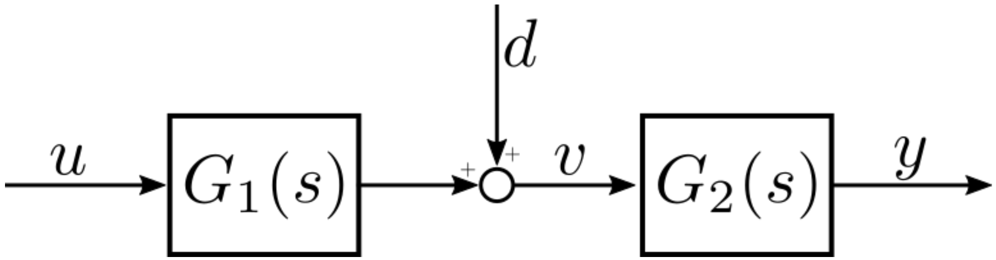
The requirements for the reference model are:

- unitary gain
- relative degree greater or equal to the relative degree of $G(s)$
- include the zeros of $G(s)$ that lie in the right half plane

Cascaded control (I)

In many real applications the process can be separated into two subsystems, thanks to an intermediate variable v that can be measured.

The system can be thus represented as a series of two transfer functions with a possible disturbance in between.



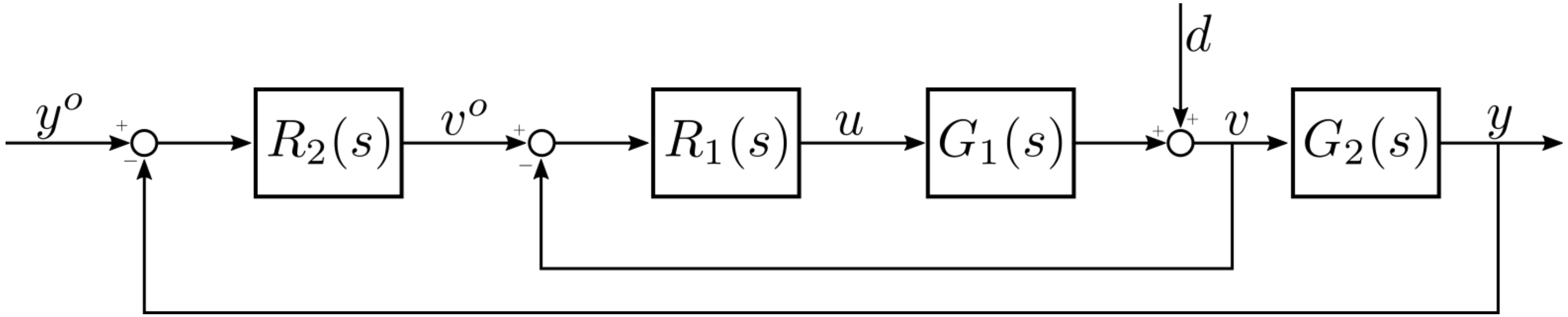
This subdivision can be exploited to simplify the design (and improve the performance of) the control system if:

- G_1 is minimum phase and G_2 is non minimum phase
- G_1 and G_2 are minimum phase systems, but the response time of G_2 is definitely greater than the response time of G_1

These conditions are satisfied, for example, when G_1 is the actuator and G_2 the process.

Cascaded control (II)

When one of the previous conditions holds we can adopt a control architecture called cascaded control.



The inner regulator R_1:

- is designed considering only the inner system $G_1(s)$
- ensures set point tracking at maximum allowable crossover frequency
- ensures high-bandwidth disturbance rejection

The outer regulator R_2:

- has a crossover frequency definitely lower than the inner loop
- the inner loop is seen by the outer regulator as a unitary transfer function
- is designed considering only the outer system $G_2(s)$

PID regulators (I)

PID regulators are characterized by the following control law

$$u(t) = K_P e(t) + K_I \int_0^t e(\tau)\mathrm{d}\tau + K_D \frac{\mathrm{d}e(t)}{\mathrm{d}t}$$

Derivative gain

Integral gain

Proportional gain

or equivalently

$$u(t) = K_P \left[e(t) + \frac{1}{T_I} \int_0^t e(\tau)\mathrm{d}\tau + T_D \frac{\mathrm{d}e(t)}{\mathrm{d}t} \right]$$

where the integral time and the derivative time are defined as

$$T_I = \frac{K_P}{K_I} \qquad T_D = \frac{K_D}{K_P}$$

PIDs are the most common industrial regulators, in particular in mechatronic applications.

The most common combination of PID actions are P, PD, PI, and PID.

PID regulators (II)

Why are PIDs so common in industrial applications to be called industrial regulators?

The most important reasons of PIDs' success are:

- they can be easily implemented using different technologies (hydraulic, pneumatic, electronic)
- they allow to control with good performance many different industrial processes
- they have been standardized (cheapness and reliability)
- they can be easily tuned (only 3 parameters to be selected)
- well-established auto-tuning techniques exist

In the following we will study both analytical tuning rules (based on Bode plots) and auto-tuning techniques.

PID regulators (III)

As any LTI system, PIDs can be represented by a transfer function

$$R(s) = \frac{E(s)}{U(s)} = K_P\left(1 + \frac{1}{sT_I} + sT_D\right) = K_P\frac{T_I T_D s^2 + T_I s + 1}{sT_I}$$

In the present form a PID is an a-causal system (has more zeros than poles), due to the presence of the derivative action.

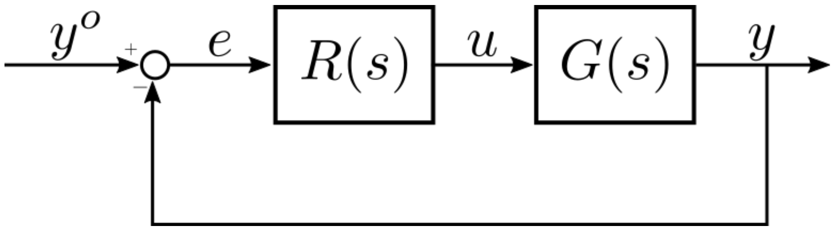
In order to make the system causal an high frequency pole is usually added to the derivative action.

Apart from making the system causal, this high-frequency pole has the following characteristics:

- it acts as a low-pass filter on the derivative action
- it has a negligible influence on the tuning of the controller parameters

PID regulators – Example (I)

Let's first introduce, with an example, a tuning methodology based on Bode plots.



$$G(s) = 0.1 \frac{e^{-3s}}{(1+5s)(1+20s)}$$

Requirements:

- $|e_\infty| = 0$ when $y^o(t) = \text{sca}(t)$
- $\varphi_m \geq 40°$
- maximize ω_c

We can write the PID transfer function as follows

$$R(s) = \mu_R \frac{(1+sT_1)(1+sT_2)}{s}$$

and select the zeros of the PID so as to cancel the poles of the process, obtaining

$$R(s) = \mu_R \frac{(1+5s)(1+20s)}{s} \quad \Rightarrow \quad L(s) = R(s)G(s) = \frac{0.1\mu_R}{s} \mathrm{e}^{-3s}$$

The crossover frequency is thus

$$\omega_c = 0.1\mu_R$$

and the phase margin

$$\varphi_m = 180^\circ - \left| -90^\circ - \omega_c \tau \frac{180^\circ}{\pi} \right| = 90^\circ - 0.3\mu_R \frac{180^\circ}{\pi} \geq 40^\circ$$

Solving the previous inequality we can find the maximum value of the crossover frequency that is compatible with the other specifications.

Solving with respect to the PID gain we obtain

$$\mu_R \leq \frac{50\pi}{0.3 \cdot 180^\circ} = 2.91$$

We can now write the transfer function of the regulator

$$R(s) = 2.9\frac{(1+5s)(1+20s)}{s} = 2.9\frac{100s^2+25s+1}{s} = K_P + \frac{K_I}{s} + K_D s$$

and determine the corresponding parameters

$$K_P = 72.5 \quad K_I = 2.9 \quad K_D = 290$$

We will now introduce auto-tuning rules, a set of techniques that allow to automatically determine the regulator parameters using the information obtained through a few experiments on the process.

Auto-tuning rules do not require any knowledge about the process model (or the model is implicitly identified from experimental data).

There are a huge number of different auto-tuning rules in the scientific literature and in commercial products.

We will introduce two classical tuning rules developed in 1942 by John Ziegler and Nathaniel Nichols.

PID regulators (V)

We start from the so called "closed-loop" rule.

The rule is composed of the following steps:

1. the regulator is started with all the gains (proportional, integral, derivative) set to zero
2. the proportional gain is slightly increased and a step response is performed
3. the proportional gain is continuously increased and the step response experiment repeated, until undamped oscillations appear in the controlled variable ($\bar{K}_P$ is the proportional gain that causes the undamped oscillations)
4. the period $\bar{T}$ of the oscillations is measured
5. PID parameters are selected following a table

	K_P	T_I	T_D
P	$0.5\bar{K}_P$		
PI	$0.45\bar{K}_P$	$\bar{T}/1.2$	
PID	$0.6\bar{K}_P$	$\bar{T}/2$	$\bar{T}/8$

PID regulators (VI)

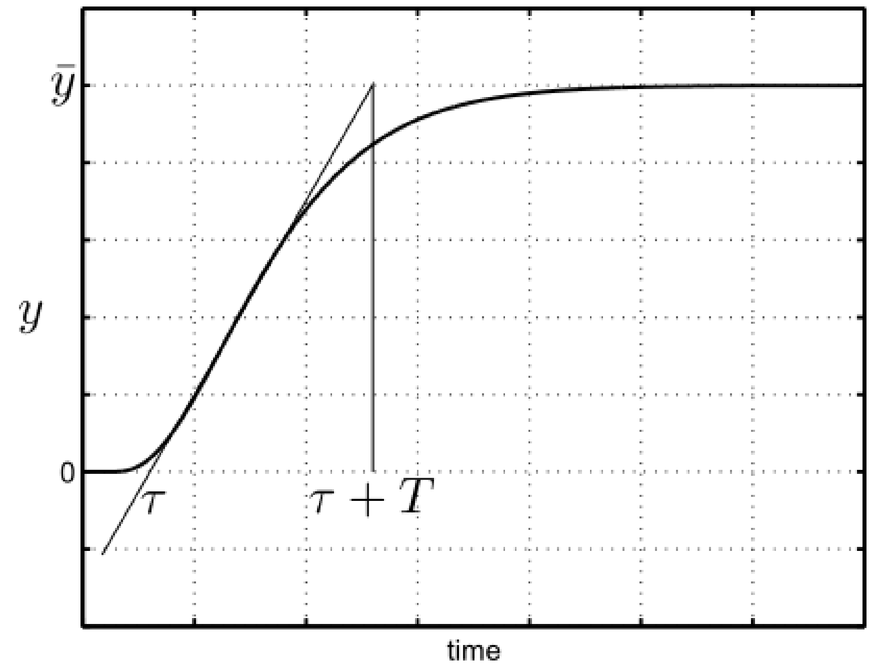
We conclude the “closed-loop” rule with some observations:

- there are systems that never generate undamped oscillations, the rule does not work with these systems
- in practice, bringing a system close to its stability limit is usually dangerous and not acceptable

The “open-loop” rule is composed of the following steps:

1. a step response is performed on the process (open-loop experiment)
2. if the step response is non oscillating and monotonically increasing, we can draw the tangent to the step response at the inflection point and compute the following parameters (graphically or numerically)
   a. gain μ, given by $\bar{y}/\bar{u}$
   b. τ and T, from the intersection of the tangent with the x-axis and the steady-state line
3. PID parameters are selected following a table

	K_P	T_I	T_D
P	$T/\mu\tau$		
PI	$0.9T/\mu\tau$	3τ	
PID	$1.2T/\mu\tau$	2τ	0.5τ



PID regulators (VIII)

We conclude the “open-loop” rule with some observations:

- if the step response is oscillating, or it is not monotonically increasing, or it does not have an inflection point, the rule cannot be applied
- in practice, operating a process in open-loop or performing a step response is not always acceptable