**Department of Mechanical and Production Engineering**

Aarhus University

# Lecture Notes

**Author:** Noah Rahbek Bigum Hansen

**Student ID:** 202405538

**Course:** Control Theory — 290231U008

**Date:** May 7, 2026

# Contents

**Lecture 1: Introduction** 17. Marts 2026

# 1 Course Information

These are the personal notes, written by Noah R. B. Hansen for the course in Control Theory at Aarhus University taught by Xuping Zhang, PhD. as part of the BSc in Mechanical Engineering.

# 2 Basic concepts and introduction of control theory

**Definition 1: Control System**

A *control system* is a device, or a set of devices, that manages, commands, directs, or regulates the behavior (output) of other device(s)/system(s)/processes by adjusting the input. For an example of a control system refer to the bottom of Figure 2.1

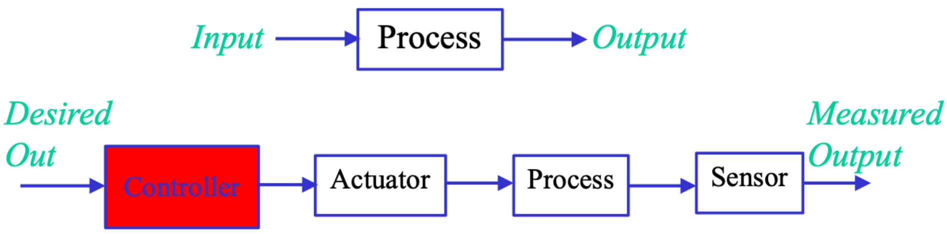



Figure 2.1: Comparison of a simple uncontrolled process and an example control system.

In Figure 2.1 two systems are shown. The top figure is an uncontrolled process that receives an input and outputs something without concern for anything other than how it is initially set up.

In the bottom diagram, the process is part of a control system. Here the desired output is fed to a controller that determines what action should be taken to match the real output to the desired one. The system is then fed to an actuator that applies the commands of the controller to the system, e.g. a motor, valve, heater, pump, brake, etc. The system is then fed to the process that is being controlled. This gives a signal to a sensor measuring the state of the process, e.g. a thermometer, encoder, tachometer, barometer, etc. The output of this sensor is then (ideally) the actual output of the system.

When talking about control systems we often differentiate between open-loop and closed-loop control systems.

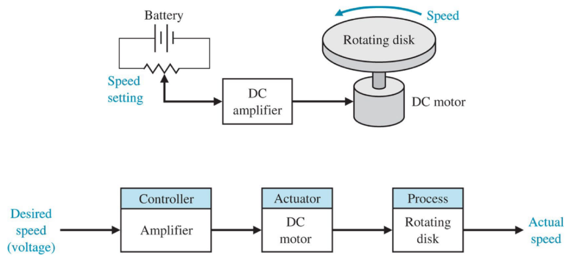
**Definition 2: Open-loop Control System**

An open-loop control system utilizes an actuating device to control the process directly without using feedback. An example of this is shown on Figure 2.2a.

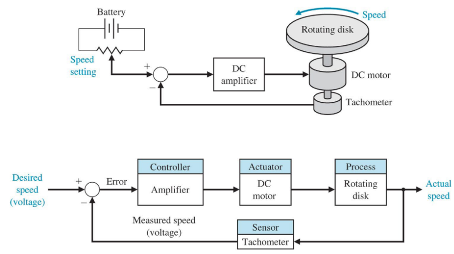
**Definition 3: Closed-loop Control System**

A closed-loop control system uses a measurement of the output and feedback of this signal to compare it with the desired output (a reference or command). An example closed-loop control system is depicted on Figure 2.2b

In general an open-loop control system are much simpler than as they do not require any sensors and do not introduce any stability problems on their own. However, their performance is also rather poor compared to closed-loop systems to match the desired output well.




(a) Open-loop control system.




(b) Closed-loop control system.

Figure 2.2: Comparison of an open and closed loop control system.

Closed-loop systems excel in that they do not need human adjustment and are able to respond to disturbance. These, however, may experience steady-state error or cause stability problems.

# 3 Description of the dynamic behavior of physical systems by Differential Equations

Many different physical systems are governed by differential equations.

**Example: Spring-Mass-Damper Mechanical system**

On Figure 3.1 a mass $M$ can move vertically up or down. It is connected to a spring with stiffness $k$, a damper with damping/friction coefficient $b$, and an applied external force $r(t)$.
The acceleration of the mass is related to the net force by Newton's second law:

$$M\ddot{y} = r(t) - b\dot{y} - ky \implies M\ddot{y} + b\dot{y} + ky = r(t).$$

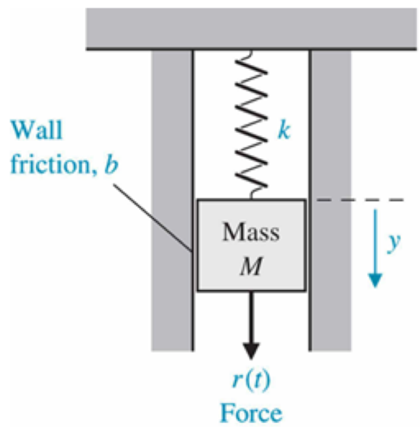



Figure 3.1: Spring-mass-damper system.

**Example: Resister-inductor-capacitor Circuit Electrical System**

On Figure 3.2 a parallel RLC circuit is driven by an input current source $r(t)$ and the output is the voltage $v(t)$ across the branches. Here, the resister gives a current proportional to voltage:

$$i_R = \frac{v}{R},$$

the capacitor has a current of:

$$i_C = C\frac{\mathrm{d}v}{\mathrm{d}t},$$

and the voltage-current relation of the inductor is:

$$v = L\frac{\mathrm{d}i_L}{\mathrm{d}t}.$$

If we apply Kirchoff's current law at the top node:

$$r(t) = i_R + i_L + i_C$$

and then substitute into the branch relations:

$$r(t) = \frac{v}{R} + i_L + C\frac{\mathrm{d}v}{\mathrm{d}t}.$$

Taking the time-derivative yields:

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \frac{1}{R}\frac{\mathrm{d}v}{\mathrm{d}t} + \frac{\mathrm{d}i_L}{\mathrm{d}t} + C\frac{\mathrm{d}^2v}{\mathrm{d}t^2} = \frac{1}{R}\frac{\mathrm{d}v}{\mathrm{d}t} + \frac{v}{L} + C\frac{\mathrm{d}^2v}{\mathrm{d}t^2}.$$

Which can be rearranged to

$$C\frac{\mathrm{d}^2v}{\mathrm{d}t^2} + \frac{1}{R}\frac{\mathrm{d}v}{\mathrm{d}t} + \frac{1}{L}v = \frac{\mathrm{d}r(t)}{\mathrm{d}t}.$$

Note that this is analogous to the result from the mass-spring-damper system above.

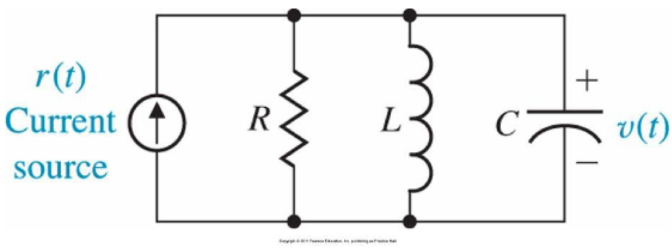



Figure 3.2: Resistor-inductor-capacitor system.

In general, many dynamic elements can be broken into inductive storage, capacitive storage, and energy dissipaters. A table is shown on Figure 3.3.

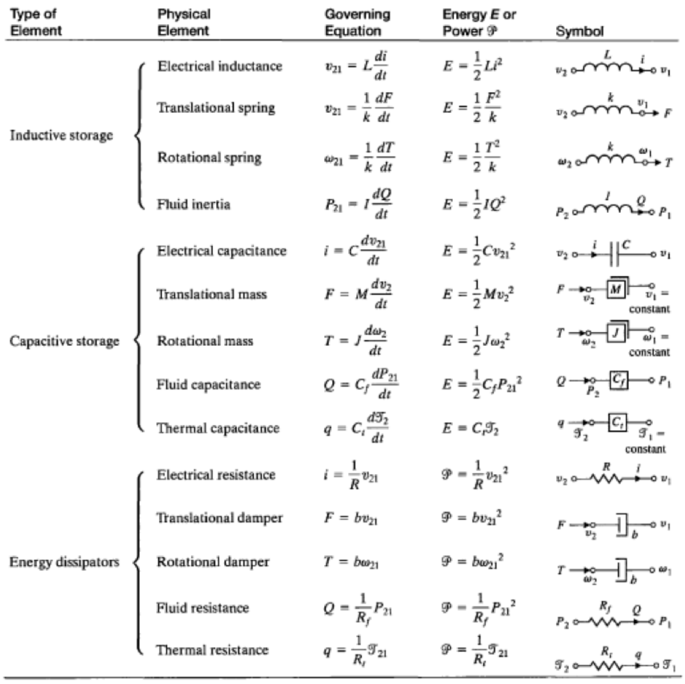
# 4 Linear Systems

A linear system is any system that satisfies superposition. I.e. if input $u_1$ gives output $y_1$ and input $u_2$ gives output $y_2$, then input $au_1 + b_u 2$ gives output $ay_1 + by_2$ for constants $a$ and $b$. This means

$$m\ddot{y} + b\dot{y} + ky = r(t)$$

is a linear equation as all $y$, $\dot{y}$ and $\ddot{y}$ appear linearly, but

$$m\ddot{y} + b\dot{y} + ky^3 = r(t)$$



| Type of Element | Physical Element | Governing Equation | Energy $E$ or Power $\mathcal{P}$ | Symbol |
|---|---|---|---|---|
| Inductive storage | Electrical inductance | $v_{21} = L\frac{di}{dt}$ | $E = \frac{1}{2}Li^2$ | |
| | Translational spring | $v_{21} = \frac{1}{k}\frac{dF}{dt}$ | $E = \frac{1}{2}\frac{F^2}{k}$ | |
| | Rotational spring | $\omega_{21} = \frac{1}{k}\frac{dT}{dt}$ | $E = \frac{1}{2}\frac{T^2}{k}$ | |
| | Fluid inertia | $P_{21} = I\frac{dQ}{dt}$ | $E = \frac{1}{2}IQ^2$ | |
| Capacitive storage | Electrical capacitance | $i = C\frac{dv_{21}}{dt}$ | $E = \frac{1}{2}Cv_{21}^2$ | |
| | Translational mass | $F = M\frac{dv_2}{dt}$ | $E = \frac{1}{2}Mv_2^2$ | |
| | Rotational mass | $T = J\frac{d\omega_2}{dt}$ | $E = \frac{1}{2}J\omega_2^2$ | |
| | Fluid capacitance | $Q = C_f\frac{dP_{21}}{dt}$ | $E = \frac{1}{2}C_fP_{21}^2$ | |
| | Thermal capacitance | $q = C_t\frac{d\mathcal{T}_2}{dt}$ | $E = C_t\mathcal{T}_2$ | |
| Energy dissipators | Electrical resistance | $i = \frac{1}{R}v_{21}$ | $\mathcal{P} = \frac{1}{R}v_{21}^2$ | |
| | Translational damper | $F = bv_{21}$ | $\mathcal{P} = bv_{21}^2$ | |
| | Rotational damper | $T = b\omega_{21}$ | $\mathcal{P} = b\omega_{21}^2$ | |
| | Fluid resistance | $Q = \frac{1}{R_f}P_{21}$ | $\mathcal{P} = \frac{1}{R_f}P_{21}^2$ | |
| | Thermal resistance | $q = \frac{1}{R_t}\mathcal{T}_{21}$ | $\mathcal{P} = \frac{1}{R_t}\mathcal{T}_{21}$ | |

Figure 3.3: Summary of Analogous Dynamic Elements

is nonlinear, due to the $y^3$-term. An equation like

$$\ddot{y} + y\dot{y} = u$$

is also non-linear as $y$ and $\dot{y}$ are multiplied together.

Many non-linear systems can be linearised using Taylor-transforms.

# 5 Fundamentals and application of Laplace transforms

Laplace transforms are used frequently in Control theory to turn differential equations in the time-domain into algebraic equations in the Laplace-domain. In general, for a time signal $f(t)$, its Laplace transform is

$$F(s) = \mathscr{L}\left\{f(t)\right\} = \int_0^\infty f(t)e^{-st}\,\mathrm{d}t$$

where

$$s = \sigma + j\omega$$

is a complex variable.

To go from the Laplace-domain back to the time-domain the inverse Laplace transform, $\mathscr{L}^{-1}$ is used:

$$f(t) = \mathscr{L}^{-1}\{F(s)\} = \frac{1}{2\pi j}\int_{\sigma-j\infty}^{\sigma+j\infty} F(s)e^{st}\,\mathrm{d}t.$$

A table of the most important Laplace Transform Pairs is given in Figure 5.1.

It is also worth noting that an integral in the time-domain corresponds to division by $s$ in the Laplace-domain:

$$\mathscr{L}\left\{\int_0^\infty f(t)\,\mathrm{d}t\right\} = \frac{1}{s}\int_0^\infty f(t)e^{-st}\,\mathrm{d}t = \frac{1}{s}F(s)$$

| $f(t)$ | $F(s)$ |
|---|---|
| Step function, $u(t)$ | $\frac{1}{s}$ |
| $e^{-at}$ | $\frac{1}{s+a}$ |
| $\sin \omega t$ | $\frac{\omega}{s^2+\omega^2}$ |
| $\cos \omega t$ | $\frac{s}{s^2+\omega^2}$ |
| $t^n$ | $\frac{n!}{s^{n+1}}$ |
| $f^{(k)}(t) = \frac{d^k f(t)}{dt^k}$ | $s^k F(s) - s^{k-1} f(0^-) - s^{k-2} f'(0^-) - \ldots - f^{(k-1)}(0^-)$ |
| $\int_{-\infty}^{t} f(t)\, dt$ | $\frac{F(s)}{s} + \frac{1}{s}\int_{-\infty}^{0} f(t)\, dt$ |
| Impulse function $\delta(t)$ | $1$ |
| $e^{-at}\sin \omega t$ | $\frac{\omega}{(s+a)^2+\omega^2}$ |
| $e^{-at}\cos \omega t$ | $\frac{s+a}{(s+a)^2+\omega^2}$ |
| $\frac{1}{\omega}[(\alpha-a)^2+\omega^2]^{1/2} e^{-at}\sin(\omega t+\phi)$, $\phi = \tan^{-1}\frac{\omega}{\alpha - a}$ | $\frac{s+\alpha}{(s+a)^2+\omega^2}$ |
| $\frac{\omega_n}{\sqrt{1-\zeta^2}} e^{-\zeta\omega_n t}\sin \omega_n\sqrt{1-\zeta^2}t,\ \zeta < 1$ | $\frac{\omega_n^2}{s^2+2\zeta\omega_n s+\omega_n^2}$ |
| $\frac{1}{a^2+\omega^2} + \frac{1}{\omega\sqrt{a^2+\omega^2}} e^{-at}\sin(\omega t - \phi)$, $\phi = \tan^{-1}\frac{\omega}{-a}$ | $\frac{1}{s[(s+a)^2+\omega^2]}$ |
| $1 - \frac{1}{\sqrt{1-\zeta^2}} e^{-\zeta\omega_n t}\sin\left(\omega_n\sqrt{1-\zeta^2}t+\phi\right)$, $\phi = \cos^{-1}\zeta,\ \zeta < 1$ | $\frac{\omega_n^2}{s(s^2+2\zeta\omega_n s+\omega_n^2)}$ |
| $\frac{\alpha}{a^2+\omega^2} + \frac{1}{\omega}\left[\frac{(\alpha-a)^2+\omega^2}{a^2+\omega^2}\right]^{1/2} e^{-at}\sin(\omega t+\phi)$. $\phi = \tan^{-1}\frac{\omega}{\alpha-a} - \tan^{-1}\frac{\omega}{-a}$ | $\frac{s+\alpha}{s[(s+a)^2+\omega^2]}$ |

Figure 5.1: Important Laplace Transform Pairs

and that a derivative in the time-domain corresponds to multiplication by $s$ in the Laplace-domain

$$
\begin{aligned}
\mathscr{L}\left\{\frac{\mathrm{d}f(t)}{\mathrm{d}t}\right\} &= sF(s) - f(0) \\
\mathscr{L}\left\{\frac{\mathrm{d}f(t)^2}{\mathrm{d}t^2}\right\} &= s^2F(s) - sf(0) - f'(0) \\
\mathscr{L}\left\{\frac{\mathrm{d}f(t)^3}{\mathrm{d}t^3}\right\} &= s^3F(s) - s^2f(0) - sf'(0) - f''(0) \\
&\vdots \\
\mathscr{L}\left\{\frac{\mathrm{d}f(t)^n}{\mathrm{d}t^n}\right\} &= s^nF(s) - s^{n-1}f(0) - s^{n-2}f'(0) - \ldots - f^{(n-1)}(0).
\end{aligned}
$$

Further, addition/scaling of Laplace transforms follows the relation

$$\mathscr{L}\left\{af_1(t) \pm bf_2(t)\right\} = aF_1(s) \pm bF_2(s),$$

convolutions are defined as

$$\int_0^t f_1(t-\tau) f_2(\tau)\,\mathrm{d}\tau = F_1(s)F_2(s),$$

the initial-value theorem is

$$f(0+) = \lim_{s\to\infty} sF(s),$$

and the final-value theorem is

$$\lim_{t\to\infty} f(t) = \lim_{s\to 0} sF(s).$$

## 5.1 Transfer Functions

A transfer function is often used to describe the dynamics of a system. These are defined only for linear, stationary (i.e. constant parameter) systems, and may not apply for time-varying systems having one or mare time-varying parameters. A transfer function does not include any information about the internal structure of the system.

**Definition 4: Transfer Function**

A transfer function is defined as the ratio of the Laplace transform of the output variable to the Laplace transform of the input variable, with all initial conditions assumed to be zero. I.e. the transfer function $G(s)$ of the Laplace transform $Y(s)$ of the output $y(t)$ and the Laplace transform $R(s)$ of the input $r(t)$ is

$$G(s) = \frac{Y(s)}{R(s)}.$$

**Example: Transfer function of a Mass-spring-damper**

For a mass-spring-damper such as that on Figure 3.1 we have Newton's second law:

$$M\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + b\frac{\mathrm{d}y}{\mathrm{d}t} + ky = r(t).$$

Now, taking the Laplace transform with zero initial conditions:

$$Ms^2Y(s) + bsY(s) + kY(s) = R(s) \implies Y(s)\left(Ms^2 + bs + k\right) = R(s).$$

So:

$$G(s) = \frac{Y(s)}{R(s)} = \frac{1}{Ms^2 + bs + k}.$$

In general for an $n$-th order linear differential equation:

$$\frac{\mathrm{d}^n y}{\mathrm{d}t^n} + q_{n-1}\frac{\mathrm{d}^{n-1}}{\mathrm{d}t^{n-1}} + \ldots + q_0 y = p_{n-1}\frac{\mathrm{d}^{n-1} r}{\mathrm{d}t^{n-1}} + p_{n-2}\frac{\mathrm{d}^{n-2} t}{\mathrm{d}t^{n-2}} + \ldots + p_0 r,$$

after taking the Laplace transform the transfer function will look like:

$$G(s) = \frac{Y(S)}{R(s)} = \frac{p_{n-1}s^{n-1} + p_{n-2}s^{n-2} + \ldots + p_0}{s^n + q_{n-1}s^{n-1} + \ldots + q_0} = \frac{p(s)}{q(s)}.$$

So in general, a transfer function is a ratio of two polynomials in $s$.

### 5.1.1 Poles and Zeros

From

$$G(s) = \frac{p(s)}{q(s)}$$

we define zeros to be the roots of $p(s) = 0$ and the poles to be the roots of $q(s) = 0$.

This means that zeros are values of $s$ that make the transfer function become zero. These change how the system responds to input and affect transient behavior, frequency response, overshoot, etc.

The poles are more fundamental for system dynamics and determine stability, oscillation, decay rate and speed of responds. Poles with positive real part cause growth and instability as these lead to positive exponentials $e^t$ after taking the inverse Laplace.

# 6 Partial Fraction Decomposition

Given a fraction such as

$$\frac{s+1}{(s+2)(s+3)}$$

we wish to split this into a sum of two fractions with simple denominators that we can easily take the inverse Laplace transform of

$$\frac{A}{s+2}+\frac{B}{s+3}$$

with some (yet) unknown constants $A$ and $B$.

To do this we:

1. Expand into a term for each factor in the denominator
2. Recombine the RHS
3. Equate terms in $s$ and constant terms, and have algebraic equations of the coefficients
4. Solve algebraic Equations
5. Each term is in a form so that inverse Laplace transforms can be applied.

Starting with:

$$\frac{s+1}{(s+2)(s+3)}=\frac{A}{s+2}+\frac{B}{s+3}.$$

We put the RHS over a common denominator:

$$\frac{A}{s+2}+\frac{B}{s+3}=\frac{A(s+3)+B(s+2)}{(s+2)(s+3)},$$

so now

$$\frac{s+1}{(s+2)(s+3)}=\frac{A(s+3)+B(s+2)}{(s+2)(s+3)}.$$

Since the denominators are equal, the numerators must also be equal:

$$\begin{aligned} s+1 &= A(s+3)+B(s+2) \\ &= As+3A+Bs+2B \\ &= (A+B)s+(3A+2B). \end{aligned}$$

From this we see that $A+B=1$ and $B=1-A$, so $A=-1$ and $B=2$. I.e.

$$\frac{s+1}{(s+2)(s+3)}=-\frac{1}{s+2}+\frac{2}{s+3}.$$

Now, taking the inverse Laplace transform is easy:

$$\mathscr{L}^{-1}\left\{\frac{1}{s+a}\right\}=e^{-at}$$

$$\mathscr{L}^{-1}\left\{\frac{s+1}{(s+2)(s+3)}\right\}=-e^{-2t}+2e^{-3t}.$$

**Example: Partial Fraction Decomposition**

$$
\begin{aligned}
G(s) &= \frac{5s+3}{(s+1)(s+2)(s+3)} \\
\frac{5s+3}{(s+1)(s+2)(s+3)} &= \frac{K_{s1}}{s+1} + \frac{K_{s2}}{s+2} + \frac{K_{s3}}{s+3} \\
K_{-1} &= \frac{p(-1)}{(-1+2)(-1+3)} = -1 \\
K_{-2} &= \frac{p(-2)}{(-2+1)(-2+3)} = 7 \\
K_{-3} &= \frac{p(-3)}{(-3+1)(-3+2)} = -6 \\
\frac{5s+3}{(s+1)(s+2)(s+3)} &= -\frac{1}{s+1} + \frac{7}{s+2} - \frac{6}{s+3} \\
G(t) &= -e^{-t} + 7e^{-2t} - 6e^{-3t}.
\end{aligned}
$$

**Example: Conversion of ODE to Algebraic Equation using Laplace Transforms**

Given the ODE

$$
\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} + 6\frac{\mathrm{d}y}{\mathrm{d}t} + 8y = 2 \qquad y(0) = y'(0) = 0
$$

we want to find a solution using the Laplace transform to turn the above ODE into an easily-solvable algebraic equation.

---

We start by finding the Laplace transform of each term.

$$
\begin{aligned}
\frac{\mathrm{d}^2 y}{\mathrm{d}t^2} &= s^2 Y(s) - sy(0) - y'(0) \\
6\frac{\mathrm{d}y}{\mathrm{d}t} &= 6\big(sY(s) - y(0)\big) \\
8y &= 8Y(s) \\
2 &= \frac{2}{s} \\
s^2 Y(s) - 0 - 0 + 6(sY(s) - 0) + 8Y(s) &= \frac{2}{s} \\
s^2 Y(s) + 6sY(s) + 8Y(s) &= \frac{2}{s} \\
Y(s)\left(s^2 + 6s + 8\right) &= \frac{2}{s} \\
Y(s) &= \frac{2}{s(s+2)(s+4)}.
\end{aligned}
$$

We have now found the solution in the Laplace-domain. To find the corresponding time-domain solution we take the inverse Laplace transform. To do this we use Partial Fraction Decomposition:

$$\begin{aligned}
\frac{2}{s\,(s+2)\,(s+4)} &= \frac{K_{s1}}{s} + \frac{K_{s2}}{s+2} + \frac{K_{s3}}{s+4} \\
K_0 &= \frac{2}{2\cdot 4} = \frac{1}{4} \\
K_{-2} &= \frac{2}{-2\cdot 2} = -\frac{1}{2} \\
K_{-4} &= \frac{2}{-4\cdot -2} = \frac{1}{4} \\
\frac{2}{s(s+2)(s+4)} &= \frac{1}{4s} - \frac{1}{2(s+2)} + \frac{1}{4(s+4)}
\end{aligned}$$

.

Now, that a partial fraction decomposition has been made we can easily take the inverse Laplace of the solution as

$$\begin{aligned}
g(t) &= \frac{1}{4} - \frac{1}{2}e^{-2t} + \frac{1}{4}e^{-4t} \\
&= \frac{1}{4}e^{-4t}\left(e^{2t}-1\right)^2.
\end{aligned}$$

**Example: In-Class Exercise: Controlling a Laser Printer**

A laser printer uses a laser beam to print copy rapidly for a computer. The laser is positioned by a control input $r(t)$, so that we have

$$Y(s) = \frac{4(s+50)}{s^2+30s+200}R(s).$$

The input $r(t)$ represents the desired position of the laser beam.

(a) If $r(t)$ is a unit step input find the output $y(t)$.

(b) What is the final value of $y(t)$?

---

(a) For a unit step $r(t)$, $R(s) = \frac{1}{s}$, so

$$\begin{aligned}
Y(s) &= \frac{4(s+50)}{s^2+30s+200}\,\frac{1}{s} \\
&= \frac{4(s+50)}{s^3+30s^2+200s} \\
&= \frac{4(s+50)}{s(s+10)(s+20)} \\
\frac{4(s+50)}{s(s+10)(s+20)} &= \frac{K_0}{s} + \frac{K_{10}}{s+10} + \frac{K_{20}}{s+20} \\
K_0 &= \frac{200}{10\cdot 20} = 1 \\
K_{-10} &= \frac{160}{-10\cdot 10} = -\frac{8}{5} \\
K_{-20} &= \frac{120}{-20\cdot -10} = \frac{3}{5} \\
Y(s) &= \frac{1}{s} - \frac{8}{5(s+10)} + \frac{3}{5(s+20)} \\
y(t) &= 1 - \frac{8}{5}e^{-10t} + \frac{3}{5}e^{-20t}.
\end{aligned}$$

(b) We solve:

$$\lim_{t\to\infty} 1 - \frac{8}{5}e^{-10t} + \frac{3}{5}e^{-20t} = 1.$$

## Lecture 2: Mathematical Models of Systems

20. Marts 2026

### 6.1 Block Diagrams

Though mathematical models are useful in describing the complex behavior of a control system they are often rather unintuitive. A more graphical alternative to the rigorous mathematical models are Block Diagrams. An example of one of these is shown on Figure 6.1

**Definition 5: Block Diagram**

A *Block Diagram* is a schematic representation of the cause-and-effect relationshop throughout the system together with the transfer functions of the variables of interest.

Typical elements of block diagrams include but are not limited to:

- Comparators
- Blocks representing individual component transfer functions (sensors, actuators, controllers, plants, etc.)
- Input or reference signals
- Output signals
- Disturbance signals
- Feedback loops

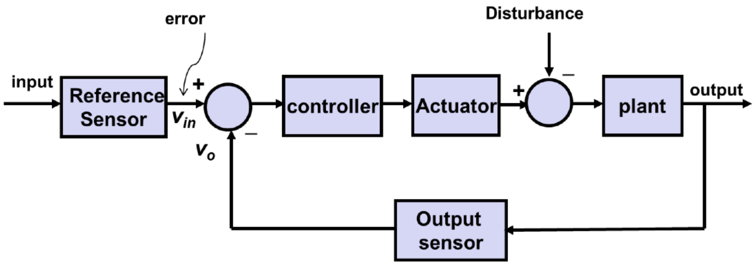



Figure 6.1: Example of a block diagram.

An example showing how different components of a system might be represented as blocks of a block diagram is shown in Figure 6.2.

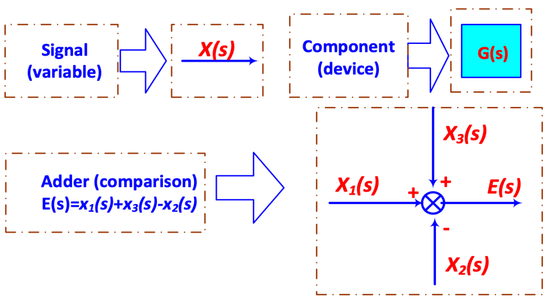



Figure 6.2: Different components represented as blocks.

For multiple input or output systems the subscript convention is such that $G_{ij}$ is the transfer function of output $i$ and input $j$.

### 6.1.1 Transformation and Reduction of Block Diagrams

Different Block diagrams can be transformed and reduced into more simple forms. This is especially useful for large and complex systems.

For a simple cascade block diagram as shown on Figure 6.3 one is able to reduce the block diagram into a single $G_1 \cdot G_2$-block.


Figure 6.3: Cascade Block diagram.

For a parallel block diagram as shown on Figure 6.4 the resulting transformation is a single $G_1 - G_2$-block.

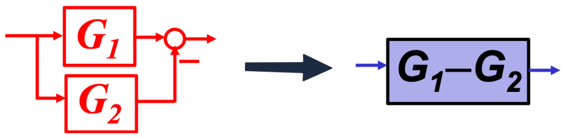



Figure 6.4: Parallel block diagram.

For a negative feedback Block diagram as shown on Figure 6.5 the reduction is a single block:

$$\frac{G_1}{1 + G_1 G_2}.$$

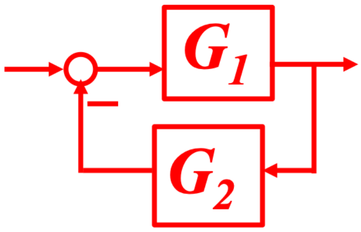



Figure 6.5: Feedback block diagram.

A summing node is able to be moved across a block as shown in Figure 6.6.

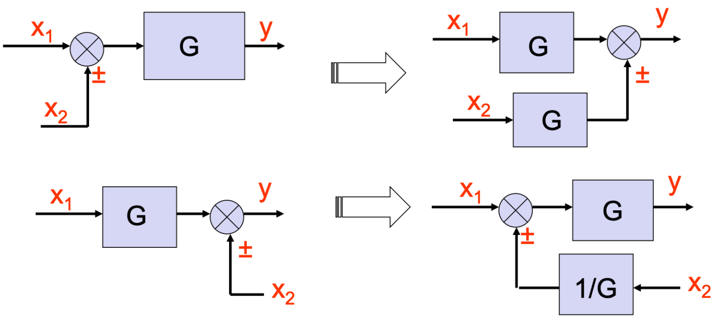



Figure 6.6: Movement of a summing node across a block.

Pickoff points can also be moved across blocks. The manner this is done by is shown in Figure 6.7.

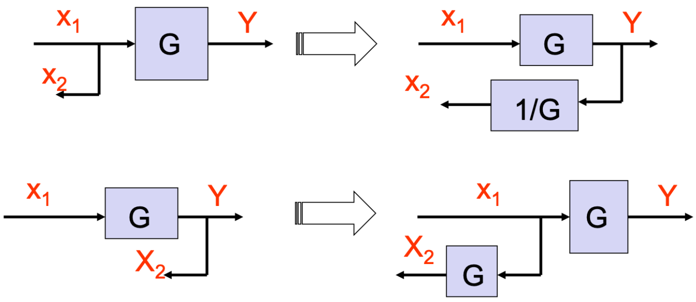



Figure 6.7: Movement of a pickoff point across a block.

It is also possible to interchange neighboring summing points or pickoff points as shown in Figure 6.8.

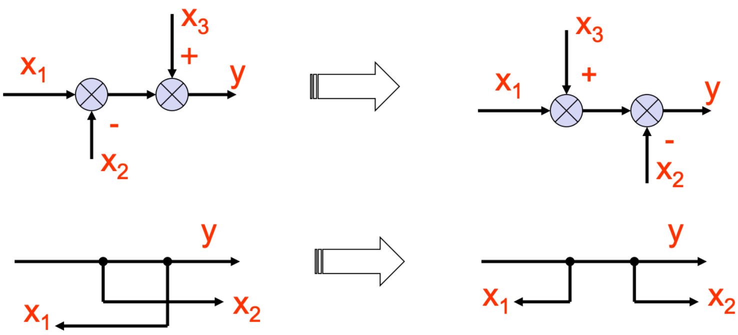



Figure 6.8: Interchanging neighboring summing and pickoff points

**Example: Transfer function of an interacting system**

Given the system shown find the transfer function.

We combine all the feedback loops using:

$$G_i(s)||H_i(s) = \frac{G_i(s)}{1 - G_i(s)H_i(s)}.$$

This gives:

$$
\begin{aligned}
G_2(s)||H_2(s) &= \frac{G_2(s)}{1 - G_2(s)H_2(s)} \\
G_3(s)||H_3(s) &= \frac{G_3(s)}{1 - G_3(s)H_3(s)} \\
G_6(s)||H_6(s) &= \frac{G_6(s)}{1 - G_6(s)H_6(s)} \\
G_7(s)||H_7(s) &= \frac{G_7(s)}{1 - G_7(s)H_7(s)}.
\end{aligned}
$$

Now the two branches are both cascading, so using $G_1G_2$ for series we get:

$$
\begin{aligned}
B_1(s) &= G_1(s) \cdot \frac{G_2(s)}{1 - G_2(s)H_2(s)} \cdot \frac{G_3(s)}{1 - G_3(s)H_3(s)} \cdot G_4(s) \\
B_2(s) &= G_5(s) \cdot \frac{G_6(s)}{1 - G_6(s)H_6(s)} \cdot \frac{G_7(s)}{1 - G_7(s)H_7(s)} \cdot G_8(s).
\end{aligned}
$$

Now, we are left with the two parallel $B_1(s)$ and $B_2(s)$ branches which are combined using $G_1||G_2 = G_1 + G_2$ as:

$$
B_1(s) + B_2(s) = G_1(s) \cdot \frac{G_2(s)}{1 - G_2(s)H_2(s)} \cdot \frac{G_3(s)}{1 - G_3(s)H_3(s)} \cdot G_4(s) + G_5(s) \cdot \frac{G_6(s)}{1 - G_6(s)H_6(s)} \cdot \frac{G_7(s)}{1 - G_7(s)H_7(s)} \cdot G_8(s).
$$

Our transfer function is therefore:

$$
t(s) = G_1(s) \cdot \frac{G_2(s)}{1 - G_2(s)H_2(s)} \cdot \frac{G_3(s)}{1 - G_3(s)H_3(s)} \cdot G_4(s) + G_5(s) \cdot \frac{G_6(s)}{1 - G_6(s)H_6(s)} \cdot \frac{G_7(s)}{1 - G_7(s)H_7(s)} \cdot G_8(s).
$$

So

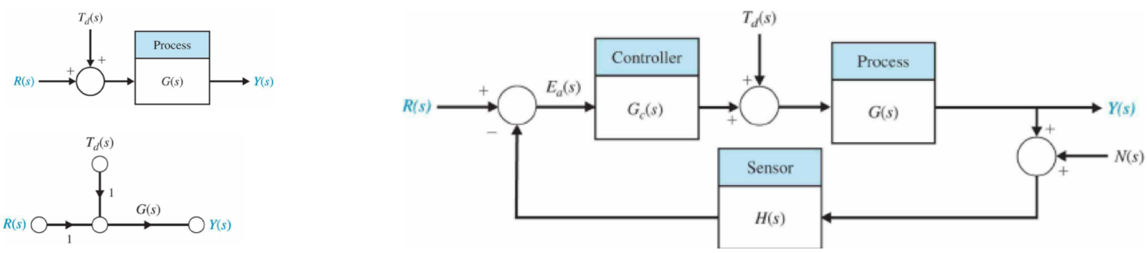
$$
Y(s) = R(s)\left(G_1(s) \cdot \frac{G_2(s)}{1 - G_2(s)H_2(s)} \cdot \frac{G_3(s)}{1 - G_3(s)H_3(s)} \cdot G_4(s) + G_5(s) \cdot \frac{G_6(s)}{1 - G_6(s)H_6(s)} \cdot \frac{G_7(s)}{1 - G_7(s)H_7(s)} \cdot G_8(s)\right)
$$

**Lecture 3: Feedback Control System Characteristics** 24. March 2026t

# 7 Feedback Control System Characteristics

## 7.1 Basic Concepts and Structure

Whereas an open-loop control system operates without feedback (see Figure 7.1a, a closed-loop system (Figure 7.1b) uses a measurement of the outout signal and a comparison with the desired output to generate an error signal that is used by the controller to adjust the actuator, which has certain robustness against disturbances and parameter changes.




(a) Open-loop control system.

(b) Closed-loop control system.

Figure 7.1: Comparison of an open and closed loop control system.

### 7.2 Error Signal Analysis

On Figure 7.2 we define the tracking error $E(s)$ to be

$$E(s) = R(s) - Y(s)$$

because with feedback $H(s) = 1$, the measured output is just the output itself. I.e. if $Y(S)$ follows $R(s)$ perfectly, then $E(s) = 0$. On Figure 7.2 the system is influenced by three inputs the desired reference $R(s)$, the disturbance $T_d(s)$ and the measurement noise $N(s)$.

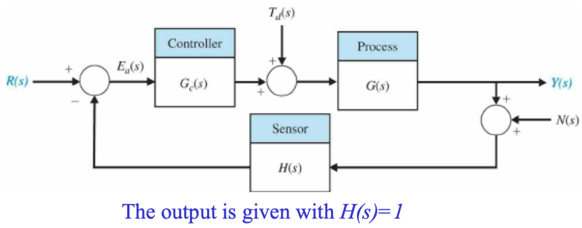



Figure 7.2: Tracking error with unity feedback, $H(s) = 1$.

With feedback $H(s) = 1$, the output $Y(s)$ can be computed as a sum of three contributions. One from the reference $R(s)$, which is the usual closed-loop transfer function from reference to output:

$$Y_R = \frac{G_c G}{1 + G_c G} R,$$

one from the disturbance $T_d(s)$, which shows how much the disturbance affects the output

$$Y_T = \frac{G}{1 + G_c G} T_D,$$

and one from the measurement noise

$$Y_N = -\frac{G_c G}{1 + G_c G} N.$$

Note the minus sign in the above appears as the noise is in the feedback signal, thus corrupting the measured output and thereby altering the error in the opposite direction. Which means the output is

$$Y(s) = Y_R + Y_T + Y_N = \frac{G_c(s)G(s)}{1 + G_c(s)G(s)} R(s) + \frac{G(s)}{1 + G_c(s)G(s)} T_D(s) - \frac{G_c(s)G(s)}{1 + G_c(s)G(s)} N(s).$$

The error $E(s)$ is then

$$E(s) = R(s) - Y(s) = \frac{1}{1 + G_c(s)G(s)} R(s) - \frac{G(s)}{1 + G_c(s)G(s)} T_d(s) + \frac{G_c(s)G(s)}{1 + G_c(s)G(s)} N(s).$$

We now define the loop gain $L(s)$ as the total gain you get by going around a feedback loop

$$L(s) = G_c(s)G(s)H(s)$$

and since $H(s) = 1$,

$$L(s) = G_c(s)G(s).$$

This simplifies the formula for the error to:

$$E(s) = \frac{1}{1 + L(s)} R(s) - \frac{G(s)}{1 + L(s)} T_d(s) + \frac{L(s)}{1 + L(s)} N(s).$$

If we now define the sensitivity function $S(s)$ as

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

and the complementary sensitivity function $C(s)$ as

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

then the error becomes

$$E(s) = S(s)R(s) - S(s)G(S)T_d(s) + C(s)N(s).$$

Here, the sensitivity function $S(s) = 1/(1+L(s))$ tells how much of the reference mismatch and disturbance gets through. From the formula we see that large $L(s)$ leads to small $S(s)$ so large loop gain reduces tracking error $R(s)$ and the effect of disturbance $T_d(s)$ On the other hand, the complementary sensitivity $C(s) = L(s)/(1+L(s))$ tell how much measurement noise gets through into the error/output. If $L(s)$ is large, then $C(s) \approx 1$, hence high gain is bad for noise sensitivity. Further, we know that

$$S(s) + C(s) = 1$$

hence both cannot be small at the same frequency. Making $S(s)$ small helps tracking and disturbance rejection and making $C(s)$ small helps noise rejection, but improving one tends to worsen the other.

Generally, disturbances are more important at low frequencies and measurement noise is usually more important at high frequencies. Therefore, a good controller typically will try to make the loop gain $L(s)$ large at low frequencies and smaller at higher frequencies. At low frequencies, if $L(s) \gg 1$, then

$$S(s) \approx 0, \qquad C(s) \approx 1$$

so the tracking error and disturbance effects are small. At high frequencies, if $L(s) \ll 1$, then

$$S(s) \approx 1, \qquad C(s) \approx 0$$

so measurement noise is suppressed. I.e. We want the controller to fight slow changes strongly, but ignore very fast noisy fluctuations.

## 7.3 Sensitivity to Parameter Variations

The input $G(s)$ is time-varying in nature due to uncertainties such as changing environments, aging, and other factors. These uncertainties will, for open-loop systems result in inaccurate output or low performance, however this can be mitigated in a closed-loop system. As the process $G(s)$ undergoes a change $\Delta G(s)$ a resulting change in the tracking error $\Delta E(s)$ will occur. For an open-loop system this will be

$$\Delta E(s) = \Delta G(s) G_c(s) R(s)$$

and for a closed loop system it will be:

$$\begin{aligned}
\Delta E(s) &= \frac{-G_c(s)\Delta G(s)}{(1+G_c(s)G(s)+G_c(S)\Delta G(s))\,(1+G_c(s)G(s))}R(s)\\
&\approx \frac{-G_c(s)\Delta G(s)}{(1+G_c(s)G(s))^2}R(s)\\
&= \frac{-G_c(s)\Delta G(s)}{(1+L(s))^2}R(s)\\
&\approx -\frac{1}{L(s)}\frac{\Delta G(s)}{G(s)}R(s).
\end{aligned}$$

I.e. for a closed-loop system the error change caused by plant uncertainties is proportional to the relative plant change $\Delta G/G$ divided by the loop gain $L(s)$. So bigger changes in input result in bigger changes in output but a bigger loop gain results in a smaller error change. If e.g. $|L(s)| = 100$, then the effect of plant variation is suppressed by about a factor 100 at that frequency.

### 7.3.1 Sensitivity

We can define system sensitivity $S$ as

$$S = \frac{\Delta T(s)/T(s)}{\Delta G(s)/G(s)}$$

where

$$T(s) = \frac{Y(s)}{R(s)}$$

is the closed-loop transfer function from reference to output. Sensitivity tells how much the transfer function changes, in percentage, when the plant changes by a percentage amount. So if $S = 0{,}1$, then a 10% change in $G$ causes only about a 1% change in $T$. This also means that a smaller sensitivity is generally better than a higher one.

For very small changes, percent changes become derivatives:

$$S = \frac{\partial T(s)/T(s)}{\partial G(s)/G(s)} = \frac{\partial \ln T}{\partial \ln G}$$

which is the infinitesimal version of the definition of sensitivity.

We also define $S_G^T$ as the sensitivity of $T$ with respect to $G$. For an open-loop system, then

$$T(s) = G_c(s)G(s) \implies S_G^T = 1$$

as $T$ is directly proportional to $G$. Meaning a 1% change in $G$ results in a 1% change in the system transfer function $T$, so the loop is fully sensitive. For a closed-loop with unity feedback

$$T(s) = \frac{G_c(s)G(s)}{1+G_c(s)G(s)} \implies S_G^T = \frac{1}{1+G_c(s)G(s)} = \frac{1}{1+L(s)}$$

which is exactly the same as the usual sensitivity function.

Instead of asking how much $T$ changes with the whole plant $G$, we can also ask how $T$ changes with some parameter $\alpha$ such as mass, damping, stiffness, time constant, gain constant, etc. Then:

$$S_\alpha^T = S_G^T S_\alpha^G$$

following the chain-rule. This can be read as the first parameter, $\alpha$ changes the plant $G$ which then changes the system transfer $T$. I.e. total sensitivity from $\alpha$ to $T$ is

$$(\text{sensitivity of } T \text{ to } G) \cdot \big(\text{sensitivity of } G \text{ to } \alpha\big).$$

We can also write

$$S_\alpha^T = \frac{\partial \ln T}{\partial \ln \alpha}$$

and if

$$T(s, \alpha) = \frac{N(s, \alpha)}{D(s, \alpha)}$$

then

$$S_\alpha^T = \frac{\partial \ln N}{\partial \ln \alpha} - \frac{\partial \ln D}{\partial \ln \alpha} = S_\alpha^N - S_\alpha^T.$$

## 7.4 Disturbance Rejection in Feedback Systems

A disturbance signal is an unwanted signal that affects the output signal. From

$$E(s) = \frac{1}{1 + L(s)} R(s) - \frac{G(s)}{1 + L(s)} T_d(s) + \frac{L(s)}{1 + L(s)} N(s)$$

with $R(s) = N(s) = 0$ we see the disturbance contribution is

$$E_T(s) = -S(s)G(s)T_d(s) = -\frac{G(s)}{1 + L(s)} T_d(s)$$

as mentioned before, large loop gain $L(s) = G_c(s)G(s)$ leads to good disturbance rejection, and disturbance signals are often low frequency so it is desirable to design $G_c(s)$ such that $S(s)$ is small at low frequencies.

To imagine how disturbance arises imagine e.g. a steel bar being moved and rolled by a motor trying to maintain constant speed for the steel bar. However when the load changes (e.g. due to adding weight to the bar, or drying out of lubricant) the required torque changes too. That change acts like a disturbance torque $T_d$. In this scenario, if the controller command is unchanged the rotational speed $\omega$ can drop.

### 7.4.1 Open-loop Speed Under Disturbance

We now want to determine what the speed change of a motor might be due to load disturbance. For the open-loop motor model on Figure 7.4, the error becomes

$$E(s) = -\omega(s) = \frac{1}{Js + b + \frac{K_m K_b}{R_a}} T_d(s).$$

The sign here is negative because a positive load torque opposes motion, so it reduces speed.

If the disturbance is a step torque

$$T_d(s) = \frac{D}{s}$$

then by the final value theorem

$$\lim_{t \to \infty} E(t) = \lim_{s \to 0} sE(s)$$

which gives

$$E(\infty) = \frac{D}{b + \frac{K_m K_b}{R_a}}$$

and since $E = -\omega$

$$\omega(\infty) = -\frac{D}{b + \frac{K_m K_b}{R_a}}.$$

This means that a constant extra load torque creates a constant steady speed drop. I.e. in open-loop, the motor cannot fully reject the step disturbance, it just settles to a lower speed.

### 7.4.2 Closed-loop Speed Under Disturbance

Now, we introduce a feedback controller to reduce the loading disturbance effect as shown on Figure 7.3. Here, we also reorganize the system as shown on Figure 7.3, into

$$G_1(s) = \frac{K_a K_m}{R_a}, \qquad G_2(s) = \frac{1}{Js+b}, \qquad H(s) = K_t + \frac{K_b}{K_a}.$$

The disturbance $T_d(s)$ is injected between $G_1$ and $G_2$ and the output is fed back through $H(s)$.

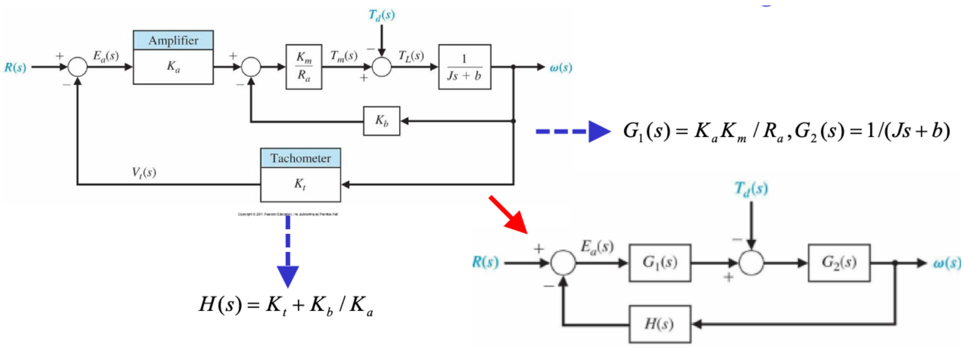



Figure 7.3: Closed-loop model.

For disturbance only, the steady-state error expression for the closed-loop system becomes:

$$E(s) = -\omega(s) = \frac{G_2(s)}{1 + G_1(s)G_2(s)H(s)} T_d(s)$$

which has the same structure as before:

$$\text{disturbance effect} = \frac{\text{plant term}}{1 + \text{loop gain}}.$$

If

$$G_1(s)G_2(s)H(S) \gg 1$$

then

$$E(s) \approx \frac{1}{G_1(s)H(s)} T_d(s)$$

i.e. for large feedback the disturbance is not much dominated by $G_2$; it is instead suppressed roughly in inverse proportion to the loop gain. Also

$$G_1(s)H(s) = \frac{K_a K_m}{R_a}\left(K_t + \frac{K_b}{K_a}\right)$$

and for $K_a \gg K_b$

$$G_1(s)H(s) \approx \frac{K_1 K_m K_t}{R_a}$$

so increasing amplifier gain $K_a$ improves disturbance rejection.

With $R(s) = 0$, the closed-loop speed under disturbance becomes

$$\omega_c(s) = -\frac{1}{Js + b + \frac{K_m}{R_a}(K_t K_a + K_b)} T_d(s).$$

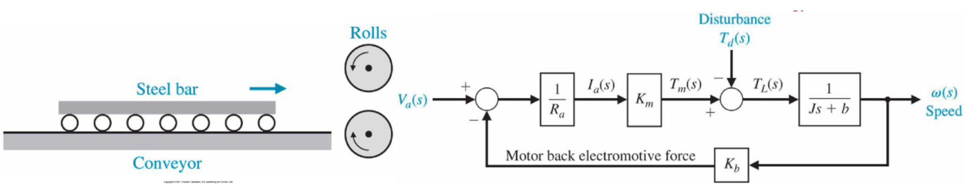



Figure 7.4: Open-loop motor model.

Comparing this with the open-loop version

$$\omega_o(s) = -\frac{1}{Js + b + \frac{K_m K_b}{R_a}} T_d(s)$$

we notice that the closed-loop version has an extra stabilizing term $K_t K_a K_m / R_a$ in the denominator.

For a step disturbance $T_d(s) = D/s$, the final value theorem gives

$$\omega(\infty) = -\frac{D}{b + \frac{K_m}{R_a}(K_t K_a + K_b)}$$

for sufficiently large $K_a$, this is

$$\omega_c(\omega) \approx -\frac{R_a}{K_a K_m K_t} D.$$

Meaning the steady speed drop is now roughly proportional to $1/K_a$.

Taking the quotient $\omega_c/\omega_o$ we get

$$\frac{\omega_c(\infty)}{\omega_o(\infty)} = \frac{R_a b + K_m K_b}{K_a K_m K_t}$$

which is usually less than 0,02. That means the disturbance effect in closed loop systems can be less than 2% of what it would be in open loop.

## 7.5 Measurement Noise Attenuation

Measurement noise is noise added in the sensor or measurement path. The plant output may be fine, but the controller does not see the true output. Instead it sees

$$y_{\text{sensor}} = y + n$$

there $n$ is the noise. That noise is then fed back, so the controller may react to the noise as if it were a real signal. Only looking at the noise contribution to the error we see

$$E_n(s) = C(s)N(s) = \frac{L(s)}{1 + L(s)} N(s)$$

where $C(s) = L(s)/(1 + L(s))$ is the complementary sensitivity function. This is complementary to the sensitivity function from above and a trade-off between them is therefore made at a given frequency.

## 7.6 The Transient Response Control

The transient response is the response of a system as a function of time, and must be adjusted until it is satisfactory. For an open-loop system, the process $G(s)$ must be replaces with a more suitable process, or design the cascade controller $G_c(s)$. For a closed-loop system, feedback loop parameters can often be adjusted to yield the desired response.

## 7.7 Steady-State Error

The steady state error is the error after the transient response has decayed, leaving only the continuous response.

$$e_{ss} = \lim_{t\to\infty} e(t).$$

In an open loop, the output is just

$$Y(s) = G(s)R(s)$$

so the error is

$$E_o(s) = R(s) - Y(s) = (1 - G(s))R(s)$$

for a unit step input, $R(s) = 1/s$, so

$$E_o(s) = (1 - G(s))\frac{1}{s}.$$

Now, applying the final value theorem we get:

$$e_o(\infty) = \lim_{s\to 0} sE_o(s) = \lim_{s\to 0} s(1 - G(s))\frac{1}{s} = 1 - G(0).$$

I.e. the steady-state error depends on the DC gain $G(0)$. If $G(0) = 1$, then $e_o(\infty) = 0$, however $G(0) \neq 1$ leads to a nonzero steady state error.

For a closed loop system, the feedback means a much smaller error is produced. For unity feedback $H(s) = 1$, the error transfer function is

$$E_c(s) = \frac{1}{1 + G_c(s)G(s)}R(S).$$

For a unit step input

$$E_c(s) = \frac{1}{1 + G_c(s)G(s)}\frac{1}{s}.$$

Now, applying the final value theorem, we get

$$e_c(\infty) = \lim_{s\to 0} sE_c(s) = \lim_{s\to 0} s\left(\frac{1}{1 + G_c(s)G(s)}\frac{1}{s}\right)$$

so

$$e_c(\infty) = \frac{1}{1 + G_c(0)G(0)}$$

or using the DC loop gain:

$$e_c(\infty) = \frac{1}{1 + L(0)}.$$

So if the DC loop gain is large, then the steady-state error is small.

## 7.8 The Cost of Feedback

In general, feedback is very useful to make more accurate systems, however one must always keep the cost of adding said feedback in mind before deciding if it is worth the investment. Adding feedback leads to increased complexity, as sensors, which are expensive, noisy, and inaccurate, need to be added. The closed loop gain of a feedback system will also decrease by a factor of $1/(1 + G_c(s)G(s))$. Further, feedback introduces possible instability of the system, which does not arise for open-loop systems.

**Lecture 4: Performance of Feedback Control Systems** 27. March 2026

# 8 Performance of Feedback Control Systems

To really understand the following part on Performance of Feedback Control Systems, a few concepts must first be defined.

**Definition 6: Transient Response**

The part of the response that disappears with time.

**Definition 7: Steady-State Response**

The part of the response that exists for a long time following an input signal initiation.

**Definition 8: Design Specifications**

The measures of performance indicating the quality of the system to the designer.

## 8.1 Standard Test Input Signals

In control theory, we seldom know the exact real-world input a system will see. Instead, we opt to test the system using simple "benchmark" inputs whose responses are easy to analyze. This helps us answer a lot general questions about the response of a system. A few of the most common standard test input signals are:

- Step Input
- Ramp Input
- Parabolic Input
- Sinusoidal Input
- Impulse Input

The three most common time-domain test inputs are the step-input, the ramp-input, and the parabolic inputs. These are all depicted on Figure 8.1.

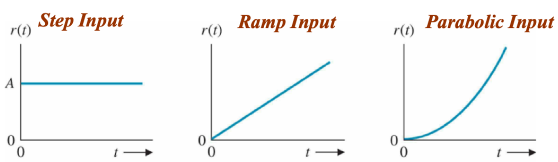



Figure 8.1: Common input signals.

### 8.1.1 Step Input

A step input jumps suddenly from 0 to a constant value $A$

$$r(t) = A, t > 0$$

and its Laplace transform is

$$R(s) = \frac{A}{s}.$$

This corresponds to a sudden physical command. Such as setting a motor to a high signal.

### 8.1.2 Ramp Input

A ramp input grows linearly with time

$$r(t) = At, t > 0$$

and its Laplace transform is

$$R(s) = \frac{A}{s^2}.$$

This corresponds to a steadily changing command, such as asking a vehicle to move forward at constant speed in position coordinates.

### 8.1.3 Parabolic Input

A parabolic input grows proportionally to $t^2$

$$r(t) = At^2, t > 0$$

and its Laplace transform is

$$R(s) = \frac{2A}{s^3}.$$

This corresponds to a constantly accelerating command, such as a position reference under constant acceleration.

Worth noting is that the Ramp Input is simply the integral of the Step Input; and the Parabolic Input is the integral of the Ramp Input.

### 8.1.4 Unit Impulse Response

The unit impulse response is defined with regard to the impulse $\delta(t)$. The ideal impulse is infinitely narrow and infinitely tall, but has area 1 (like an infinitely thin standard normal distribution). We define this approximately as:

$$\delta(t) = \begin{cases} \frac{1}{\epsilon}, & -\frac{\epsilon}{2} \leq t \leq \frac{\epsilon}{2} \\ 0, & \text{otherwise} \end{cases}$$

as $\epsilon \to 0$.

If the transfer function is $G(s)$, then the system's impulse response is:

$$g(t) = \mathscr{L}^{-1}\{G(s)\}.$$

This is an extremely important result, as it states that for an LTI system, the impulse response completely characterizes the system.

A convolution is defined as:

$$y(t) = \int_{-\infty}^{t} g(t-\tau)r(\tau)\,\mathrm{d}\tau.$$

I.e. the output is the convolution of the input $r(t)$ with the impulse response $g(t)$. In other words, any input can be built from many tiny impulses, and the output is the sum of the system's response to all those tiny impulses. In Laplace form this is

$$Y(s) = G(s)R(s)$$

which is the transform-domain version of convolution.

### 8.1.5 Sinusoidal Input

The sinusoidal input is defined by

$$r(t) = \begin{cases} A\sin\omega t, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

with Laplace transform:

$$R(s) = \frac{A\omega}{s^2+\omega^2}.$$

### 8.1.6 Standard Test Signal Input

In general, step, ramp, and parabolic inputs can be written as

$$r(t) = t^n$$

with Laplace transform

$$R(s) = \frac{n!}{s^{n+1}}$$

## 8.2 Performance of Second-Order Systems

We consider the general second-order system on Figure 8.2. Here, we have the open-loop transfer function

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

with unity negative feedback. This results in the closed-loop transfer function becoming

$$\frac{Y(s)}{R(s)} = \frac{G(s)}{1+G(s)} = \frac{\omega_n^2}{s^2+2\zeta\omega_n s+\omega_n^2}.$$

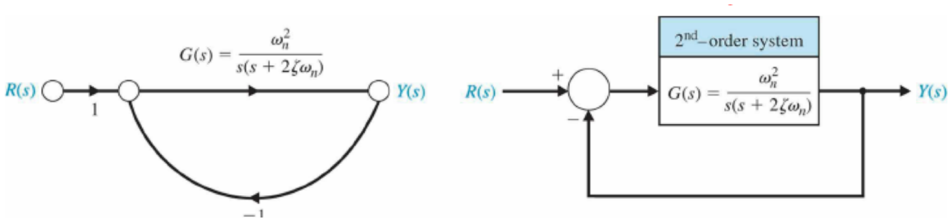



Figure 8.2: Second order system.

Now, we provide the system with a unit step input

$$R(s) = \frac{1}{s}$$

so

$$Y(s) = \frac{\omega_n^2}{s\left(s^2+2\zeta\omega_n s+\omega_n^2\right)}$$

and in time domain, for the underdamped case $0 < \zeta < 1$,

$$y(t) = 1 - \frac{1}{\beta}e^{-\zeta\omega_n t}\sin\left(\omega_n\beta t+\theta\right)$$

where

$$\beta = \sqrt{1-\zeta^2}, \qquad \theta = \cos^{-1}(\zeta).$$

This corresponds to a final value of 1 plus an oscillation whose amplitude decays with the exponential term. The factor

$$e^{-\zeta\omega_n t}$$

tells how fast the oscillation dies out, so larger $\zeta$ results in faster decay and less overshoot, whilst smaller $\zeta$ results in slower decay and more ringing.

For the above case, the poles are

$$s_{1,2} = -\zeta\omega_n \pm j\omega_n\sqrt{1-\zeta^2}.$$

Here, $-\zeta\omega_n$ determines how fast the oscillations die out; a more negative real part results in faster decay, and a real part closer to 0 results in slower decay. The imaginary part

$$\omega_n\sqrt{1-\zeta^2}$$

determines the oscillation frequency, i.e. it is the damped natural frequency

$$\omega_d = \omega_n\sqrt{1-\zeta^2}.$$

The pole plots on Figure 8.3, show how the roots move as $\zeta$ changes.

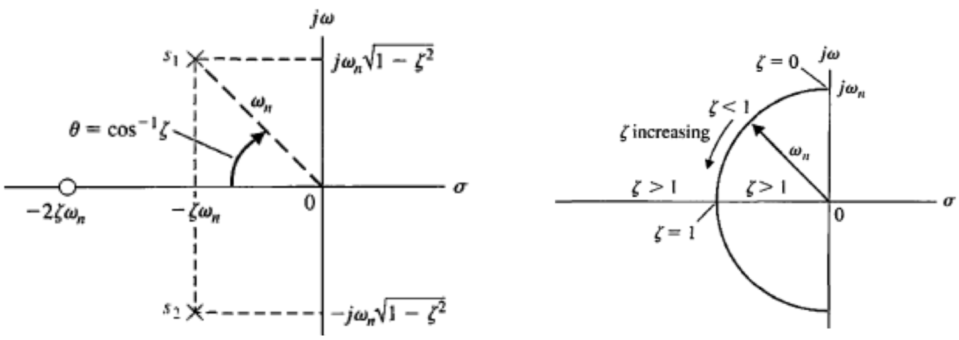



Figure 8.3: Pole plots

When $0 < \zeta < 1$, the system is underdamped, meaning the poles are complex conjugates

$$-\zeta\omega_n \pm j\omega d.$$

When $\zeta = 1$ the system is critically damped. Here, the two poles meet on the real axis.

When $\zeta > 1$, the system is overdamped. Here, the poles are real and distinct, and the responds has no oscillation.

For $\zeta = 0$, the system is undamped. This gives poles on the imaginary axis $\pm j\omega_n$, which correspond to sustained oscillation with no decay.

Also on Figure 8.3, a geometric relationship in the $s$-plane is shown. For poles of a standard second-order system the distance from the origin is $\omega_n$ and the angle is related to $\zeta$ by

$$\cos\theta = \zeta$$

so for fixed $\omega_n$, changing $\zeta$ moves the poles along a circle of radius $\omega_n$.

On Figure 8.4 the step time response is shown for different damping ratios. For the small damping ratios, we observe fast initial rise, large overshoot, and strong oscillation. For the medium damping ratio, we have some overshoot, and less oscillation. For critical damping, we have no oscillation, and no overshoot. For the overdamped case, we have no overshoot, and no overshoot, but with a sluggish response.

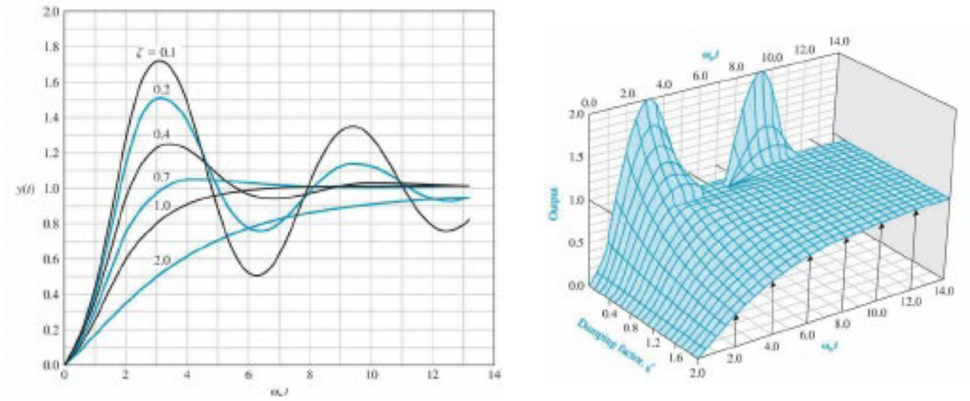


Figure 8.4: Time responses for different damping ratios

In general, we can observe that as the damping ratio decreases, the closed-loop roots approach the imaginary axis, and the response becomes increasingly oscillatory.

## 8.3 Performance Specification Definition

On Figure 8.5 a typical step response $y(t)$ approaching its final value $y(\infty)$ is depicted.

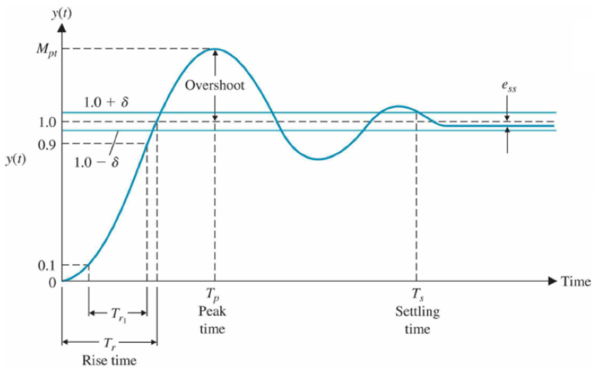



Figure 8.5: Typical step response $y(t)$ going toward a final value $y(\infty)$.

### 8.3.1 Rise Time

We define the rise time on Figure 8.5 as the time it takes the output to "rise" to its target value. For an underdamped system ($\zeta < 1$), the rise time $T_r$ is the first time the response reaches the final value $y(\infty)$. For an overdamped system ($\zeta > 1$), the response never overshoots, so rise time is often defined as the 10% to 90% time:

$$T_{r1} = T_2 - T_1$$

with

$$y(T_1) = 0{,}1\, y(\infty), \qquad y(T_2) = 0{,}9\, y(\infty).$$

### 8.3.2 Peak Time

The peak time on Figure 8.5, is defined as the time, when the response reaches its first maximum. At the peak, the slope is zero

$$\left. \frac{\mathrm{d}y(t)}{\mathrm{d}t} \right|_{t=T_p} = 0.$$

This metric only makes sense for an underdamped response, because overdamped responses do not have a peak above the final value.

### 8.3.3 Percent Overshoot

On Figure 8.5, the percent overshoot is defined as

$$\sigma_p\% = \frac{y(T_p) - y(\infty)}{y(\infty)} \cdot 100\%.$$

I.e. how much higher than the final value did the first peak go, as a percentage.

For the standard second-order system

$$\%\mathrm{OS} = 100e^{-\zeta\pi\frac{1}{\sqrt{1-\zeta^2}}}, \qquad 0 < \zeta < 1)$$

I.e. small $\zeta$ leads to lots of oscillation and a large overshoot and vice versa.

### 8.3.4 Settling Time

For an underdamped second-order system, the oscillations decay with an exponential envelope like

$$e^{-\zeta\omega_n t}.$$

A few different conventions exist for the settling time, the two most common of which being the 2% criterion and the 5% criterion. If you want your response to be within 2%, you use

$$e^{-\zeta\omega_n T_s} < 0{,}02$$

which gives the approximation

$$T_s \approx \frac{4}{\zeta\omega_n} = 3\tau$$

for the 2% criterion, with $\tau = 1/(\zeta\omega_n)$ being the exponential decay time constant.

Likewise, for the 5% criterion, a common approximation is

$$T_s \approx \frac{3}{\zeta\omega_n} = 3\tau.$$

### 8.3.5 Delay Time

The delay time, $T_d$ on Figure 8.5, is defined by

$$y(T_d) = 0{,}5y(\infty)$$

so $T_d$ is the time it takes the output to reach 50% of its final value.

### 8.3.6 Performance Compromise

Two important formulas from the above is the formula for the peak time and for the percentage overshoot:

$$T_P = \frac{\pi}{\omega_n\sqrt{1-\zeta^2}} \qquad\qquad \%\mathrm{OS} = 100e^{-\zeta\pi\frac{1}{\sqrt{1-\zeta^2}}}.$$

These are depicted on Figure 8.6. Here we see that, as the damping ratio $\zeta$ increases the percent overshoot decreases and the normalized peak time $\omega_n T_p$ increases, so more damping leads to less overshoot but also a slower response.

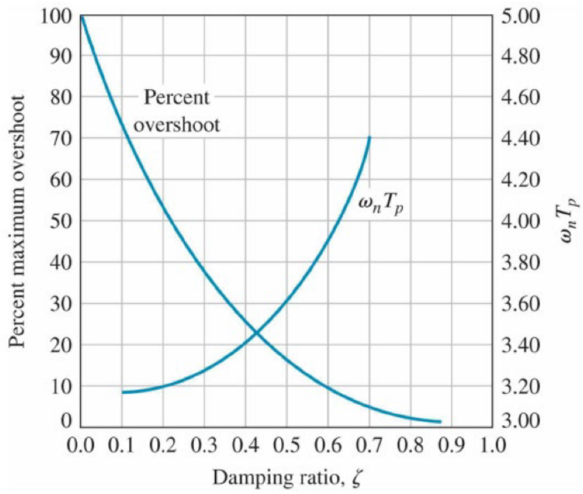



Figure 8.6: Performance compromise of control systems.

This is why a response cannot be made both extremely fast and extremely smooth with no overshoot without changing other elements of the design.

### 8.3.7 Swiftness of the System

On Figure 8.7 the normalized rise time $\omega_n T_{r1}$ is plotted versus the damping ratio $\zeta$. Also present, is a linear approximation of $T_{r1}$ stating:

$$T_{r1} \approx \frac{2{,}16\zeta + 0{,}60}{\omega_n}$$

I.e. increasing $\zeta$ increases rise time and increasing $\omega_n$ decreases rise time.

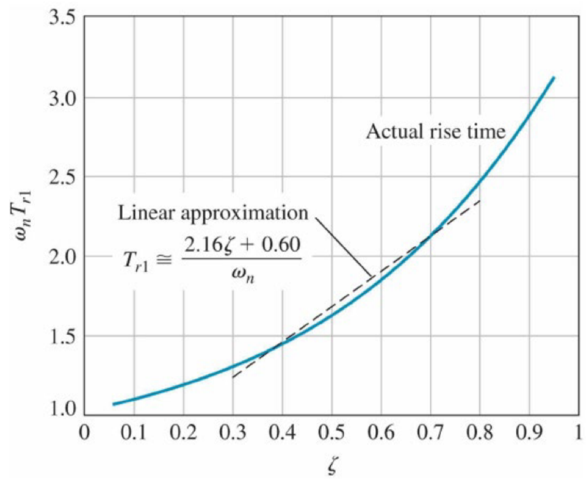



Figure 8.7: Normalized rise time $\omega_n T_{r1}$ versus $\zeta$ with linear approximation.

## 8.4 Effect of a Third Pole

We consider a third-order system with a closed-loop transfer function:

$$T(s) = \frac{1}{\left(s^2 + 2\zeta s + 1\right)\left(\gamma s + 1\right)}.$$

This is basically a second-order part with a complex pole pair and an extra third pole at

$$s = -\frac{1}{\gamma}.$$

If that third pole is far enough left in the $s$-plane, it will die out fast and not affect the response much, meaning the system behaves approximately like the second order part. In the above example, the dominant second-order poles have real part of approximately $-\zeta\omega_n$, and to be able to disregard the influence of the third pole, the dominant pole rule states that

$$\left|\frac{1}{\gamma}\right| \geq |10\zeta\omega_n|$$

i.e. the third pole, should be at least 10 times farther left than the real part of the dominant pole pair. The size of $1/\gamma$ and therefore the location of the third root changes the settling time and overshoot quite a bit as shown on Figure 8.8.

| $\gamma$ | $\frac{1}{\gamma}$ | Percent Overshoot | Settling Time* |
|---|---|---|---|
| 2.25 | 0.444 | 0 | 9.63 |
| 1.5 | 0.666 | 3.9 | 6.3 |
| 0.9 | 1.111 | 12.3 | 8.81 |
| 0.4 | 2.50 | 18.6 | 8.67 |
| 0.05 | 20.0 | 20.5 | 8.37 |
| 0∞ | 20.5 | 8.24 | |

*Note: Settling time is normalized time, $\omega_n T_s$ and uses a 2% criterion.

Figure 8.8: Effect of a third pole depending on its location.

## 8.5 Effect of a Zero

We now consider a system with an added zero, with transfer function

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

Here, the extra zero is at $s = -a$.

A zero near the dominant poles can strongly change the transient response. In other words, if finite zeros are located relatively near the dominant complex poles, they materially affect the transient response.

A zero near the poles typically makes the step response rise faster, overshoot more, and become more "peaky', as shown on Figure 8.9.

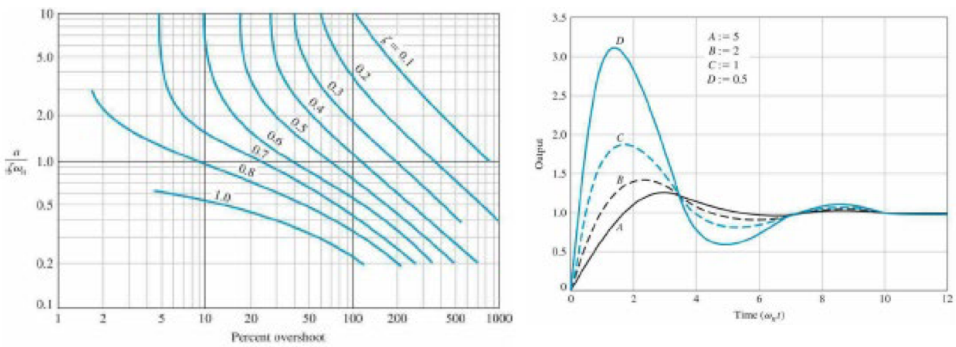



Figure 8.9: Effect of adding a zero.

Intuitively, a pole stores/slows energy, and a zero tends to add a kind of anticipatory or differentiating effect to this.

## 8.6 Specification Selections

We consider the system on Figure 8.10 with forward path

$$G(s) = \frac{K}{s(s+p)}$$

and unity feedback, yielding the closed-loop transfer function

$$T(s) = \frac{G(s)}{1+G(s)} = \frac{K}{s^2+ps+K}$$

which is on the same form as a standard second-order system

$$T(s) = \frac{\omega_n^2}{s^2+2\zeta\omega_n s+\omega_n^2}.$$

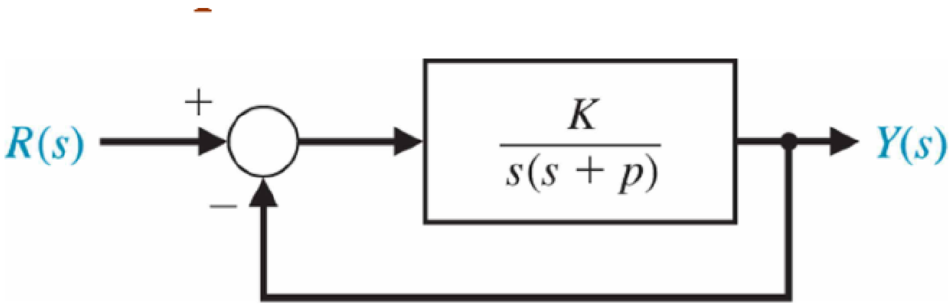



Figure 8.10: Example system.

Given this system, we want the step response to have less than 5% overshoot and a 2% settling time lower than 4 s.

For a standard second-order system, percent overshoot depends on $\zeta$ by

$$\%\text{OS} = 100e^{-\pi\zeta\frac{1}{\sqrt{1-\zeta^2}}}.$$

From Figure 8.11, an overshoot of about 4,3 % corresponds to $\zeta = 0{,}707$.

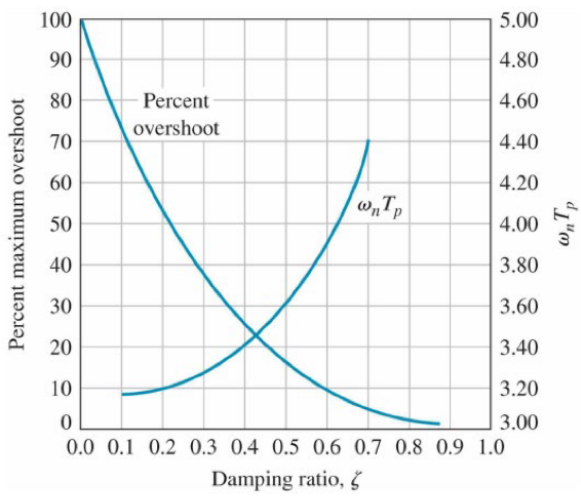



Figure 8.11: Percentage overshoot versus damping ratio.

For the 2% settling time, the standard approximation is

$$T_s \approx \frac{4}{\zeta\omega_n}.$$

We require $T_s \le 4\,\mathrm{s}$, so

$$\frac{4}{\zeta\omega_n} \le 4 \implies \zeta\omega_n \ge 1.$$

This gives the two conditions $\zeta \ge 0{,}707$ and $\zeta\omega_n \ge 1$. For a standard underdamped second-order system, the poles are

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

i.e. $\zeta$ controls the angle of the poles and $\zeta\omega_n$ controls the real part.

On the pole plot on Figure 8.12 the 45° lines correspond to $\zeta = 0{,}707$ and the vertical line is $\zeta\omega_n = 1$, meaning real part $-1$. The acceptable region is therefore to the left of $\Re(s) = -1$ inside the wedge corresponding to $\zeta \ge 0{,}707$

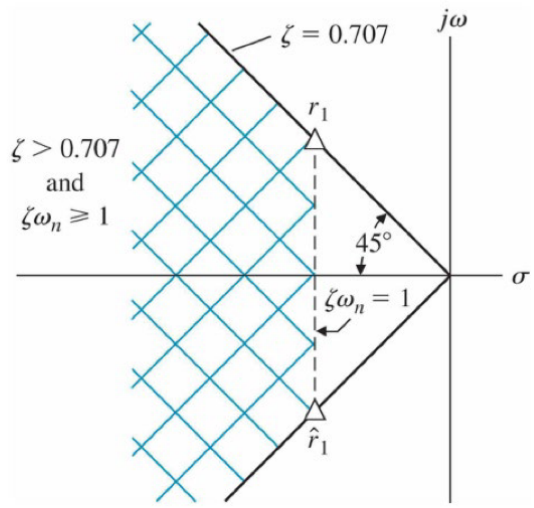



Figure 8.12: Pole plot

We choose the boundary case $\zeta = 0{,}707, \zeta\omega_n = 1$, so

$$\omega_n = \frac{1}{\zeta} = \frac{1}{0{,}707} = \sqrt{2}.$$

From earlier

$$T(s) = \frac{K}{s^2 + ps + K}.$$

Comparing this with

$$\frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2}.$$

Matching coefficients gives

$$K = \omega_n^2, \qquad p = 2\zeta\omega_n.$$

Substituting, we get:

$$K = \left(\sqrt{2}\right)^2 = 2$$

and since $\zeta\omega_n = 1$, $p = 2\zeta\omega_n = 2$, so

$$K = 2, \qquad p = 2.$$

### 8.6.1 Estimating the Damping Ratio

The dampind ratio, $\zeta$, can be estimated, just by looking at a step response. In general the number of visible oscillations is:

$$\text{Cycles visible} = \frac{4\beta}{2\pi\zeta} \approx \frac{0{,}55}{\zeta}.$$

So approximately

$$\zeta \approx \frac{0{,}55}{\text{cycles visible}}.$$

## 8.7 s-Plane Root Location and Transient Response

A transfer-function response can be decomposed into terms like

$$\frac{A_i}{s+\sigma_i} \quad \text{and} \quad \frac{B_k s + C_k}{s^2 + 2\alpha_k s + \left(\alpha_k^2 + \omega_k^2\right)}.$$

In time domain these become real exponential terms like $e^{-\sigma t}$ and oscillatory ones like $e^{-\alpha t}\sin(\omega t + \theta)$. The step-response of different root locations in the $s$-plane is shown on Figure 8.13.

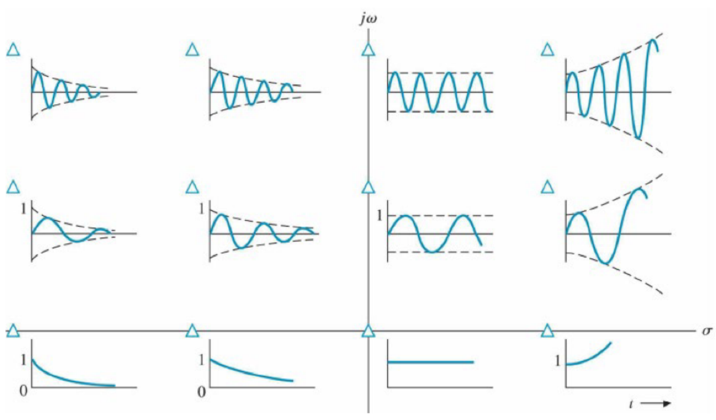



Figure 8.13: Impulse response for various root locations in $s$-plane

On the $s$-plane, poles are at $s = \sigma + j\omega$, the real part $\sigma$ decides whether the response decays ($\sigma < 0$), stays constant ($\sigma = 0$), or grows ($\sigma > 0$). The imaginary part gives oscillation frequency.

## 8.8 Steady-state Error

We consider the system on Figure 8.14 with:

- Reference input $R(s)$
- Error $E_a(s)$
- Controller $G_C(s)$
- plant $G(s)$
- Output $Y(s)$

Because of unity feedback,

$$E(s) = R(s) - Y(s)$$

and since

$$Y(s) = G_C(s)G(s)E(s)$$

we get

$$E(s) = \frac{1}{1 + G_C(s)G(s)} R(s).$$

This is the error-transfer formula.

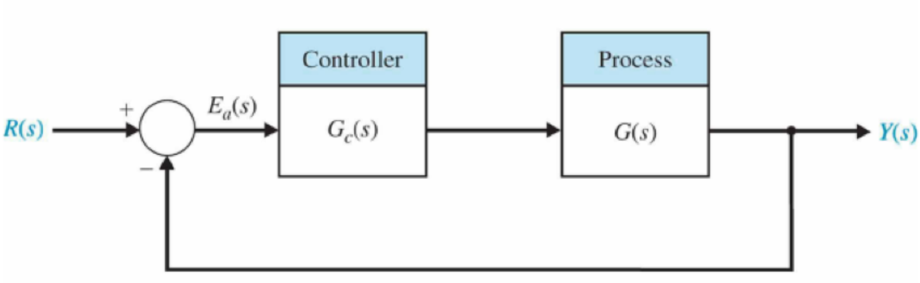



Figure 8.14: Example unity-feedback control system.

The steady-state error is the error after the transient has died out:

$$e_{ss} = \lim_{t\to\infty} e(t).$$

Using the final value theorem:

$$e_{ss} = \lim_{s\to 0} sE(s).$$

So, with the previous expression for $E(s)$

$$e_{ss} = \lim_{s\to 0} s \cdot \frac{1}{1 + G_c(s)G(s)} R(s). \tag{1}$$

Here, it is worth noting that $s \to 0$ means low-frequency / DC-behavior, so steady-state error depends on the system's behavior near $s = 0$.

### 8.8.1 Steady-State Error for Step Input

For a stem of magnitude $A$,

$$r(t) = A \qquad \Longrightarrow \qquad R(s) = \frac{A}{s}.$$

Plugging this into Equation (1) we get:

$$e_{ss} = \lim_{s\to 0} s \frac{1}{1 + G_C(s)G(s)} \cdot \frac{A}{s} = \frac{A}{1 + \lim_{s\to 0} G_C(s)G(s)}.$$

So, the step error depends on the DC loop gain.

The loop transfer function is written generally as:

$$G_C(s)G(s) = \frac{K \prod_{i=1}^{M} (s + z_i)}{S^N \prod_{k=1}^{Q} (s + p_k)}.$$

The $S^N$ factor in the denominator ensures we have $N$ poles at the origin, i.e. $N$ *pure integrators.* This $N$ is called the type number.

For a type 0 system ($N = 0$), $G_C(s)G(s)$ stays finite as $s \to 0$. We define the position error constant as

$$K_p = \lim_{s\to 0} G_c(s)G(s)$$

then

$$e_{ss} = \frac{A}{1+K_p}.$$

Meaning a type-0 system usually has a nonzero constant error to a step. Increasing low-frequency gain increases $K_p$, which reduces error, but it does not make it zero.

If the system has one or more integrators ($N \geq 1$), then $\lim_{s\to 0} G_C(s)G(s) = \infty$, so

$$e_{ss} = 0$$

for a step input. So, type-0 will have a finite step error and type-1 or higher will have zero step error.

### 8.8.2 Steady-State Error for Ramp Input

For a ramp with slope $A$,

$$r(t) = A(t) \qquad \Longrightarrow R(s) = \frac{A}{s^2}$$

then Equation (1) becomes

$$e_{ss} = \lim_{s\to 0} s \cdot \frac{1}{1+G_C(s)G(s)} \cdot \frac{A}{s^2},$$

which simplifies to a behavior controlled by $\lim_{s\to 0} sG_C(s)G(s)$. We now define the velocity error constant

$$K_v = \lim_{s\to 0} sG_C(s)G(s)$$

then for a type-1 system,

$$e_{ss} = \frac{A}{K_v}.$$

A type-0 system has no integrator, then as $sG_C(s)G(s) \to 0$, $e_{ss} = \infty$, meaning a type-0 system cannot follow a ramp forever, the error keeps growing infinitely.

For a type-1 system $K_v$ is finite, so

$$e_{ss} = \frac{A}{K_v},$$

meaning the output follows the ramp with the same slope, but with a constant offset.

A type-2 or higher system will mean $K_v \to \infty$, so

$$e_{ss} = 0$$

meaning type-2+ can track a ramp perfectly in steady-state.

### 8.8.3 Steady-State Error for Accelerating/Parabolic Input

Here, the input is

$$r(t) = \frac{At^2}{2} \qquad \Longrightarrow \qquad R(s) = \frac{A}{s^3}$$

meaning the relevant term is

$$\lim_{s\to 0} s^2 G_C(s)G(s).$$

We now define the acceleration error constant $K_a = \lim_{s\to 0} s^2 G_C(s)G(s)$, and for a type-2 system we get, by insertion into Equation (1):

$$e_{ss} = \frac{A}{K_a}.$$

For a type-0 or type-1 system we do not have enough integrators to track a parabolic input so $e_{ss} = \infty$.

A type-2 system gives two integrators, so

$$e_{ss} = \frac{A}{K_a}$$

corresponding a constant finite error.

Type-3 or higher systems, gives

$$e_{ss} = 0$$

corresponding to a perfect steady-state tracking of a parabolic input.

## 8.9 Performance Indices

A performance index is a single number used to measure how good or bad a system response is. Usually in control, we care about how large the error is, how long the error lasts, how much overshoot there is, how quickly the system settles, etc. Instead of judging all of these from a graph, we define a mathematical quantity based on the error $e(t)$, and then try to minimize it, so a small performance index corresponds to a better response and a large performance index corresponds to a worse response.

An optimum control system is one where the system parameters are chosen so that the chosen performance index becomes as small as possible. E.g. one might tune controller gains, $K_p$, $K_i$, $K_d$, so that one of these errors is minimized.

All of the indices are based on the tracking error

$$e(t) = r(t) - y(t)$$

Some common performance indices are presented in the following parts.

### 8.9.1 Integral of the Square of the Error (ISE)

The integral of the square of the error, is defined as

$$\text{ISE} = \int_0^T e^2(t)\,\mathrm{d}t$$

meaning you square the error and add it up over time. This tends to favor responses that reduce large errors quickly, but as small errors are not weighted as much oscillation is not impeded much.

### 8.9.2 Integral of the Absolute Error (IAE)

The integral of the absolute error, is defined as

$$\text{IAE} = \int_0^T |e(t)|\,\mathrm{d}t$$

meaning you take the absolute value of the error and add it over time. This does not exaggerate large errors as much as ISE and generally treats errors more “linearly”.

### 8.9.3 Integral of Time Multiplied by Absolute Error (ITAE)

The integral of the time multiplied by absolute error, is defined as

$$\text{ITAE} = \int_0^T t\,|e(t)|\,\mathrm{d}t$$

meaning the absolute error is now multiplied by the time $t$ before being integrated. I.e. errors at the beginning count less and errors later count more. This generally helps to reduce settling time and long-lasting oscillation.

#### 8.9.4 Integral of Time multiplied by the Squared Error (ITSE)

The integral of time multiplied bu the squared error is defined as

$$\mathrm{ITSE} = \int_0^T t e^2(t)\,\mathrm{d}t.$$

This combines the ideas of squaring the error to emphasise large errors, and multiplying by the time to emphasise late errors.

#### 8.9.5 The factor T

The factor $T$ in all the above integrals is usually chosen as the settling time $T_s$. This means the integral is often evaluated only over a practical time window. In theory, people sometimes also use

$$\int_0^\infty \ldots\,\mathrm{d}t$$

if the response dies out and the integral remains finite.

#### 8.9.6 Simplification of Linear Systems

We have two main methods if we want to approximate a high-order transfer function as a lower-order one that keeps the most important behavior.

The first method is deleting an insignificant pole. E.g.

$$G(s) = \frac{K}{s(s+2)(s+30)}$$

with poles at $s = 0, -2, -30$. Here the pole at $-30$ corresponds to a very fast mode that dies out much faster than the other two, so for the slower response, we approximate

$$s + 30 \approx 30$$

which gives

$$G(s) \approx \frac{K}{30s(s+2)} = \frac{K/30}{s(s+2)}.$$

This is called a *dominant-pole approximation.*

The second method is matching the frequency response. Instead of just deleting a pole, you can instead build a lower-order model whose frequency response is close to the original one. Say for a high-order system

$$G_H(s) = K\frac{a_m s^m + a_{m-1}s^{m-1} + \ldots + a_1 s + 1}{b_n s^n + b_{n-1}s^{n-1} + \ldots + b_1 s + 1}$$

we can create a lower-order approximation

$$G_L(s) = K\frac{c_p s^p + \ldots + c_1 s + 1}{d_q s^q + \ldots + d_1 s + 1}$$

by choosing the $c_i$ and $d_i$ so that $G_L(s)$ behaves as much like $G_H(s)$ as possible in frequency. In general, to match frequency response, the ratio

$$\frac{G_H(s)}{G_L(s)} = \frac{M(s)}{\Delta(s)} \to 1$$

I.e. if $G_L(s)$ is a good approximation, then the ratio $G_H(s)/G_L(s)$ should be close to 1. So the goal is to make the numerator and denominator of that ratio behave similarly. We can match using,

$$M_{2q} = \Delta_{2q}$$

for $q = 0, 1, 2, \ldots$ I.e. you compare certain coefficients derived from $M(s)$ and $\Delta(s)$ and force them to be equal.

**Example: Frequency Response Matching**

We seek to approximate the system

$$G_H(s) = \frac{6}{s^3 + 6s^2 + 11s + 6}$$

by a second-order model

$$G_L(s) = \frac{1}{1 + d_1 s + d_2 s^2}.$$

We start by dividing the denominator by 6

$$G_H(s) = \frac{1}{1 + \frac{11}{6}s + s^2 + \frac{1}{6}s^3}.$$

Now, we define

$$\Delta(s) = 1 + \frac{11}{6}s + s^2 + \frac{1}{6}s^3$$

and for the approximate model

$$M(s) = 1 + d_1 s + d_2 s^2$$

because

$$\frac{G_H(s)}{G_L(s)} = \frac{1/\Delta(s)}{1/M(s)} = \frac{M(s)}{\Delta(s)}.$$

Now we compute the derivatives at $s = 0$. For $\Delta(s)$:

$$\Delta^{(0)}(0) = 1, \quad \Delta^{(1)}(0) = \frac{11}{6}, \quad \Delta^{(2)}(0) = 2, \quad \Delta^{(3)}(0) = 1$$

and for $M(s)$:

$$M^{(0)}(0) = 1, \quad M^{(1)}(0) = d_1, \quad M^{(2)}(0) = 2d_2, \quad M^{(3)}(0) = 0.$$

Now, we match the coefficients, by first computing

$$M_2 = d_1^2 - 2d_2$$

and

$$\Delta_2 = \frac{49}{36}.$$

So matching gives

$$d_1^2 - 2d_2 = \frac{49}{36}.$$

The next matching condition gives

$$M_4 = \Delta_4$$

which becomes

$$d_2^2 = \frac{7}{18} \implies d_2 \approx 0{,}624 \implies d_1 1{,}62.$$

So

$$G_L(s) \approx \frac{1}{1 + 1{,}62s + 0{,}624s^2}.$$

**Lecture 5: Stability of Linear Feedback Systems** 7. April 2026

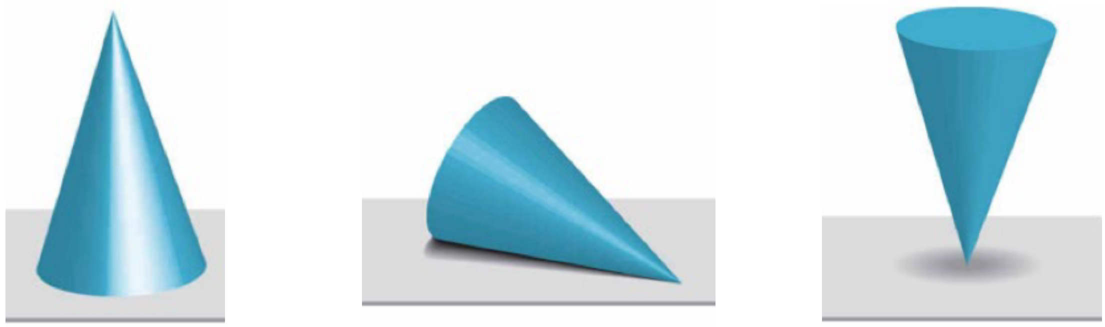
# 9 Stability of Linear Feedback Systems

## 9.1 Stability Definition

Stability is often the most important concept in control systems. A system might have a nice speed, accuracy, and disturbance rejection, but none of that matters if the system is unstable, because the output can then grow uncontrollably.

**Definition 9: Stable System**

A stable system is any system that gives a bounded output for any bounded input. This is also called *BIBO*-stability (Bounded-Input, Bounded-Output). Stability is illustrated by the different cones on Figure 9.1.



(a) Stable (b) Neutral (c) Unstable

Figure 9.1: a cone in different configurations corresponding to a stable, neutral, and unstable system.

On Figure 9.1, a cone is shown in three different configurations. Figure 9.1a shows a cone standing on its base. If you disturb it a little, it stays in a reasonable state (on its base). Here, the system does not "run away" due to a small input. On Figure 9.1b a cone lying on its side is shown. Here, if you move it, it stays where you put it rather than returning or diverging (ideally). This corresponds to marginal/neutral stability. Figure 9.1c represents an unstable system by a cone balanced on its tip. Here, any small disturbance makes the cone fall away.

## 9.2 Stability Analysis with s-plane Approach

On Figure 8.4 a correlation between different pole locations and the corresponding time responses was shown. This is summarized on Figure 9.2.

Here, we see that a system is stable if all poles have negative real parts, and the system is unstable if any pole has positive real part. Marginally stable systems are the special case, where poles are on the imaginary axis (real part equal to zero) and they are simple poles, while all other poles have negative real parts. This happens, as a pole contributes a term like

$$e^{st} = e^{(\sigma + j\omega)t} = e^{\sigma t} e^{j\omega t}$$

so, if $\sigma < 0$, $e^{\sigma t}$ decays with time and vice versa. For $\sigma = 0$, $e^{\sigma t} = 1$ so the response will have sustained oscillation with no decay or growth.

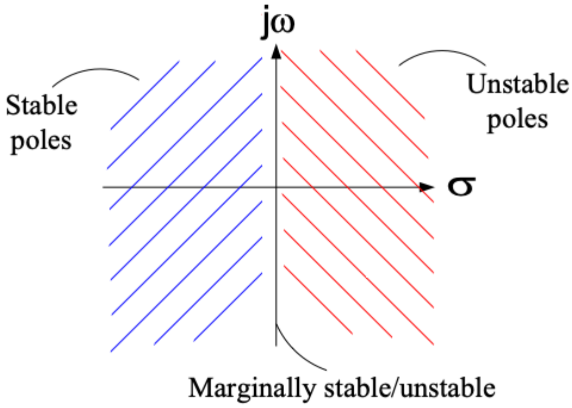



Figure 9.2: Stability of poles depending on their location in the $s$-plane

### 9.2.1 Relative Stability of Feedback Control Systems

Stability is rarely binary. In general, we work with two different ideas: Absolute stability and relative stability.

Absolute stability is a simply yes/no question: "Is the system stable or not?". If all poles are in the left half-plane, the system is absolutely stable.

Relative stability instead concerns itself with how stable a system is, how close the poles are to instability, and how oscillatory the response is.

These two concepts are compared on Figure 9.3.

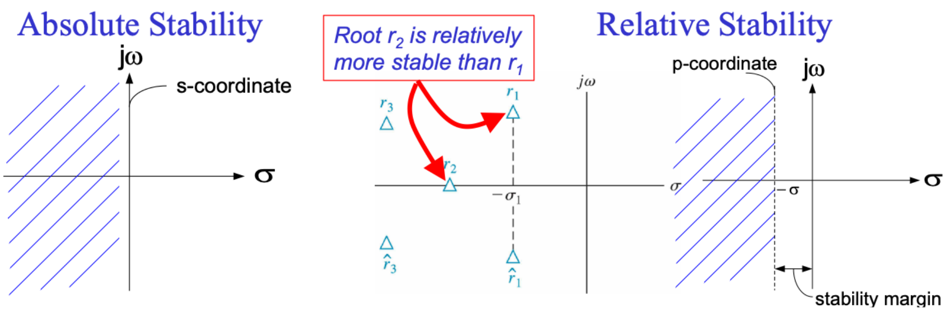



Figure 9.3: Comparison of different ideas of stability.

On Figure 9.3, it is stated that $r_2$ is relatively more stable than $r_1$. This is true, as $r_2$ is farther left in the $s$-plane, so its real part is more negative, meaning it is associated with a faster response decay.

The right figure on Figure 9.3 shows a vertical line left of the imaginary axis. The distance from that line to the imaginary axis is called the stability margin. Here, we do not just want poles barely in the left half-plane. We want them *comfortably* to the left, so the system remains well-behaved even with modelling errors or parameter variations.

For complex poles, relative stability is also described by the damping ratio. Higher damping ratios leads to less oscillation and better settling corresponding to better relative stability.

## 9.3 Routh-Hurwitz Stability Criterion

For a closed-loop LTI system, stability is determined by the roots of the characteristic polynomial

$$q(s) = a_n s^n + a_{n-1} s^{n-1} + \ldots + a_1 s + a_0$$

these roots are the closed loop poles.

The Routh-Hurwitz method is a way to determine whether the roots of the characteristic equation have negative real parts just by looking at the coefficients of said equation. Here, we use the Hurwitz's necessary condition stating that if $a_n > 0$, then for stability all coefficients must be positive. This is only a necessary and not a sufficient condition. I.e. a stable system will necessarily have positive coefficients to the characteristic equation, but having positive coefficients does not mean a system is necessarily stable. This is exactly why Routh-Hurwitz is needed as it gives a stronger test.

### 9.3.1 Routh Array

To use the Routh-Hurwitz method, we start by building a so-called Routh array from the polynomial coefficients. For

$$a_n s^n + a_{n-1} s^{n-1} + a_{n-2} s^{n-2} + \ldots + a_0.$$

The Routh array is defined as:

$$\begin{array}{c|cccc} s^n & a_n & a_{n-2} & a_{n-4} \\ s^{n-1} & a_{n-1} & a_{n-3} & a_{n-5} \\ s^{n-2} & b_{n-1} & b_{n-3} & b_{n-5} \\ s^{n-3} & c_{n-1} & c_{n-3} & n_{n-5} \\ \vdots & \vdots & \vdots & \vdots \\ s^0 & h_{n-1} & & \end{array}$$

where

$$\begin{aligned} b_{n-1} &= \frac{a_{n-1}a_{n-2} - a_n a_{n-3}}{a_{n-1}} = -\frac{1}{a_{n-1}} \begin{vmatrix} a_n & a_{n-2} \\ a_{n-1} & a_{n-3} \end{vmatrix} \\ b_{n-3} &= -\frac{1}{a_{n-1}} \begin{vmatrix} a_n & a_{n-4} \\ a_{n-1} & a_{n-5} \end{vmatrix} \\ c_{n-1} &= -\frac{1}{b_{n-1}} \begin{vmatrix} a_{n-1} & a_{n-3} \\ b_{n-1} & b_{n-3} \end{vmatrix}. \end{aligned}$$

I.e. the first two rows are arranged as $a_n, a_{n-2}, a_{n-4}$ and $a_{n-1,a_{n-3}}, a_{n-5}$, respectively and the next rows are then computed from the previous two.

The important result here, is then that the number of roots of $q(s)$ with positive real parts equals the number of sign changes in the first column of the Routh array.

**Example: Use of the Routh-Hurwitz Stability Criterion**

We are given a characteristic polynomial in factored form as

$$q(s) = \left(s - 1 + j\sqrt{7}\right)\left(s - 1 - j\sqrt{7}\right)(s+3).$$

From this we see that the roots are $s = 1 + \pm j\sqrt{7}, -3$ so the first two have real part $1 > 0$, so two roots are in the right-half plane, meaning the system is unstable. Expanding the polynomial we get:

$$q(s) = s^3 + s^2 + 2s + 24.$$

Here, all coefficients are positive, yet the system is still unstable, which directly proves the statement that the Hurwitz's necessary condition is not sufficient but only necessary.

For a cubic

$$a_3 s^3 + a_2 s^2 + a_1 s + a_0$$

the Routh array is

$$\begin{array}{c|cc} s^3 & a_3 & a_1 \\ s^2 & a_2 & a_0 \\ s^1 & b_1 & 0 \\ s^0 & c_1 & 0 \end{array}.$$

Here, $a_3 = 1, a_2 = 1, a_1 = 2, a_0 = 24$, so the first two rows are:

$$\begin{array}{c|cc} s^3 & 1 & 2 \\ s^2 & 1 & 24 \end{array}.$$

Now we compute

$$b_1 = \frac{a_2 a_1 - a_3 a_0}{a_2} = \frac{1 \cdot 2 - 1 \cdot 24}{1} = -22$$

so the last row becomes

$$c_1 = a_0 = 24$$

meaning the array is

$$\begin{array}{c|cc} s^3 & 1 & 2 \\ s^2 & 1 & 24 \\ s^1 & -22 & 0 \\ s^0 & 24 & 0 \end{array}.$$

Here, there are two sign changes, so we have two 2 roots with positive real part meaning the system is unstable, which matches the expected result from the introduction.

**Lecture 6: Root Locus Method, I** 9. April 2026

# 10 The Root Locus Method

## 10.1 Introduction and Concepts

For the unity-feedback system,

$$\frac{Y(s)}{R(s)} = \frac{KG(s)}{1 + KG(s)}$$

the closed-loop poles are the roots of the denominator, so they satisfy

$$1 + KG(s) = 0.$$

This equation is the characteristic equation.

So root locus means that we pick a value of $K$, solve $1 + KG(s) = 0$, plot the resulting closed-loop poles in the $s$-plane and then vary $K$ from 0 to $\infty$. The curve traced by these poles is the root locus.

We can rewrite the above characteristic equation as

$$KG(s) = -1$$

which is useful because $-1$ has magnitude 1 and angle 180° modulo 360°. This means we can split the condition into two parts:

$$|KG(s)| = 1$$
$$\angle KG(s) = 180^\circ + k360^\circ.$$

The key idea then is that changing $G$ changes the magnitude, but not the angle if $K > 0$. So the angle condition tells where the root locus can exist and the magnitude condition tells which value of $K$ places a pole at that point.

**Example: Example system**

We consider the example system on Figure 10.1 with open-loop transfer function

$$G(s) = \frac{1}{s(s+2)}$$

and gain $K$, so

$$KG(s) = \frac{K}{s(s+2)}.$$

Which means that the open-loop poles are at $s = 0$ and $s = -2$. There are no zeros. The two roots are where the root-locus branches start when $K = 0$.
Using

$$1 + KG(s) = 0 = 1 + \frac{K}{s(s+2)}$$

we multiply by $s(s+2)$ to get

$$s(s+2) + K = 0$$

so

$$s^2 + 2s + K = 0$$

which is the closed-loop characteristic equation. We can compare this to the standard second-order form

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

Matching coefficients gives:

$$2\zeta\omega_n = 2 \quad \text{and} \quad \omega_n^2 = K.$$

So

$$\zeta\omega_n = 1, \qquad \omega_n = \sqrt{K}.$$

Which means

$$\zeta = \frac{1}{\sqrt{K}}.$$

So as $K$ increases $\omega_n$ also increases, but $\zeta$ decreases.
Solving

$$s^2 + 2s + K = 0$$

gives

$$s_{1,2} = -1 \pm \sqrt{1-K}.$$

When $0 < K < 1$, $\sqrt{1-K}$ is real, so both poles are real. One will move rightward from $-2$ and one will move leftward from 0 approaching each other along the real axis.
When $K = 1$ $s = -1$ so both poles meet at $-1$ which is the breakaway point.
When $K > 1$, $\sqrt{1-K} = j\sqrt{K-1}$, so the poles are $s = -1 \pm j\sqrt{K-1}$, from where they move vertically upward

and downward, with constant real part $-1$.
Further, for this system

$$KG(s) = \frac{K}{s(s+2)}.$$

At a candidate point $s = s_1$, the angle condition is

$$\angle\frac{K}{s(s+2)} = -\angle s - \angle(s+2) = 180^\circ \mod 360^\circ.$$

Because there are no zeros, the total angle comes only from the poles. On the vertical line through $-1$, the geometry is symmetric, with one pole contributing angle $\theta$ and one contributing angle $180^\circ - \theta$, which means the total denominator angle is $180^\circ$, which means the angle condition is satisfied.
After we know a point is on the root locus from the angle condition, the magnitude condition gives the corresponding gain

$$\left|\frac{K}{s(s+2)}\right| = 1.$$

so

$$K = |s|\,|s+2|.$$

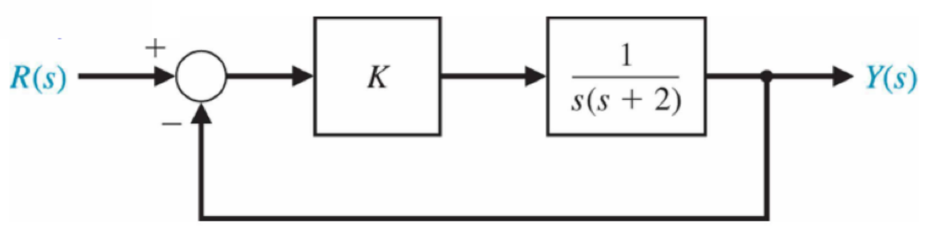



Figure 10.1: Example 2nd order control system

In general, if the loop transfer function is

$$F(s) = K\frac{\prod_{i=1}^{M}(s+z_i)}{\prod_{k=1}^{n}(s+p_k)}$$

then the characteristic equation is

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

and the root locus conditions become

$$K\left|\frac{\prod(s+z_i)}{\prod(s+p_k)}\right| = 1$$
$$\sum\angle(s+z_i) - \sum\angle(s+p_k) = 180^\circ + k360^\circ.$$

So zeros add angle, whilst poles subtract angle.

## 10.2 The Root Locus Procedure

The following part will outline the general procedure for drawing a root-locus sketch.

### 10.2.1 Prepare the Root Locus Sketch

We should start by rewriting the characteristic equation so the gain appears as a multiplier

$$1 + KP(s) = 0.$$

Then factor $P(s)$ into poles and zeros.

The branches will always start at poles and end at zeros as

$$1 + K\frac{\prod(s+z_i)}{\prod(s+p_k)} = 0$$

then, when $K = 0$ the equation reduces to the denominator being zero, so the roots are the open-loop poles, and when $K \to \infty$, the numerator dominates, so the roots approach the open-loop zeros. So the root locus begins at poles and ends at zeros. If there are fewer finite zeros than poles, then the extra branches go to zeros at infinity. Also, because the coefficients are real, the locus is symmetri about the real axis.

**Example: Preparing the Locus sketch**

We are given an example system

$$1 + \frac{K}{s^4 + 12s^3 + 64s^2 + 128s} = 0.$$

And we want to apply the above procedure to prepare a root locus sketch.

We factor the denominator:

$$\begin{aligned} s^4 + 12s^3 + 64s^2 + 128s &= s(s+4)\left(s^2+8s+32\right) \\ &= s\,(s+4))\left(s+4+4j\right)\left(s+4-4j\right). \end{aligned}$$

So the open loop poles are $s = 0, -4, -4+4j, -4-4j$, and there are no finite zeros. So the number of poles $n = 4$, the number of zeros $M = 0$, and the number of branches is 4.

### 10.2.2 Find the real-axis part of the locus

We have a general rule stating that a point on the real axis is on the root locus if the number of poles and zeros to its right is odd.

For the above example, we have 0 singularities to the right of 0, so anything to the right of 0 is not on the locus. Between $-4$ and 0, we have 1 singularity on the right so this part of the real axis is on the locus. To the left of $-4$, we have 2 singularities to the right so this part of the real axis is not on the locus. This means the only real-axis segment on the root locus is

$$-4 < s < 0.$$

### 10.2.3 Asymptotes to Infinity

Since there are more poles than zeros, some branches must go to infinity. The number of asymptotes is

$$n - M = 4 - 0 = 4.$$

The asymptote angles are

$$\phi_A = \frac{(2k+1)180^\circ}{n-M}, \qquad k = 0,1,2,3.$$

So here:

$$\phi_A = 45^\circ, 135^\circ, 225^\circ, 315^\circ.$$

The asymptote centroid is found as

$$\sigma_A = \frac{\sum \text{poles} - \sum \text{zeros}}{n-M}.$$

Here, the poles sum to $-12$ and the zeros sum to 0, so

$$\sigma_A = \frac{-12}{4} = -3.$$

So the four asymptotes all cross the real axis at $s = -3$, with angles $45°, 135°, 225°, 315°$.

#### 10.2.4 Crossing Between Root Locus and the Imaginary Axis

Here, we use the Routh-Hurwitz criterion. We start by rewriting the characteristic equation as a polynomial:

$$s^4 + 12s^3 + 64s^2 + 128s + K = 0.$$

So our Routh array is

$$\begin{array}{c|ccc} s^4 & 1 & 64 & K \\ s^3 & 12 & 128 & 0 \\ s^2 & \frac{12\cdot 64 - 1\cdot 128}{12} & K & 0 \\ s^1 & \frac{\frac{640}{12}\cdot 128 - 12K}{\frac{640}{12}} & 0 & 0 \\ s^0 & K & & \end{array}.$$

The $s^2$ row first term is $(768-128)/12 = 53{,}3$ and the $s^1$ ros first term becomes zero at the crossing

$$\frac{53{,}3 \cdot 128 - 12K}{53{,}3} = 0$$

which gives $K \approx 569$.

Now we use

$$53{,}3s^2 + 569 = 0$$

so $s^2 = -10{,}7 \implies s \pm j3{,}27$.

So the root locus crosses the imaginary axis at $s \pm j3{,}27$ when $K = 569$.

#### 10.2.5 Breakaway Points

A breakaway point is where two branches meet on the real axis and then leave it into the complex plane.

Since the real-axis locus is only between $-4$ and 0, the poles at 0 and $-4$ move toward each other along that segment, meet, and then break away.

From

$$1 + \frac{K}{D(s)} = 0 \implies D(s) + K = 0 \implies K = -D(s)$$

where

$$D(s) = s(s+4)(s+4+4j)(s+4-4j).$$

The breakaway points satisfy

$$\frac{\mathrm{d}K}{\mathrm{d}s} = 0$$

which is equivalent to differentiating $D(s)$. The real solution on the locus is $s \approx -1{,}58$, so the branches from 0 and $-4$ meet around $s = -1{,}58$ and leave the real axis there.

#### 10.2.6 Angle of departure from the complex poles

Because there are complex poles at

$$-4 \pm 4j$$

the branches do not leave those poles in arbitrary directions. They must satisfy the angle condition

$$\angle P(s) = (2k+1)180°.$$

For the upper pole $-4+4j$, the angles contributed by the other poles are used to determine the departure angle. We can find one equivalent angle as 225°, and because angles are modulo 360°, that is the same as $-135°$. So the key-idea is that the departure direction is forced by the angle condition, and that the two complex-pole branches leave symmetrically.

#### 10.2.7 Complete the Sketch

Now all the pieces are known. We have poles at 0, $-4$, and $-4 \pm 4j$, no finite zeros and the only part of the real axis on the locus is $(-4, 0)$. The asymptotes are through $-3$ at 45° , 135° , 225° , and 315°. The breakaway is at $-1{,}58$ and the imaginary crossing is at $\pm j3{,}27$ when $K = 569$.

### 10.3 Parameter Design by the Root Locus Method

Suppose a system has a characteristic equation on the form

$$a_n s^n + a_{n-1} s^{n-1} + \ldots + a_1 s + a_0 = 0$$

and we then want to know what happens when one coefficient, say $a_1$, changes.

We can rearrange the above into the root-locus form:

$$1 + \frac{a_1 s}{a_n s^n + a_{n-1} s^{n-1} + \ldots + a_2 s^2 + a_0} = 0$$

so now we can interpret it as a gain-like parameter $K = a_1$, and an open-loop part:

$$P(s) = \frac{s}{a_n s^n + a_{n-1} s^{n-1} + \ldots + a_2 s^2 + a_0}.$$

So even though $a_1$ is not the loop gain, it can still be treated as the parameter that "scales" the transfer function. Then the root locus tells us where the closed-loop poles go as $a_1$ changes, whether stability improves or worsens, or how damping and speed change.

#### 10.3.1 Varying Multiple Parameters

If we instead have a characteristic equation like

$$s^3 + s^2 + \beta s + \alpha = 0$$

and we want to adjust both the $\alpha$ and $\beta$ parameters, then we cannot directly make an ordinary root locus that simultaneously sweeps both $\alpha$ and $\beta$ as this method can only vary one scalar parameter at a time. Instead we repeat the procedure twice.

We start by studying $\alpha$ by setting $\beta = 0$. Then the characteristic equation becomes

$$s^3 + s^2 + \alpha = 0 \implies 1 + \frac{\alpha}{s^3 + s^2} = 0.$$

Now this is a root locus with parameter $K_2 = \alpha$ and tranfer function:

$$P_2(s) = \frac{1}{s^3 + s^2}.$$

We then use this root locus to choose a useful value of $\alpha$ based on the pole locations we want.

Then we fix $\alpha$ at the chosen value and study $\beta$, by rewriting the equation as

$$1+\frac{\beta s}{s^3+s^2+\alpha}=0$$

so now the parameter is $K_1=\beta$ and the transfer function is

$$P_1(s)=\frac{s}{s^3+s^2+\alpha}.$$

This, in turn, lets us see how the poles move as $\beta$ changes.

**Lecture 7: Root Locus Method, II** 14. April 2026

# 11 PID Controllers

A PID-controller (Proportional-Integral-Derivative) takes the error signal

$$e(t)=r(t)-y(t)$$

and turns it into a control input $u(t)$ for a plant/process. A block diagram for how this process might look is shown on Figure 11.1.

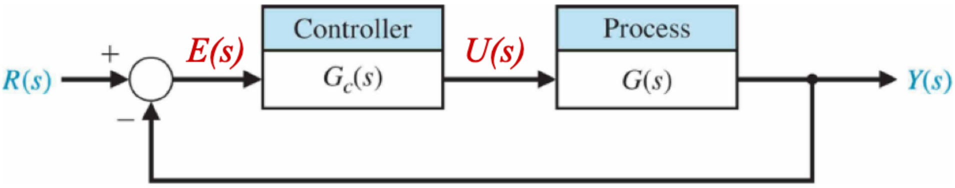



Figure 11.1: Block diagram of a simple process with a controller and feedback.

The controller is

$$u(t)=K_P e(t)+K_I\int e(t)\,\mathrm{d}t+K_D\frac{\mathrm{d}e(t)}{\mathrm{d}t}$$

or in the $s$-domain

$$G_c(s)=K_P+\frac{K_I}{s}+K_D s. \tag{2}$$

Here, the three terms $K_P$, $K_I$, and $K_D$ each correspond to different elements as shown in Table 1. In this way, the

Table 1: Meaning of the different terms in a PID-controller.

| Term | Meaning | Effect |
|---|---|---|
| $K_P e(t)$ | Proportional | Reacts to present error |
| $K_I\int e(t)\,\mathrm{d}t$ | Integral | Reacts to past error; Removes steady-state error |
| $K_D\frac{\mathrm{d}e(t)}{\mathrm{d}t}$ | Derivative | Reacts to change in error; Adds damping |

output of a PID-controller is based on both the present error, the past accumulated error, and the rate in change of the error.

## 11.1 Forms of PID Controllers

The ideal derivative term in Equation (2) is written as $K_D s$, however, a pure differentiator is usually not implemented directly, as it amplifies high frequency noise. In practice, the derivative part is often filtered as

$$G_d(s) = \frac{K_D s}{\tau_d s + 1}$$

which is approximately equal to $K_D s$ when $\tau_d$ is very small compared with the process time constants.

It is also possible to leave out the derivative part entirely leaving you with a PI-controller with transfer function:

$$G_c(s) = K_P + \frac{K_I}{s}.$$

Or a PD-controller with transfer function:

$$G_c(s) = K_P + K_D s.$$

A PI-controller is common when steady-state error matters, whereas a PD-controller is more common when damping/overshoot reduction is important but the integral action is not as useful.

### 11.1.1 PID as a Cascade of PI and PD

A PID-controller can be interpreted as a PI controller followed by a PD controller:

$$G_c(s) = G_{PI}(s)G_{PD}(s)$$

where

$$\begin{aligned} G_{PI}(s) &= \hat{K}_P + \frac{\hat{K}_I}{s} \\ G_{PD}(s) &= \overline{K}_P + \overline{K}_D s. \end{aligned}$$

Multiplying these (as we do when combining cascading blocks) gives

$$G_C(s) = \left(\hat{K}_P + \frac{\hat{K}_I}{s}\right)\left(\overline{K}_P + \overline{K}_D s\right) = \hat{K}_P\overline{K}_P + \hat{K}_P\overline{K}_D s + \frac{\hat{K}_I\overline{K}_P}{s} + \hat{K}_I\overline{K}_D.$$

We see that this is a PID-controller with

$$\begin{aligned} K_P &= \overline{K}_P\hat{K}_P + \hat{K}_I\overline{K}_D \\ K_D &= \hat{K}_P\overline{K}_D \\ K_I &= \hat{K}_I\overline{K}_P. \end{aligned}$$

### 11.1.2 Pole-zero interpretation of a PID controller

We use the general form of the transfer function for a PID controller, Equation (2), and put everything over a common denominator

$$G_c(s) = \frac{K_D s^2 + K_P s + K_I}{s}.$$

Factoring out $K_D$ we get:

$$G_C(s) = \frac{K_D\left(s^2 + \frac{K_P}{K_D}s + \frac{K_I}{K_D}\right)}{s}.$$

Then the quadratic part of the numerator can be written as

$$s^2 + \alpha s + b, \qquad \alpha = \frac{K_P}{K_D}, \quad b = \frac{K_I}{K_D},$$

which can be factored to $(s+z_1)(s+z_2)$. So:

$$G_c(s) = \frac{K_D(s+z_1)(s+z_2)}{s}.$$

This means a PID-controller contributes one pole at $s = 0$ and two zeros determined by $K_P$, $K_I$, and $K_D$. The pole at the origin comes from the integral term, $K_I/s$. The two zeros come from the numerator:

$$K_D s^2 + K_p s + K_I.$$

By choosing $K_P$, $K_I$, and $K_D$, you change where these zeros are located. In root locus design this is powerful as these zeros can then be used to pull the root locus toward the desired regions of the $s$-plane.

## 11.2 PID Tuning

As mentioned in Section 11.1.2, the values of $K_P$. $K_I$, and $K_D$ can be tuned to change the root locus. This can generally be done in two different ways, namely via manual trial-and-error tuning with minimal analytic analysis using step responses and via Ziegler-Nichols tuning, which is a more analytic tuning method based on closed- and open-loop response to a step input.

### 11.2.1 Manual PID Tuning

Manual tuning starts by simplifying the controller as much as possible by setting $K_I = K_D = 0$, so initially we only have

$$G_c(s) = K_P.$$

Which means, first we must tune a proportional controller. We then increase $K_P$ until the closed-loop system is just about unstable, before we reduce $K_P$ until the response has roughly quarter-amplitude decay. Then we increase $K_I$ and $K_D$ manually to improve the step response. Here, “quarter-amplitude decay” means that oscillations die out such that each overshoot peak os about one quarter of the previous overshoot peak. This is a classic tuning target as it is neither too sluggish nor too oscillatory for most applications.

**Effect of each PID gain** Table 2 summarizes the usual qualitative effects of increasing each PID gain. The

Table 2: Qualitative effect of increasing each PID gain

| Increasing | Overshoot | Settling Time | Steady-state error |
|---|---|---|---|
| $K_P$ | Increases | Small/minimal effect | Decreases |
| $K_I$ | Increases | Increases | Removes steady-state error |
| $K_D$ | Decreases | Decreases | Little/no direct effect |

intuition here is that $K_P$ pushes harder when the error is large. This usually makes the system respond faster, but it can also lead to more overshoot. $K_I$ keeps accumulating error over time, which is excellent for removing steady-state error, but it often makes the system more oscillatory. $K_D$ reacts to the rate of change of the error acting like a damper. It tends to reduce overshoot and make oscillations die out faster.

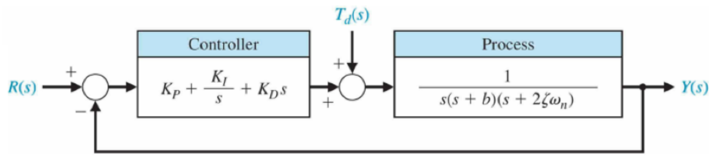



Figure 11.2: Example system used for the manual tuning example.

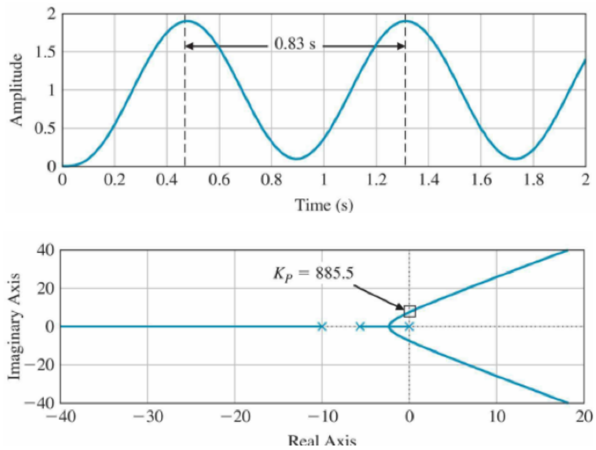



Figure 11.3: Root locus of P-controller at marginal stability.

**Example: Manual Tuning of a PID-controller**

We consider the system on Figure 11.2. Here, the plant has the transfer function

$$G(s) = \frac{1}{s(s+b)(s+2\xi\omega_n}$$

with $b = 10, \xi = 0{,}707, \omega_n = 4$. Hence, the process is approximately

$$G(s) = \frac{1}{s(s+10)(s+5{,}66)}.$$

So the open-loop system has poles at $s = 0, s = -10, s = -5{,}66$. The pole at the origin means the plant already contains an integrator

We start by setting $K_I = K_D = 0$, leaving us with

$$G_c(s) = K_P.$$

This gives the closed-loop characteristic equation:

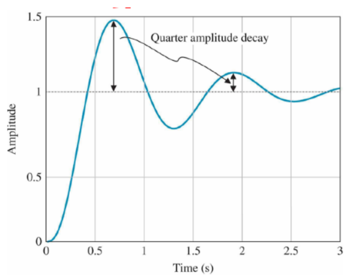
$$1 + K_P \frac{1}{s(s+10)(s+5{,}66)} = 0$$

or

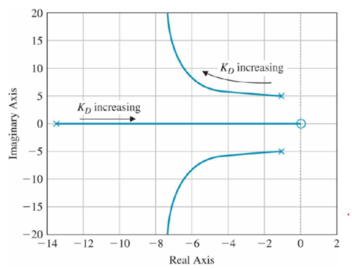
$$s(s+10)(s+5{,}66) + K_P = 0.$$

When $K_P = 886$, the closed-loop poles are exactly at $s = \pm 7{,}5j$, which is exactly on the imaginary axis. This is shown on Figure 11.3.

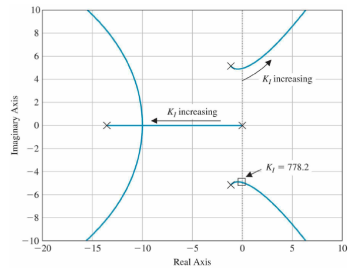
Here, the response oscillates forever without growing or decaying – This is also known as marginal stability. The oscillation period is approximately $T = 2\pi/7{,}5 = 0{,}84\,\text{s}$ Since $K_P = 886$ gives sustained oscillation, this




(a) Quarter-Amplitude decay case.




(b) Varying $K_D$ in the PD-controller.




(c) Varying $K_I$ in the PI-controller.

is too aggressive. We therefore reduce $K_P$ to about 370 to obtain quarter-amplitude decay. This is shown on Figure 11.4a.

We now fix $K_P = 370, K_I = 0$ and vary $K_D$, which means the controller is now a PD-controller:

$$G_c(s) = K_P + K_D s.$$

The root locus for this is shown on Figure 11.4b, where we see that increasing $K_D$ moves the complex poles left, corresponding to faster decay.

If we now fix $K_P = 370, K_D = 0$ and instead vary $K_I$ the controller is instead a PI-controller:

$$G_c(s) = K_P + \frac{K_I}{s}.$$

On Figure 11.4c, we see that increasing $K_I$ moves the complex poles to the right. At $K_I = 778$, the system becomes marginally stable with poles at $s = \pm 4{,}86 j$.

Finally, we choose $K_P = 370, K_I = 100, K_D = 75$ giving the PID-controller

$$G_c(s) = 370 + \frac{100}{s} + 60s = \frac{60s^2 + 370s + 100}{s}.$$

This has the step response depicted on Figure 11.5 with overshoot of approximately 8,78 % and settling time of approximately 3,24 s.

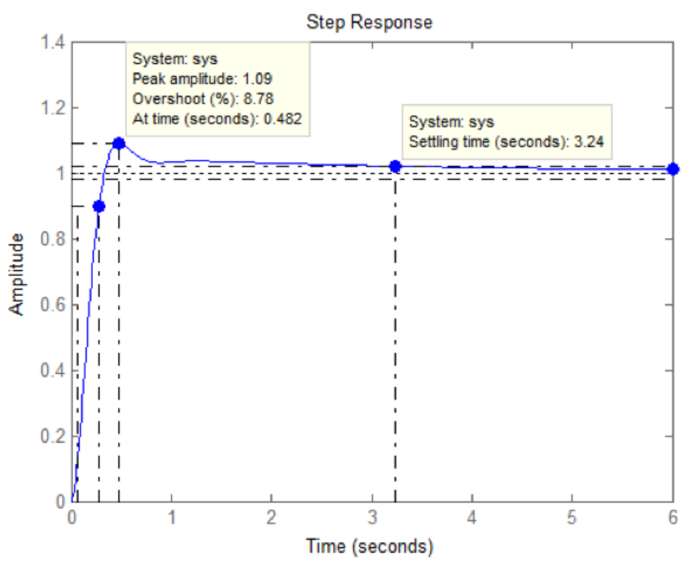



Figure 11.5: Manually tuned PID-controller.

### 11.2.2 Closed-loop Ziegler-Nichols PID Tuning

Ziegler-Nichols tuning is a rule-based way to get initial PID gains without needing a perfect mathematical model of the process. We generally distinguish between closed-loop and open-loop Ziegler Nichols tuning. Closed-loop Ziegler-Nichols tuning uses the closed-loop system and finds the ultimate gain and ultimate period, whereas open-loop Ziegler-Nichols tuning uses the open-loop step response, also called the reaction curve. Here, we focus on the closed-loop Ziegler-Nichols method.

The method starts like manual tuning with $K_I = K_D = 0$, so the controller is initially only proportional:

$$G_c(s) = K_P.$$

Then $K_P$ is increased, until the closed-loop system is on the boundary of instability, i.e. marginal stability. At this point we measure two things, $K_U$ which is the ultimate gain (the value of $K_P$ that gives sustained oscillations), and the ultimate period $T_U$ (the period of the sustained oscillations).

The closed-loop Ziegler-Nichols rules are shown on Table 3.

Table 3: Closed-loop Ziegler-Nichols rules

| **Controller** | $K_P$ | $K_I$ | $K_D$ |
|---|---|---|---|
| P | $0{,}5K_U$ | – | – |
| PI | $0{,}45K_U$ | $\frac{0{,}54K_U}{T_U}$ | – |
| PID | $0{,}6K_U$ | $\frac{1{,}2K_U}{T_U}$ | $\frac{0{,}6K_U T_U}{s}$ |

**Example: Ziegler-Nichols Tuning of a PID-controller**

We use the same process as before:

$$G(s) = \frac{1}{s(s+b)(s+2\xi\omega_n)} = \frac{1}{s(s+10)(s+5{,}66)}.$$

---

With $K_I = K_D = 0$, the closed-loop characteristic equation is

$$1 + K_P \frac{1}{s(s+10)(s+5{,}66)} = 0.$$

When $K_P = 886$ the system is at marginal stability and the closed-loop poles are at $s = \pm 7{,}5j$ as shown on Figure 11.3. Therefore $K_U = 886$ and the period of oscillation here is $T_U = 0{,}84\,\mathrm{s}$.
Using the PID-row of Table 3, we see that:

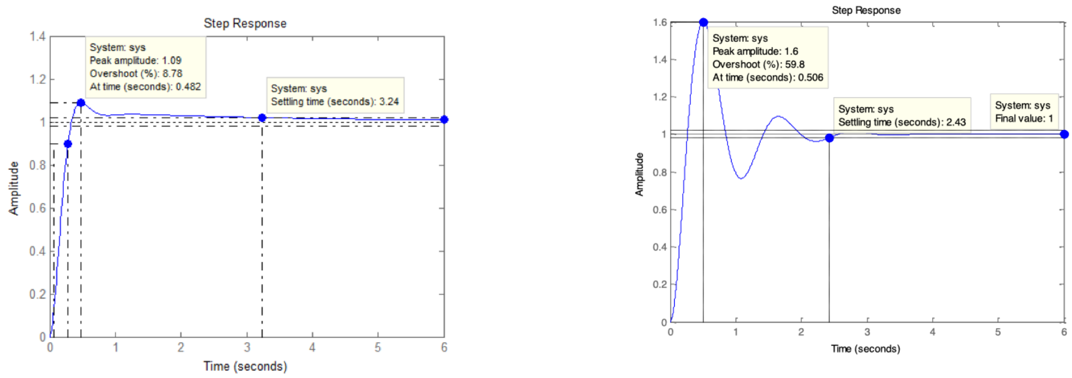
$$K_P = 0{,}6K_U = 531$$
$$K_I = \frac{1{,}2K_U}{T_U} = 1280$$
$$K_D = \frac{0{,}6K_U T_U}{8} = 55{,}1.$$

So using Ziegler-Nichols tuning the PID controller becomes

$$G_c(s) = 531 + \frac{1280}{s} + 55{,}1s = \frac{55{,}1s^2 + 531s + 1280}{s}.$$

This gives the step response on Figure 11.6b, where it is compared with the result from the manual tuning on Figure 11.6a.

On Figure 11.6 the step responses from the manually tuned on the Ziegler-Nichols tuned PID controllers are




(a) Step response from Manual tuning.

(b) Step response from Ziegler-Nichols tuning.

Figure 11.6: Comparison of step responses between manually tuned and Ziegler-Nichols tuned PID controller.

shown. Here, we see that the manual tuning has much less overshoot but a bit longer settling time than the Ziegler-Nichols tuned PID controller.

**Example: Automobile Velocity Control**

We define $y(t)$ to be the relative velocity between two vehicles, and our plant is modeled as

$$G(S) = \frac{1}{(s+2)(s+8)}$$

so the open-loop system has two stable poles at $s = -2$ and $s = -8$. The controller must satisfy four requirements:

1. Zero steady-state error to a step input
2. Steady-state error to a ramp input less than 25%
3. Percent overshoot less than 5%
4. Settling time less than 1,5 s

We start by considering the first design spec: zero steady state error to a step. The plant is

$$G(s) = \frac{1}{(s+2)(s+8)}$$

so we have no poles at the origins and hence a type 0 system. For zero steady state error to a step input, the open-loop system must be at least type 1, meaning it needs one integrator $1/s$. Therefore, the controller must be a type 1 controller to provide this integrator. For this, we choose a PI controller:

$$G_c(s) = K_P + \frac{K_I}{s} = \frac{K_P s + K_I}{s}.$$

For the second spec; the ramp error condition, we remember that steady-state error to a ramp input is related to the velocity error constant:

$$K_v = \lim_{s\to 0} sG_c(s)G(s).$$

We require that the error is less than 25% so $K_v > 1/0{,}25 = 4$. Now, using the PI controller and the plant:

$$G_C(s) = \frac{K_P s + K_I}{s} \qquad G(s) = \frac{1}{(s+2)(s+8)}.$$

Then:

$$K_v = \lim_{s\to 0} s\frac{K_P s + K_I}{s}\frac{1}{(s+2)(s+8)} = \lim_{s\to 0}\frac{K_P s + K_I}{(s+2)(s+8)}.$$

At $s = 0$:

$$K_v = \frac{K_I}{2\cdot 8} = \frac{K_I}{16} \implies \frac{K_I}{16} > 4 \implies K_I > 64.$$

For the overshoot condition, we look at the damping ratio $\xi$. The requirement here is $M_P < 5\%$, which corresponds approximately to $\xi \geq 0{,}69$. So the closed-loop dominant poles should lie in the region of the $s$-plane corresponding to a damping ratio of at least 0,69.
To look at the settling time condition we remember that 2% settling time is approximately:

$$T_s \approx \frac{4}{\xi\omega_n}.$$

We require $T_s \leq 1{,}5$, so

$$\frac{4}{\xi\omega_n} \leq 1{,}5$$
$$\xi\omega_n \geq 2{,}67.$$

The real part of a complex pole pair is $\Re(s) = -\xi\omega_n$, so the dominant poles should be to the left of $s \approx -2{,}6$.
The open-loop transfer function is:

$$G_c(s)G(s) = \frac{K_P s + K_I}{s(s+2)(s+8)}.$$

The closed-loop transfer function is:

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

Substituting gives:

$$T(s) = \frac{K_P s + K_I}{s(s+2)(s+8) + K_P s + K_I} = \frac{K_P s + K_I}{s^3 + 10s^2 + (16 + K_P)s + K_I}.$$

We now apply the Routh-Hurwitz criterion to the characteristic equation:

$$s^3 + 10s^2 + (16 + K_P)s + K_I = 0.$$

The Routh table gives the stability conditions:

$$K_I > 0$$
$$K_P > \frac{K_I}{10} - 16.$$

So to be stable $K_I$ must be positive and $K_P$ must be large enough relative to $K_I$. From spec 2 we have $K_I > 64$, so

$$K_I > 64$$
$$K_P > \frac{K_I}{10} - 16.$$

The open-loop transfer function can be written as

$$G_c(s)G(s) = K_P\frac{s + \frac{K_I}{K_P}}{s(s+2)(s+8)}.$$

Which has poles at $s = 0$, $s = -2$, $s = -8$ and a zero at $s = -K_I/K_P$. The asymptote angles are $\phi_A = +90°, -90°$. The asymptote centroid is

$$\sigma_A = \frac{\sum \text{poles} - \sum \text{zeros}}{n_p - n_z} = \frac{0-2-8-\left(-\frac{K_I}{K_P}\right)}{2} = -5 + \frac{K_I}{2K_P}.$$

We now use spec 4 to choose a location for our zero. To meet the settling time requirement, the dominant poles should be left of $s = -2{,}6$. We use the asymptote centroid as a root-locus design guide and require:

$$\begin{aligned} \sigma_A &< -2{,}6 \\ -5 + \frac{K_I}{2K_P} &< 2{,}4 \\ \frac{K_I}{K_P} &< 4{,}8. \end{aligned}$$

We choose $K_I/K_P = 2{,}5$, which means the PI zero is located at $s = -2{,}5$ so the root locus becomes:

$$1 + K_P \frac{s+2{,}5}{s(s+2)(s+8)} = 8.$$

With the zero fixed at $-2{,}5$, the root locus can be plotted as shown on Figure 11.7a, where the desired region is shaded. We conclude that $K_P < 30$ is required.
Therefore our final gain conditions are

$$\begin{aligned} K_I &> 64 \\ K_P &< 30 \\ \frac{K_I}{K_P} &= 2{,}5. \end{aligned}$$

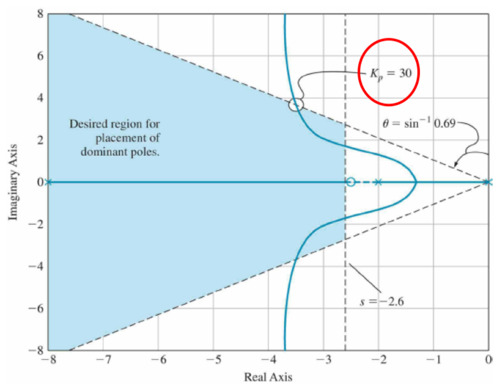
Therefore we need $2{,}5K_P > 64 \implies K_P > 25{,}6$. Together with $K_P < 30$, a reasonable choice is $K_P = 26$, and then $K_I = 2{,}5 \cdot 26 = 65$, so our final PID controller is

$$G_c(s) = 26 + \frac{65}{s} = \frac{26s+65}{s}.$$

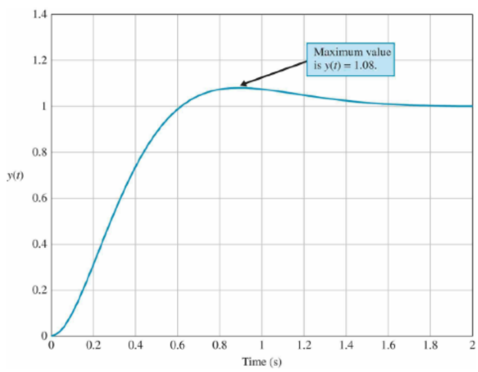
The final closed-loop transfer function is then:

$$T(s) = \frac{26s+65}{s^3+10s^2+42s+65}.$$

The final step response is shown on Figure 11.7b.




(a) Root Locus with fixed zero at $-2{,}5$.




(b) Final step response.

**Lecture 8: Frequency Response Methods** 21. April 2026

# 12 Frequency Response Methods

## 12.1 Introduction and Concepts

A frequency response describes how a system responds in steady-state when the input is sinusoidal. So if the input is:

$$r(t) = A\sin(\omega t)$$

then, after the transient part has died away, the output of a linear time-invariant system will be sinusoidal. The output here will have the *same frequency*, but generally *different amplitude and different phase.* So the system does not change the frequency of the sinusoidal signal. It only scales it and shifts it in time.

For example, if you input a slow sine wave into a mechanical system, the output may follow it almost perfectly. If you input a very fast sine wave, the output may be much smaller and delayed. This is exactly this dependency on frequency that frequency response studies.

In general, for a sinusoidal input

$$r(t) = A\sin(\omega t)$$

the output is

$$y(t) = \left|T(j\omega)\right| A\sin(\omega t + \theta).$$

Where $\left|T(j\omega)\right|$ is the gain of the system at frequency $\omega$, which tells hoe much the amplitude is magnified, and $\theta = j\omega$ is the phase angle, which tells how much the output is shifted relative to the input. Normally $\theta < 0$ in physical systems, meaning they lag behind as they cannot react instantly.

### 12.1.1 Laplace vs. Fourier transform

We have previously mainly been using Laplace transforms, defined as:

$$F(s) = \mathscr{L}\left\{f(s)\right\} = \int_0^\infty f(t)e^{-st}\,\mathrm{d}t.$$

Now, we introduce the Fourier transform, defined by:

$$F(\omega) = \mathscr{F}\left\{f(t)\right\} = \int_{-\infty}^\infty f(t)e^{-j\omega t}\,\mathrm{d}t.$$

The key difference between these is the complex variable. For Laplace $s = \sigma + j\omega$, while for Fourier, we only use the imaginary axis $s = j\omega$. So, the Fourier transform can be viewed as a special case of the Laplace transform, where we evaluate along the imaginary axis. This is the reason frequency response uses $s = j\omega$, i.e. we are asking, "what does the transfer function look like for pure sinusoidal motion at frequency $\omega$?". This happens as sinusoids are associated with complex exponentials, $e^{j\omega t}$, and therefore with the imaginary axis in the $s$-plane. The inverse Fourier transform is given as:

$$y(t) = \frac{1}{2\pi}\int_{-\infty}^\infty Y(j\omega)e^{j\omega t}\,\mathrm{d}\omega.$$

This integral is usually difficult to compute directly, so instead of reconstructing $y(t)$ exactly, graphical and experimental methods are often used – This is where Bode plots, and polar plots come in. These lets us study the magnitude and phase of $T(j\omega)$ without needing to solve the full time-domain response every time.

In earlier control theory, the Laplace transform lets us study the pole-zero structure of $T(s)$, meaning we can study stability, transient behavior, damping, natural frequency, and so on. But in frequency response methods, we evaluate the transfer function at $s = j\omega$, so we get $T(j\omega)$, which is now a frequency dependant complex number. In the frequency domain, the transfer function $T(j\omega)$ is defined similarly as for the time domain

$$Y(j\omega) = T(j\omega)R(j\omega).$$

For a closed-loop system this becomes:

$$T(j\omega) = \frac{G(j\omega)}{1 + G(j\omega)H(j\omega)}.$$

Which is just the usual closed-loop transfer function, but with $s$ replaced by $j\omega$.

### 12.1.2 Advantages and Disadvantages of Frequency Response Methods

Frequency response methods are useful in variety of ways. Often, the frequency response can be obtained directly by testing the real system. For example, you can apply sinusoidal inputs at different frequencies:

$$r(t) = A\sin(\omega t)$$

then measure the output amplitude and phase, from where $\left|T(j\omega)\right|$ and $j\omega$ can then be estimated.

The second advantage is analytical. If you already know the transfer function $T(s)$, you can get the frequency response by simply replacing $s$ with $j\omega$, so $T(s) \to T(j\omega)$, then you study how the complex number $T(j\omega)$ changes as $\omega$ increases.

The third advantage is graphical. Magnitude and phase can be represented clearly using e.g. Bode plots, which show magnitude and phase versus frequency.

The disadvantages are also important. Frequency response is not a time-domain method, so it does not show the exact time response to e.g. a step input at each frequency. So in this way there is an indirect link between frequency domain information and time domain behavior. For example, bandwidth, phase margin, gain margin, and resonance tell us a lot about speed, overshoot and stability, but the connection is not always as visually immediate as looking at a step response.

## 12.2 Polar Plots

Once we replace $s$ by $j\omega$, the transfer function becomes a complex-valued function of frequency, so instead of writing only $G(s)$, we now write $G(j\omega)$, and because it is complex, it can be split into a real and an imaginary part

$$G(j\omega) = R(\omega) + jX(\omega).$$

So, for every frequency $\omega$, the tranfer function gives one complex number. That complex number can be drawn as a point in the complex plane on Figure 12.1.

The horizontal axis of Figure 12.1 is the real part, $R(\omega)$ and the vertical axis is the imaginary part $X(\omega)$.

The same complex number can also be written in polar form:

$$G(j\omega) = \left|G(j\omega)\right| e^{j\phi(\omega)}$$

or

$$G(j\omega) = \left|G(j\omega)\right| \angle\phi(\omega).$$

The magnitude is:

$$\left|G(j\omega)\right| = \sqrt{R(\omega)^2 + X(\omega)^2}$$

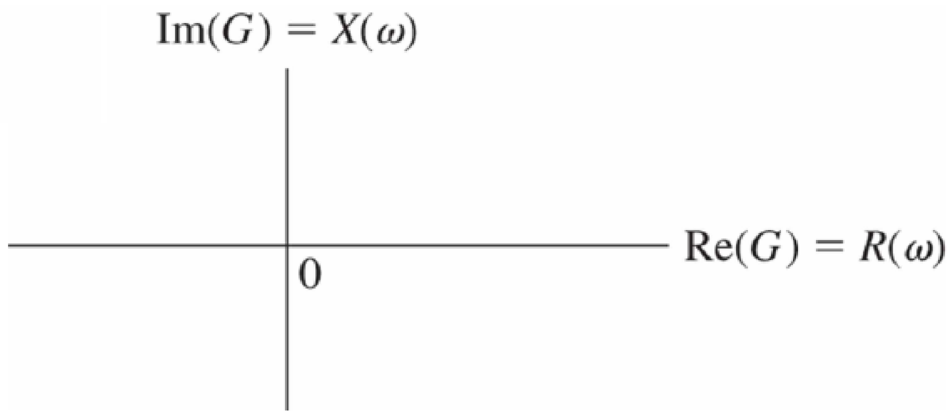



Figure 12.1: The complex plane

and the phase is:

$$\phi(\omega) = \tan^{-1}\left(\frac{X(\omega)}{R(\omega)}\right).$$

So the polar plot is simply the path traced by $G(j\omega)$ in the complex plane as $\omega$ changes.

#### 12.2.1 Example of Polar Plotting

We will now show this for a concrete example: an RC filter, shown on Figure 12.2. The circuit has a resistor $R$ in the series and a capacitor $C$ to ground. The input voltage is $V_1(s)$, and the output voltage $V_2(s)$ is measured across the capacitor. This is a standard first-order low-pass filter.

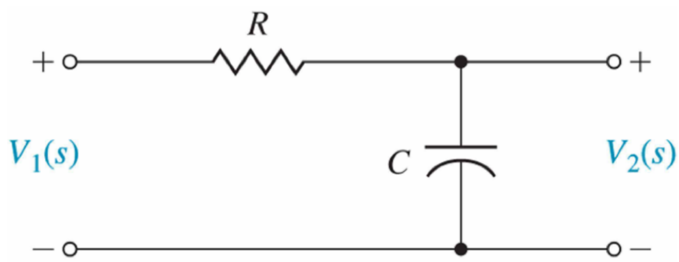



Figure 12.2: RC filter for example of polar plotting.

At low frequency, the capacitor behaves almost like an open circuit. So almost no current flows, meaning almost no voltage is lost across the resistor. Therefore $V_2 \approx V_1$. At high frequency, the capacitor behaves almost like a short circuit to ground. So the output node is pulled close to ground, meaning $V_2 \approx 0$. So physically, we would expect this circuit to pass low frequencies and attenuate high frequencies.

To draw the pole plot, we must first derive the frequency domain transfer function of the system. The circuit is just a voltage divider

$$G(s) = \frac{V_2(s)}{V_1(s)} = \frac{Z_C}{R + Z_C}.$$

Substituting the impedance of a capacitor, $Z_C = 1/(sC)$, we obtain

$$G(s) = \frac{\frac{1}{sC}}{R + \frac{1}{SC}} = \frac{1}{RCs + 1}.$$

Now, we move into frequency response by replacing $s = j\omega$, so:

$$G(j\omega) = \frac{1}{j\omega RC + 1}.$$

Multiplying with the complex conjugate we obtain:

$$G(j\omega) = \frac{1 - j\omega RC}{1 + (\omega RC)^2} = \frac{1}{1 + (\omega RC)^2} - j\frac{\omega RC}{1 + (\omega RC)^2}.$$

So, now we can identify:

$$R(\omega) = \frac{1}{1 + (\omega RC)^2}$$

and

$$X(\omega) = -\frac{\omega RC}{1 + (\omega RC)^2}.$$

We also define the corner frequency:

$$\omega_1 = \frac{1}{RC}$$

so that

$$\omega RC = \frac{\omega}{\omega_1}.$$

Which makes the equations cleaner:

$$G(j\omega) = \frac{1}{1 + \left(\frac{\omega}{\omega_1}\right)^2} - j\frac{\frac{\omega}{\omega_1}}{1 + \left(\frac{\omega}{\omega_1}\right)^2}.$$

The magnitude is

$$\left|G(j\omega)\right| = \frac{1}{\sqrt{1 + \left(\frac{\omega}{\omega_1}\right)^2}}$$

and the phase is:

$$\phi(\omega) = -\tan^{-1}\left(\frac{\omega}{\omega_1}\right).$$

We can now turn the real and imaginary parts into a polar plot. We have:

$$R(\omega) = \frac{1}{1 + \left(\frac{\omega}{\omega_1}\right)^2} \qquad X(\omega) = -\frac{\frac{\omega}{\omega_1}}{1 + \left(\frac{\omega}{\omega_1}\right)^2}.$$

Now, we look at a few important frequencies. At $\omega = 0$, we get $R(0) = 1$, and $X(0) = 0$, so $G(j0) = 1$. As the polar plot is in the complex plane (the $R$-$X$-plane in this context), it starts at the point $(1, 0)$. Physically, this means DC passes through the low-pass filter unchanged.

At $\omega = \omega_1$ we get $R(\omega_1) = 1/2$ and $X(\omega_1) = -1/2$. So the point corresponding to $\omega = \omega_1$ is $(1/2, -1/2)$- The magnitude is:

$$\left|G(j\omega_1)\right| = \sqrt{\left(\frac{1}{2}\right)^2 + \left(-\frac{1}{2}\right)^2} = \frac{1}{\sqrt{2}}$$

and the phase is:

$$\phi(\omega_1) = -45°.$$

This is the familiar cutoff point of a first-order RC low-pass filter.

At $\omega \to \infty$, we get $R(\omega) \to \infty$ and $X(\omega) \to \infty$, so the plot approaches the origin. Physically, very high-frequency signals are blocked.

For positive frequencies, the curve goes through the lower half of the complex plane because the imaginary part is negative, for negative frequencies, the curve mirrors the upper half-plane. This is all shown on Figure 12.3.

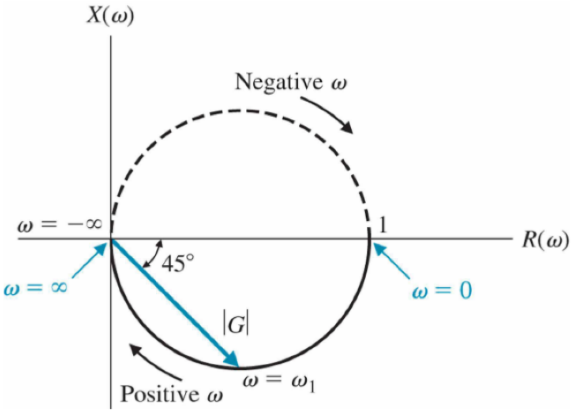



Figure 12.3: Polar plot of the RC filter

Here, we see that the locus is a circle with centre $(1/2, 0)$ and radius $1/2$. We see this by:

$$\left(R - \frac{1}{2}\right)^2 + X^2 = \left(\frac{1}{2}\right)^2.$$

So the RC filter's polar plot is a circle passing through $(1,0)$ and $(0,0)$ with its center at $(1/2, 0)$.

#### 12.2.2 The Limitations of Polar Plots

Polar plots are useful because they show the complex value of $G(j\omega)$ directly. They show both magnitude and phase in one curve, but they have limitations. The first problem is that if you add another pole or zero to the system, the whole polar plot changes.

The second problem is that polar plots do not clearly show the individual effect of each pole and zero. A pole contributes magnitude reduction and phase lag. A zero contributed magnitude increase and phase lead. But in a polar plot, these effects are mixed together in one curve.

### 12.3 Bode Plots

An alternative to polar plots introduced in Section 12.2 are Bode plots. For polar plots, we represented the frequency response as a complex number

$$G(j\omega) = \left|G(j\omega)\right| e^{j\phi(\omega)}.$$

So $G(j\omega)$ contains both information about magnitude ($|\ldots|$) and phase angle ($\phi(\omega)$). A Bode plot separates these into two plots: a magnitude plot in decibels versus frequency and a phase plot which shows the phase angle in degrees versus frequency. The magnitude is not usually plotted directly as $\left|G(j\omega)\right|$. Instead, it is plotted in decibels:

$$20\log_{10}\left|G(j\omega)\right|$$

also known as the logarithmic gain.

The reason for using decibels is that multiplication becomes addition. This is very useful when transfer functions contain many factors, e.g. if

$$G(j\omega) = G_1(j\omega)G_2(j\omega)$$

then

$$20\log_{10}\left|G(j\omega)\right| = 20\log_{10}\left|G_1(j\omega)\right| + 20\log_{10}\left|G_2(j\omega)\right|.$$

So complicated transfer functions become much easier to sketch.

### 12.3.1 Example of Bode Plotting

Now, we move to the same RC low-pass filter as in Section 12.2.1. The transfer function was:

$$G(j\omega) = \frac{1}{j\omega RC + 1}.$$

We now define the time constant $\tau = RC$, so the transfer function may be written as

$$G(j\omega) = \frac{1}{j\omega t + 1}.$$

The magnitude is:

$$\left|G(j\omega)\right| = \frac{1}{\sqrt{1+(\omega\tau)^2}}.$$

So the logarithmic gain is:

$$20\log\left|G(j\omega)\right| = 20\log\left(\frac{1}{\sqrt{1+(\omega\tau)^2}}\right) = -10\log\left(1+(\omega\tau)^2\right).$$

The phase angle is then:

$$\phi(\omega) = -\tan^{-1}(\omega\tau).$$

So the RC filter has magnitude decreasing with frequency, and phase becoming more negative with frequency. This is shown on Figure 12.4.

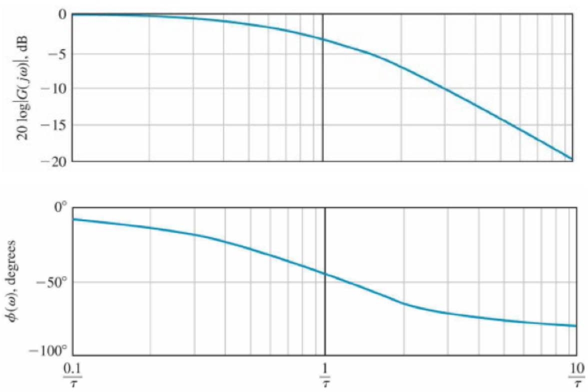



Figure 12.4: Bode plot of the RC-filter

We not analyze the magnitude expression in three important regions. When $\omega \ll 1/\tau$, then $\omega\tau \ll 1$, so $1+(\omega\tau)^2 \approx 1$, therefore, for $\omega \ll 1$:

$$20\log\left|G(j\omega)\right| \approx -10\log(1) = 0\,\text{dB}.$$

I.e. the filter passes low frequencies almost unchanged.

When $\omega \gg 1/\tau$, then $1+(\omega\tau)^2 \approx (\omega\tau)^2$, so

$$20\log\left|G(j\omega)\right| \approx -10\log\left((\omega\tau)^2\right) = -20\log(\omega\tau).$$

So after the corner frequency, the magnitude falls linearly on the logarithmic plot.

At the corner frequency, $\omega_c = 1/\tau$, we get $\omega\tau = 1$, so $|G(j\omega_c)| = 1/\sqrt{2}$. In decibels this is:

$$20\log\left(\frac{1}{\sqrt{2}}\right) = -3{,}01\,\text{dB}.$$

Hence, why the corner frequency is also known as the $-3$ dB frequency. So for a first-order RC low-pass filter: $\omega_c = 1/(RC)$ is the frequency where the amplitude has dropped to about 70,7 % of the input amplitude.

**Hand Sketching Bode Plots and Straight-Line Approximation** Instead of drawing the exact curve, we can approximate it using straight-line asymptotes. For the RC low-pass filter from before, at low frequency $\omega \ll 1/\tau$, the magnitude is approximately 0 dB, so the first asymptote is a horizontal line.

At high frequencies, $\omega \gg 1/\tau$, the magnitude is approximately $-20\log(\omega\tau)$, which is a straight line with slope $-20\,\text{dB/decade}$, where a decade is a factor 10 in frequency. I.e. going from $\omega$ to $10\omega$ is one decade. This is shown on Figure 12.5.

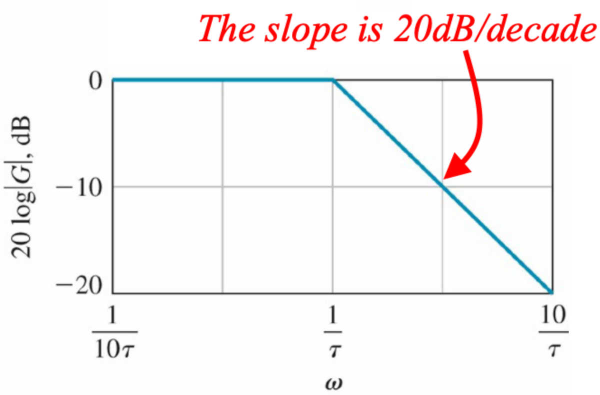



Figure 12.5: Straight line approximation of the magnitude part of the bode plot on Figure 12.4.

.

**General Bode Plot Construction** A general transfer function can be written as a product of simpler factors:

$$G(j\omega) = \frac{K_b \prod_{i=1}^{Q}(1 + j\omega\tau_i)}{(j\omega)^N \prod_{m=1}^{M}(1 + j\omega\tau_m)\prod_{k=1}^{R}\left(1 + 2\frac{\xi_k}{\omega_{nk}}j\omega + \left(\frac{j\omega}{\omega_{nk}}\right)^2\right)}.$$

This looks horrible, but the point is that it is built from simpler pieces. There may be:

- Constant gain $K_b$
- Zeros on the real axis
- Poles at the origin
- Poles on the real axis
- Complex conjugate pole pairs

The bode magnitude is bades on $20\log|G(j\omega)|$. Because of the logarithm, products become sums and divisions become subtractions. So the total magnitude plot is:

$$\text{gain contributions} + \text{zero contributions} - \text{pole contributions}.$$

The same idea applies to the phase, where the total phase angle is the sum of the phase angles from each factor. So instead of trying to plot the whole transfer function at once, Bode plots let us break it down factor by factor.

### 12.3.2 The Four Basic Kinds of Factors

Almost all transfer functions we meet in this course can be decomposed into four types of factors:

- Constant gain, $K_b$. This just shifts the magnitude plot up or down.
- Poles or zeros at the origin. E.g. $j\omega$ or $1/(j\omega)$. A zero at the origin behaves like a differentiator and a pole at the origin behaves like an integrator.
- Real-axis poles or zeros. E.g. $1 + j\omega\tau$ or $1/(1 + j\omega\tau)$. These are first-order factors, exactly like the RC filter from the previous examples are.
- Complex conjugate poles or zeros. E.g.

$$1 + 2\xi\frac{j\omega}{\omega_n} + \left(\frac{j\omega}{\omega_n}\right)^2.$$

  These are second-order factors and are important for systems with damping, resonance, overshoot, oscillation, and natural frequency.

The practical recipe is to first split the transfer function into the above factors. Then find the magnitude and phase contributions from each factor. Then, add all the magnitude together and add all the phase curves together.

### 12.3.3 Bode Plot for Constant Gain

This is the easiest factor, i.e. a constant gain. If the transfer function is just:

$$G(j\omega) = K_b$$

then the magnitude is constant

$$20\log|K_b|.$$

So on a Bode magnitude plot, this is a horizontal line. E.g. $K_b = 10 \implies 20\log(10) = 20\,\mathrm{dB}$ and $K_b = 0{,}1 \implies 20\log(0{,}1) = -20\,\mathrm{dB}$.

If $K_b$ is positive, the phase angle is $\phi(\omega) = 0°$, because multiplying by a positive number does not shift the signal in phase. If the gain is negative, e.g.

$$G(j\omega) = -K_b$$

then, the magnitude is still:

$$20\log|K_b|$$

but the phase is shifted by $-180°$.

### 12.3.4 Poles and Zeros at the Origin

Factors like

$$\left(j\omega\right)^N$$

and

$$\frac{1}{\left(j\omega\right)^N}$$

contribute zeros or poles at the origin because the corresponding factor is located at $s = 0$.

**Pole at the origin** For a pole at the origin:

$$G(j\omega) = \frac{1}{\left(j\omega\right)^N}$$

the magnitude is:

$$20\log\left|\frac{1}{(j\omega)^N}\right| = -20N\log\omega$$

so the slope is $-20N$ dB/decade. I.e. for one pole at the origin the slope is $-20$ dB/decade whilst two poles at the origin gives a slope of $-40$ dB.

The phase is:

$$\phi(\omega) = -90^\circ N.$$

So one integrator gives $-90^\circ$, and two integrators give $-180^\circ$.

**Zero at the Origin** For a zero at the origin:

$$G(j\omega) = (j\omega)^N.$$

The magnitude is

$$20\log\left|(j\omega)^N\right| = 20N\log\omega$$

so the slope is $20N$ dB/decade.

The phase is:

$$\phi(\omega) = +90^\circ N.$$

So poles at the origin makes the magnitude slope downward, while zeros at the origin make it slope upward.

The effects of poles and zeros at the origin are shown graphically on Figure 12.6.

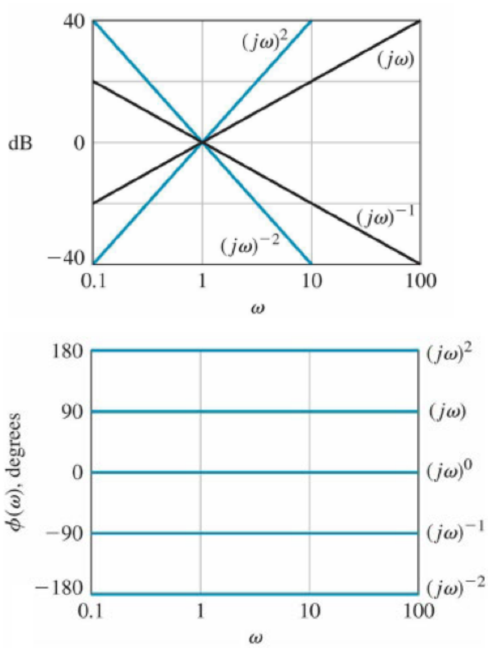



Figure 12.6: Effect of poles and zeros at the origin on Bode plots

### 12.3.5 Real-Axis Poles and Real-Axis Zeros

For a first-order factor:

$$\frac{1}{1 + j\omega\tau}$$

we get a real-axis pole.

The magnitude of this is

$$20\log\left|\frac{1}{1+j\omega\tau}\right| = -10\log\left(1+(\omega\tau)^2\right)$$

and the phase is

$$\phi(\omega) = -\tan^{-1}(\omega\tau).$$

The corner frequency is $\omega_c = 1/\tau$. For frequencies much smaller than the corner frequency $\omega \ll 1/\tau$, we have $1 + j\omega\tau \approx 1$, so the magnitude is approximately 0 dB. For frequencies much larger than the corner frequency, $\omega \gg 1/\tau$, the magnitude falls with slope $-20$ dB/decade. So the asymptotic approximation is flat before the corner frequency and then slopes downward after it.

The phase goes from $0° \to 90°$, with the midpoint at the corner frequency, where $\phi = -45°$.

For a zero:

$$1 + j\omega\tau$$

the magnitude after the corner frequency is $+20$ dB/decade, and the phase goes from $0° \to 90°$. So the rule is a real pole bends the magnitude downward and adds phase lag, whilst a real zero bends the magnitude upward and adds phase lead. This is all summarized graphically for a pole on the real axis on Figure 12.7.

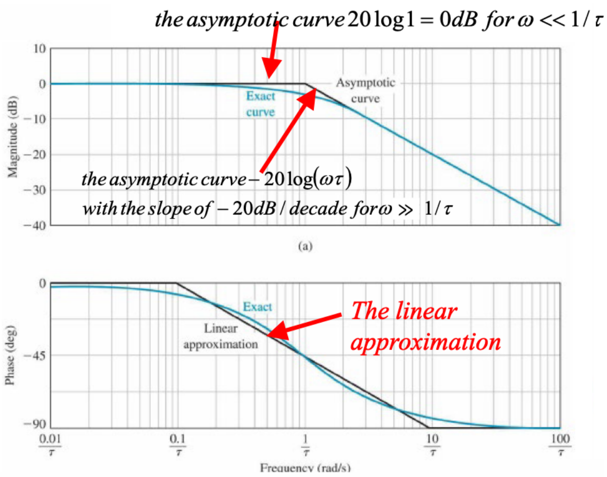



Figure 12.7: Graphical representation of the effect of a pole on the real axis on a Bode plot.

### 12.3.6 Complex Conjugate Poles or Zeros

Now, we move to second-order factors. A standard second-order pole pair has the form:

$$G(j\omega) = \frac{1}{1 + 2\xi\frac{j\omega}{\omega_n} + \left(\frac{j\omega}{\omega_n}\right)^2}.$$

We now introduce the frequency ratio:

$$u = \frac{\omega}{\omega_n}.$$

Then the magnitude becomes:

$$20\log|G(j\omega)| = -10\log\left(\left(1-u^2\right)^2 + 4\xi^2u^2\right).$$

And the phase becomes:

$$\phi(\omega) = -\tan^{-1}\left(\frac{2\xi u}{1-u^2}\right).$$

The important new feature for this case is resonance. For low damping, the magnitude can rise above 0 dB near the natural frequency. This peak is called the resonant peak. The resonant frequency is:

$$\omega_r = \omega_n\sqrt{1-2\xi^2}.$$

This only exists as a real resonant *peak* when the damping is sufficiently small. The maximum magnitude at this peak is:

$$M_{p\omega} = \left|G(j\omega_r)\right| = \left(2\xi\sqrt{1-\xi^2}\right)^{-1}.$$

So when $\xi$ is small (a lightly damped system) the resonant peak becomes large. The magnitude plot on Figure 12.8 shows this clearly. For small damping ratios, the curve has a tall peak near $\omega_n$, whilst for larger ratios, the peak becomes smaller or even disappears.

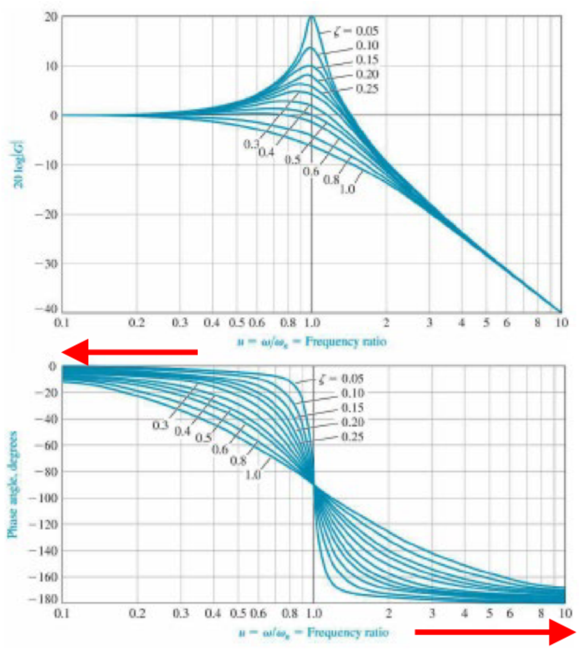



Figure 12.8: Graphical representation of the effect of complex conjugate poles on the Bode plots.

The phase plot on Figure 12.8 shows that a second-order pole pair changes phase from $0° \to -180°$, with the transition being sharper for smaller damping. So compared with a first-order pole, a second-order pole pair contributed twice as much final phase lag and twice the high-frequency magnitude slope: $-40\,\text{dB/decave}$.

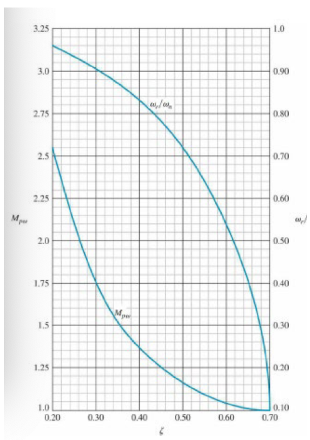
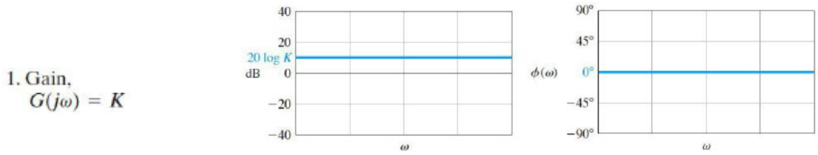


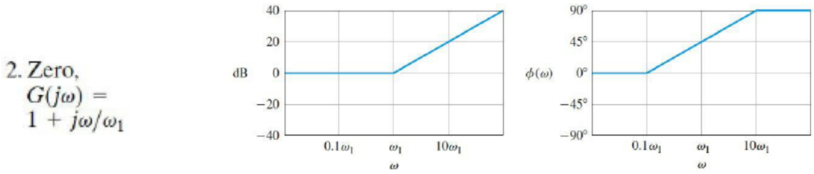
Figure 12.9: Effect of varying $\xi$ on the resonant frequency and the resonant peak.

Second-order frequency response is very useful experimentally. As the resonant peak and resonant frequency depend on the damping ratio $\xi$, if you can experimentally test a system and obtain its frequency response you can also estimate the damping ratio. For example, if the Bode magnitude plot has a large resonant peak, the system is lightly damped, and if the peak is small the system is more heavily damped. The graph on Figure 12.9 shows how varying the damping ratio changes the resonant peak and resonant frequency.

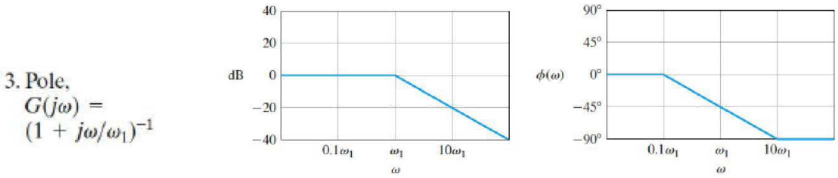
All of the above results are also shown in Figure 12.10.



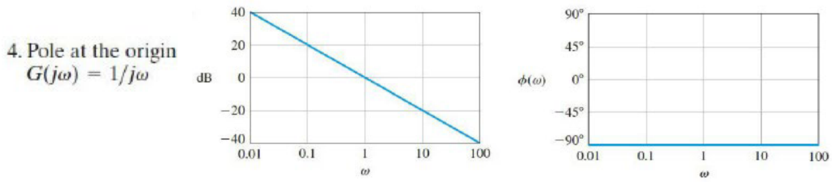

(a) Asymptotic curve for the gain term of a transfer function.



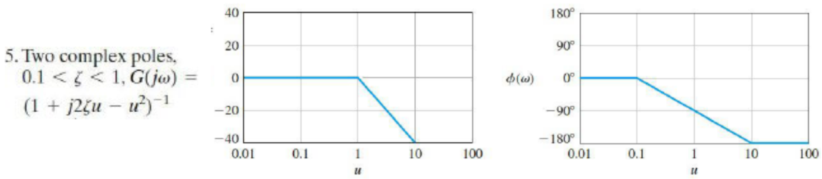

(b) Asymptotic curve for a zero term of a transfer function.




(c) Asymptotic curve for a pole term of a transfer function.




(d) Asymptotic curve for a pole at the origin term of a transfer function.




(e) Asymptotic curve for two complex conjugate pole term of a transfer function.

Figure 12.10: Asymptotic curves for basic terms of a transfer function.

### 12.3.7 Bode Plots in MatLab

Creating a Bode plot in MatLab is easy, you simply use the `bode` command. E.g. for a system with transfer function:

$$G(j\omega) = \frac{5(1 + j0{,}1\omega)}{j\omega(1 + j0{,}5\omega)\left(1 + j0{,}6\left(\frac{\omega}{50}\right) + \left(\frac{j\omega}{50}\right)^2\right)}.$$

We can sketch the Bode plots using the following code:

```
num=5*[0.1 1];
f1=[1 0];
f2=[0.5 1];
```

```
f3=[1/2500 .6/50 1];
den=conv(f1, conv(f2, f3));
sys=tf(num,den)
bode(sys)
```

Which yields the plot on Figure 12.11.

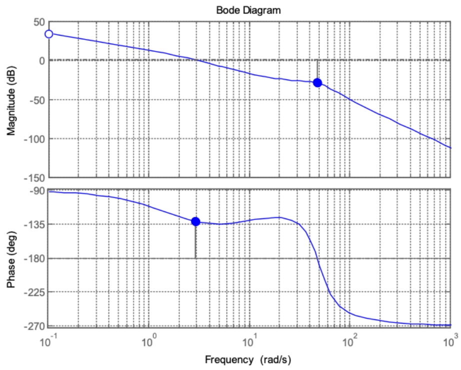



Figure 12.11: Example Bode plot produced in MatLab.

### 12.3.8 Basic Steps for Sketching Bode Plots

**Step 1: Rewrite in Proper Form** We want each factor to look like on of the standard Bode factors:

$$K$$
$$j\omega$$
$$1+\frac{j\omega}{\omega_1}$$
$$\left(1+\frac{j\omega}{\omega_1}\right)^{-1}$$
$$1+2\xi\frac{s}{\omega_n}+\left(\frac{s}{\omega_n}\right)^2.$$

To do this we first want to make the lowest-order term unity. For example,

$$s+10$$

should be rewritten as

$$10\left(1+\frac{s}{10}\right)$$

because now the term inside the brackets has constant term 1, and the break frequency is visible. Similarly,

$$s+3$$

should be rewritten as

$$3\left(1+\frac{s}{3}\right).$$

This is important as Bode sketching depends on seeing break frequencies directly.

**Step 2: Replace $s$ by $j\omega$** Once the transfer function is on the proper form per step 1, set

$$s=j\omega$$

because the Bode plot describes the steady-state response to sinusoidal inputs.

**Step 3: Separate into constituent parts** We now decompose the transfer function into:

- constant gain
- zeros
- poles at the origin
- real poles
- complex poles

where each one has a known Bode contribution.

**Step 4: Draw each part** Now, the magnitude and phase contribution of each individual term is drawn.

**Step 5: Add the results** The final Bode plot is obtained by adding all magnitude curves in dB and adding all phase curves in degrees.

**Example: Example of Sketching a Bode Plot**

Given the transfer function:

$$G(s) = \frac{2500(10+s)}{s(s+2)\left(s^2+30s+2500\right)}.$$

We want to sketch the Bode plot of the system

---

The first step is to rewrite this so all factors have unity lowest-order term. First, we rewrite the numerator:

$$10+s = 10\left(1+\frac{s}{10}\right).$$

Now, we rewrite the first real pole:

$$s+2 = 2\left(1+\frac{s}{2}\right).$$

Then, we rewrite the second-order denominator:

$$s^2+30s+2500 = 2500\left(1+\frac{30}{2500}s+\frac{s^2}{2500}\right) = 2500\left(s+0{,}6\frac{s}{50}+\left(\frac{s}{50}\right)^2\right).$$

Now we collect the constants in the transfer function:

$$G(s) = \frac{5\left(1+\frac{s}{10}\right)}{s\left(1+\frac{s}{2}\right)\left(1+0{,}6\frac{s}{50}+\left(\frac{s}{50}\right)^2\right)}.$$

Now, we replace $s$ with $j\omega$ leading to:

$$G(j\omega) = \frac{5\left(1+\frac{j\omega}{10}\right)}{j\omega\left(1+\frac{j\omega}{2}\right)\left(1+j0{,}6\frac{\omega}{50}+\left(\frac{j\omega}{50}\right)^2\right)}.$$

This makes all the Bode factors visible.
Now we determine the magnitude contribution of each part. From the transfer function we identify a constant gain $K=5$.

$$20\log(5) \approx 14\,\text{dB}$$

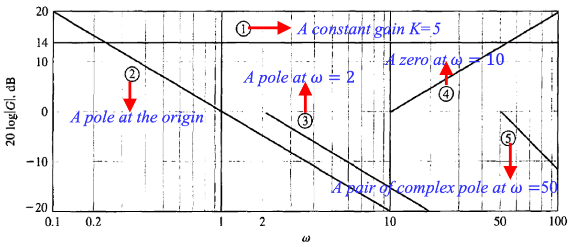
so this contributes a horizontal line at 14 dB.

We also have a pole at the origin:
$$\frac{1}{/j\omega}.$$
This contributes −20 dB/dec over the entire frequency range. It crosses 0 dB at $\omega = 1$ because
$$20\log\left(\frac{1}{1}\right) = 0.$$
We also have a real pole at $\omega = 2$ given by
$$\frac{1}{1+\frac{j\omega}{2}}.$$
This contributes 0 dB before $\omega = 2$ and −20 dB/dec after $\omega = 2$
The zero at $\omega = 10$ caused by
$$1+\frac{j\omega}{10}$$
contributes 0 dB before $\omega = 10$ and 20 dB/dec after $\omega = 10$.
We also have a complex pole pair at $\omega = 50$ caused by:
$$1 + j0{,}6\frac{\omega}{50} + \left(\frac{j\omega}{50}\right)^2.$$
This contributes 0 dB before $\omega = 50$ and −40 dB/dec after $\omega = 50$. These individual contributions are shown on Figure 12.12a.
We can now add the magnitude curves. At low frequency, before $\omega = 2$, the only frequency dependent term is the pole at the origin. So the slope is −20 dB/dec. At $\omega = 2$, the real pole begins to contribute another −20 dB/dec leading the total slope to become −40 dB/dec. At $\omega = 10$ the zero contributes 20 dB/dec so the slope changes to −20 dB/dec. At $\omega = 50$, the complex pole pair contributes −40 dB/dec so the final slope becomes −60 dB/dec. This is shown on Figure 12.12b.
Now we do the same thing for the phase. The constant gain $K = 5$ contributes 0°. The pole at the origin $1/(j\omega)$ gives −90° for all frequencies, so the phase starts at −90°.
The real pole at $\omega = 2$ is a first order pole contributing from 0° to −90°. This transition is sketched on Figure 12.12c, from $0{,}1 \cdot 2 = 0{,}2$ to $10 \cdot 2 = 20$. At the break frequency $\omega = 2$, the contribution is −45°.
The zero at $\omega = 10$ is a first order zero contributing from 0° to 90°. This is sketched from $0{,}1 \cdot 10 = 1$ to $10 \cdot 10 = 100$. At $\omega = 10$, the contribution is 45°.
The complex pole pair at $\omega = 50$ contributes from 0° to −180°. The transition is from $0{,}1 \cdot 50 = 5$ to $10 \cdot 50 = 500$, and at $\omega = 50$, the contribution is −90°. The final phase at very high frequency is therefore $-90° - 90° + 90° - 180° = -270°$. This is shown on Figure 12.12d.
The final Bode plot is the same as Figure 12.11.

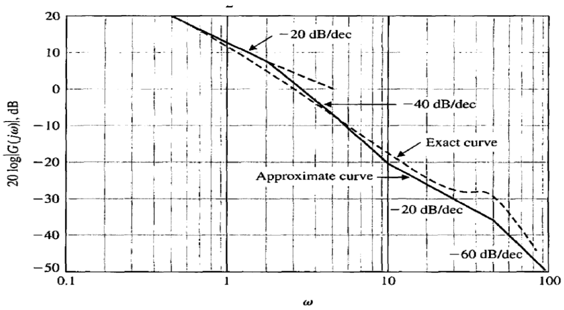
## 12.4 Performance Specification

### 12.4.1 Bandwidth

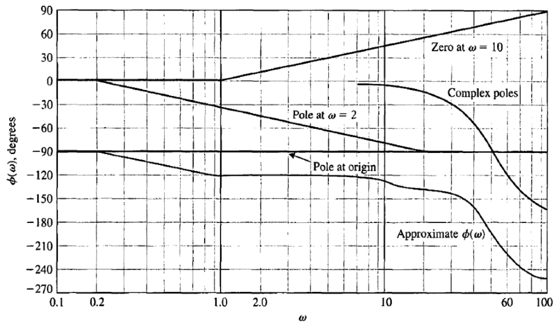
The bandwidth $\omega_B$ is defined as the frequency has dropped by 3 dB from its low-frequency value. So if the low-frequency value is 0 dB, the bandwidth occurs at −3 dB. In ordinary magnitude, −3 dB corresponds to
$$10^{-\frac{3}{20}} \approx 0{,}707 \approx \frac{1}{\sqrt{2}}.$$
So the bandwidth is the frequency, where the system output amplitude has fallen to about 70,7% of the low-frequency amplitude. This often functions as a useful cutoff frequency. Below $\omega_B$, the system reproduces inputs fairly well, whilst above $\omega_B$, the system attenuates systems more strongly. I.e. the bandwidth of the closed-loop



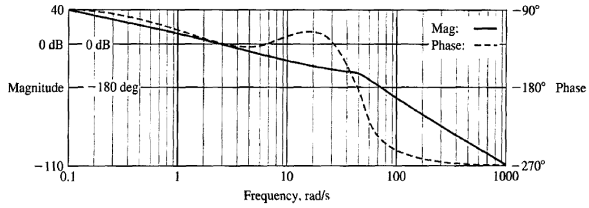

(a) Contributions of the individual terms in the example Bode magnitude plot.




(b) Bode magnitude plot of example system.




(c) Contributions of the individual terms in the example Bode phase plot.




(d) Bode phase plot of example system.

control system measures the range of frequencies over which the system can faithfully reproduce the input.

For example, suppose the reference signal contains both slow and fast components. If the system has low bandwidth, it may follow the slow variations but fail to track the fast ones. If the system has high bandwidth it can respond to faster variations. So in this manner, bandwidth is a measure of tracking ability. In general, a higher $\omega_B$ corresponds to faster step response and shorter settling time.

**Example: First-order Example: Larger Bandwidth Gives Faster Response**

We compare the two first-order closed loop systems:

$$T_1(s) = \frac{1}{s+1}$$
$$T_2(s) = \frac{1}{5s+1}.$$

---

We start by rewriting $T_2$ slightly as:

$$T_2(s) = \frac{1}{5s+1} = \frac{0{,}2}{s+0{,}2}.$$

So $T_1$ has a pole at $s=-1$ and $T_2$ has a pole at $s=-0{,}2$. For a first-order system,

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

the bandwidth is approximately $\omega_B = 1/\tau$. So for $T_1$,

$$\tau_1 = 1 \implies \omega_B = 1$$

and for $T_2$,

$$\tau_2 = 5 \implies \omega_B = 0{,}2.$$

So $T_1$ has a 5x larger bandwidth meaning $T_1$ is faster. This is also shown on the step input plot and ramp input plots on Figure 12.13.

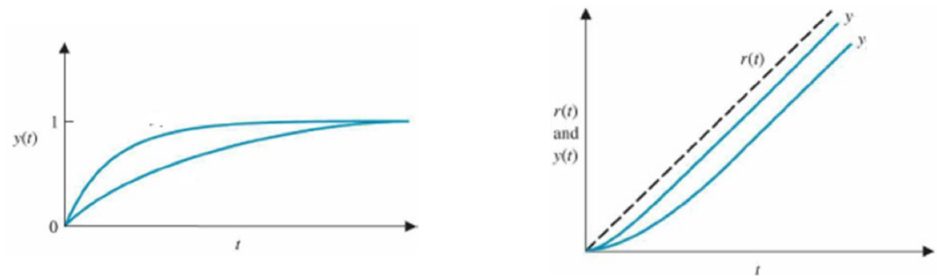



Figure 12.13: Step and ramp responses for the two first-order systems with different bandwidths.

**Example: Second-Order Example: Same Damping; Different Natural Frequency**

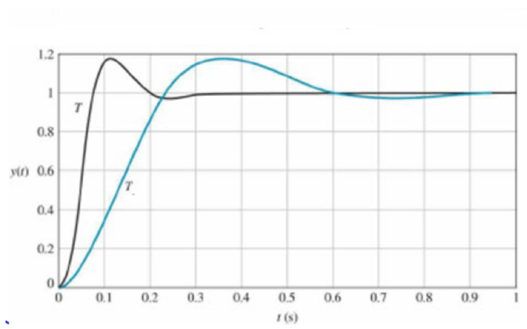
Now we are concerned with two second-order systems:

$$T_3(s) = \frac{100}{s^2 + 10s + 100}$$
$$T_4(s) = \frac{900}{s^2 + 30s + 900}.$$

Comparing these with the standard second-order form:

$$T(s) = \frac{\omega_n^2}{s^2 + 2\xi\omega_n s + \omega_n^2}.$$

For $T_3$, $\omega_n^2 = 100 \implies \omega_n = 10$ and $2\xi\omega_n = 10 \implies \xi = 0{,}5$. And for $T_4$, $\omega_n^2 = 900 \implies \omega_n = 30$ and $2\xi\omega_n = 30 \implies \xi = 0{,}5$. So the two systems have the same damping ratio, but different natural frequencies. Because $T_4$ has a natural frequency three times larger, it is basically a faster version of $T_3$. This is also evident on Figure 12.14, where the table shows that the bandwidth triples. The overshoot however stays the same as this depends mainly on the damping ratio $\xi$, which is equal for both systems. However both the peak time and settling time is approximately three times quicker for the system with the higher natural frequency.




| Performance | T3 | T4 |
|---|---|---|
| $\xi$ | 0.5 | 0.5 |
| $\omega_n$ | 10 | 30 |
| $\omega_B$ | 12.7 | 38.1 |
| P.O. | 16% | 16% |
| $T_p$ | 0.36 | 0.12 |
| $T_s$ | 0.8 | 0.27 |

Figure 12.14: Comparison of the two second-order systems for a step input and for various important parameters.

In general, the bandwidth $\omega_B$ of a second order system:

$$T(s) = \frac{\omega_n^2}{s^2 + 2\xi\omega_n s + \omega_n^2}$$

is related to the natural frequency $\omega_n$. A plot of $\omega_B/\omega_n$ against damping ratio $\xi$ is shown on Figure 12.15.

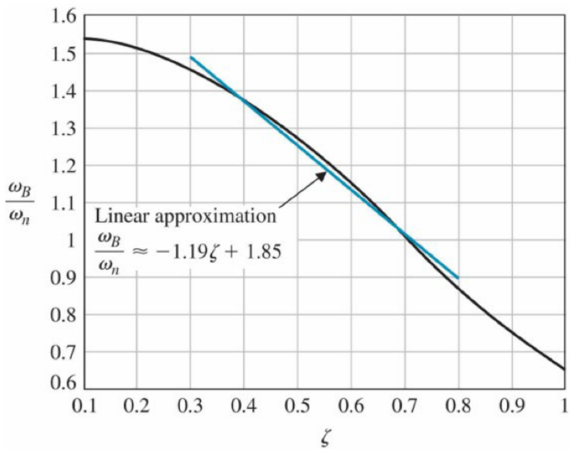



Figure 12.15: Bandwidth against damping ratio.

So instead of bandwidth being equal to the natural frequency we have that:

$$\omega_B \approx \text{constant depending on } \xi \cdot \omega_n.$$

On Figure 12.15 we see that for $\xi = 0{,}5$, $\omega_B/\omega_n \approx 1{,}27$, so for $\xi = 0{,}5$ we have that $\omega_B \approx 1{,}27\omega_n$.

### 12.4.2 Relative Stability in the Frequency Domain

We are now looking at the open-loop transfer function

$$L(j\omega) = G_c(j\omega)G(j\omega)H(j\omega)$$

and not the closed loop transfer function $T(j\omega)$. E.g.

$$L(j\omega) = \frac{K}{j\omega\left(j\omega+1\right)\left(0{,}2j\omega+2\right)}.$$

The critical point for closed-loop stability is the point $-1 + j0$ in the complex $L(j\omega)$-plane. On a Bode plot, that same point corresponds to $\left|L(j\omega)\right| = 1$, which is $20\log\left|L(j\omega)\right| = 0\,\text{dB}$ and phase $\angle L(j\omega) = -180°$. So the dangerous situation is 0 dB and $-180°$. I.e. if the open-loop system reaches magnitude 1 while having $-180°$ phase, then the negative feedback effectively becomes positive feedback, which is why this point is critical.

On Figure 12.16 two important margins are shown. First is the phase margin, i.e. how far the phase is from $-180°$ at the frequency corresponding to a magnitude of 0 dB and the gain margin i.e. how far the magnitude is from 0 dB when the phase is $-180°$. These margins tell us not just whether the system is stable, but how safely stable it is.

The same margins idea can be applied to a log-magnitude-phase diagram (essentially a Nichols-style plot). Instead of plotting magnitude and phase separately against frequency, we plot

$$20\log\left|L(j\omega)\right|$$

against $\angle L(j\omega)$. So the horizontal axis is the phase, and the vertical axis is magnitude in dB. The critical points become very easy to see at $-180°$ and 0 dB. This is all shown on Figure 12.17.

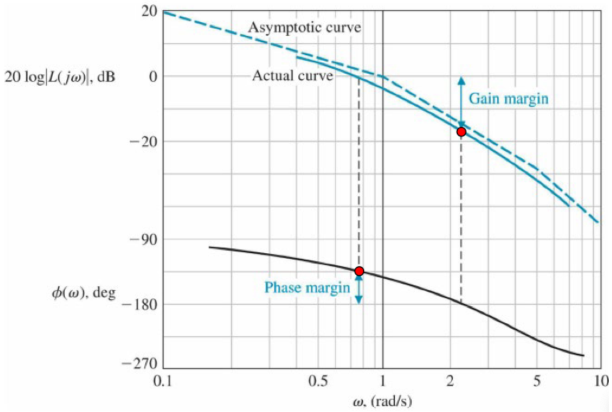



Figure 12.16: Phase and gain margin.

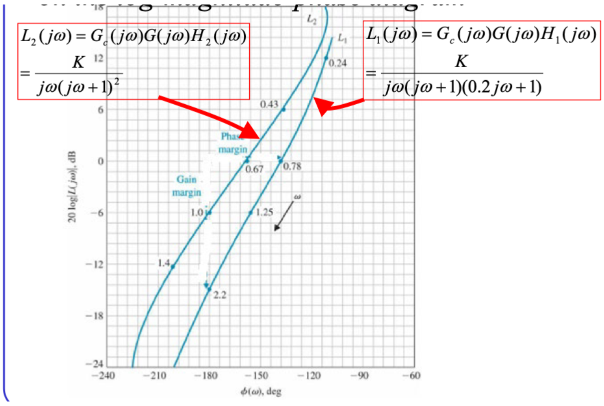



Figure 12.17: Phase and gain margin on a log-magnitude-phase diagram.

The two curves on Figure 12.17 correspond to two loop transfer functions:

$$L_1(j\omega) = \frac{K}{j\omega\left(j\omega+1\right)\left(0{,}2j\omega+1\right)}$$

and

$$L_2(j\omega) = \frac{K}{j\omega\left(j\omega+1\right)^2}.$$

The closer the curve passes to the point $(-180°, 0\,\text{dB})$, the less stable the system is.

On Figure 12.17 the actual margins are, for $L_1$:

$$\text{GM} = 15\,\text{dB}$$
$$\text{PM} = 43°$$

and for $L_2$:

$$\text{GM} = 5{,}7\,\text{dB}$$
$$\text{PM} = 20°.$$

So $L_2$ is less stable than $L_1$ as its margins are smaller. It can tolerate less extra gain and less extra phase lag before becoming unstable.

**Phase Margins and Damping Ratios** We now consider a second-order open-loop transfer function:

$$L(j\omega) = \frac{\omega_n^2}{s(s+2\xi\omega_n)}.$$

This open-loop form gives a standard second-order closed-loop system.

The exact relationship between phase margin and damping ratio is shown on Figure 12.18 and given by:

$$\phi_{\text{pm}} = \tan^{-1}\frac{2}{\sqrt{\sqrt{4+\frac{1}{\xi^4}}-2}}.$$

However an important approximation is $\xi = 0{,}01\phi_{\text{pm}}$ for $\xi \leq 0{,}7$, where $\phi_{\text{pm}}$ is measured in degrees. So roughly:

$$\begin{aligned} \text{PM} = 30^\circ &\implies \xi &&\approx 0{,}3 \\ \text{PM} = 45^\circ &\implies \xi &&\approx 0{,}45 \\ \text{PM} = 60^\circ &\implies \xi &&\approx 0{,}6. \end{aligned}$$

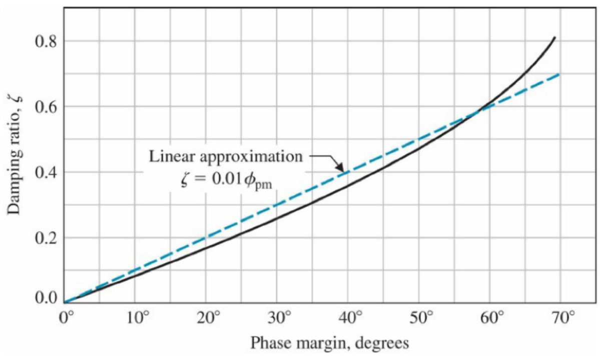



Figure 12.18: Dependence of phase margins on damping ratio.

This is useful because damping ratio tells us about transient behavior. A low damping ratio means more oscillation and overshoot and vice versa.

So phase-margins becomes a frequency-domain way of estimating transient response quality.

**Margins versus gain $K$** To show how stability margins change as the loop gain $K$ changes we use the example

$$L(s) = G_c(s)G(s)H(s) = \frac{K}{s(s+4)^2}.$$

As $K$ increases, the magnitude plot moves upwards.

That usually causes the gain crossover frequency to move to the right, where the system often has more phase lag, therefore, increasing $K$ tends to reduce the phase margin. This is shown on the plots on Figure 12.19. Here we see the gain margin versus the gain $K$, the phase margin versus the gain $K$ and the gain margin versus the phase margin.

On Figure 12.19, there is a critical value:

$$K^* = 128.$$

At this value, the system is marginally stable. The gain margin is then 0 dB and the phase margin is approximately $0^\circ$. Here:

$$\text{GM} = 20\log\left(\frac{K^*}{K}\right).$$

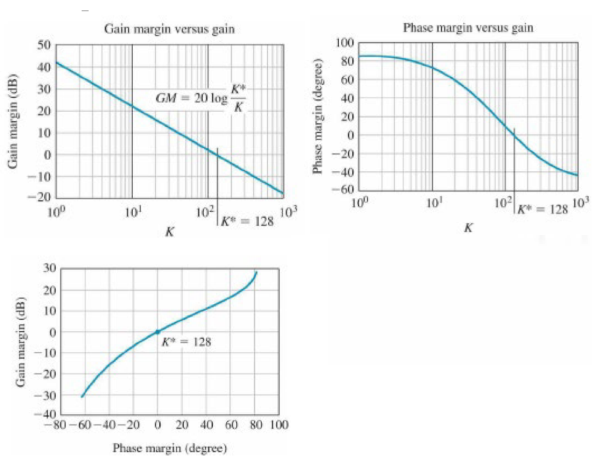



Figure 12.19: Influence of gain $K$ on stability margins.

So if the current gain $K$ is much smaller than $K^*$ you have a large gain margin and vice versa.

**Desired Frequency-Domain Specifications** A good control system should have:

1. Small resonant magnitude. E.g. $M_{p\omega} < 1{,}5$, as a small resonant peak means less overshoot and better damping.
2. Large bandwidth. A larger bandwidth usually means a faster time response, as this leads to transients dying out more quickly.
3. A large phase margin. This helps keep the system away from the dangerous $-180°$, 0 dB point, meaning the system is less likely to become unstable.

### 12.4.3 Stability with Time Delays

Now, we introduce time delays. A time delay is the time between an action and its effect being observed somewhere in the system.

A pure time delay is represented by:

$$G_d(s) = e^{-sT}$$

where $T$ is the delay time. Substituting $s = j\omega$, we get

$$G_d(j\omega) = e^{-j\omega T}.$$

The magnitude is:

$$\left|e^{-j\omega T}\right| = 1$$

so time delay does not change the magnitude plot.

But the phase is:

$$\phi(\omega) = -\omega T$$

in radians, and

$$\phi(\omega) = -\omega T \frac{180°}{\pi}$$

in degrees. So time delay is dangerous because it adds phase lag without changing the magnitude. On the Bode magnitude plot, nothing obvious happens, but on the phase plot, the system gets closer to $-180°$, so the phase margin decreases.

**Lecture 9: The Design of Feedback Control Systems** 28. April 2026

# 13 The Design of Feedback Control Systems

## 13.1 Introduction and Concepts

Often, it is not enough to only adjust one system parameter. For example, suppose your motor system is too slow, overshoots too much, or has steady-state error. You might try increasing the gain $K$, to make the response faster, however this will, in turn, create more overshoot thereby worsening one of the other problems.

To alleviate this, a compensator might be added. Such a component can either be electrical, mechanical, hydraulic, pneumatic, or even purely software-based. A compensator is often just written as a transfer function $G_c(s)$, implemented digitally or analogically. The compensator transfer function is defined as all other transfer functions as

$$G_c(s) = \frac{E_o(s)}{E_{\text{in}}(s)}.$$

The diagrams on Figure 13.1 depict different places where the compensator can be inserted.

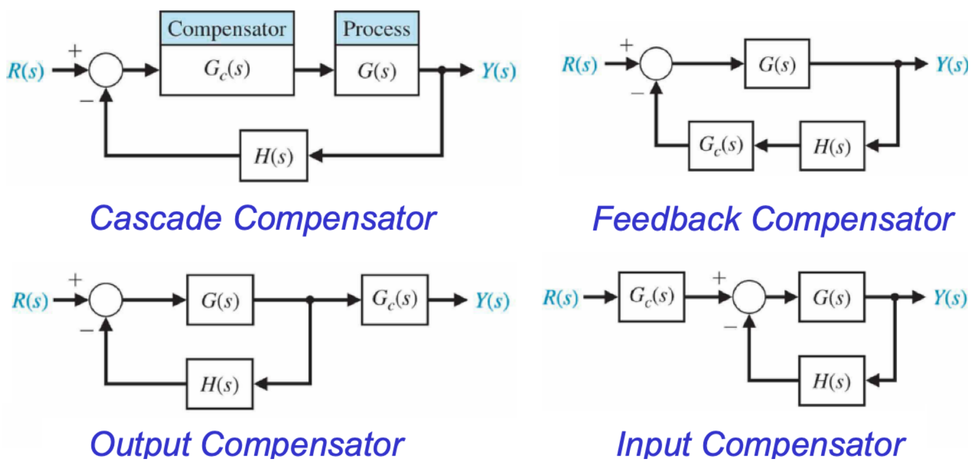



Figure 13.1: Different types of compensator placements.

The most common form of a compensator is a cascade compensator, which is just placed in series with the plant. In this case, the loop transfer function becomes:

$$L(s) = G_c(s)G(s)H(s).$$

Another option is the feedback controller, where you are modifying the feedback signal before it is compared with the reference. This can be useful if you want to shape sensor dynamics, filter feedback or change how the output is fed back, but it is usually less intuitive than a cascade controller.

One may also place the compensator after the plant output, thereby creating an output compensator, meaning the feedback loop may regulate an internal signal and then $G_c(s)$ modifies the final output signal. This can be useful for output filtering or signal shaping, but it doe not affect the main feedback loop in the same direct way as a cascade or feedback compensator does.

Placing the compensator before the summing junction acting on the reference input creates an input compensator, which is essentially a prefilter. It changes how the command enters the system, but it does not directly change the feedback loop stability. E.g. one might choose to smooth a step input before feeding it to the controller so the motor does not get a harsh command.

There are two major classical design approaches in this field. The first of these is the root locus method, which focuses on the closed-loop poles, as these determine much of the transient response. This is especially connected to time-domain specs such as settling time, overshoot, damping ratio and natural frequency.

The other major method is the frequency response method, which uses Bode plots. Instead of directly looking at pole movement, you instead reshape the magnitude and phase of the loop transfer function. This lets you tune things like bandwidth, phase margin, gain margin, and noise sensitivity.

The general form of a compensator is

$$G_c(S) = \frac{K\prod_{i=1}^{M}(s+z_i)}{\prod_{j=1}^{n}\left(s+p_j\right)}.$$

Meaning the compensator is representable by a gain $K$, zeros at $s = -z_i$ and poles at $s = -p_j$. The design problem then becomes choosing appropriate values of $K, z_i, p_j$ to obtain the desired closed-loop system behaviour.

From the above, the simplest compensator is a so-called first-order compensator:

$$G_c(s) = \frac{K(s+z)}{s+p}.$$

This has one zero at $s = -z$, one pole at $s = -p$ and a gain $K$. So here, the design problem is simply choosing a value of $K, z, p$.

## 13.2 Phase-Lead Compensation

From above, the basic form of a first-order compensator is

$$G_c(s) = \frac{K(s+z)}{s+p}.$$

For a phase-lead compensator, the zero is closer to the origin than the pole, i.e. $|z| < |p|$ or, since $z$ and $p$ are positive constants $z < p$. This is shown on the pole-zero plot on Figure 13.2.

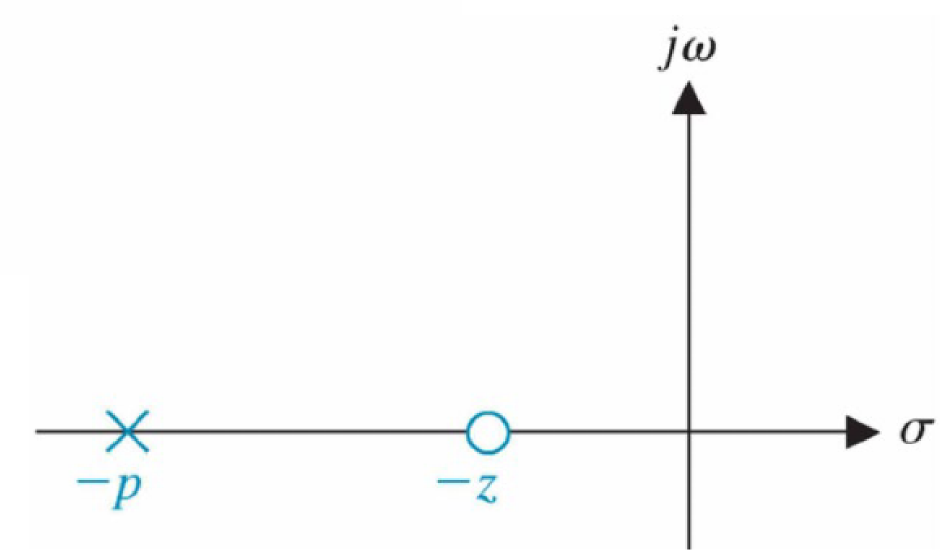



Figure 13.2: Pole-zero plot for a phase-lead compensator.

A phase-lead compensator adds positive phase over some frequency range, thereby making the system behave as if it has less delay. This is very useful as phase lead can improve the phase margin and thereby improve stability margin, damping, overshoot, and speed. So a phase-lead compensator is usually used when the system is too sluggish, too oscillatory, or has too little phase margin.

We can rewerite

$$G_c(j\omega) = \frac{K(j\omega + z)}{j\omega + p}$$

in a more useful frequency-response form. We start by factoring $p$ and $z$ as:

$$G_c(j\omega) = \frac{Kz}{p}\frac{1 + j\frac{\omega}{z}}{1 + j\frac{\omega}{p}}.$$

Then, we define $\alpha = p/z$, so for phase lead $\alpha > 1$. We also define $\tau = 1/p$, and since $p = \alpha z$ we get $1/z = \alpha\tau$. Therefore

$$G_c(j\omega) = K_1\frac{1 + j\alpha\omega\tau}{1 + j\omega\tau}, \quad K_1 = \frac{K}{\alpha} \quad \alpha > 1.$$

This is the standard lead-compensator form.

### 13.2.1 Bode Plot of a Phase-Lead Compensator

The Bode plot on Figure 13.3 depicts the magnitude and phase as functions of frequency for the standard phase-lead compensator. For the normalized lead compensator

$$\frac{G_c(j\omega)}{K_1} = \frac{1 + j\alpha\omega\tau}{1 + j\omega\tau}$$

the zero occurs at

$$\omega = z = \frac{1}{\alpha\tau}$$

and the pole occurs at

$$\omega = p = \frac{1}{\tau}.$$

Since $\alpha > 1$, the zero comes before the pole:

$$\frac{1}{\alpha\tau} < \frac{1}{\tau}.$$

Which is why the magnitude plot is flat at low frequencies, then rises with slope 20 dB/dec between the zero and pole and then flattens out again. The high-frequency gain increase is $20\log\alpha$, which is the aforementioned extra high-frequency gain caused by a lead compensator. Therefore a phase-lead compensator may be very useful for improving speed and bandwidth, but it can also start to amplify high-frequency noise if overdone.

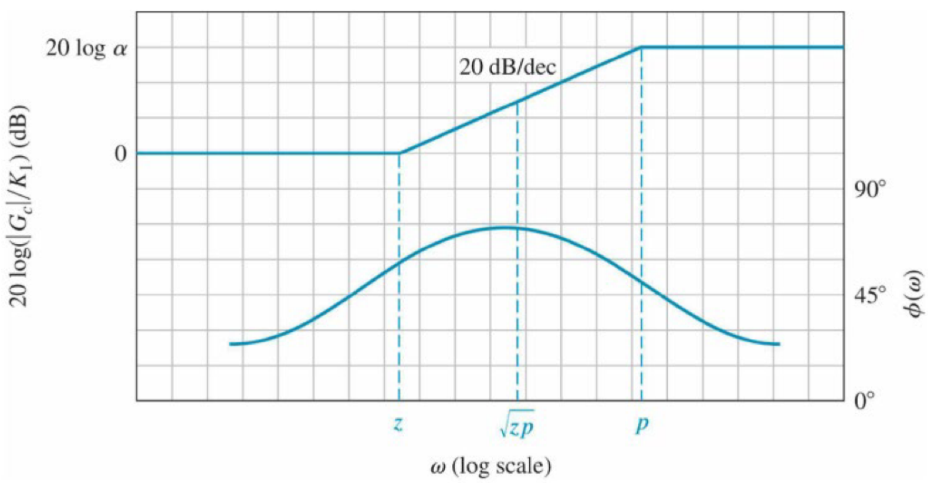



Figure 13.3: Bode plot of a phase-lead compensator.

The phase angle of the lead compensator on Figure 13.3 is given by

$$\phi(\omega) = \tan^{-1}(\alpha\omega\tau) - \tan^{-1}(\omega\tau).$$

The numerator zero contributes positive phase, and the denominator pole contributes negative phase, but because the zero happens first, the positive phase contribution appears before the negative pole contribution catches up. So for a range of frequencies $\phi(\omega) > 0$; that is the phase lead. The curve rises, reaches a maximum, and then falls back toward zero.

The maximum phase lead occurs halfway between the zero and the pole on a logarithmic frequency scale:

$$\omega_m = \sqrt{zp} = \frac{1}{\tau\sqrt{\alpha}}.$$

This is very important in compensator design. Usually, you place $\omega_m$ near the desired gain crossover frequency, as that is where we want the most extra phase margin.

We also have that

$$\sin\phi_m = \frac{\alpha - 1}{\alpha + 1}$$

which tells how much maximum phase lead you get from a chosen pole-zero ratio $\alpha$. This can also be solved as

$$\alpha = \frac{1 + \sin\phi_m}{1 - \sin\phi_m}$$

so if you know how much extra phase you need, you can calculate the required $\alpha$. In theory, as $\alpha \to \infty$, the phase lead approaches 90°, as shown on Figure 13.4.

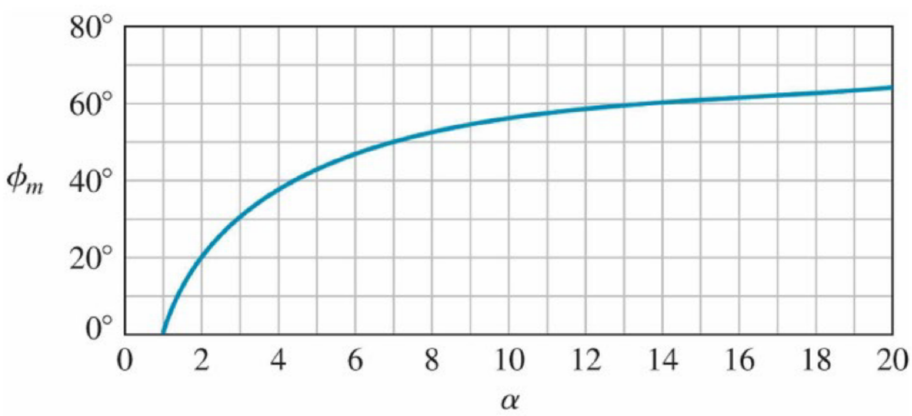



Figure 13.4: Maximum phase lead $\phi_m$ and pole-zero ratio $\alpha$

However, in practice, single lead stages are usually kept below roughly 60° to 70°, because large $\alpha$ give too much gain boost and noise sensitivity. If more phase lead is needed, you normally use two lead compensators instead of one extreme one.one extreme one.

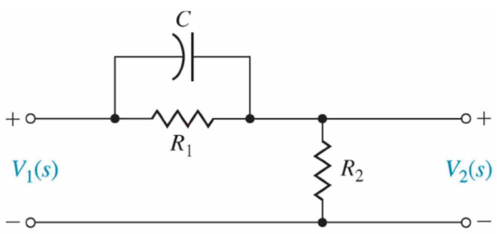



Figure 13.5: Simple RC-realization of a phase-lead compensator.

The RC-network on Figure 13.5 realizes the transfer function

$$G_c(s) = \frac{1+\alpha\tau s}{\alpha(1+\tau s)}.$$

The important part here is not the exact circuit algebra, but the structure. For the circuit shown:

$$\tau = \frac{R_1 R_2}{R_1 + R_2} C$$

and

$$\alpha = \frac{R_1 + R_2}{R_2}.$$

Since $R_1 + R_2 > R_2$, we get $\alpha > 1$, which is the condition for phase lead. As the circuit is passive, it also includes attenuation, hence the $1/\alpha$ factor in front of the transfer function.

## 13.3 Phase-Lag Compensation

The transfer function for a phase-lag compensator is normally written as

$$G_c(s) = \frac{1+\tau s}{1+\alpha\tau s}$$

with $\alpha > 1$. The above can also be written as

$$G_c(s) = \frac{1}{\alpha}\frac{s+z}{s+p}.$$

Where $z = 1/\tau$ and $p = 1/\alpha\tau$. Since $\alpha > 1$, $p < z$, so the pole is closer to the origin than the zero as shown on Figure 13.6.

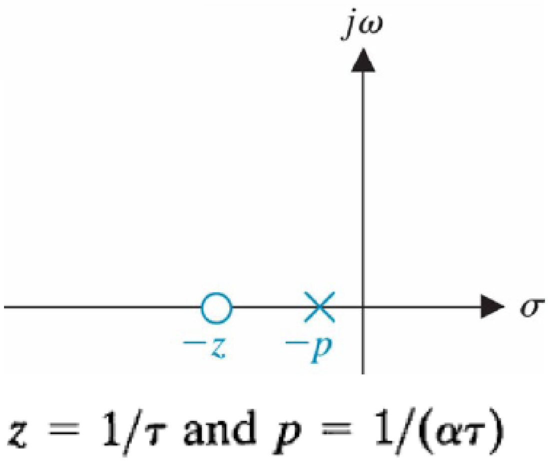



Figure 13.6: Pole-zero plot for a phase-lag compensator

A phase-lag compensator adds negative phase over some frequency range. This seems contrary to our goal of stability as this seemingly decreases phase margin, however phase-lag compensators are mainly used to improve steady-state accuracy without strongly changing the crossover frequency. This is because a lag-compensator can increase the effective low-frequency gain compared with high-frequency gain, which helps reduce steady-state error.

A phase-lag network behaves like an integrator over a finite range of frequencies. An ideal integrator has transfer function $1/s$ and gives $-20$ dB/dec with phase $-90°$. A lag-compensator is not a true integrator, but between its pole and zero, its Bode magnitude falls with approximately $-20$ dB/dec and its phase becomes negative. So in this way a phase-lag compensator behaves integrator-like over a limited frequency range.

### 13.3.1 Bode Plot of a Phase-Lag Compensator

The Bode plot for lag compensation, shown on Figure 13.7 is basically the mirror image of the phase-lead case. The pole comes first at

$$\omega = p = \frac{1}{\alpha\tau}$$

and then the zero comes later

$$\omega = z = \frac{1}{\tau}$$

so the magnitude starts flat, then drops $-20$ dB/dec, between the pole and the zero and then becomes flat again at a lower gain.

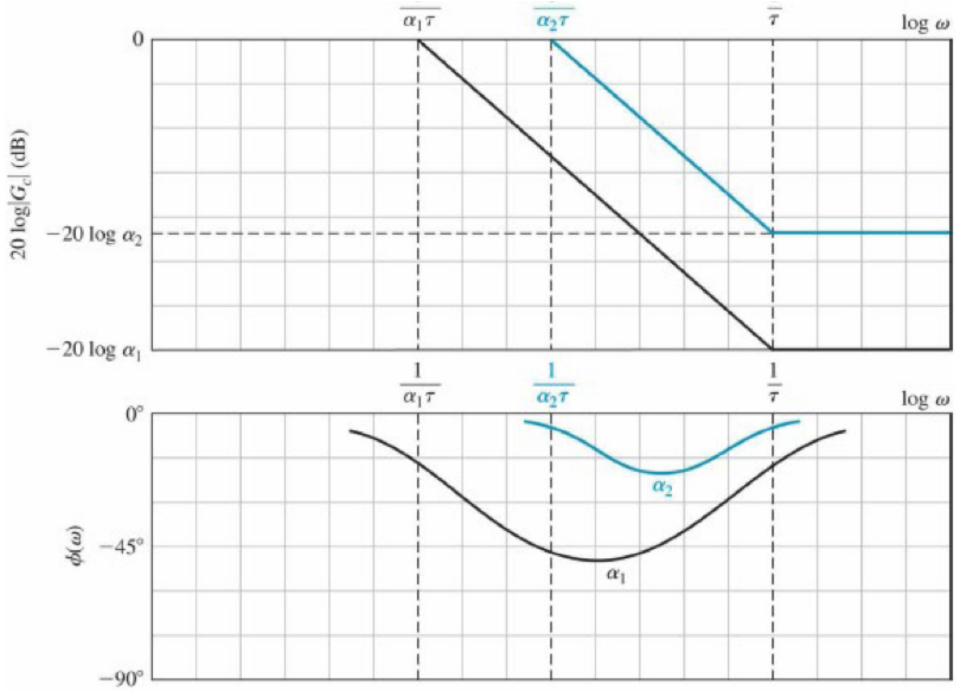



Figure 13.7: Bode plot of a phase-lag compensator

The phase becomes negative between the pole and zero on Figure 13.7, and the maximum phase lag occurs at

$$\omega_m = \sqrt{zp}$$

just like the lead case, but now it is a maximum negative phase angle instead of a positive one.

## 13.4 Phase-Lead Design

In the following we will be going through the phase-lead design process first through the Bode diagram method in Section 13.4.1 and then the root locus method in Section 13.4.2 through two worked examples.

### 13.4.1 Phase-Lead Design Using the Bode Diagram

The overall goal in the following section is to take an uncompensated system with poor stability margin and add a compensator $G_c(s)$ to give extra positive phase near the gain crossover frequency.

The general Bode design procedure in this context is to

1. start with the uncompensated open-loop system,
2. check its current phase margin,
3. decide how much extra phase is needed,
4. choose $\alpha$ from the required phase lead,
5. plade the compensator zero and pole around the new crossover frequency
6. check the final Bode plot.

The important trick here is that at the frequency of maximum phase lead, the compensator adds a magnitude boost of $10\log\alpha$, so if we want the new crossover to happen at $\omega_n$, you simply choose $\omega_m$, where the uncompensated magnitude is $-10\log\alpha$ dB. Then once the compensator adds $10\log\alpha$, the magnitude becomes 0 dB and that frequency becomes the new gain crossover frequency.

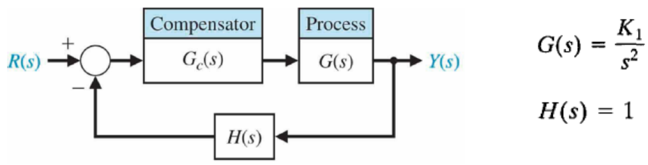



Figure 13.8: Example system for Bode-phase-lead design.

We consider the system on Figure 13.8 with $G(s) = K_1/s^2$ and $H(s) = 1$. This is a type-two system as the open-loop transfer function has two pure integrators $1/s^2$. The specifications are a settling time $T_s \leq 4$ s and a damping constant $\xi \geq 0{,}45$.

Without the compensator the closed loop transfer function of the system on Figure 13.8 is

$$T(s) = \frac{Y(s)}{R(s)} = \frac{K_1}{s^2 + K_1}.$$

Comparing this with the standard second order form:

$$T(s) = \frac{\omega_n^2}{s^2 + 2\xi\omega_n + \omega_n^2}$$

i.e. $2\xi\omega_n = 0 \implies \xi = 0$, so the system is unddampened and the poles are purely imaginary $s = \pm j\sqrt{K_1}$. This means the system oscillates without settling properly, and since the system is undamped the requirements are not fulfilled. Changing $K_1$ alone changes the natural frequency, but it doe not add damping.

For a second-order system, our settling time is given as:

$$T_s \approx \frac{4}{\xi\omega_n}.$$

We require $T_s \leq 4$ and $\xi \geq 0{,}45$ so

$$4 \geq \frac{4}{0{,}45\omega_n} \implies \omega_n \geq \frac{1}{0{,}45} = 2{,}22\,\text{rad/s}.$$

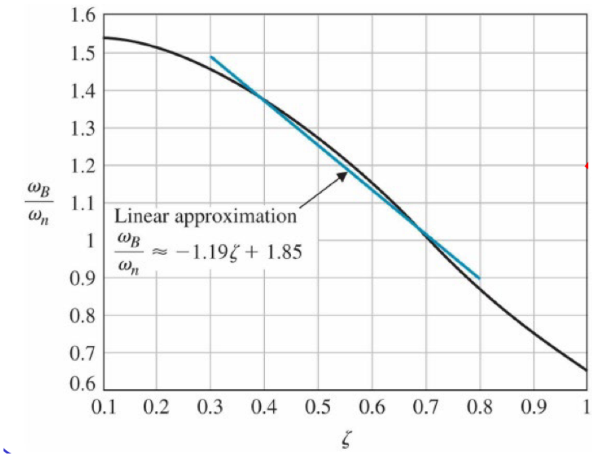



Figure 13.9: Closed-loop bandwidth $\omega_B$ compared to natural frequency $\omega_n$

For $\xi = 0{,}45$, Figure 13.9 gives approximately:

$$\begin{aligned} \frac{\omega_B}{\omega_n} &\approx 1{,}33 \\ \omega_B &= 1{,}33\omega_n \\ &= 1{,}33 \cdot 2{,}22\,\mathrm{rad/s} \\ &= 3{,}0\,\mathrm{rad/s}. \end{aligned}$$

So the closed-loop bandwidth of the system should be roughly $\omega_B \approx 3{,}0\,\mathrm{rad/s}$.

For the uncompensated system,

$$T(s) = \frac{K_1}{s^2 + K_1}$$

so $\omega_n = \sqrt{K_1}$. From $\omega_B = 1{,}33\omega_n \geq 3{,}0\,\mathrm{rad/s}$ we see that $K_1 = 5$ is enough for the bandwidth, but we opt for $K_1 = 10$ to provide a safety margin. So the uncompensated open-loop transfer function becomes

$$G(s) = \frac{10}{s^2}$$

but thi still has a phase of $-180°$ at all frequencies because

$$\frac{1}{(j\omega)^2} = -\frac{1}{\omega^2}$$

so the phase margin is zero. The gain is therefore now large enough, but the damping/stability margin is still not acceptable.

The required damping ratio is $\xi \geq 0{,}45$. A rough design rule is that $\xi \approx 0{,}01\phi_{pm}$, where $\phi_{pm}$ is in degrees. So

$$\phi_{pm} \approx \frac{\xi}{0{,}01} = \frac{0{,}45}{0{,}01} = 45°.$$

The uncompensated system has approximately zero phase margin, so the compensator should add roughly 45° of phase lead.

The maximum phase lead is given by

$$\sin\phi_m = \frac{\alpha - 1}{\alpha + 1}.$$

Here $\phi_m = 45°$, so

$$\sin 45° = \frac{\alpha - 1}{\alpha + 1} \implies \alpha \approx 5{,}8.$$

We choose $\alpha = 6$ to provide a little safety margin.

Now we find the new crossover frequency by computing $10 \log \alpha$ with $\alpha = 6$, we get

$$10 \log 6 = 7{,}78\,\mathrm{dB}.$$

The design rule says to find the frequency, where the uncompensated magnitude is $-7{,}78$ dB. For the uncompensated open-loop system:

$$G(j\omega) = \frac{10}{(j\omega)^2}.$$

The magnitude is

$$|G(j\omega)| = \frac{10}{\omega^2}.$$

Or in dB:

$$20 \log |G(j\omega)| = 20 \log \left(\frac{10}{\omega^2}\right).$$

This equals $-7{,}78$ dB at about $\omega_n = 4{,}95\,\mathrm{rad/s}$. So we place the maximum phase lead at $\omega_m = 4{,}95\,\mathrm{rad/s}$.

For a lead compensator:

$$\omega_m = \sqrt{zp}.$$

and

$$\alpha = \frac{p}{z}.$$

So

$$p = \omega_m \sqrt{a}$$
$$z = \frac{p}{\alpha}.$$

Using

$$\omega_m = 4{,}95$$
$$\alpha = 6.$$

Gives

$$p = 4{,}95\sqrt{6} \approx 12{,}0 \qquad z = \frac{12{,}0}{6} = 2{,}0.$$

So the compensator zero is at $s = -2$ and the compensator is at $s = -12$. Using the pole and zero values the lead compensator can be written as

$$G_c(s) = \frac{1 + \frac{s}{2}}{1 + \frac{s}{12}} = \frac{6(s+2)}{s+12}.$$

The uncompensated plant was

$$G(s) = \frac{10}{s^2}$$

and the compensator is

$$G_c(s) = \frac{1 + \frac{s}{2}}{1 + \frac{s}{12}},$$

so the compensated loop transfer function is:

$$\begin{aligned} L(s) &= G_c(s)G(s) \\ &= \frac{10}{s^2}\frac{1+\frac{s}{2}}{1+\frac{s}{12}} \\ &= \frac{60(s+2)}{s^2(s+12)}. \end{aligned}$$

For unity feedback,

$$T(s) = \frac{L(s)}{1+L(s)} = \frac{60(s+2)}{s^3+12s^2+60s+120} \approx \frac{60(s+2)}{\left(s^2+6s+20\right)(s+6)}.$$

The dominant second-order part is approximately $s^2+6s+20$. Comparing this with the standard second-order form $s^2+2\xi\omega_n s+\omega_n^2$, we see that $\omega_n = 4{,}47$ and $\xi \approx 0{,}67$. Hence, the damping ratio $\xi \approx 0{,}67 \geq 0{,}45$ exceeds the requirement. The approximate settling time is

$$T_s \approx \frac{4}{\xi\omega_n} = \frac{4}{0{,}67 \cdot 4{,}47} = 1{,}34\,\mathrm{s}.$$

So the requirement for the settling time $T_s \leq 4\,\mathrm{s}$ is also satisfied.

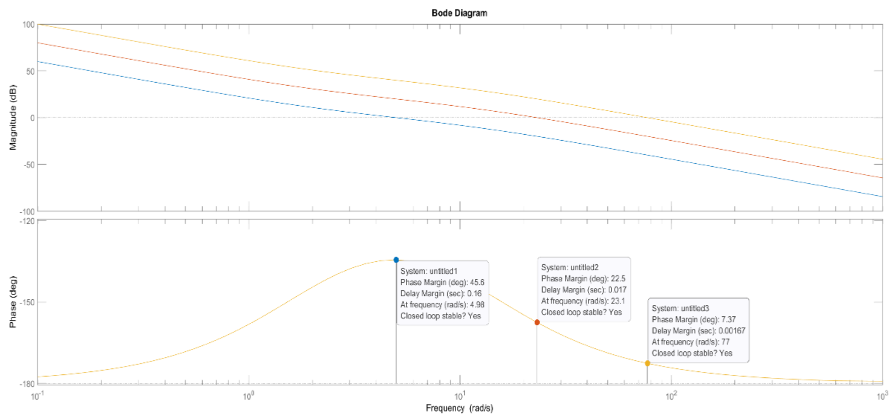



Figure 13.10: Bode plot for the compensated loop transfer function.

On Figure 13.10 the Bode plots for the compensated loop transfer function is shown. The phase margin here is roughly 45,6° at around $\omega \approx 4{,}98\,\mathrm{rad/s}$, that matches the design goal of $\phi_{pm} = 45°$ almost exactly.

### 13.4.2 Phase-Lead Design using Root Locus

The overall goal here, is to choose $G_c(s)$ so that the closed-loop poles move to a desired location in the $s$-plane. For this example we will continue working with the system from before with $G(s) = K_1/s^2$ and unity feedback as depicted on Figure 13.11.

The compensator is chosen as a first-order lead compensator:

$$G_c(s) = \frac{s+z}{s+p}.$$

For phase lead, we have that $|z| < \left|p\right|$, so the zero is closer to the origin than the pole. The purpose of adding this pole-zero pair is to reshape the root locus so that it passes through the desired closed-loop pole location.

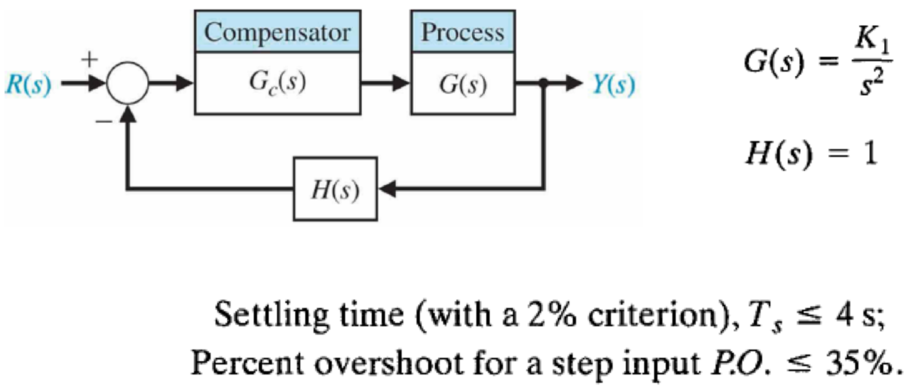



Figure 13.11: Example system for root locus phase-lead design.

The general root-locus design recipe is, to start from performance requirements, e.g.:

$$T_s \leq 4\,\text{s}$$
$$P.O. \leq 35\%.$$

Then you translate those into a desired dominant pole location:

$$s_d = -\sigma \pm j\omega_d$$

where the real part $-\sigma$ controls the settling speed, and the imaginary part $\omega_d$ controls oscillation frequency. Then, you

1. draw the uncompensated root locus,
2. check whether the desired point lies on it,
3. if not, add a phase-lead compensator,
4. place the compensator zero,
5. choose the compensator pole so that the angle condition is satisfied,
6. compute the required gain using the magnitude condition.

The root locus condition is

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

so the point on the root locus must satisfy

$$\angle L(s) = (2k+1)\,180^\circ$$

and $|L(s)| = 1$. The angle condition tells you where the pole and zero should go, and the magnitude condition tells you the required gain.

The root locus is prepared as explained in Section 10, however, for completenes, the most important rules will be reiterated here. First of all, the root locus starts at open-loop poles and ends at open-loop zeros. If there are more poles than zeros, some branches go to infinity. The number of branches is equal to the number of open-loop poles. The locus is symmetric about the real axis. On the real axis, the root locus exists to the left of an odd number of real poles and zeros. For poles going to infinity, the asymptote centroid is:

$$\sigma_A = \frac{\sum \text{poles} - \sum \text{zeros}}{n - M}$$

and the asymptote angles are

$$\phi_A = \frac{(2k+1)\,180^\circ}{n-M}.$$

As mentioned above and on Figure 13.11, the example system is $G(s) = K_1/s^2$, $H(s) = 1$, where we require $T_s \leq 4\,\text{s}$ and $P.O. \leq 35\%$. Currently, the poles of the closed-loop system are $s = \pm j\sqrt{K_1}$, so the closed-loop poles are purely imaginary. That means the root locus lies on the $j\omega$-axis, meaning the uncompensated system is undamped and does not meet the requirements.

The overshoot requirement is $P.O. \leq 35\%$. For a second-order system, overshoot is related to damping ratio $\xi$. A 35% overshoot corresponds roughly to $\xi \geq 0{,}32$. The settling-time requirement is

$$T_s \approx \frac{4}{\xi\omega_n} \implies \xi\omega_n \geq 1.$$

We choose the desired dominant roots as $s_d = -1 \pm j2$. For this point:

$$\omega_n = \sqrt{1^2 + 2^2} = \sqrt{5} \approx 2{,}23$$

and

$$\xi = \frac{1}{\sqrt{5}} \approx 0{,}45.$$

So this chosen point satisfies both requirements

$$\xi = 0{,}45 > 0{,}32$$

and $T_s \approx 4/(0{,}45 \cdot 2{,}23) \approx 4$, so the desired dominant closed-loop poles are:

$$s = -1 \pm j2.$$

We choose to place the compensator zero directly below the desired root location, meaning the compensator zero is places at $s = -1 \implies z = 1$. So at this point

$$G_c(s) = \frac{s+1}{s+p}.$$

We still need to find the compensator pole $p$.

Now, we need the desired point $s_d = -1 + j2$ to lie on the compensated root locus. The compensated transfer function will be:

$$L() = \frac{K_1(s+1)}{s^2(s+p)}.$$

The angle condition says that $\angle L(s_d) = -180^\circ$. The contributions come from two poles at the origin, one zero at $-1$, and one unknown pole at $-p$. At the desired point, the angle from each pole to the origin is about $116^\circ$. There are two of them, so the pole contribution from these is $-232^\circ$. The zero at $-1$ points vertically upward to the desired point, so its angle contribution is $90^\circ$. Therefore without the unknown pole:

$$\phi = -232^\circ + 90^\circ = -142^\circ.$$

So the missing pole contribution must be

$$-180^\circ = -142^\circ - \theta_p \implies \theta_p = 38^\circ.$$

So we draw a line from the desired root location at an angle of 38° down to the real axis as shown on Figure 13.12. This line intersects the real axis at approximately $s = -3{,}6 \implies p = 3{,}6$.

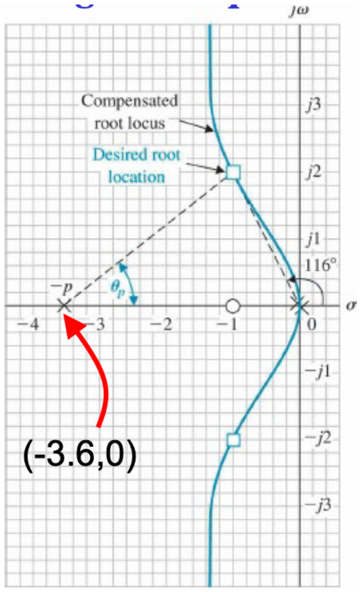



Figure 13.12: Compensated root locus with desired root locations and location of pole.

So the compensator is

$$G_c(s) = \frac{s+1}{s+3{,}6}.$$

As $1 < 3{,}6$, this is a phase-lead compensator. The compensated loop transfer function is then:

$$\begin{aligned} L(s) &= G_c(s)G(s) \\ &= \frac{K_1(s+1)}{s^2(s+3{,}6)}. \end{aligned}$$

Now the root locus has been reshaped so that it passes through the desired point $s = -1 \pm j2$.

Now that the desired pole location lies on the root locus, the gain is found from the magnitude condition

$$|L(s_d)| = 1.$$

Using

$$L(s) = \frac{K_1(s+1)}{s^2(s+3{,}6)}$$

we get

$$K_1 = \frac{|s_d|^2\,|s_d+3{,}6|}{|s_d+1|}.$$

At $s_d = -1 \pm j2$, we have

$$\begin{aligned} |s_d| &= \sqrt{(-1)^2+2^2} = \sqrt{5} = 2{,}23 \\ |s_d+3{,}6| &= \left|2{,}6 + j2\right| \approx 3{,}25 \\ |s_d+1| &= \left|j2\right| = 2. \end{aligned}$$

So

$$K_1 = \frac{(2{,}23)^2\,(3{,}25)}{2} \approx 8{,}1.$$

Therefore, the final open-loop compensated transfer function is

$$L(s) = \frac{8{,}1(s+1)}{s^2\,(s+3{,}6)}.$$

Because the open-loop transfer function still has two integrators, this is still a type-two system. The acceleration error constant is

$$K_a = \lim_{s\to 0} s^2 L(s) = \frac{8{,}1 \cdot 1}{3{,}6} = 2{,}25.$$

So the parabolic-input steady-state error would be $e_{ss} = 1/K_a \approx 0{,}44$.

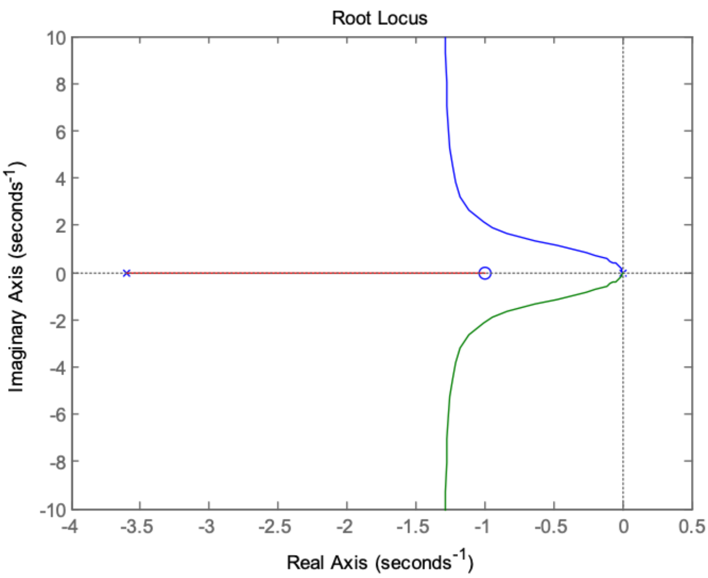



Figure 13.13: Final root-locus plot for the example system.

The final root-locus plot for the example system is shown on Figure 13.13.