2 Classical feedback control

2.1 Frequency response

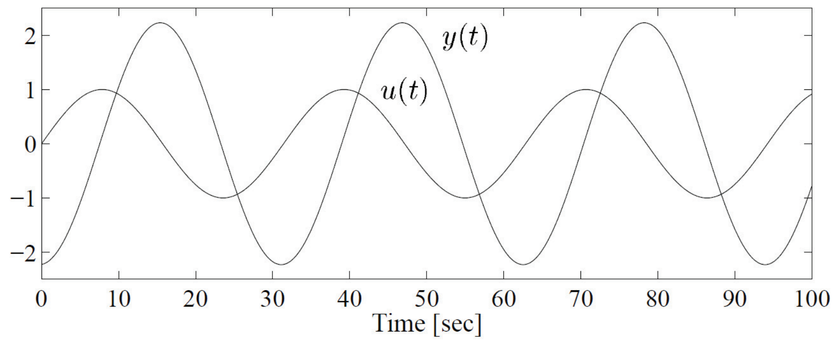


Figure 1: Sinucoidal response

The response of a stable linear system:

$$y = G(s)u$$

with sinucoidal input:

$$u(t) = u_0 \sin(\omega t + \alpha)$$

is:

$$y = y_0 \sin(\omega t + \beta)$$

where $y_0/u_0 = |G(j\omega)|$ and $\beta - \alpha = \angle G(j\omega)$.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ UNIVERSITY OF PATRAS

Bode plot

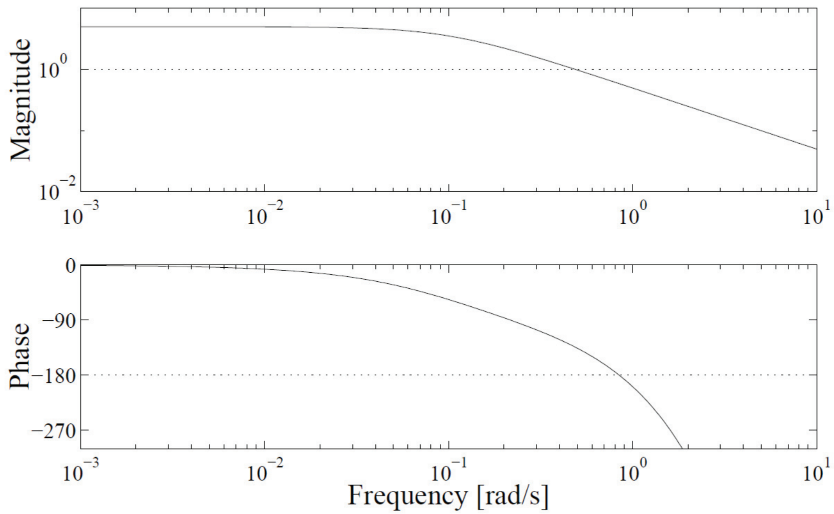


Figure 2: Frequency response

The Bode plot describes how the system responds to sinucoidal inputs of frequency ω.

- The magnitude $|G(j\omega)|$ is also referred to as system gain. Sometimes it is given in units dB (decibel) defined as $20\log_{10}|G(j\omega)|$, e.g., $|G| = 2$ corresponds to 6.02 dB, $|G| = 1$ corresponds to 0 dB.
- We employ a log-scale for frequency and gain and linear scale for the phase.

2.2 Feedback control

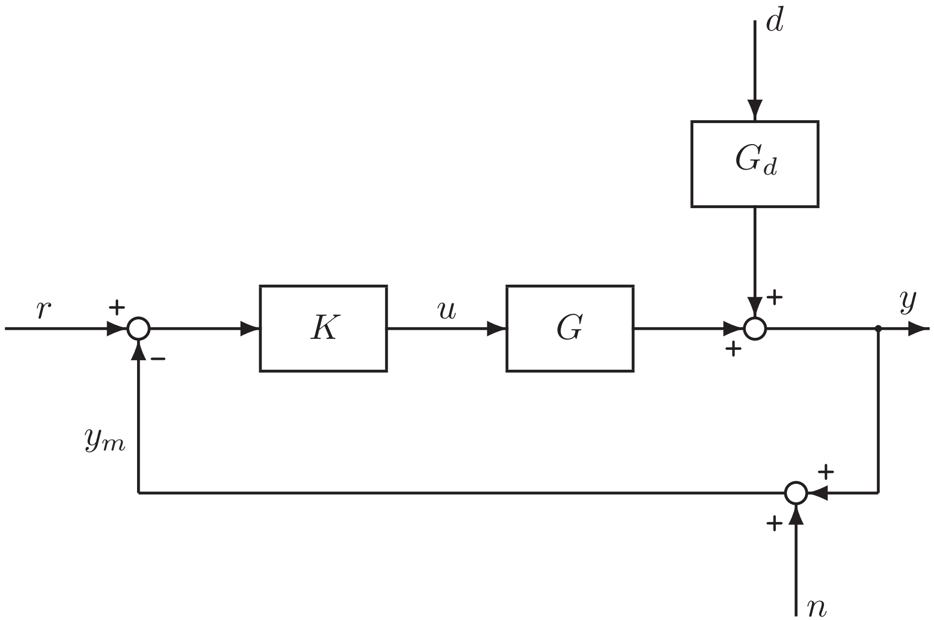


Figure 3: Block diagram of one degree-of-freedom feedback control system

Output response:

$$y = GK(r - y - n) + G_d d \tag{2.1}$$

or

$$(I + GK)y = GKr + G_d d - GKn \tag{2.2}$$

Closed-loop response:

$$y = \underbrace{(I + GK)^{-1}GK}_{T}\, r \tag{2.3}$$

$$+ \underbrace{(I + GK)^{-1}}_{S}\, G_d d \tag{2.4}$$

$$- \underbrace{(I + GK)^{-1}GK}_{T}\, n \tag{2.5}$$

Control error:

$$e = y - r = -Sr + SG_d d - Tn \tag{2.6}$$

Plant input:

$$u = KSr - KSG_d d - KSn \tag{2.7}$$

Note that:

$$L = GK \tag{2.8}$$

$$S = (I + GK)^{-1} = (I + L)^{-1} \tag{2.9}$$

$$T = (I + GK)^{-1}GK = (I + L)^{-1}L \tag{2.10}$$

$$S + T = I \tag{2.11}$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

Notation :

$$L = GK \quad \text{loop transfer function}$$

$$S = (I+L)^{-1} \quad \text{sensitivity function}$$

$$T = (I+L)^{-1}L \quad \text{complementary sensitivity function}$$

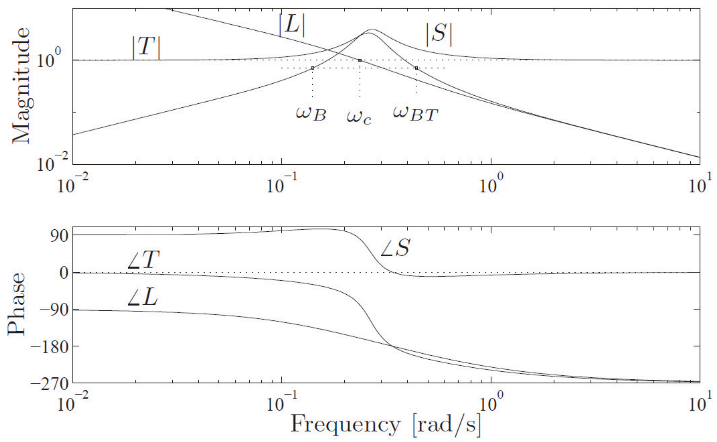


Figure 4: Bode magnitude and phase plots.

Why feedback?

Applying the inverse control:

$$u = G^{-1}(s)r - G^{-1}(s)G_d(s)d$$

fails in:

- Signal uncertainty - unknown d
- Model uncertainty
- Internal instability

High-gain control

Adopting high gain K leads in large $L = GK$. Thus, $S \to 0$ and $T \to 1$. Moreover, $KS = G^{-1}T$, which for high gain leads in $u \approx G^{-1}(r - G_d d)$.

High-gain feedback may induce **instability** and sensitivity to measurement noise:

1. Use high-gain only over a limited range of frequencies.
2. Ensure that the gain “rolls off” at high frequencies.
3. The design is critical around the bandwidth frequency.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ UNIVERSITY OF PATRAS

2.3 Closed-loop stability

Methods to determine closed-loop stability:

- The poles of the closed-loop system, i.e., the roots of $1 + L(s) = 0$. If all poles lie in the open left half plane then the closed-loop system is stable.
- Nyquist stability criterion. The number of encirclements of -1 equals to the number of open-loop unstable poles.
- Bode stability criterion (open-loop stable systems):

$$\text{Stability} \Longleftrightarrow |L(j\omega_{180})| < 1$$

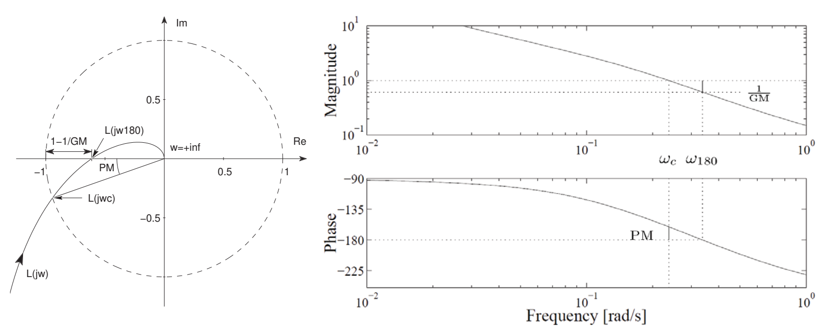


Figure 5: Typical Nyquist and Bode plots of $L(j\omega)$.

2.4 Closed-loop performance

Time domain performance

Step response analysis

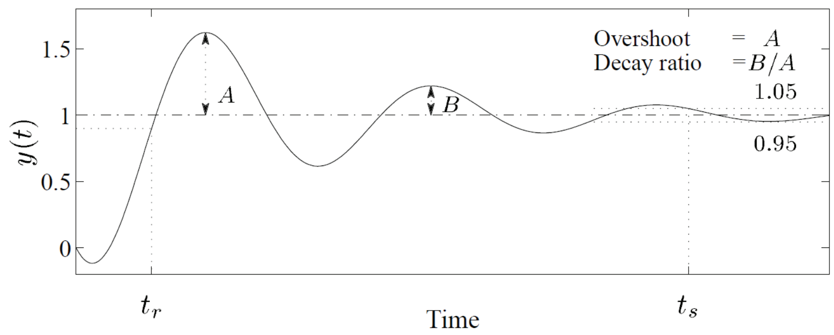


Figure 6: Closed-loop response to step in reference

- Rise time (t_r): time it takes to reach 90% of its final value.
- Settling time (t_s): the time after which the output remains 5% close to its final value.
- Overshoot: peak value, less than 20%.
- Decay ratio: ration of the first two peaks, less than 0.3
- Steady-state offset: difference of the final and reference value.
- Total variation: the total up and down movement of the signal.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

Frequency domain performance

Gain and phase margins

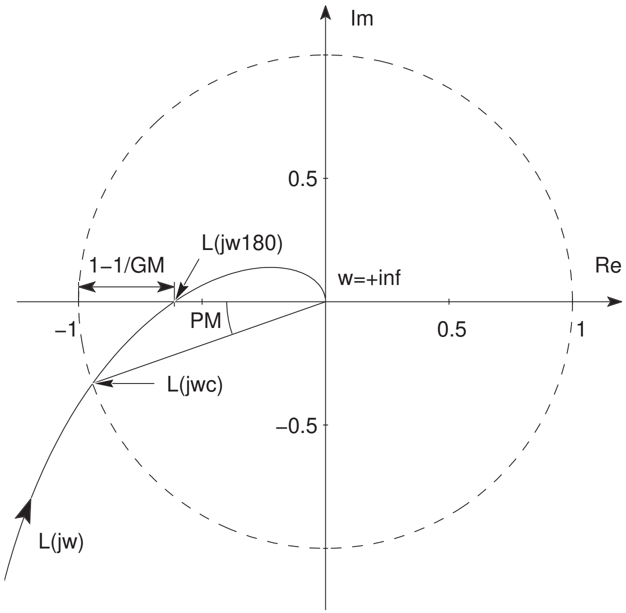


Figure 7: Typical Nyquist plot of $L(j\omega)$ for stable plant with PM and GM indicated. Closed-loop instability occurs if $L(j\omega)$ encircles the critical point -1

$$
\begin{aligned}
\text{GM} &= 1/|L(j\omega_{180})| \; > \; 2 \\
\text{PM} &= \angle L(j\omega_c) + 180^o \; > \; 30^o
\end{aligned}
$$

Maximum tolerant time delay: $\tau = \text{PM}/\omega_c$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ UNIVERSITY OF PATRAS

Maximum peak criteria

Maximum peaks of sensitivity and complementary sensitivity functions:

$$M_S \triangleq \max_\omega |S(j\omega)|; \qquad M_T \triangleq \max_\omega |T(j\omega)| \tag{2.12}$$

Typically :

$$M_S \leq 2 \quad (6dB) \tag{2.13}$$

$$M_T < 1.25 \quad (2dB) \tag{2.14}$$

Note :

$$\text{GM} \geq \frac{M_S}{M_S - 1} \tag{2.15}$$

$$\text{PM} \geq 2 \arcsin\left(\frac{1}{2M_S}\right) \geq \frac{1}{M_S} \; [\text{rad}] \tag{2.16}$$

For example, for $M_S = 2$ we are guaranteed

$$\text{GM} \geq 2 \text{ and } \text{PM} \geq 29.0^\circ.$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

Bandwidth and crossover frequency

Bandwidth is defined as the frequency range $[\omega_1, \omega_2]$ over which control is "effective". Usually $\omega_1 = 0$, and then $\omega_2 = \omega_B$ is the bandwidth.

Definition The (closed-loop) bandwidth, ω_B, is the frequency where $|S(j\omega)|$ first crosses $1/\sqrt{2} = 0.707(\approx -3 \text{ dB})$ from below.

The bandwidth in terms of T, ω_{BT}, is the highest frequency at which $|T(j\omega)|$ crosses $1/\sqrt{2} = 0.707(\approx -3 \text{ dB})$ from above. (Usually a poor indicator of performance).

The *gain crossover frequency*, ω_c, is the frequency where $|L(j\omega_c)|$ first crosses 1 from above. For systems with PM $< 90^\circ$ we have

$$\omega_B < \omega_c < \omega_{BT} \tag{2.17}$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

2.5 Controller design

Three main approaches:

1. **Shaping of transfer functions.**
   (a) **Loop shaping.** Classical approach in which the magnitude of the open-loop transfer function, $L(j\omega)$, is shaped.
   (b) **Shaping of closed-loop transfer functions, such as S, T and KS**
   $\Rightarrow \mathcal{H}_\infty$ optimal control
2. **The signal-based approach.** One considers a particular disturbance or reference change and tries to optimize the closed-loop response
   $\Rightarrow$ Linear Quadratic Gaussian (LQG) control.
3. **Numerical optimization.** Multi-objective optimization to optimize directly the true objectives, such as rise times, stability margins, etc. Computationally difficult.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

2.6 Loop shaping

Shaping of open loop transfer function $L(j\omega)$:

$$e = -\underbrace{(I+L)^{-1}}_{S} r + \underbrace{(I+L)^{-1}}_{S} G_d d - \underbrace{(I+L)^{-1} L}_{T} n \tag{2.18}$$

Fundamental trade-offs:

1. Good disturbance rejection: L large.
2. Good command following: L large.
3. Mitigation of measurement noise on plant outputs: L small.
4. Small magnitude of input signals: K small and L small.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

Fundamentals of loop-shaping design

Specifications for desired loop transfer function:

1. Gain crossover frequency, ω_c, where $|L(j\omega_c)| = 1$.
2. The shape of $L(j\omega)$, e.g. slope of $|L(j\omega)|$ in certain frequency ranges:
   $N = \frac{d\,ln\,|L|}{d\,ln\,w}$
   Typically, a slope $N = -1$ (-20 dB/decade) around crossover, and a larger roll-off at higher frequencies. The desired slope at lower frequencies depends on the nature of the disturbance or reference signal.
3. The system type, defined as the number of pure integrators in $L(s)$.

Note:

1. *for offset-free tracking, $L(s)$ should contain at least one integrator for each integrator in $r(s)$.*

2. *slope and phase are dependent. For example:*

$$\angle \frac{1}{s^n} = -n\,\frac{\pi}{2}$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

2.6.1 Inverse-based controller

Note: $L(s)$ must contain all RHP-zeros of $G(s)$.
Idea for minimum phase plant:

$$L(s) = \frac{\omega_c}{s} \tag{2.19}$$

$$K(s) = \frac{\omega_c}{s} G^{-1}(s) \tag{2.20}$$

i.e. controller inverts plant and adds integrator $(1/s)$.

BUT:

This is *not* generally desirable as:

1. RHP-zeros or time delay in $G(s)$ cannot be inverted,
2. the controller will not be implementable if $G(s)$ has a pole excess and may yield large input signals (such problems can be partly fixed by adding high frequency dynamics to the controller),
3. it works for step references and disturbances.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

Example: Disturbance process.

$$\boxed{G(s) = \frac{200}{10s+1}\frac{1}{(0.05s+1)^2}, \quad G_d(s) = \frac{100}{10s+1}} \tag{2.21}$$

Objectives are:

1. Command tracking: rise time (to reach 90% of the final value) less than 0.3 s and overshoot less than 5%.
2. Disturbance rejection: response to unit step disturbance should stay within the range $[-1, 1]$ at all times, and should return to 0 as quickly as possible ($|y(t)|$ should at least be less than 0.1 after 3 s).
3. Input constraints: $u(t)$ should remain within $[-1, 1]$ at all times.

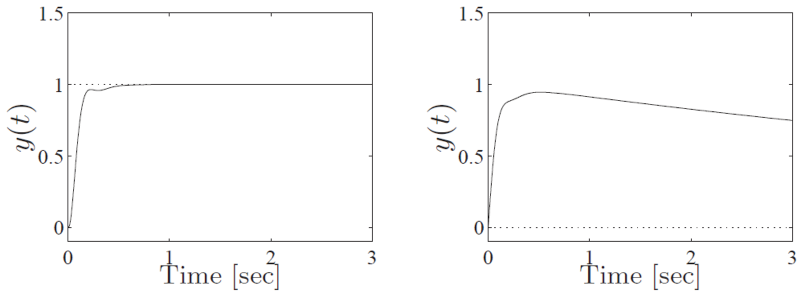
Analysis. $|G_d(j\omega)|$ remains larger than 1 up to $\omega_d \approx 10$ rad/s $\Rightarrow$ $\underline{\omega_c \approx 10 \text{ rad/s}}$.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ UNIVERSITY OF PATRAS

Inverse-based controller design.

$$
\begin{aligned}
K_0(s) &= \frac{\omega_c}{s}\frac{10s+1}{200}(0.05s+1)^2 \\
&\approx \frac{\omega_c}{s}\frac{10s+1}{200}\frac{0.1s+1}{0.01s+1},
\end{aligned}
$$

$$
L_0(s) = \frac{\omega_c}{s}\frac{0.1s+1}{(0.05s+1)^2(0.01s+1)}, \;\; \omega_c = 10 \quad (2.22)
$$



(a) Tracking response (b) Disturbance response

Figure 8: Responses with "inverse-based" controller $K_0(s)$ for the disturbance process. Note poor disturbance responce

2.6.2 Loop shaping for disturbance rejection

$$e = y = SG_d d, \tag{2.23}$$

to achieve $|e(\omega)| \leq 1$ for $|d(\omega)| = 1$ (the worst-case disturbance) we require $|SG_d(j\omega)| < 1, \forall \omega$, or

$$|1 + L| \geq |G_d| \quad \forall \omega \tag{2.24}$$

or approximately:

$$|L| \geq |G_d| \quad \forall \omega \tag{2.25}$$

Initial guess:

$$|L_{\min}| \approx |G_d| \tag{2.26}$$

or:

$$|K_{\min}| \approx |G^{-1}G_d| \tag{2.27}$$

Controller contains the model of the disturbance.

To improve low-frequency performance

$$|K| = |\frac{s + \omega_I}{s}||G^{-1}G_d| \tag{2.28}$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

Summary:

- Controller contains the dynamics (G_d) of the disturbance and inverts the dynamics (G) of the inputs.
- For disturbance at plant output, $G_d = 1$, we get $|K_{\text{min}}| = |G^{-1}|$.
- For disturbances at plant input we have $G_d = G$ and we get $|K_{\text{min}}| = 1$.

Loop-shape $L(s)$ may be modified as follows:

1. Around crossover make slope N of $|L|$ to be about -1 for transient behaviour with acceptable gain and phase margins.
2. Increase the loop gain at low frequencies to improve the settling time and reduce the steady-state offset $\rightarrow$ add an integrator
3. Let $L(s)$ roll off faster at higher frequencies (beyond the bandwidth) in order to reduce the use of manipulated inputs, to make the controller realizable and to reduce the effects of noise.

Example: Loop-shaping design for the disturbance process.

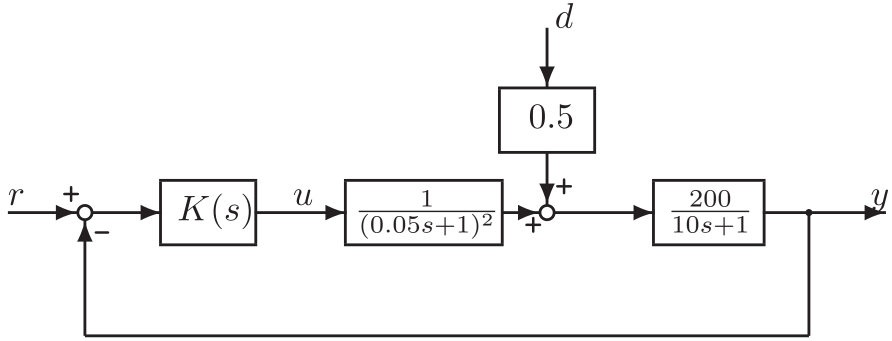
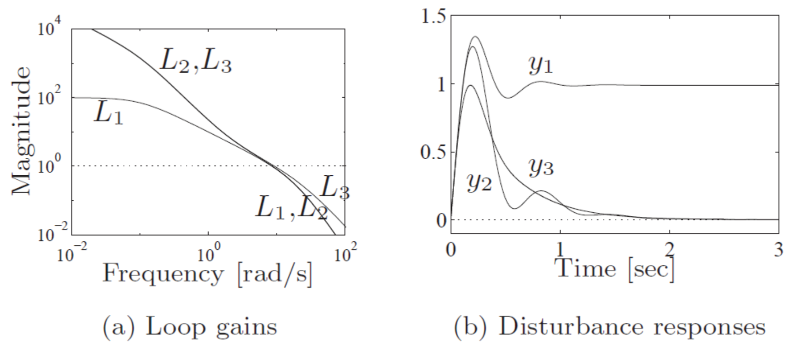


Figure 9: Block diagram representation of the disturbance process in (2.21)



(a) Loop gains (b) Disturbance responses

Figure 10: Loop shapes and disturbance responses for controllers K_1, K_2 and K_3 for the disturbance process

Step 1. Initial design.
$K(s) = G^{-1}G_d = 0.5(0.05s+1)^2$.

Make proper:

$$K_1(s) = 0.5 \tag{2.29}$$

$\Longrightarrow$ offset!

Step 2. More gain at low frequency. To get integral action multiply the controller by the term $\frac{s+\omega_I}{s}$. For $\omega_I = 0.2\omega_c$ the phase contribution from $\frac{s+\omega_I}{s}$ is $\arctan(1/0.2) - 90° = -11°$ at ω_c. For $\omega_c \approx 10$ rad/s, select the following controller

$$K_2(s) = 0.5\frac{s+2}{s} \tag{2.30}$$

$\Longrightarrow$ response exceeds 1, oscillatory, small phase margin

Step 3. High-frequency correction. Supplement with "derivative action" by multiplying $K_2(s)$ by a lead-lag term effective over one decade starting at 20 rad/s:

$$K_3(s) = 0.5\frac{s+2}{s}\,\frac{0.05s+1}{0.005s+1} \tag{2.31}$$

$\Longrightarrow$ poor reference tracking (simulation)

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

2.6.3 Two degrees of freedom design

In order to meet both **tracking** $(K \approx \frac{\omega_c}{s} G^{-1})$ and **regulator** $(K \approx \frac{s+\omega_I}{s} G^{-1} G_d)$ performance, use K_r (= "prefilter"):

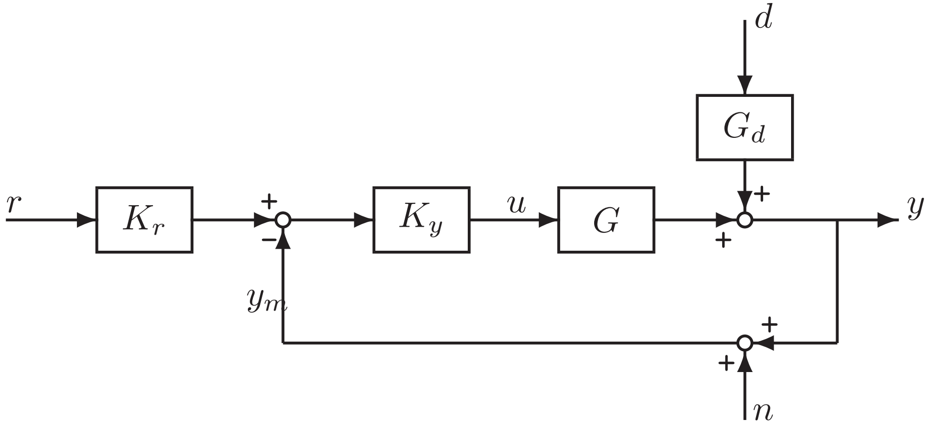


Figure 11: Two degrees-of-freedom controller

Idea:

- Design K_y
- $T = L(I+L)^{-1}$ with $L = GK_y$
- Desired $y = T_{ref} r$

$$\Longrightarrow K_r = T^{-1} T_{ref} \tag{2.32}$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

Remark:

Practical choice of prefilter is the lead-lag network

$$K_r(s) = \frac{\tau_{\text{lead}} s + 1}{\tau_{\text{lag}} s + 1} \tag{2.33}$$

$\tau_{\text{lead}} > \tau_{\text{lag}}$ to speed up the response, and $\tau_{\text{lead}} < \tau_{\text{lag}}$ to slow down the response.

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

Example Two degrees-of-freedom design for the disturbance process.

$K_y = K_3$. Approximate response by inpection of y_3:

$T(s) \approx \frac{1.5}{0.1s+1} - \frac{0.5}{0.5s+1} = \frac{(0.7s+1)}{(0.1s+1)(0.5s+1)}$ $\quad T_{ref}(s) = \frac{1}{0.1s+1}$

which yields:

$K_r(s) = \frac{0.5s+1}{0.7s+1}$.

By closed-loop simulations:

$$K_{r3}(s) = \frac{0.5s+1}{0.65s+1} \cdot \frac{1}{0.03s+1} \tag{2.34}$$

where $1/(0.03s+1)$ included to avoid initial peaking of input signal $u(t)$ above 1.

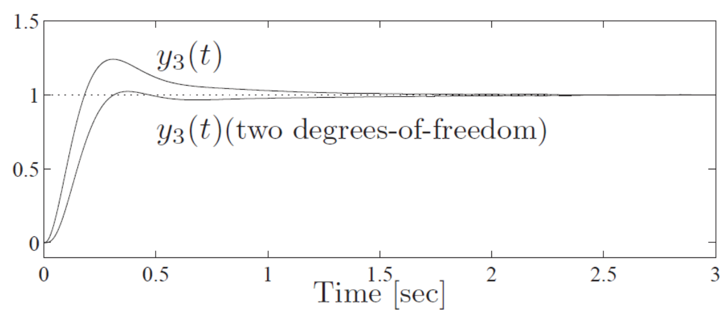


Figure 12: Tracking responses with the one degree-of-freedom controller (K_3) and the two degrees-of-freedom controller (K_3, K_{r3}) for the disturbance process

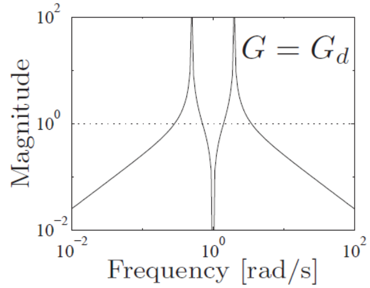
ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

Example: Loop shaping for a flexible structure.

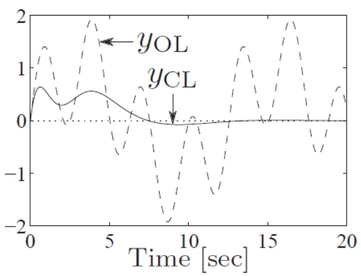
$$G(s) = G_d(s) = \frac{2.5s(s^2+1)}{(s^2+0.5^2)(s^2+2^2)} \qquad (2.35)$$

$$|K_{\min}(j\omega)| = |G^{-1}G_d| = 1 \Rightarrow$$

$$K(s) = 1 \qquad (2.36)$$



(a) Magnitude plot of $|G| = |G_d|$



(b) Open-loop and closed-loop disturbance responses with $K = 1$

Figure 13: Flexible structure in (2.35)

2.7 Closed-loop shaping

Why ?

We are interested in S and T:

$$
\begin{array}{lll}
|L(j\omega)| \gg 1 & \Rightarrow & S \approx L^{-1}; \quad T \approx 1 \\
|L(j\omega)| \ll 1 & \Rightarrow & S \approx 1; \quad\quad\; T \approx L
\end{array}
$$

but in the crossover region where $|L(j\omega)|$ is close to 1, one cannot infer anything about S and T from $|L(j\omega)|$.

Alternative:

Directly shape the magnitudes of closed-loop $S(s)$ and $T(s)$.

The term $\mathcal{H}_\infty$

The $\mathcal{H}_\infty$ norm of a stable scalar transfer function $f(s)$ is simply the peak value of $|f(j\omega)|$ as a function of frequency, that is,

$$\|f(s)\|_\infty \triangleq \max_\omega |f(j\omega)| \tag{2.37}$$

The symbol ∞ comes from:

$$\max_\omega |f(j\omega)| = \lim_{p\to\infty} \left(\int_{-\infty}^{\infty} |f(j\omega)|^p d\omega \right)^{1/p}$$

The symbol $\mathcal{H}$ stands for "Hardy space", and $\mathcal{H}_\infty$ is the set of transfer functions with bounded ∞-norm, which is simply the set of *stable and proper* transfer functions.

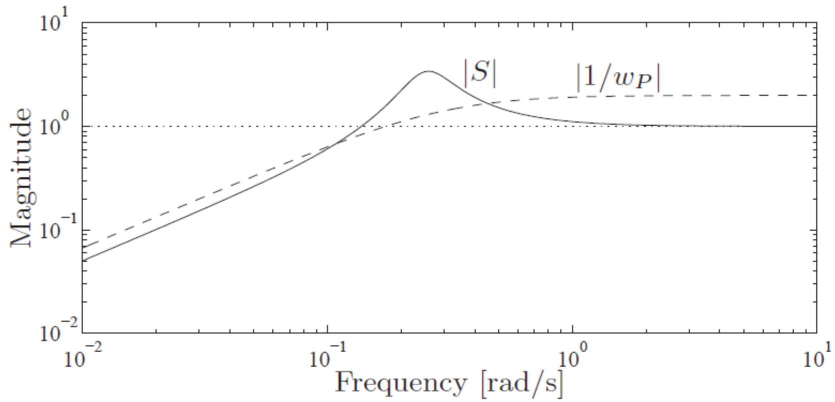
ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

2.7.1 Weighted sensitivity

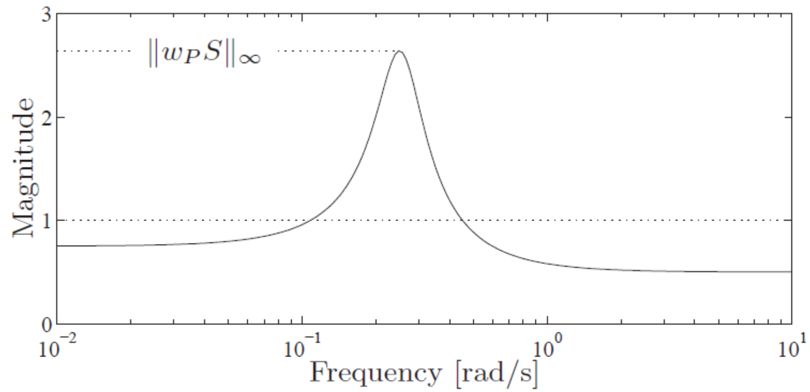
Typical specifications in terms of S:

1. Minimum bandwidth frequency ω_B^*.
2. Maximum tracking error at selected frequencies.
3. System type, or alternatively the maximum steady-state tracking error, A.
4. Shape of S over selected frequency ranges.
5. Maximum peak magnitude of S, $\|S(j\omega)\|_\infty \leq M$.

Specifications may be captured by an upper bound, $1/|w_P(s)|$, on $\|S\|$.



(a) Sensitivity S and performance weight w_P



(b) Weighted sensitivity $w_P S$

Figure 14: Case where $|S|$ exceeds its bound $1/|w_P|$, resulting in $\|w_P S\|_\infty > 1$

$$|S(j\omega)| < 1/|w_P(j\omega)|, \ \forall\omega \tag{2.38}$$

$$\Leftrightarrow \quad |w_P S| < 1, \ \forall\omega \quad \Leftrightarrow \quad \boxed{\|w_P S\|_\infty < 1} \tag{2.39}$$

Typical performance weight:

$$\boxed{w_P(s) = \frac{s/M + \omega_B^*}{s + \omega_B^* A}} \tag{2.40}$$

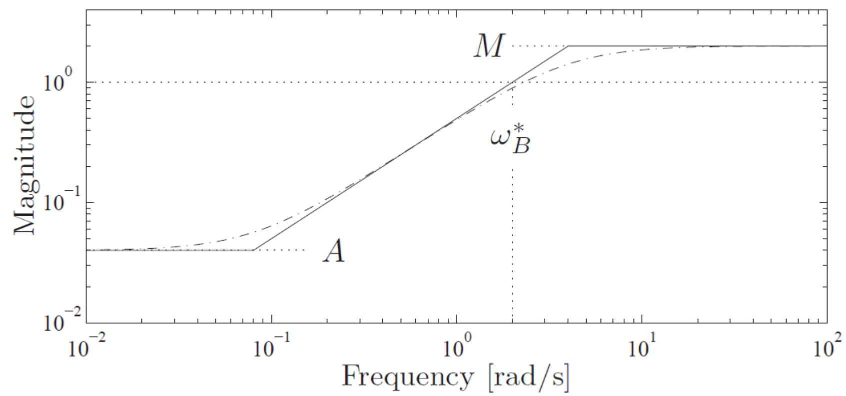


Figure 15: Inverse of performance weight. Exact and asymptotic plot of $1/|w_P(j\omega)|$ in (2.40)

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

To get a steeper slope for L (and S) below the bandwidth:

$$\boxed{w_P(s) = \frac{(s/M^{1/2} + \omega_B^*)^2}{(s + \omega_B^* A^{1/2})^2}} \tag{2.41}$$

2.7.2 Stacked requirements: mixed sensitivity

In order to enforce specifications on other transfer functions:

$$\|N\|_\infty = \max_\omega \bar{\sigma}(N(j\omega)) < 1; \quad N = \begin{bmatrix} w_P S \\ w_T T \\ w_u KS \end{bmatrix} \tag{2.42}$$

N is a vector and the maximun singular value $\bar{\sigma}(N)$ is the usual Euclidean vector norm:

$$\bar{\sigma}(N) = \sqrt{|w_P S|^2 + |w_T T|^2 + |w_u KS|^2} \tag{2.43}$$

The $\mathcal{H}_\infty$ optimal controller is obtained from

$$\min_K \|N(K)\|_\infty \tag{2.44}$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

Example: $\mathcal{H}_\infty$ mixed sensitivity design for the disturbance process.

Consider the plant in (2.21), and an $\mathcal{H}_\infty$ mixed sensitivity S/KS design in which

$$N = \begin{bmatrix} w_P S \\ w_u KS \end{bmatrix} \tag{2.45}$$

Selected $w_u = 1$ and

$$w_{P1}(s) = \frac{s/M + \omega_B^*}{s + \omega_B^* A}; \quad M = 1.5,\ \omega_B^* = 10, \quad A = 10^{-4} \tag{2.46}$$

$\Longrightarrow$ poor disturbance response

To get higher gains at low frequencies:

$$w_{P2}(s) = \frac{(s/M^{1/2} + \omega_B^*)^2}{(s + \omega_B^* A^{1/2})^2}, \quad M = 1.5, \omega_B^* = 10, A = 10^{-4} \tag{2.47}$$

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ
UNIVERSITY OF PATRAS

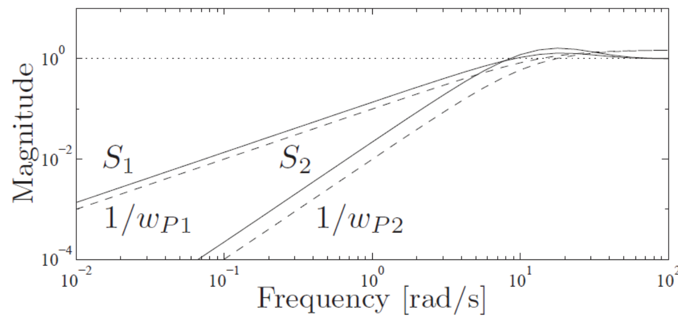
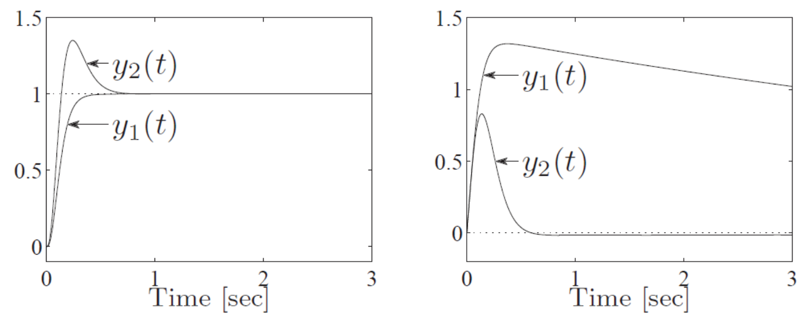


Figure 16: Inverse of performance weight (dashed line) and resulting sensitivity function (solid line) for two $\mathcal{H}_\infty$ designs (1 and 2) for the disturbance process



(a) Tracking response

(b) Disturbance response

Figure 17: Closed-loop step responses for two alternative $\mathcal{H}_\infty$ designs (1 and 2) for the disturbance process

ΠΑΝΕΠΙΣΤΗΜΙΟ ΠΑΤΡΩΝ UNIVERSITY OF PATRAS