Automatic Control 2

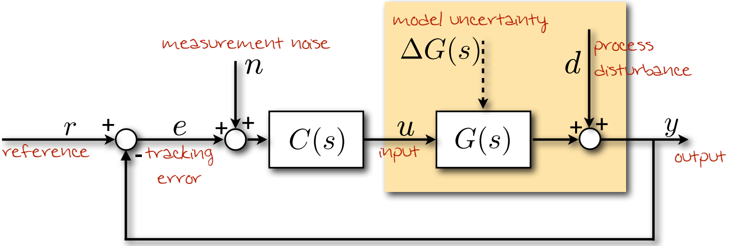
Loop shaping

Prof. Alberto Bemporad

University of Trento

Academic year 2010-2011

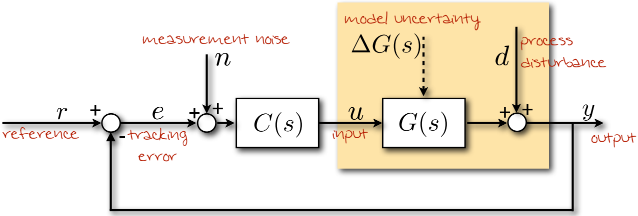
Feedback control problem



Objective: make the tracking error $e(t) = r(t) - y(t) \simeq 0$, despite

- the dynamics of the open-loop system $G(s)$ (slow, unstable, etc.)
- disturbance $d(t)$ affecting the process
- model uncertainty $\Delta G(s)$ (the system is not exactly as we modeled it)
- measurement noise $n(t)$

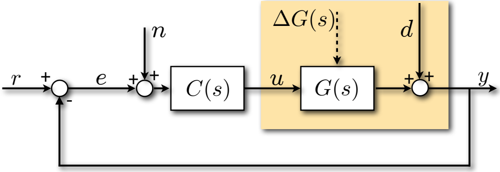
Feedback control problem



To achieve the objective, we want to design a feedback control law satisfying a number of requirements:

- stability in nominal conditions ($\Delta G(s) = 0$, $d(t) = 0$, $n(t) = 0$)
- stability in perturbed conditions
- static performances (tracking error for constant $r(t)$)
- dynamic performances (transients and frequency response)
- noise attenuation (be insensitive to measurement noise $n(t)$)
- feasibility of the controller (strictly proper transfer function $C(s)$)

Closed-loop function



- The *closed-loop function* is the transfer function from r to y

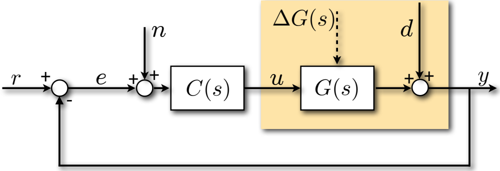
$$W(s) = H_{yr}(s) = \frac{C(s)G(s)}{1 + C(s)G(s)}$$

- We call *loop function* the transfer function

$$L(s) = C(s)G(s)$$

- For $|L(j\omega)| \gg 1$ we get $|W(j\omega)| \simeq 1$, $\angle W(j\omega) \simeq 0$
- For $|L(j\omega)| \ll 1$ we get $|W(j\omega)| \simeq 0$
- Given $G(s)$, we will choose $C(s)$ to *shape* the *loop* function response $L(j\omega)$

Sensitivity functions



- The *sensitivity function* $S(s)$ is the function

$$S(s) = \frac{\partial W(s)/W(s)}{\partial G(s)/G(s)} = \frac{\partial W(s)}{\partial G(s)} \frac{G(s)}{W(s)} = \frac{1}{1+L(s)}$$

- The sensitivity function $S(s)$ is also the transfer function from d to y

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

- The *complementary sensitivity function* $T(s)$ is the transfer function from n to y

$$T(s) = \frac{Y(s)}{N(s)} = \frac{L(s)}{1+L(s)} = W(s)$$

Performance specifications: frequency domain

sensitivity

transfer function from d to y

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

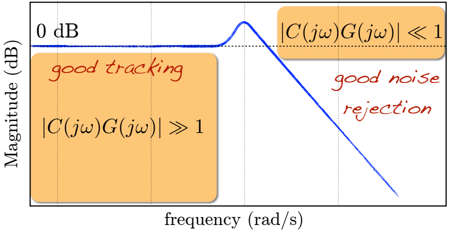
complementary sensitivity

transfer function from n to y

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

- We want to be immune (=small tracking errors) to both process disturbances and measurement noise → both S and T small
- **Problem**: $S(s) + T(s) = 1$! How to make both them small ?

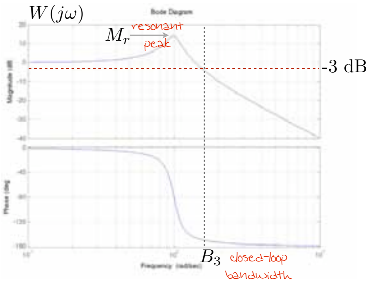
- **Solution**:
    - keep S small at low frequencies, and hence $W \simeq 1$ (=good tracking)
    - keep T small at high frequencies (=good noise rejection)



Constraint on performance: tracking cannot be too good at high frequencies !

Performance specifications: frequency domain

- We usually refer to “regular” closed-loop responses (i.e., similar to 2nd-order underdamped systems)



$$W(s) = \frac{1}{1 + \frac{2\zeta}{\omega_n}s + \frac{1}{\omega_n^2}s^2}$$

- the *resonant peak* M_r is the max value of $|W(j\omega)|$
- the *bandwidth* of the closed loop system is the range of frequencies at which $|W(j\omega)| \geq \sqrt{2}$ ($\frac{|W(j\omega)|}{|W(0)|} \geq \sqrt{2}$, in general)
- note that $\left|\frac{1}{\sqrt{2}}\right|_{dB} = 20\log_{10} 2^{-\frac{1}{2}} \simeq -3dB$

- For 2nd order systems $W(s)$: $M_r^W = \frac{1}{2\zeta\sqrt{1-\zeta^2}}$, $B_3^W = \omega_n\sqrt{1 - 2\zeta^2 + \sqrt{2 - 4\zeta^2 + 4\zeta^4}}$

Performance specifications: steady-state tracking

- We look at tracking of constant set-points, ramps, etc., $R(s) = \dfrac{1}{s^k}$
- Write the loop function in Bode form

$$L(s) = \frac{K}{s^h} L_1(s), \text{ with } L_1(0) = 1$$

- Compute $\lim_{t \to +\infty} e(t)$ using final value theorem for $h \geq k-1$:

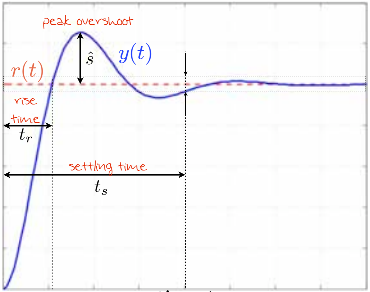
$$\begin{aligned}
\lim_{t \to +\infty} e(t) &= \lim_{s \to 0} sE(s) = \lim_{s \to 0} s(R(s) - W(s)R(s)) = \lim_{s \to 0} \frac{1}{1+L(s)} \frac{1}{s^{k-1}} \\
&= \lim_{s \to 0} \frac{s^{h-k+1}}{s^h + KL_1(s)} = \begin{cases} 0 & \text{if} \quad h \geq k \\ \dfrac{1}{1+K} & \text{if} \quad h = 0, k = 1 \\ \dfrac{1}{K} & \text{if} \quad h = k-1, k > 1 \end{cases}
\end{aligned}$$

Performance specifications: steady-state tracking

$$\lim_{t\to+\infty} e(t) = \begin{cases} 0 & \text{if} \quad h \geq k \\ \dfrac{1}{1+K} & \text{if} \quad h = 0, k = 1 \\ \dfrac{1}{K} & \text{if} \quad h = k-1, k > 1 \end{cases}$$

- Constraint on *type* h of $L(s) = C(s)G(s) = \dfrac{KL_1(s)}{s^h}$:
  - need $h \geq k-1$ to track $r(t) = t^{k-1}$
  - need $h \geq k$ to track t^{k-1} with zero asymptotic error
  - special case: need $h \geq 1$ to track constant set-points with zero offset in steady-state $\longleftarrow$ (integral action!)
- Constraint on *Bode gain* K of $L(s)$: need K sufficiently large to bound steady-state tracking error when $h = k-1$

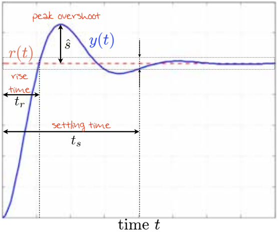
Performance specifications: time response



- We look at the shape of the transient closed-loop response due to a unit step $r(t) = 1$ for $t \geq 0$, $r(t) = 0$ for $t < 0$
- the *rise time* t_r is the time required to rise from 0 to 100% of steady-state ($10 \div 90\%$ for non-oscillating systems)
- the *settling time* t_s is the time to reach and stay within a specified tolerance band (usually 2% or 5%)
- the *peak overshoot* $\hat{s}$ is the max relative deviation from steady-state, $\hat{s} = \frac{\max_t y(t)}{y(+\infty)} - 1$

Relations between frequency and time response

- Usually we refer to the step response of a second-order closed-loop system



$$W(s) = \frac{1}{1 + \frac{2\zeta}{\omega_n}s + \frac{1}{\omega_n^2}s^2}$$

- In this case we have

$$\hat{s} = e^{-\frac{\pi\zeta}{\sqrt{1-\zeta^2}}}, \quad t_r = \frac{1}{\omega_n}\frac{1}{\sqrt{1-\zeta^2}}\left[\pi - \tan^{-1}\frac{\sqrt{1-\zeta^2}}{\zeta}\right], \quad t_{s[5\%]} \simeq \frac{3}{\zeta\omega_n} \qquad \boxed{t_{s[2\%]} \simeq \frac{4}{\zeta\omega_n}}$$

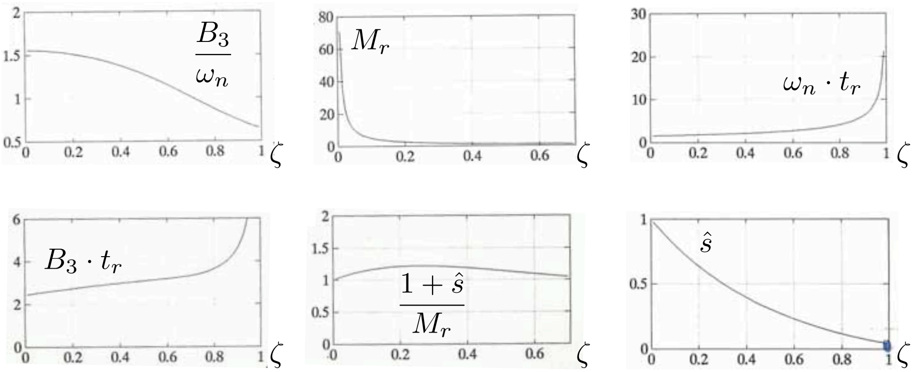
- Good average formulas for closed-loop systems $W(s)$ are

$$\boxed{t_r B_3^W \simeq 3} \qquad \boxed{\hat{s} \simeq 0.85 \cdot M_r^W - 1}$$

where t_r [s], B_3^W [rad/s], $\hat{s}$ [no unit], M_r^W [not in dB]

Relations between frequency and time response

- Relations between frequency and time response for second-order closed-loop systems as a function of the damping factor ζ



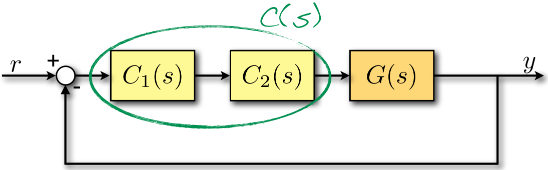
Controller synthesis via loop shaping

- Typical closed-loop specifications include static specs
  - system type h
  - tracking error $e(+\infty)$ in steady-state for $r(t) = t^k$ $\longleftarrow$ Bode gain K

and dynamic specs

  - peak overshoot $\hat{s}$ $\longleftarrow$ resonant peak M_r
  - rise time t_r $\longleftarrow$ Bandwidth B_3

- Designing a regulator that meets all specs in one shot can be a hard task
- In *loop-shaping* synthesis the controller $C(s)$ is designed in a series of steps



where $C_1(s)$ satisfies static specs, while $C_2(s)$ dynamic specs

Synthesis of $C_1(s)$ (static performance)

- The general form of $C_1(s)$ is
$$C_1(s) = \frac{K_c}{s^{h_c}}$$
where K_c is the controller gain and h_c is the type of the controller
- If the type of $G(s)$ is h_g and the desired type is h_d, then $h_c = max\{0, h_d - h_g\}$
- The gain K_c is chosen by imposing the desired steady-state tracking error

$$\frac{1}{1+K_cG(0)} \leq e_d \quad \text{if } r_d = 0$$
$$\frac{1}{K_cG(0)} \leq e_d \quad \text{if } r_d > 0$$

Example: $G(s) = \dfrac{1}{(s+1)}$

- Track a step reference with zero steady-state error $\longrightarrow r_d = 1$
- since the type of $G(s)$ is $h_g = 0$, $h_c = max\{0, 1\} = 1$, and $C_1(s) = \frac{1}{s}$
- steady-state error e_d to unit ramp reference is not required, so K_c is free

Mapping closed-loop specs to open-loop specs

- Closed-loop specifications must be translated into specifications on the loop transfer function $L(j\omega) = C(j\omega)G(j\omega)$
- In particular closed-loop specs should be translated into a desired phase margin M_p^L and desired crossover frequency ω_c^L of $L(j\omega)$
- We have some approximate formulas to do that:

$$\boxed{M_p^L \simeq \frac{2.3 - M_r^W}{1.25}}$$

where M_r is in not expressed in dB, and M_p is expressed in rad, and

$$\boxed{\omega_c^L \simeq [0.5, 0.8] B_3^W}$$

- The next step of loop shaping is to synthesize $C_2(s)$ so that $L(j\omega)$ has the right phase margin $M_p^L \geq M_{pd}$ and crossover frequency $\omega_c^L = \omega_{cd}$

Synthesis of $C_2(s)$ (dynamic performance)

- The general form of $C_2(s)$ is

$$C_2(s) = \frac{\prod_i (1+\tau_i s) \prod_i \left(1 + \frac{2\zeta_i' s}{\omega_{ni}'} + \frac{s^2}{\omega_{ni}'^{\,2}}\right)}{\prod_j (1+T_j s) \prod_j \left(1 + \frac{2\zeta_j s}{\omega_{nj}} + \frac{s^2}{\omega_{nj}^2}\right)}$$

- $C_2(s)$ must be designed to guarantee closed-loop asymptotic stability
- $C_2(s)$ must be designed to satisfy the dynamic specifications
- We focus here on controllers containing *only real* poles/zeros

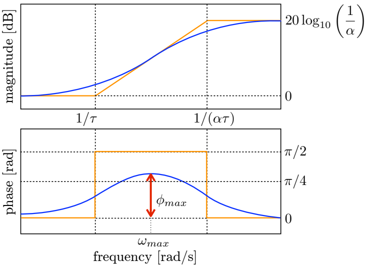
Lead network

- A *lead network* has transfer function

$$C_{Lead}(s) = \frac{1+\tau s}{1+\alpha\tau s}$$

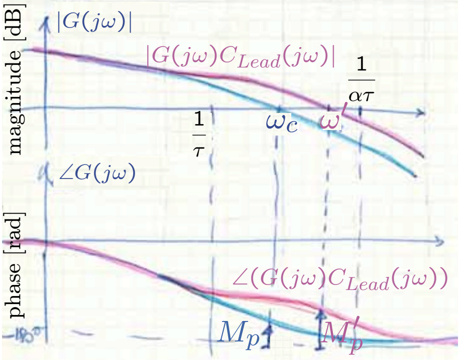
where $\tau > 0$, $0 < \alpha < 1$

- $C_{Lead}(s)$ provides phase lead in the frequency range $[\frac{1}{\tau}, \frac{\alpha}{\tau}]$



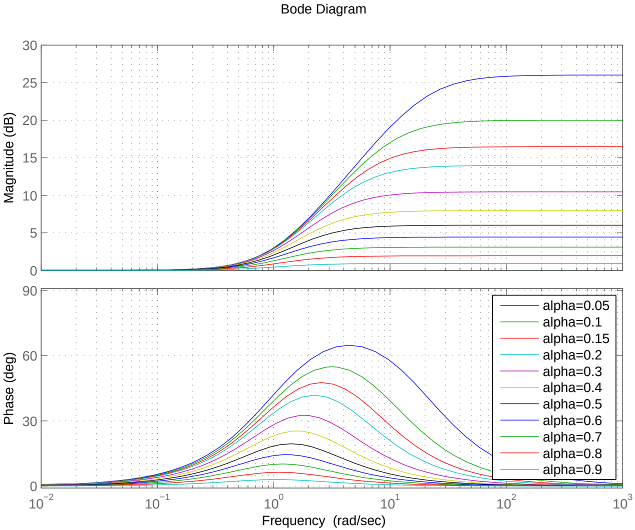
- The maximum phase lead is at $\omega_{max} = \frac{1}{\tau\sqrt{\alpha}}$ (midpoint between $\frac{1}{\tau}$ and $\frac{1}{\alpha\tau}$ in logarithmic scale)

Lead network



- The main goal of the lead network is to increase the phase margin
- As a side effect, the loop gain is increased at high frequencies (=reduced complementary sensitivity)

Lead network



$$C_{Lead}(s) = \frac{1+s}{1+\alpha s}$$

$$\tau = 1$$

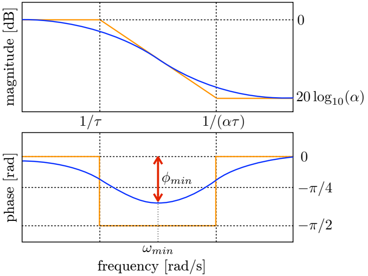
Lag network

- A *lag network* has transfer function

$$C_{Lag}(s) = \frac{1+\alpha\tau s}{1+\tau s}$$

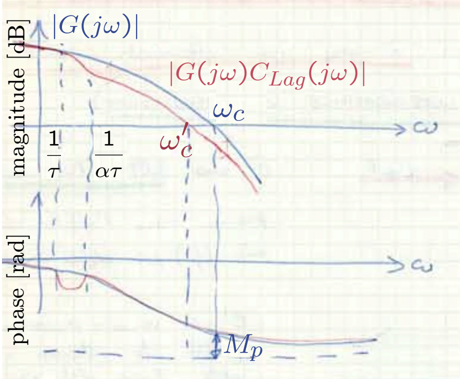
where $\tau > 0$, $0 < \alpha < 1$

- $C_{Lag}(s)$ provides phase lag in the frequency range $[\frac{1}{\tau}, \frac{\alpha}{\tau}]$



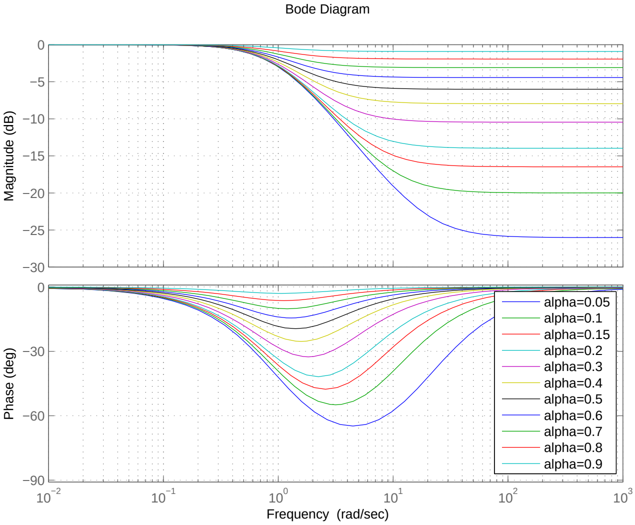
- The maximum phase lag is at $\omega_{min} = \frac{1}{\tau\sqrt{\alpha}}$ (midpoint between $\frac{1}{\tau}$ and $\frac{1}{\alpha\tau}$ in logarithmic scale)

Lag network



- The main goal of the lag network is to decrease the loop gain at high frequencies (=reduce complementary sensitivity)
- To avoid decreasing the phase margin, set ω_{min} at low frequencies, where the loop gain $|L(j\omega)|$ is still high, therefore avoiding $\omega_c \simeq \omega_{min}$

Lag network



$$C_{Lag}(s) = \frac{1+\alpha s}{1+s}$$

$$\tau = 1$$

Note that

$$\left|\frac{1+\alpha\tau j\omega}{1+\tau j\omega}\right|_{dB} = -\left|\frac{1+\tau j\omega}{1+\alpha\tau j\omega}\right|_{dB}$$

and that

$$\angle\frac{1+\alpha\tau j\omega}{1+\tau j\omega} = -\angle\frac{1+\tau j\omega}{1+\alpha\tau j\omega}$$

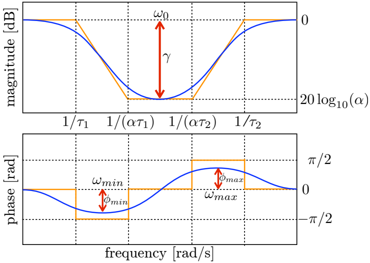
Lead-lag network

- A lead-lag network has transfer function

$$C_{Lead-Lag}(s) = \frac{(1+\alpha_1\tau_1 s)}{(1+\tau_1 s)}\frac{(1+\tau_2 s)}{(1+\alpha_2\tau_2 s)}$$

  where $\tau_1, \tau_2 > 0$, $0 < \alpha_1, \alpha_2 < 1$

- The lead-lag network $C_{Lead-Lag}(s)$ provides the coupled effect of a lead and a lag network: increase the phase margin without increasing the closed-loop bandwidth



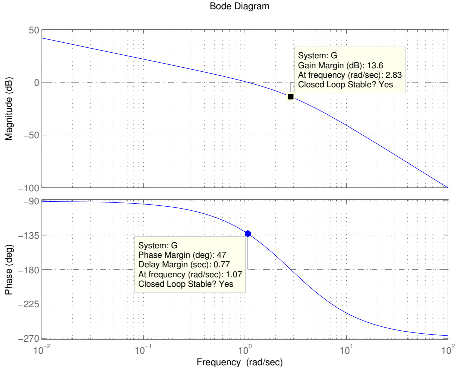
A special form of lead-lag networks most used in industrial practice are *proportional integral derivative (PID)* controllers (see more later ...)

PID temperature controller
`http://www.auberins.com`

Example of loop shaping

Open-loop transfer function of the process

$$G(s) = \frac{10}{s(s+2)(s+4)} = \frac{1.25}{s}\,\frac{1}{(1+0.5s)(1+0.25s)}$$



Closed-loop specifications

1. Track a ramp reference $r(t) = t$ with finite steady-state error $e_d \leq 0.2$
2. Rise-time of unit step response $t_r \simeq 0.4$ s
3. Overshoot of unit step reference $\hat{s} \leq 25\%$

Example of loop shaping – Static performance

1. Track a ramp reference $r(t) = t$ with finite steady-state error $e_d \leq 0.2$

- Since $G(s)$ is of type $r_g = 1$, no need to add integrators

$$C_1(s) = \frac{K_c}{s^{h_c}} = \frac{K_c}{s^0} = K_c$$

- Choose K_c by looking at steady-state tracking error of unit ramp

$$\frac{1}{K_c \cdot 1.25} \leq e_d = 0.2 \ \rightarrow \ K_c \geq 4$$

- Finally, set

$$C_1(s) = 4$$

Example of loop shaping – Dynamic performance

2. Rise-time of unit step response $t_r \simeq 0.4$

- Since $B_3 \simeq 3/t_r$, we get a closed-loop bandwidth constraint

$$B_3 \simeq \frac{3}{t_r} = \frac{3}{0.4} = 7.5 \text{ rad/s}$$

- As the desired $\omega_c \simeq [0.5, 0.8]B_3$, we get a target for the crossover frequency of the loop function $L(j\omega)$

$$\boxed{\omega_c \simeq 4.7 \text{ rad/s}}$$

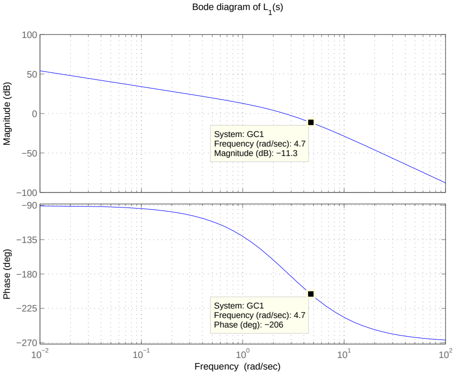
2. Overshoot of unit step reference $\hat{s} \leq 25\%$

- As $\hat{s} \simeq 0.85 \cdot M_r - 1$ and $M_p \simeq \dfrac{2.3 - M_r}{1.25}$ we get $0.25 \simeq 0.85 \cdot M_r - 1$ or $M_r = 1.47$. The resulting specification on the phase margin of $L(j\omega)$ is

$$\boxed{M_p \simeq 0.664 \text{ rad} \simeq 38 \text{ deg}}$$

Example of loop shaping – Dynamic performance

- Let's examine the current loop gain $L_1(s) = C_1(s)G(s) = \dfrac{40}{s(s+2)(s+4)}$ for $s = j\omega_c = j4.7$

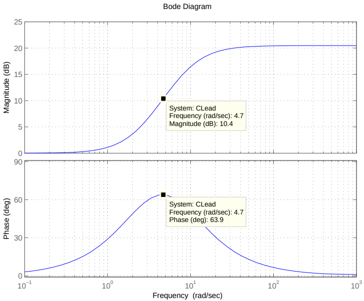


- $|L_1(j4.7)| = -11.4$ dB
- $\angle L_1(j4.7) \simeq -206$ deg
- at $\omega = 4.7$ rad/s we need to increase the gain by $\Delta M = 0 + 11.4 = 11.4$ dB
- and the phase by $\Delta\phi = 206 - (180 - 38) = 64$ deg

Example of loop shaping – Lead network

- A suitable network is a lead network. As we need to gain $\Delta\phi = 64$ deg, we choose a cascade of two identical lead networks $C_{Lead}(s) = \dfrac{1+\tau s}{1+\alpha\tau s}$
- to gain 32 deg at $\omega_c = 4.7$ rad/s, we set $\alpha = 0.3$ and $4.7 = \dfrac{1}{\tau\sqrt{\alpha}}$, or $\tau = 0.39$ s

$$C_{2Lead}(s) = \left(\frac{1+0.39s}{1+0.12s}\right)^2$$

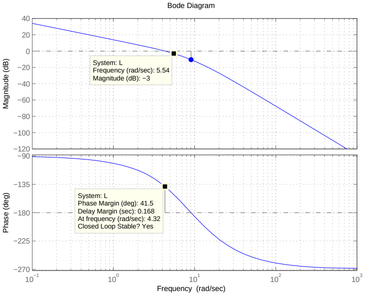


- Note that at the desired crossover frequency $\omega_c = 4.7$ rad/s, $|C_{2Lead}(j4.7)| = 10.4$ dB

Example of loop shaping – Resulting controller

- The gain difference $10.4 - 11.4 = 1$ dB at $\omega_c = 4.7$ rad/s is tolerable
- If the gain difference was too large, we should also have designed and cascaded a lag network
- The resulting feedback controller is

$$C(s) = C_1(s)C_{2Lead}(s) = 4\left(\frac{1+0.39s}{1+0.12s}\right)^2$$

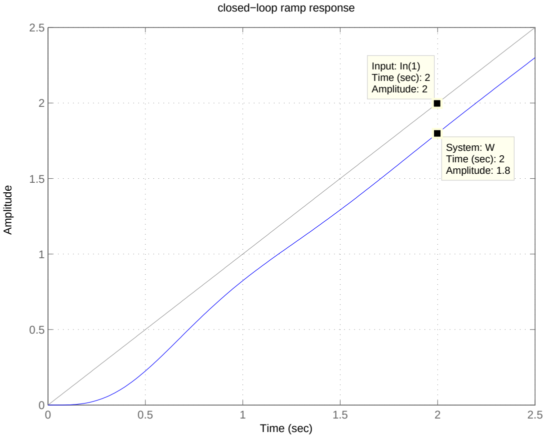


Bode plot of loop transfer function $L(j\omega) = C(j\omega)G(j\omega)$

- $B_3 = 5.54$ rad/s
- $M_p = 41.5$ deg

Example of loop shaping – Validation of the controller

1. Track a ramp reference $r(t) = t$ with finite steady-state error $e_d \leq 0.2$



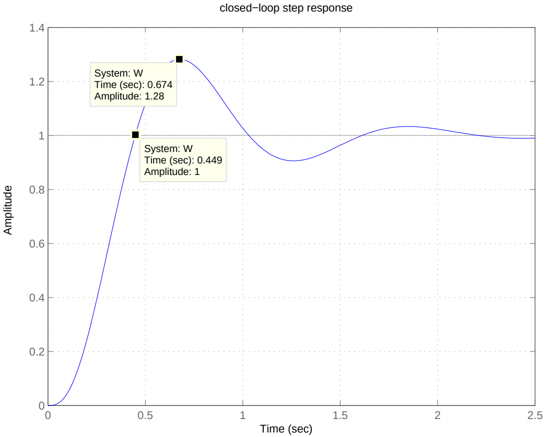
tracking error
$e(2) = 2 - 1.8 = 0.2$

MATLAB

```
»W = L/(1+L);
»t = 0:0.01:2.5;
»ramp=t;
»lsim(W,ramp,t);
```

Example of loop shaping – Validation of the controller

2. Rise-time of unit step response $t_r \simeq 0.4$
3. Overshoot of unit step reference $\hat{s} \leq 25\%$



MATLAB

```
»step = ones(1,length(t));
»lsim(W,step,t);
```

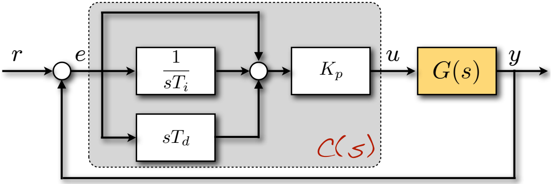
Proportional integral derivative (PID) controllers

- *PID (proportional integrative derivative) controllers* are the most used controllers in industrial automation since the '30s

$$u(t) = K_p \left[e(t) + \frac{1}{T_i} \int_0^t e(\tau) d\tau + T_d \frac{de(t)}{dt} \right]$$

  where $e(t) = r(t) - y(t)$ is the tracking error
- Initially constructed by analog electronic components, today they are implemented digitally
  - *ad hoc* digital devices
  - just few lines of C code included in the control unit

PID parameters



$$u(t) = K_p \left[e(t) + \frac{1}{T_i} \int_0^t e(\tau)d\tau + T_d \frac{de(t)}{dt} \right]$$

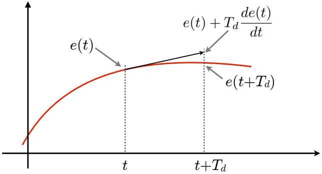
- K_p is the *controller gain*, determining the “aggressiveness” (=closed-loop bandwidth) of the controller
- T_i is the *reset time*, determining the weight of the integral action
- T_d is the *derivative time*, determining the phase lead of the controller
- we call the controller P, PD, PI, or PID depending on the feedback terms included in the control law

Derivative term

- The derivative term has transfer function sT_d, a high pass filter
- To avoid amplifying high-frequency noise (and to make the PID transfer function proper) sT_d gets replaced by

$$sT_d \approx \frac{sT_d}{1+s\frac{T_d}{N}}$$

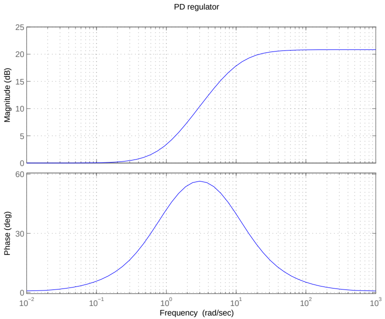
- No effect of the new pole $s=-\frac{N}{T_d}$ at low frequencies, but the high-frequency gain is limited to N (typically $N=3\div 20$)
- The derivative term has the effect of "predicting" the future tracking error $\hat{e}(t+T_d)=e(t)+T_d\dfrac{de(t)}{dt}$ (linear extrapolation)



There are more advanced controllers that use a more refined prediction, based on the mathematical model of the process (*model predictive control*, MPC – See more later ...)

Frequency response of PD controller

$$G_{PD}(s) = K_p\left(1 + \frac{sT_d}{1 + s\frac{T_d}{N}}\right) = K_p \frac{1 + \overbrace{\left(T_d\frac{N+1}{N}\right)}^{\tau} s}{1 + \underbrace{\frac{1}{N+1}}_{\alpha = \frac{1}{N+1}} \underbrace{\left(T_d\frac{N+1}{N}\right) s}_{\tau}}$$

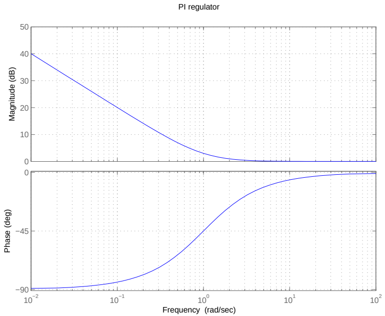


- $K_p = 1$
- $K_d = 1$
- $N = 10$

- The PD controller is equivalent to a lead network

Frequency response of PI controller

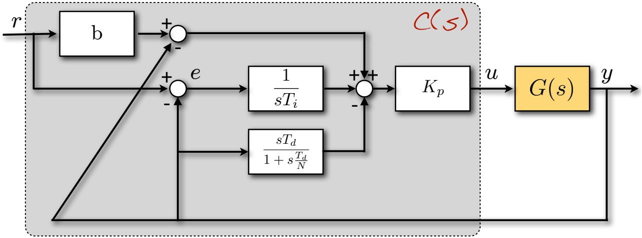
$$G_{\mathrm{PI}}(s) = K_p\left(1 + \frac{1}{sT_i}\right) = \frac{K_p/T_i}{s}(1 + T_i s)$$



- $K_p = 1$
- $T_i = 1$

- The PI controller introduces integral action
- The zero $-\dfrac{1}{T_i}$ compensates the decrease of phase margin of the integrator

Implementation of PID controller



$$u(t) = K_p \left[br(t) - y(t) + \frac{1}{sT_i}(r(t) - y(t)) - \frac{sT_d}{1 + s\frac{T_d}{N}} y(t) \right]$$

- The reference signal r is not included in the derivative term ($r(t)$ may have abrupt changes)
- The proportional action only uses a fraction $b \leq 1$ of the reference signal r.

Final remarks on loop shaping

- Loop-shaping techniques are most adequate for SISO (single-input single-output) systems
- They provide good insight in frequency domain properties of the closed loop (bandwidth, noise filtering, robustness to uncertainty, etc.)
- Most traditional single-loop industrial controllers are PID, and over 90% of PIDs are PI

Curiosity:

PID Temperature Control Retrofit KIT for Gaggia

This PID controller kit is designed for retrofitting into the Gaggia Classic, Gaggia Coffee, and Gaggia Coffee Deluxe espresso machine. By adding a PID controller to the heater control circuit, brewing water temperature can be controlled to $\pm 1\ ^\circ F$ *accuracy. Thus, it will significantly improve the taste of your espresso. Users can also easily adjust the brew water temperature to suit their own tastes.*

Auber Instruments

English-Italian Vocabulary

	
loop shaping	*sintesi per tentativi*
peak overshoot	*sovraelongazione massima*
rise time	*tempo di salita*
settling time	*tempo di assestamento*
lead network	*rete anticipatrice*
lag network	*rete attenuatrice*

Translation is obvious otherwise.