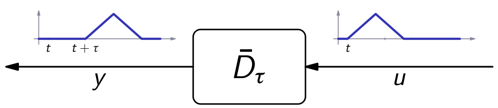
Control Theory (00350188)
lecture no. 4

Leonid Mirkin

Faculty of Mechanical Engineering
Technion — IIT

Outline

Delay element in time domain



I/O relation:

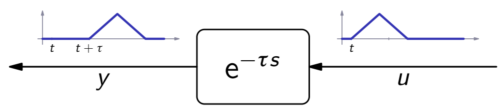
$$y = \bar{D}_\tau u \iff y(t) = u(t - \tau)$$

This system is

- linear
  $\bar{D}_\tau(\alpha_1 u_1 + \alpha_2 u_2) = \alpha_1 u_1(t-\tau) + \alpha_2 u_2(t-\tau) = \alpha_1(\bar{D}_\tau u_1) + \alpha_2(\bar{D}_\tau u_2)$
- time invariant
  $\bar{D}_{\tau_1}(\mathbb{S}_{\tau_2} u) = u(t - \tau_2 - \tau_1) = \mathbb{S}_{\tau_2}(\bar{D}_{\tau_1} u)$
- BIBO stable[1]
  $\|y\|_\infty = \|u\|_\infty$ for all $u \in L_\infty$

[1] $L_\infty := \{x : \mathbb{R} \to \mathbb{R} \mid \|x\|_\infty < \infty\}$, where $\|x\|_\infty := \sup_{t \in \mathbb{R}} |x(t)|$

Delay element in s-domain



By the time shift property of the Laplace transform:

$$y(t) = u(t - \tau) \iff Y(s) = e^{-\tau s} U(s)$$

Thus, delay element has the transfer function

$$\bar{D}_\tau(s) = e^{-\tau s}.$$

This transfer function is

- irrational,

so $\bar{D}_\tau$ is an infinite-dimensional system.

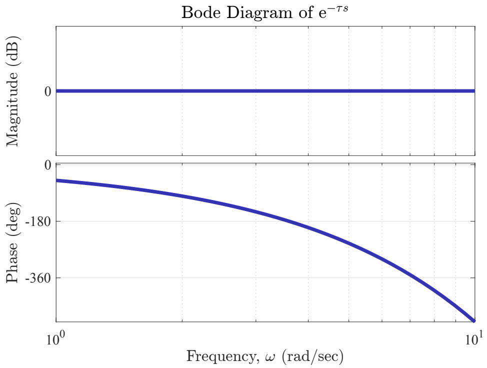
Frequency response

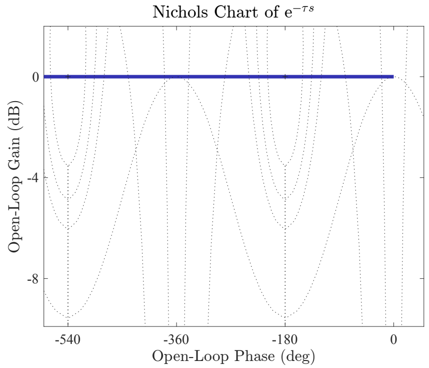
Now,

$$e^{-\tau s}\big|_{s=j\omega} = e^{-j\tau\omega} = \cos(\tau\omega) - j\sin(\tau\omega)$$

and it has

- unit magnitude ($|e^{j\tau\omega}| \equiv 1$) and
- linearly decaying phase ($\arg e^{j\tau\omega} = -\tau\omega$, in radians if ω is in rad/sec)

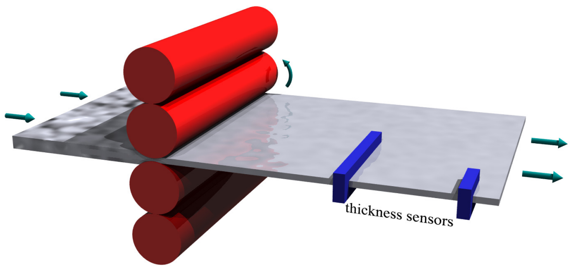




Why delays?

- Ubiquitous in physical processes
  - loop delays
  - process delays
  - ...
- Compact/economical approximations of complex dynamics
- Exploiting delays to improve performance

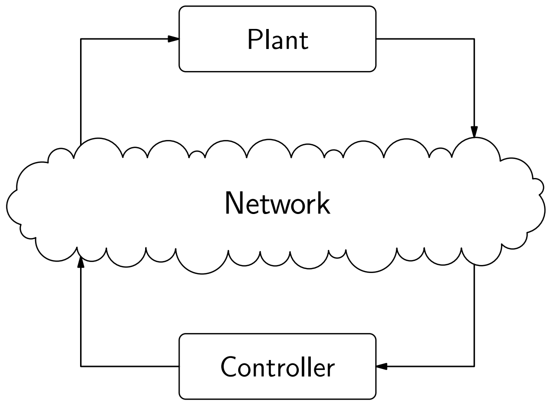
Loop delays: steel rolling



Thickness can only be measured at some distance from rolls, leading to

- measurement delays

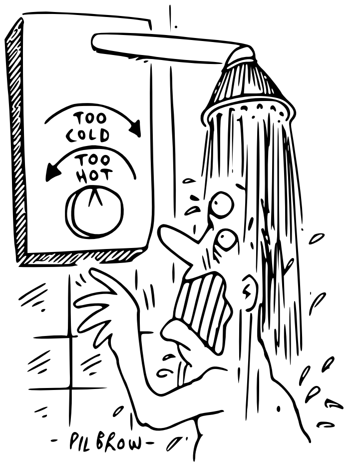
Loop delays: networked control



Sampling, encoding, transmission, decoding need time. This gives rise to

- measurement delays
- actuation delays

Loop delays: temperature control



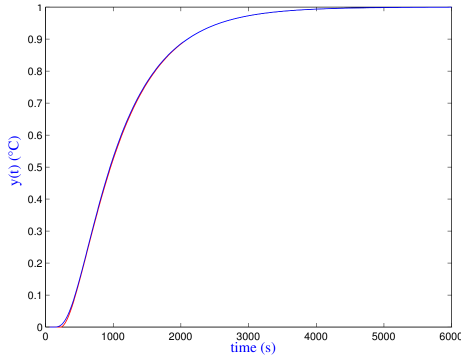
Everybody experienced this, I guess. . .

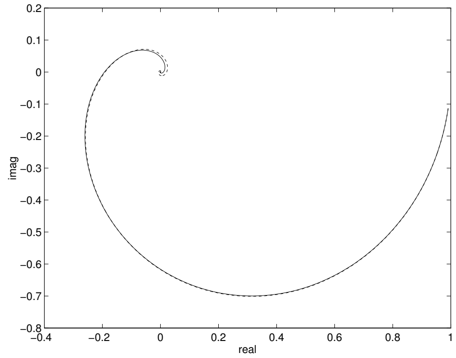
Delays as modeling tool: heating a can

Transfer function of a heated can (derived from a PDE model):

$$G(s)=\frac{1}{J_0\bigl(\sqrt{\frac{-s}{\alpha}}R\bigr)}+\sum_{m=1}^{\infty}\frac{2}{\lambda_m R\,J_1(\lambda_m R)}\,\frac{s}{s+\alpha\lambda_m^2}\,\frac{1}{\cosh\bigl(\sqrt{\frac{s}{\alpha}+\lambda_m^2}\cdot\frac{L}{2}\bigr)}.$$

Its approximation by $G_2(s)=\frac{1}{(\tau_1 s+1)(\tau_2 s+1)}\mathrm{e}^{-\tau s}$ is reasonably accurate:





Delays as modeling tool: process control

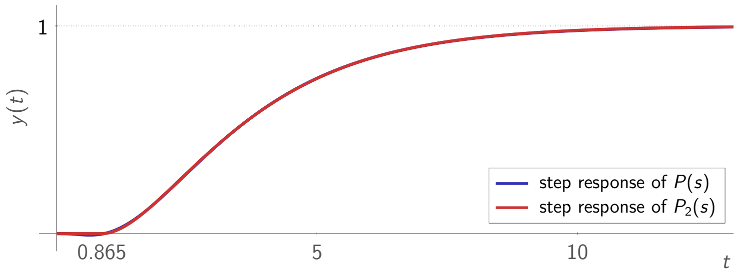
Consider a process described by

$$P(s)=\frac{(-0.3s+1)(0.08s+1)}{(2s+1)(s+1)(0.4s+1)(0.2s+1)(0.05s+1)^3}.$$

Its "second-order+delay" approximation (found by a brute-force search),

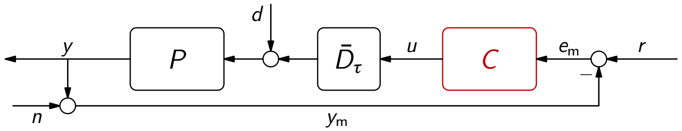
$$P_2(s)=\frac{1}{(1.19s+1)(1.91s+1)}\,\mathrm{e}^{-0.865s},$$

is quite accurate, yet includes less parameters to identify:



Outline

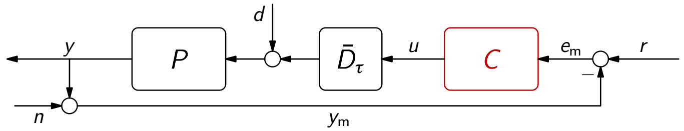
Systems with loop delays



where

- P plant, has a rational transfer function
- C controller, has a rational transfer function

Effect of loop delay on characteristic polynomial



If $P(s) = N_P(s)/D_P(s)$ and $C(s) = N_C(s)/D_C(s)$, then

$$\chi_{cl}(s) = e^{-\tau s} N_P(s) N_C(s) + D_P(s) D_C(s)$$

has infinitely many roots (this $\chi_{cl}(s)$ called quasi-polynomial).

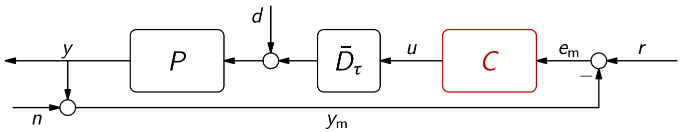
Example

Let $P(s) = 1$ and $C(s) = k_p$. Then

$$\chi_{cl}(s) = k_p e^{-\tau s} + 1$$

has roots at $\tau s = \ln k_p + j(\pi + 2\pi i)$ for all $i \in \mathbb{Z}$.

Effect of loop delay on closed-loop frequency response



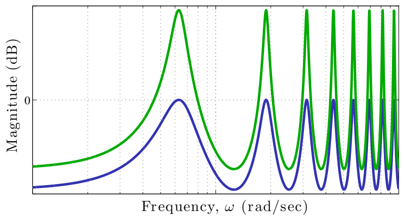
For instance,

$$T(j\omega) = \frac{P(j\omega)C(j\omega)e^{-j\tau\omega}}{1 + P(j\omega)C(j\omega)e^{-j\tau\omega}}$$

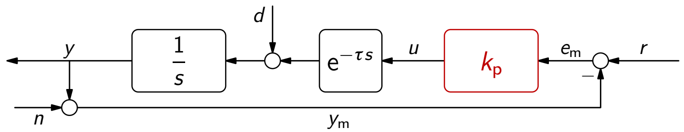
might be very complicated function of ω.

Example

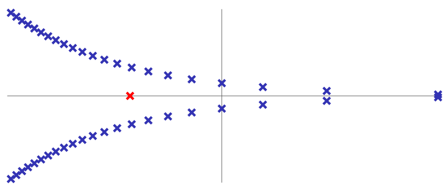
Again, let $P(s) = 1$ and $C(s) = k_p < 1$. Then for $k_p = \{0.5, 0.75\}$:

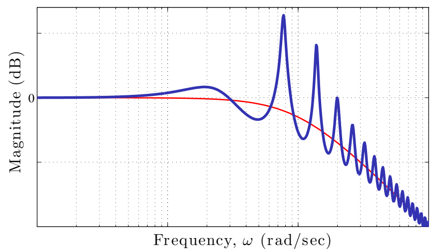


Another scary example



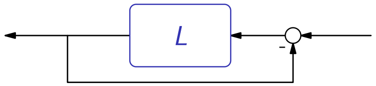
Comparing delay-free ($\tau = 0$) and delayed ($\tau > 0$) closed-loop systems





we can see that the presence of the delay considerably complicates matters.

Effect of loop delay on $L(\mathrm{j}\omega)$



Let $L(s) = L_r(s)\mathrm{e}^{-\tau s}$ for some rational $L_r(s)$. In this case

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

$$\Downarrow$$

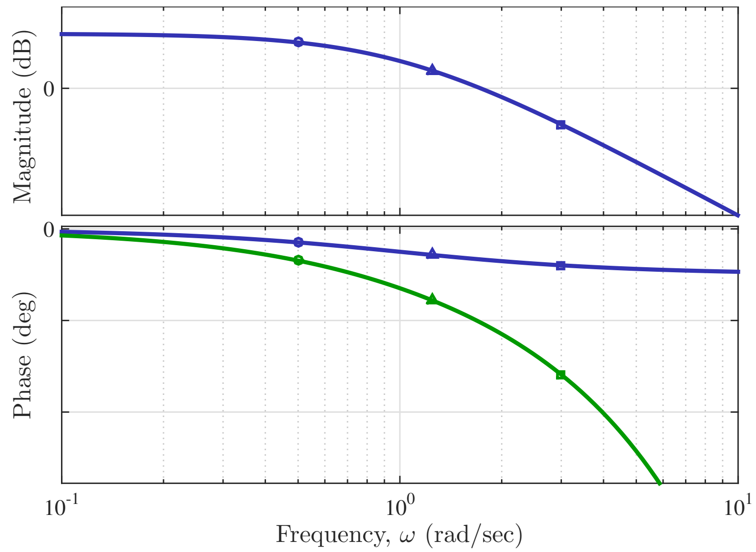
$$|L(\mathrm{j}\omega)| = |L_r(\mathrm{j}\omega)| \quad \text{and} \quad \arg L(\mathrm{j}\omega) = \arg L_r(\mathrm{j}\omega) - \tau\omega.$$

In other words, delay in this case

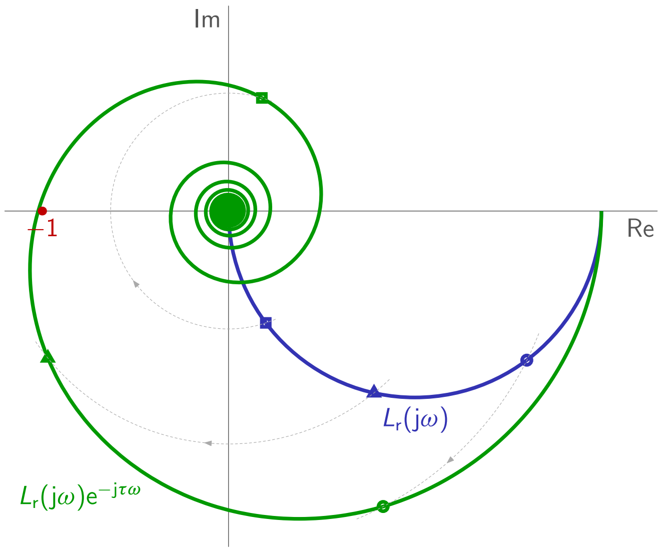
- does not change the magnitude of $L_r(\mathrm{j}\omega)$ and
- adds phase lag proportional to ω,

which is not hard to account for.

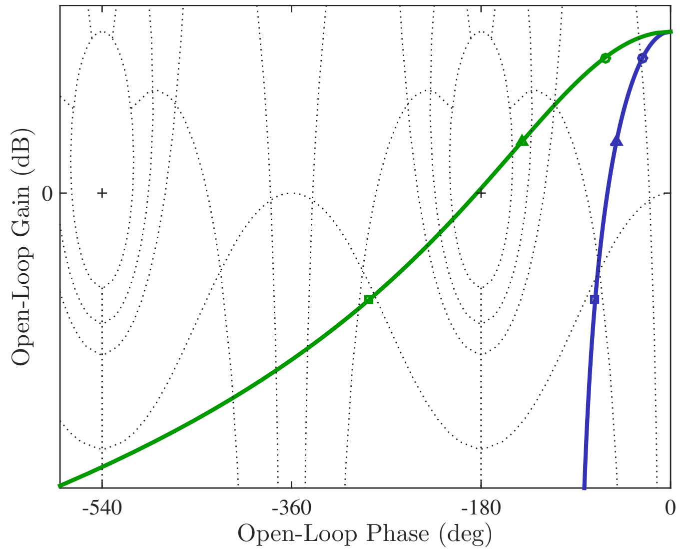
Effect of loop delay on $L(\mathrm{j}\omega)$: Bode diagram



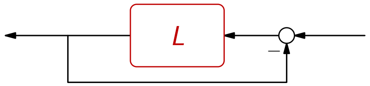
Effect of loop delay on $L(\mathrm{j}\omega)$: polar plot



Effect of loop delay on $L(\mathrm{j}\omega)$: Nichols chart



Nyquist stability criterion



The idea is to

- use plot of $L(\mathrm{j}\omega)$ to count the number of closed-loop poles in RHP.

Namely, let $L(s)$ have no unstable pole/zero cancellations and denote

n_{ol} number of unstable poles of $L(s)$

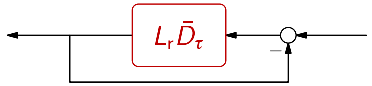
n_{cl} number of unstable poles of $\frac{1}{1+L(s)}$

$\varkappa$ number of clockwise encirclements of $-1+\mathrm{j}0$ by Nyquist plot of $L(\mathrm{j}\omega)$ as ω runs from $-\infty$ to ∞

In this case

$$n_{\mathrm{cl}} = n_{\mathrm{ol}} + \varkappa.$$

What is changed for dead-time systems ?



Nothing, at least if the high-frequency gain of $L_r(s)$ is not 1.

Remark: If $|L_r(\mathrm{j}\infty)| = 1$, the situation is quite complicated, the closed-loop system might have no RHP poles and still be unstable, like for

$$L_r(s) = \frac{s}{s+1}$$

(explanations go beyond the scope of this course). It is safe to say that in this case the closed-loop system is unstable, regardless its pole locations.

Loop delays and closed-loop stability: "rigid" loops

For systems with "rigid" loops

- delay is a destabilizing factor

as it adds phase lag, thus imposing limitations on achievable crossover ω_c.

Bode's gain-phase relation for $L(s) = L_r(s)\mathrm{e}^{-\tau s}$ (if $L_r(s)$ is minimum-phase)

$$\arg L(\mathrm{j}\omega_0) = \frac{1}{\pi}\int_{-\infty}^{\infty} \frac{\mathrm{d}\ln|L_r(\mathrm{j}\nu)|}{\mathrm{d}\nu} \ln\coth\frac{|\nu|}{2}\,\mathrm{d}\nu - \tau\omega_0, \quad \text{where } \nu := \ln\frac{\omega}{\omega_0},$$

so effect of delay is similar to effect of RHP zero.

Loop delays and closed-loop stability: "flexible" loops

For systems with "flexible" loops

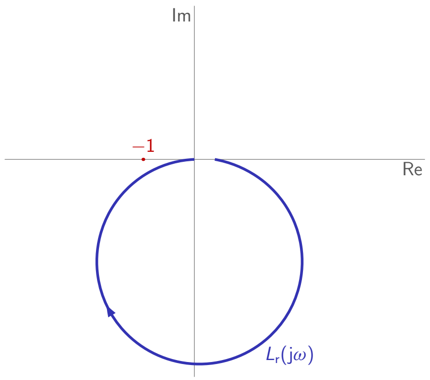
- delay in some (very special) cases may be stabilizing factor,

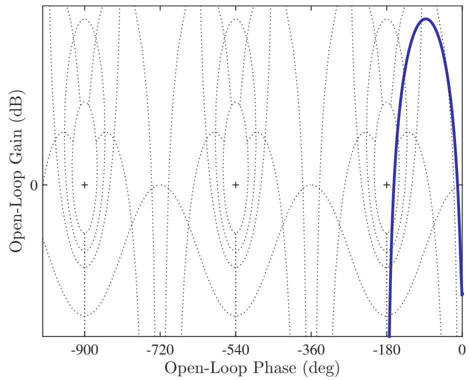
yet this property shall be used with great care[2].

[2] Don't try it at home!

Example

$L(s) = \dfrac{0.4}{s^2 + 0.1s + 1}$:

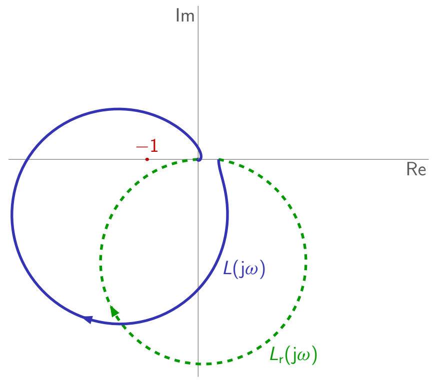


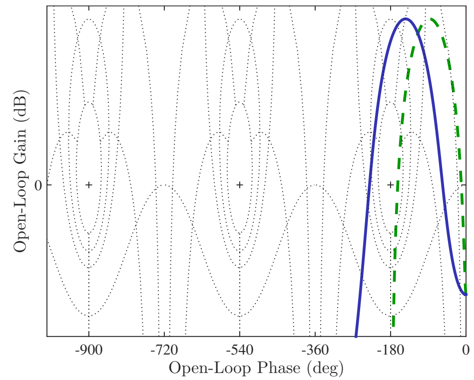


no closed-loop poles in RHP

Example (contd)

$L(s) = \dfrac{0.4}{s^2 + 0.1s + 1}\, e^{-s}$:



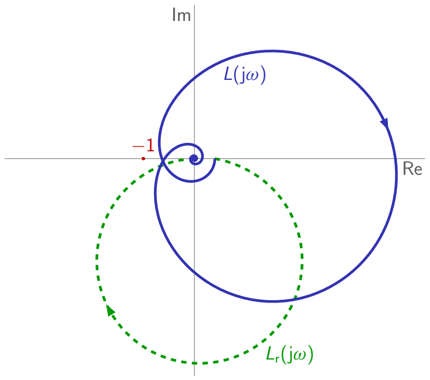


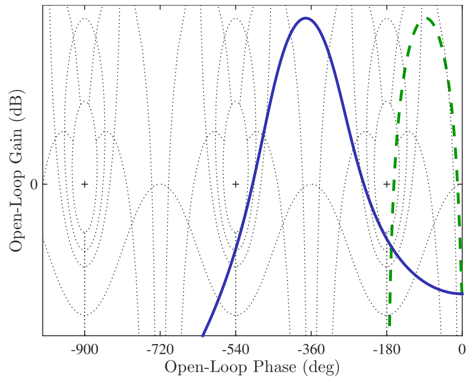
two closed-loop poles in RHP

– delay $\tau = 1$ is destabilizing

Example (contd)

$L(s) = \dfrac{0.4}{s^2 + 0.1s + 1}\, e^{-5s}$:



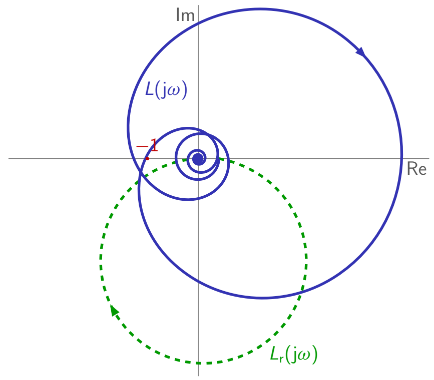


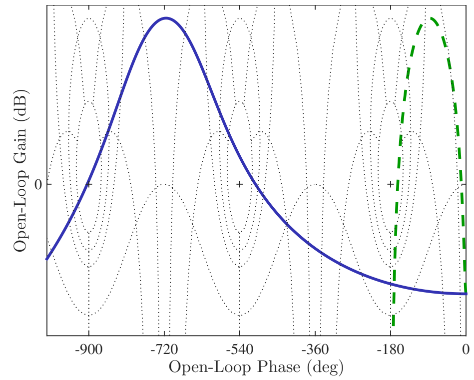
no closed-loop poles in RHP

– delay $\tau = 5$ is stabilizing (and stability margins > those with $\tau = 0$)

Example (contd)

$L(s) = \dfrac{0.4}{s^2 + 0.1s + 1}\, e^{-11s}$:





four closed-loop poles in RHP

– delay $\tau = 11$ is destabilizing again (and destabilizing for all larger τ)

Outline

Time-delay systems (mostly from IC)

Time delays and feedback (mostly from IC)

Rational approximations of time delays

Introduction to dead-time compensation

Modified Smith predictor (optional self-study)

Why to approximate

Delay element is infinite-dimensional, which complicates its treatment. It is not a surprise then that we want to approximate delay by finite-dimensional (rational) elements, say $R(s)$, to

- use standard methods in analysis and design,
- use standard software for simulations,
- ~~avoid learning new methods,~~
- ...

Approximation measure

We may consider as the approximation measure the quantity

$$\epsilon_R := \max_{\omega \in \mathbb{R}} |R(j\omega) - e^{-j\tau\omega}|,$$

assuming that $R(s)$ is stable.

What to approximate: bad news

On the one hand,

- phase lag of the delay element is not bounded (and continuous in ω).

On the other hand,

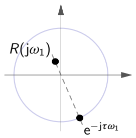
- rational systems can only provide finite phase lag.

Therefore, phase error between $e^{-\tau s}$ and any rational transfer function $R(s)$ is arbitrarily large. Moreover, for every $R(s)$ there always is ω_0 such that

- $\arg e^{-j\tau\omega} - \arg R(j\omega)$ continuously decreasing function of ω, $\forall \omega \geq \omega_0$.

Hence there always is frequency ω_1 such that

$$\arg e^{-j\tau\omega_1} - \arg R(j\omega_1) = -\pi - 2\pi k \qquad \text{i.e.}$$



and $\epsilon_R \geq 1$ for all $R(s)$. Thus,

- rational approximation of pure delay, $e^{-\tau s}$, is hopeless,

just because $\epsilon_R = 1$ already for the trivial (and senseless) choice $R(s) = 0$.

What to approximate: good news

We never work over infinite bandwidth. Hence, we

- need to approximate $e^{-\tau s}$ in finite frequency range

or, equivalently,

- approximate $F(s)e^{-\tau s}$ for some low-pass (strictly proper) $F(s)$.

This can be done, since

- phase lag of delay over finite bandwidth is finite

and

- magnitude of $F(j\omega)e^{-j\tau\omega}$ decreases as ω increases,

which implies that at frequencies where the phase lag of $F(j\omega)e^{-j\tau\omega}$ large, the function effectively vanishes.

Also, we may consider $\tau = 1$ w.l.o.g., otherwise $s \to s/\tau$ makes the trick.

Truncation-based methods

General idea is to

- truncate some power series,

which could give accurate results in a (sufficiently large) neighborhood of 0.

Naïve approach. We may try it as follows:

$$e^{-s} = \frac{e^{-s/2}}{e^{s/2}} = \frac{1 - \frac{s/2}{1!} + \frac{s^2/2^2}{2!} - \frac{s^3/2^3}{3!} + \cdots}{1 + \frac{s/2}{1!} + \frac{s^2/2^2}{2!} + \frac{s^3/2^3}{3!} + \cdots} \approx \frac{1 - \frac{s/2}{1!} + \cdots + \frac{(-1)^n s^n/n^2}{n!}}{1 + \frac{s/2}{1!} + \cdots + \frac{s^n/n^2}{n!}}$$

for some $n \in \mathbb{N}$. The problem is that this approximation is

- unstable whenever $n > 4$.

which is not convenient (e^{-s} itself is stable). For example, for $n = 5$ this approach yields

$$e^{-s} \approx \frac{-s^5 + 10s^4 - 80s^3 + 480s^2 - 1920s + 3840}{s^5 + 10s^4 + 80s^3 + 480s^2 + 1920s + 3840}$$

with poles at $\{-4.361, -3.299 \pm j3.388, 0.48 \pm j6.257\}$. Thus, more sophisticated methods are required ...

Truncation-based methods: Padé approximant

Consider approximation

$$e^{-s} \approx \frac{P_m(s)}{Q_n(s)} =: R_{[m,n]}(s),$$

where $P_m(s)$ and $Q_n(s)$ are polynomials of degrees m and n, respectively. Power series at $s = 0$ of each side are

$$e^{-s} = 1 - \frac{s}{1!} + \frac{s^2}{2!} - \frac{s^3}{3!} + \cdots$$

$$R_{[m,n]}(s) = R_{[m,n]}(0) + \frac{R'_{[m,n]}(0)s}{1!} + \frac{R''_{[m,n]}(0)s^2}{2!} + \frac{R'''_{[m,n]}(0)s^3}{3!} + \cdots$$

The idea of $[m, n]$-Padé approximant is to find coefficients of $R_{[m,n]}(s)$ via

- matching their first $n + m + 1$ power series coefficients.

It can be shown that

- $n = m \implies P_n(s) = Q_n(-s)$.

Example: [2, 2]-Padé approximant

In this case $R_{[2,2]}(s) = \frac{s^2 - q_1 s + q_0}{s^2 + q_1 s + q_0}$ and power series are

$$e^{-s} = 1 - s + \frac{s^2}{2} - \frac{s^3}{6} + \frac{s^4}{24} - \cdots$$

$$R_{[2,2]}(s) = 1 - \frac{2q_1}{q_0}s + \frac{2q_1^2}{q_0^2}s^2 - \frac{2(q_1^3 - q_1 q_0)}{q_0^3}s^3 + \frac{2(q_1^4 - 2q_1^2 q_0)}{q_0^4}s^4 - \cdots$$

from which

$$q_0 = 2q_1 \quad \text{and} \quad \frac{q_1 - 2}{4q_1} = \frac{1}{6}$$

and then $q_1 = 6$ and $q_0 = 12$, matching 5 coefficients.

Thus, [2, 2]-Padé approximant is

$$e^{-s} \approx \frac{s^2 - 6s + 12}{s^2 + 6s + 12} = \frac{1 - \frac{s}{2} + \frac{s^2}{12}}{1 + \frac{s}{2} + \frac{s^2}{12}}.$$

Truncation-based methods: Padé approximant (contd)

General formula for $[n, n]$-Padé approximant is

$$e^{-s} \approx \frac{\sum_{i=0}^n \binom{n}{i} \frac{(2n-i)!}{(2n)!} (-s)^i}{\sum_{i=0}^n \binom{n}{i} \frac{(2n-i)!}{(2n)!} s^i} = \frac{\sum_{i=0}^n \frac{(2n-i)!n!}{(2n)!(n-i)!\,i!} (-s)^i}{\sum_{i=0}^n \frac{(2n-i)!n!}{(2n)!(n-i)!\,i!} s^i}$$

This yields:

n	1	2	3	4
$e^{-\tau s} \approx$	$\frac{1-\frac{\tau s}{2}}{1+\frac{\tau s}{2}}$	$\frac{1-\frac{\tau s}{2}+\frac{\tau^2 s^2}{12}}{1+\frac{\tau s}{2}+\frac{\tau^2 s^2}{12}}$	$\frac{1-\frac{\tau s}{2}+\frac{\tau^2 s^2}{10}-\frac{\tau^3 s^3}{120}}{1+\frac{\tau s}{2}+\frac{\tau^2 s^2}{10}+\frac{\tau^3 s^3}{120}}$	$\frac{1-\frac{\tau s}{2}+\frac{3\tau^2 s^2}{28}-\frac{\tau^3 s^3}{84}+\frac{\tau^4 s^4}{1680}}{1+\frac{\tau s}{2}+\frac{3\tau^2 s^2}{28}+\frac{\tau^3 s^3}{84}+\frac{\tau^4 s^4}{1680}}$

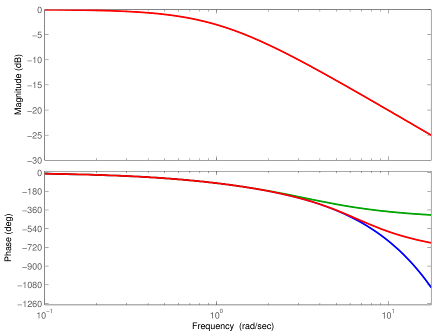
It can be proved that

- $[n, n]$-Padé approximation is stable for all $n \in \mathbb{N}$.

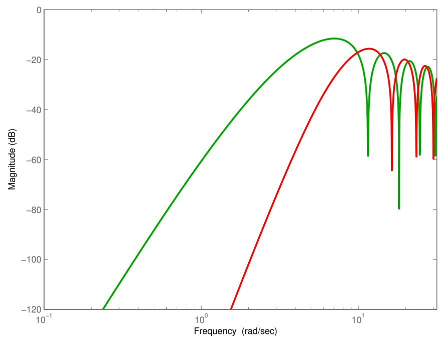
Padé approximant: example

Let $e^{-s}/(s+1)$. Its Padé approximant can be calculated with the Matlab command `pade(tf(1,[1 1],'InputDelay',1),N)`.

$\frac{e^{-s}}{s+1}$ and its 2nd and 4th order approximations



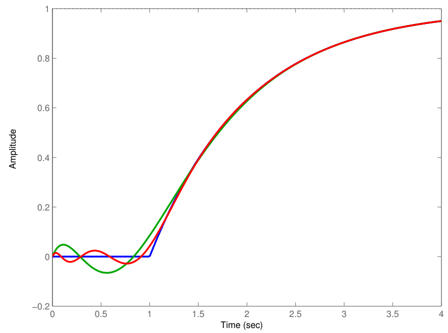
Approximation errors



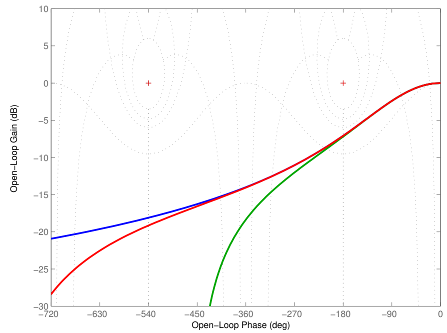
Padé approximant: example (contd)

We may also compare step responses and Nichols charts

$\frac{e^{-s}}{s+1}$ and its 2nd and 4th order approximations



$\frac{e^{-s}}{s+1}$ and its 2nd and 4th order approximations



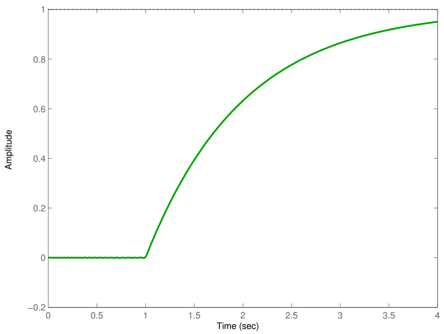
From loop shaping perspectives,

– approximation performance depends on crossover requirements.

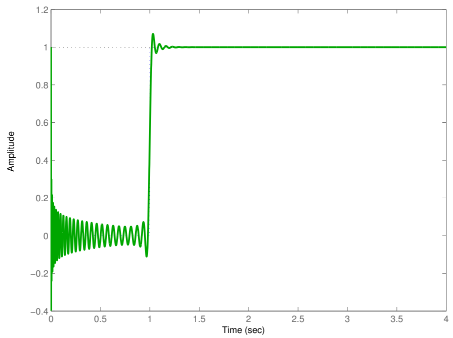
Padé approximant: example (contd)

Increasing approximants orders improves the match between step responses of $e^{-s}/(s+1)$ and its Padé approximation:

$\frac{e^{-s}}{s+1}$ and its 50th order approximation



e^{-s} and its 50th order approximation



Not true for the approximation of the pure delay e^{-s} !

Outline

Time-delay systems (mostly from IC)

Time delays and feedback (mostly from IC)

Rational approximations of time delays

Introduction to dead-time compensation

Modified Smith predictor (optional self-study)

Infinite dimensional for infinite dimensional

In many cases,

- controller complexity should be compatible to plant complexity,

since controller should, in a sense, "counteract" plant dynamics.

Plant with dead time,

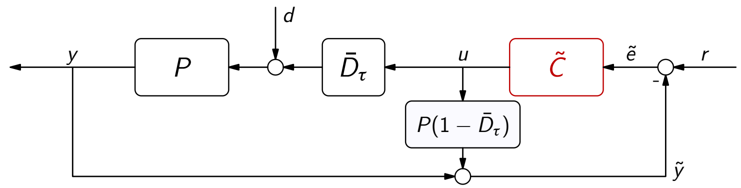
- $P\bar{D}_\tau$, is an infinite dimensional.

We may, hence, expect that we can do better by using infinite-dimensional controllers. In doing this, the following aspects are of primary importance:

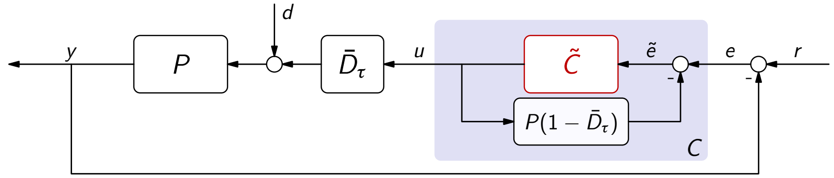
- small number of design parameters;
- implementability;
- design transparency.

Smith controller

Otto J. M. Smith (1957) proposed



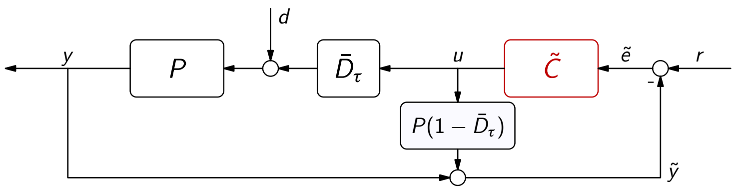
or, equivalently, in the unity-feedback form



where the overall controller $e \mapsto u$ has the irrational transfer function

$$C(s) = \frac{\tilde{C}(s)}{1 + P(s)(1 - e^{-\tau s})\tilde{C}(s)}.$$

Smith controller: rationale



The signal

$$\begin{aligned}\tilde{y} &= y + P(1-\bar{D}_\tau)u = P(d + \bar{D}_\tau u) + P(1-\bar{D}_\tau)u \\ &= P(d+u)\end{aligned}$$

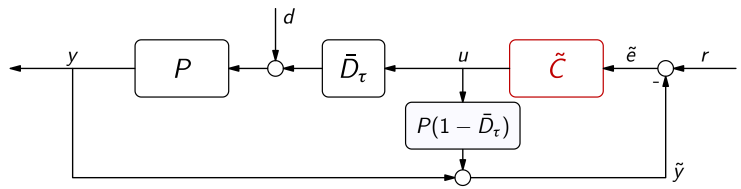
would be the output in the delay-free case. The internal feedback block

- $P(1-\bar{D}_\tau)$, dubbed the Smith predictor or dead-time compensator,

helps to predict the plant output τ time units ahead. The designed part is

- the primary controller $\tilde{C}$.

Smith controller: the trick



The transfer function of the closed-loop system $r \mapsto y$ is

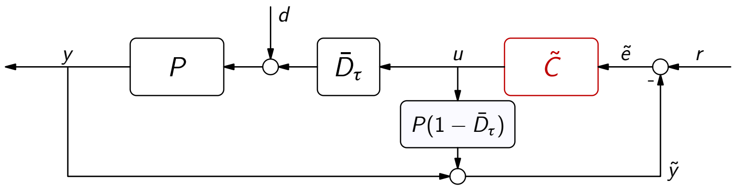
$$T(s) = \frac{P(s)\tilde{C}(s)}{1 + P(s)\tilde{C}(s)}\, e^{-\tau s} =: \tilde{T}(s)\, e^{-\tau s}.$$

Note that

- denominator of the closed-loop transfer function is delay free.

We may expect that if $\tilde{C}$ stabilizes P, then C stabilizes $P\bar{D}_\tau$.

Smith controller: design paradigm



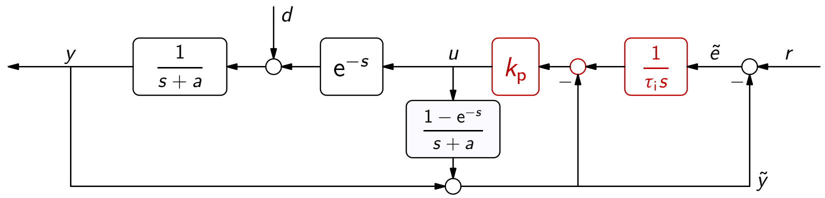
The following two-stage procedure appears natural:

1. design primary controller $\tilde{C}$ for delay-free plant, P;
2. implement primary controller in combination with the Smith predictor.

It yields a

- finite-dimensional design with an infinite-dimensional controller,
- small number of tuning parameters (those of $\tilde{C}$),
- implementability (the only infinite-dimensional part, $\bar{D}_\tau$, is a buffer)

Smith controller: design example 1



Here the primary controller is PI (implemented to avoid zeros in $r \mapsto y$)

$$\tilde{C}(s) = k_p\left(1 + \frac{1}{\tau_i s}\right) = \frac{k_p(\tau_i s + 1)}{\tau_i s}.$$

We aim at

- "good" step response and "good" disturbance attenuation

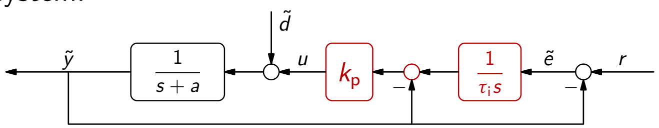
and hope for the

- transparency of tuning k_p and τ_i

(i.e. that their "good" choices result in "good" overall controller C).

Design example 1: stage 1

Delay-free system:



Its characteristic polynomial,

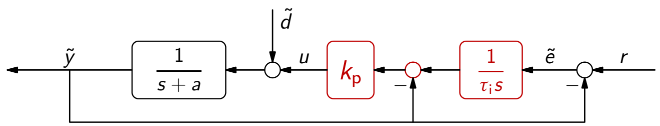
$$\tilde{\chi}_{cl}(s) = k_p(\tau_i s + 1) + \tau_i s(s+a) = \tau_i s^2 + \tau_i(a + k_p)s + k_p,$$

is Hurwitz iff either $\tau_i > 0 \wedge k_p > \max\{-a, 0\}$ or $\tau_i < 0 \wedge k_p < \min\{-a, 0\}$. The closed-loop transfer functions,

$$\begin{bmatrix} \tilde{T}_{yr}(s) \\ \tilde{T}_d(s) \end{bmatrix} = \begin{bmatrix} k_p \\ \tau_i s \end{bmatrix} \frac{1}{\tau_i s^2 + \tau_i(a + k_p)s + k_p},$$

are easy to understand (with $\tilde{T}_{yr}(0) = 1$ and $\tilde{T}_d(0) = 0$ because of "I").

Design example 1: stage 1 (contd)



Let's choose

$$k_p = \tau_i = 2 - a,$$

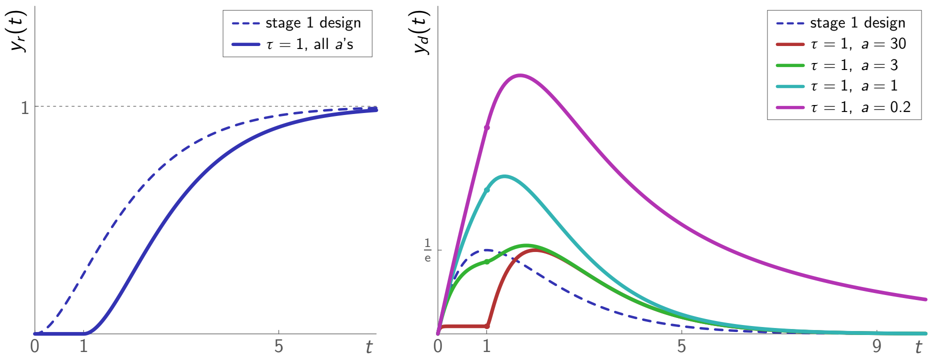
for which

$$\tilde{T}_{yr}(s) = \frac{1}{(s+1)^2} \quad \text{and} \quad \tilde{T}_d(s) = \frac{s}{(s+1)^2}$$

(if $a = 2$ we end up with the I controller $\tilde{C}(s) = 1/s$, otherwise $\tilde{C}(s)$ is PI).

Design example 1: stage 2 ($a > 0$, stable plant)

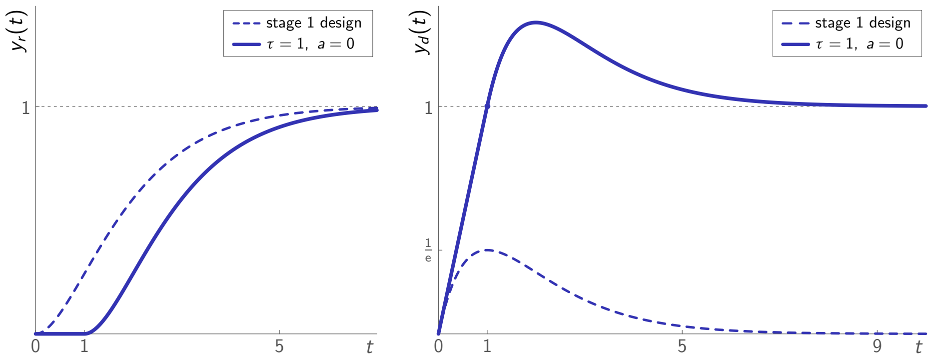
Closed-loop step responses:



– reference responses are as expected
– disturbance responses are not always (decays slow for small a) why?

Design example 1: stage 2 ($a = 0$, intergator)

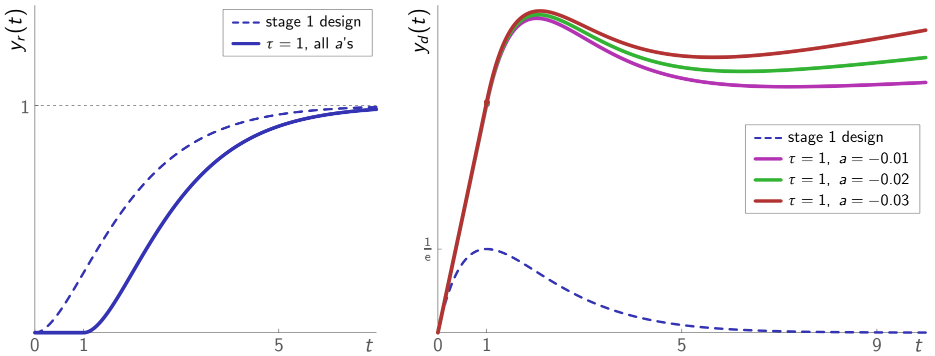
Closed-loop step responses:



– reference response is as expected
– disturbance response is not ($\lim_{t\to\infty} y_d(t) = 1 \neq 0$) why?

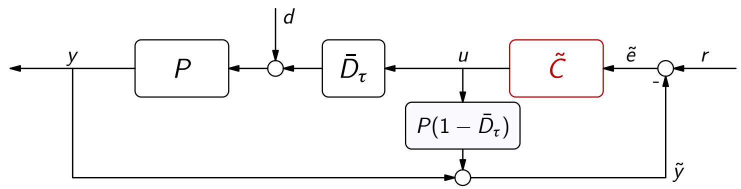
Design example 1: stage 2 ($a < 0$, unstable plant)

Closed-loop step responses:



– reference responses are as expected
– disturbance responses are not (diverge!) why?

Smith controller: pole-zero cancellations



Let

$$P(s) = \frac{N_P(s)}{D_P(s)} \quad \text{and} \quad \tilde{C}(s) = \frac{N_{\tilde{C}}(s)}{D_{\tilde{C}}(s)}$$

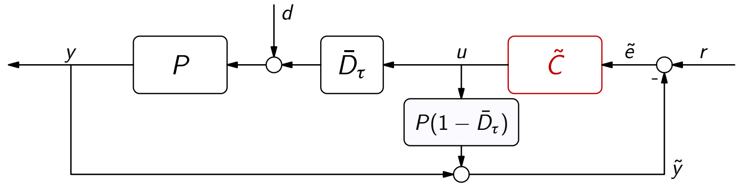
and assume that these fractions are irreducible. Then

$$C(s) = \frac{\tilde{C}(s)}{1 + P(s)\tilde{C}(s)(1 - e^{-\tau s})} = \frac{D_P(s)N_{\tilde{C}}(s)}{D_P(s)D_{\tilde{C}}(s) + N_P(s)N_{\tilde{C}}(s)(1 - e^{-\tau s})}$$

and, excluding the obvious case when $\tilde{C}(s)$ cancels poles of $P(s)$,

– poles of $P(s)$ are zeros of $C(s)$, unless they are zeros of $1 - e^{-\tau s}$ too.

Pole-zero cancellations: implications

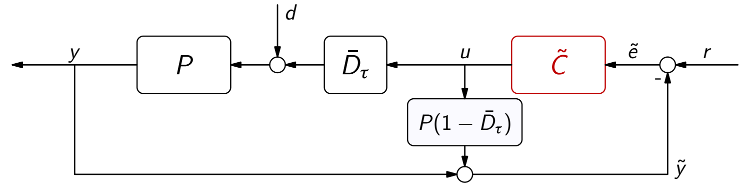


Smith controller is

- internally unstable whenever P is unstable
  (unless[3] all unstable poles of $P(s)$ are zeros of $1-\mathrm{e}^{-\tau s}$, which are at $\mathrm{j}\frac{2\pi}{h}k$, $\forall k \in \mathbb{Z}$)
- inefficient in attenuating load disturbances if P has "slow" poles
- inefficient in dampening lightly-damped dynamics of the plant

[3]This is what happened in the example with $a=0$.

Smith controller: disturbance response



Disturbance sensitivity

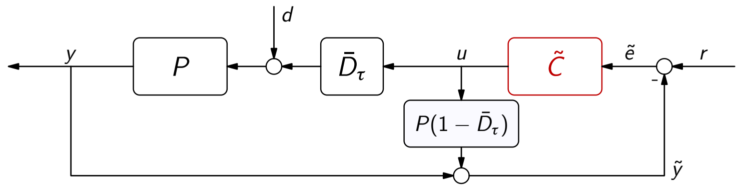
$$T_\mathrm{d}(s) = \frac{1+P(s)\tilde{C}(s)(1-\mathrm{e}^{-\tau s})}{1+P(s)\tilde{C}(s)}P(s) = \tilde{T}_\mathrm{d}(s) + \tilde{T}(s)(1-\mathrm{e}^{-\tau s})P(s)$$

is indeed unstable, unless all CRHP poles of $P(s)$ are canceled by $1-\mathrm{e}^{-\tau s}$. Also note that

- a "good" $\tilde{T}_\mathrm{d}(s)$ does not necessarily result in a "good" $T_\mathrm{d}(s)$

(because the relation between $|T_\mathrm{d}(\mathrm{j}\omega)|$ and $|\tilde{T}_\mathrm{d}(\mathrm{j}\omega)|$ is complicated, unless $|\tilde{T}(\mathrm{j}\omega)(1-\mathrm{e}^{-\mathrm{j}\omega h})P(\mathrm{j}\omega)| \ll 1$).

Smith controller: integral action in the controller



If $\tilde{C}(s) = \frac{1}{s}\tilde{C}_0(s)$ for some $\tilde{C}_0(s)$ such that $|\tilde{C}_0(0)| < \infty$, then

$$C(s) = \frac{\tilde{C}(s)}{1+(1-\mathrm{e}^{-\tau s})P(s)\tilde{C}(s)} = \frac{\tilde{C}_0(s)}{s+(1-\mathrm{e}^{-\tau s})P(s)\tilde{C}_0(s)}$$

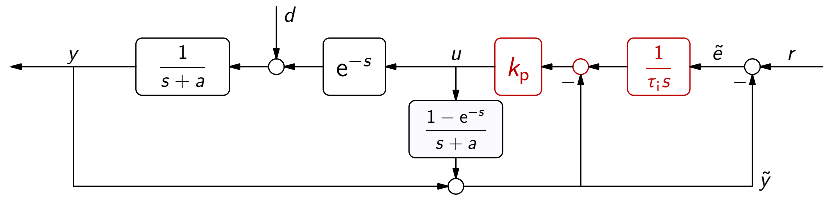
and there is an

- integrator in $C \iff \lim_{s\to 0}(1-\mathrm{e}^{-\tau s})P(s) = 0$

(i.e. the predictor part has zero static gain). As a general rule,

- design of $\tilde{C}$ is transparent at frequencies where the predictor gain is low, i.e. $|(1-\mathrm{e}^{-\mathrm{j}\tau\omega})P(\mathrm{j}\omega)| \ll 1$.

Design example 1: controller static gain



We have that

$$\lim_{s\to 0}\frac{1-\mathrm{e}^{-s}}{s+a} = \begin{cases} 1 & \text{if } a=0 \\ 0 & \text{otherwise} \end{cases} \implies \lim_{s\to 0} C(s) = \begin{cases} 1 & \text{if } a=0 \\ \infty & \text{otherwise} \end{cases}$$

This agrees with simulations, where the disturbance was rejected in steady state only with $a>0$, but not with $a=0$.

Outline

Time-delay systems (mostly from IC)

Time delays and feedback (mostly from IC)

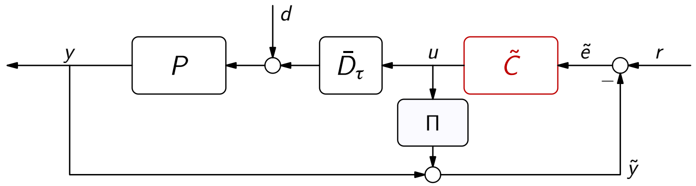
Rational approximations of time delays

Introduction to dead-time compensation

Modified Smith predictor (optional self-study)

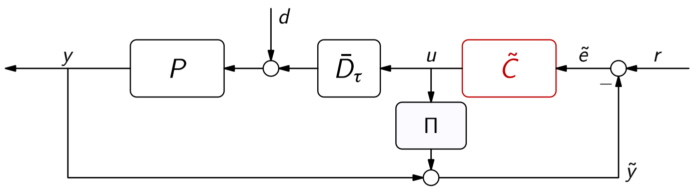
Modified Smith predictor

Some problems may be resolved in the Modified Smith Predictor (MSP):



where
$$\Pi = \tilde{P} - P\bar{D}_\tau,$$
for some $\tilde{P}$, having rational and proper $\tilde{P}(s)$, which may be different from P. This Π also compensates the delay:
$$\tilde{y} = P(d + \bar{D}_\tau u) + (\tilde{P} - P\bar{D}_\tau)u = Pd + \tilde{P}u,$$
although no longer predicts the delay-free output $P(d+u)$.

Modified Smith predictor (contd)

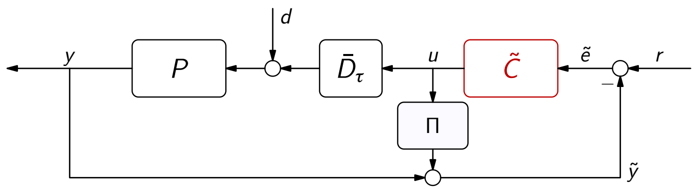


The transfer function of the closed-loop system $r \mapsto y$ is then
$$T(s) = \frac{P(s)\tilde{C}(s)}{1 + \tilde{P}(s)\tilde{C}(s)}\,\mathrm{e}^{-\tau s}$$
and its denominator is delay free (a standard polynomial if $\tilde{C}(s)$ is rational).

The two-stage design procedure may then be modified as follows:

1. design primary controller $\tilde{C}$ for $\tilde{P}$;
2. implement primary controller in combination with Π.

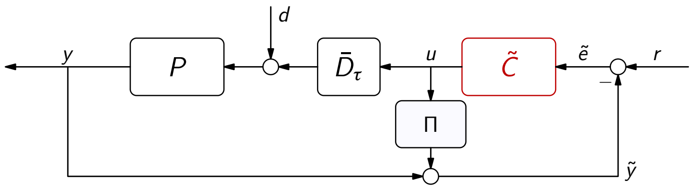
MSP: pole-zero cancellations



Let
$$P(s) = \frac{N_P(s)}{D_P(s)}, \quad \tilde{P}(s) = \frac{N_{\tilde{P}}(s)}{D_P(s)}, \quad \text{and} \quad \tilde{C}(s) = \frac{N_{\tilde{C}}(s)}{D_{\tilde{C}}(s)}$$
be irreducible ($\tilde{P}(s)$ is frequently chosen to have the same denominator as $P(s)$, although it need not). In this case
$$C(s) = \frac{\tilde{C}(s)}{1 + \tilde{C}(s)\Pi(s)} = \frac{D_P(s)N_{\tilde{C}}(s)}{D_P(s)D_{\tilde{C}}(s) + N_{\tilde{C}}(s)(N_{\tilde{P}}(s) - N_P(s)\mathrm{e}^{-\tau s})}$$
and, excluding the obvious case when $\tilde{C}(s)$ cancels poles of $\tilde{P}(s)$,

– poles of $P(s)$ are zeros of $C(s)$, unless zeros of $N_{\tilde{P}}(s) - N_P(s)\mathrm{e}^{-\tau s}$ too.

Pole-zero cancellations: implications



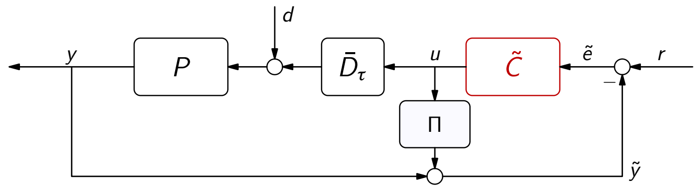
$\deg D_P(s)+1$ free parameters in $N_{\tilde{P}}(s)$ can be used to

- assign zeros of $N_\Pi(s) := N_{\tilde{P}}(s) - N_P(s)e^{-\tau s}$ at points of need.

This can be used to

- prevent unstable cancellations $\implies$ internal stability
- avoid harmful stable cancellations $\implies$ better disturbance attenuation
- render the logic in the choice of $\tilde{C}$ more streamlined (transparency)

MSP: stability



There are no unstable cancellations between $C(s)$ and $P(s)$ iff all unstable poles of $P(s)$ are zeros of $N_\Pi(s)$. Because

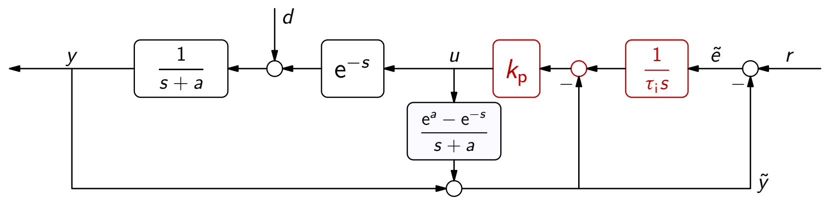
$$\Pi(s) = \frac{N_{\tilde{P}}(s)}{D_P(s)} - \frac{N_P(s)}{D_P(s)}e^{-\tau s} = \frac{N_\Pi(s)}{D_P(s)},$$

there are

- no unstable pole-zero cancellations in MSP iff Π is stable itself

and then the closed-loop system is internally stable iff $\tilde{C}$ stabilizes $\tilde{P}$.

MSP: design example 2



If

$$\tilde{P}(s) = \frac{e^a}{s+a} \qquad \implies \qquad \Pi(s) = \frac{e^a - e^{-s}}{s+a}.$$

In this case

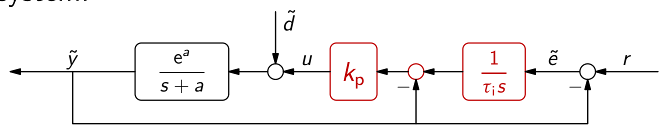
$$\lim_{s\to a}\Pi(s) = \lim_{s\to a}\frac{e^a - e^{-s}}{s+a} = \lim_{s\to a} e^{-s} = e^{-h}$$

is finite, so the singularity at $s=-a$ is removable (i.e. not a pole)[4].

[4] The implementation of this $\Pi(s)$ might not be straightforward if $a \le 0$. Yet this issue goes beyond our scope here, just know that $\Pi(s)$ can be implemented.

Design example 2: stage 1

Delay-free system:



Its characteristic polynomial,

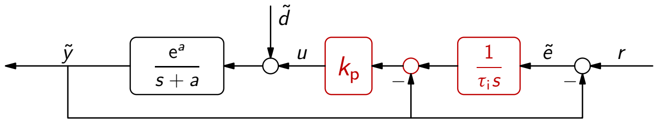
$$\tilde{\chi}_{cl}(s) = k_p e^a(\tau_i s + 1) + \tau_i s(s+a) = \tau_i s^2 + \tau_i(a + k_p e^a)s + k_p e^a,$$

is stable iff either $\tau_i > 0 \wedge k_p > \max\{-\frac{a}{e^a}, 0\}$ or $\tau_i < 0 \wedge k_p < \min\{-\frac{a}{e^a}, 0\}$. The closed-loop transfer functions,

$$\begin{bmatrix} \tilde{T}_{yr}(s) \\ \tilde{T}_d(s) \end{bmatrix} = \begin{bmatrix} k_p \\ \tau_i s \end{bmatrix} \frac{e^a}{\tau_i s^2 + \tau_i(a + k_p e^a)s + k_p e^a},$$

are still easy to understand ($\tilde{T}_{yr}(0) = 1$ and $\tilde{T}_d(0) = 0$ because of "I").

Design example 2: stage 1 (contd)



Let's choose

$$k_p = e^{-a}(2-a) \quad \text{and} \quad \tau_i = 2-a,$$

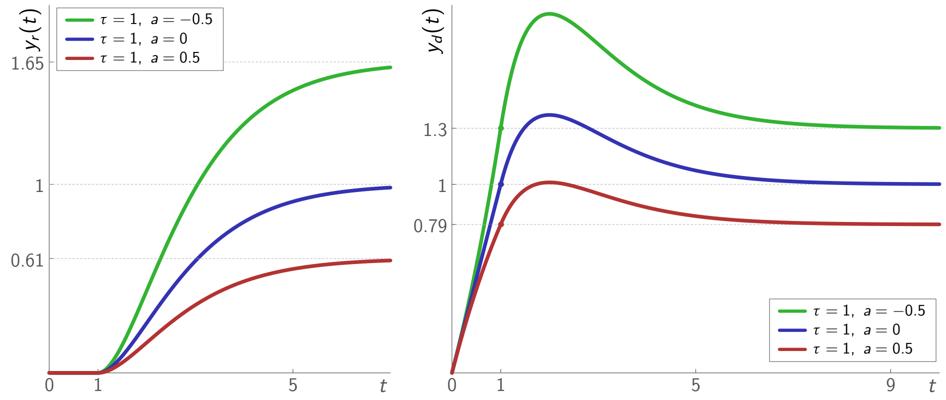
for which

$$\tilde{T}_{yr}(s) = \frac{1}{(s+1)^2} \quad \text{and} \quad \tilde{T}_d(s) = \frac{e^a s}{(s+1)^2}$$

(if $a = 2$ we end up with the I controller $\tilde{C}(s) = 1/s$, otherwise $\tilde{C}(s)$ is PI).

Design example 2: stage 2

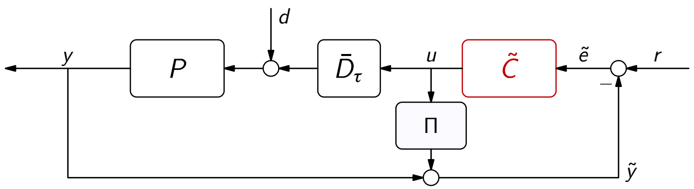
Closed-loop step responses, now converging:



- reference responses are not as expected ($\lim_{t\to\infty} y_r(t) = e^{-a} \neq 1$)
- disturbance responses are not as expected ($\lim_{t\to\infty} y_d(t) = \frac{1-e^{-a}}{a} \neq 0$)

What's wrong now?

MSP: integral action in the controller



Let $\tilde{C}(s) = \frac{1}{s}\tilde{C}_0(s)$ for some $\tilde{C}_0(s)$ such that $|\tilde{C}_0(0)| < \infty$. Then

$$C(s) = \frac{\tilde{C}(s)}{1+\Pi(s)\tilde{C}(s)} = \frac{\tilde{C}_0(s)}{s+\Pi(s)\tilde{C}_0(s)}$$

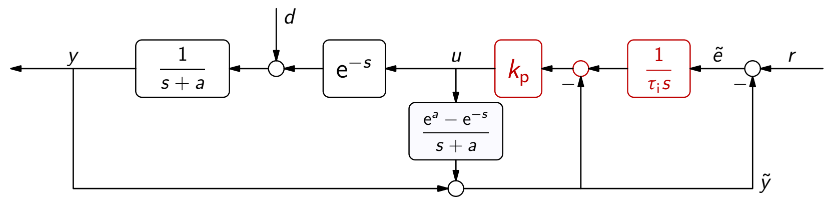
and there is an

- integrator in $C(s) \iff \lim_{s\to 0}\Pi(s) = 0$

(i.e. the predictor part has zero static gain). As a general rule, again,

- design of $\tilde{C}(s)$ is transparent at frequencies where the predictor gain is low, i.e. $|\Pi(j\omega)| \ll 1$.

Design example 2: controller static gain



We have that

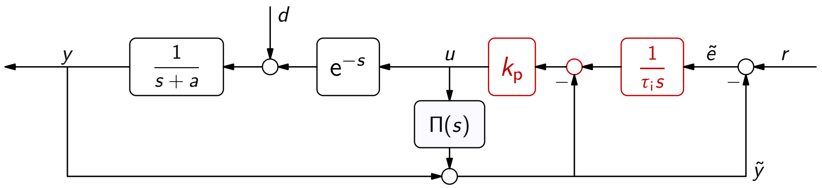
$$\lim_{s\to 0}\Pi(s) = \lim_{s\to 0}\frac{e^a - e^{-s}}{s+a} = \frac{e^a-1}{a} \implies \lim_{s\to 0} C(s) = \frac{a}{e^a-1}.$$

Because the static gain of the plant is $1/a$, we have that

$$T_d(0) = \frac{1/a}{1+1/(e^a-1)} = \frac{1-e^{-a}}{a} \neq 0$$

(monotonically decreasing function of a, in agreement with simulations).

MSP: design example 3



Consider a more general

$$\tilde{P}(s) = \frac{\alpha_1 s + \alpha_0}{s+a} \quad \Longrightarrow \quad \Pi(s) = \frac{\alpha_1 s + \alpha_0 - e^{-s}}{s+a}$$

and try to impose the following constraints:

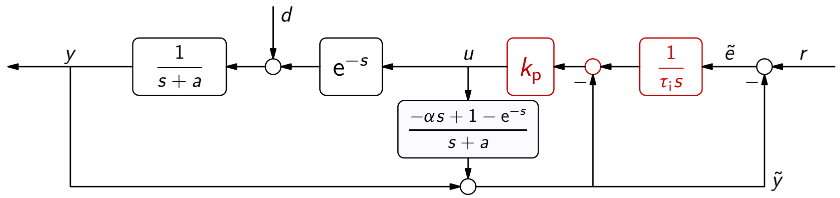
1. $|\Pi(-a)| < \infty$ if a is small enough (say $a \le 3$)
   (to prevent canceling the problematic—unstable or slow stable—pole of the plant)
2. $\Pi(0) = 0$
   (to keep integral action in the PI $\tilde{C}(s)$)

MSP: design example 3 (contd)

These conditions yield (mind that $\lim_{s\to 0}(1 - e^{-s})/s = 1 < \infty$)

$$\begin{cases} \alpha_0 - \alpha_1 a = e^a \\ \alpha_0 = 1 \end{cases} \iff (\alpha_0 = 1) \wedge \left(\alpha_1 = \begin{cases} \frac{1-e^a}{a} & \text{if } a \le 3 \\ 0 & \text{otherwise} \end{cases}\right)$$

Thus, considering only the nontrivial case of $a \le 3$, we have

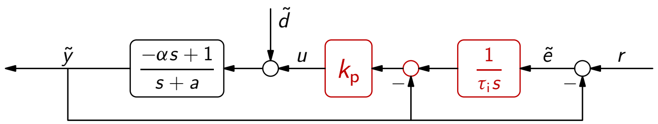


and end up with the (bi-proper and nonminimum-phase)

$$\tilde{P}(s) = \frac{-\alpha s + 1}{s+a},$$

where $\alpha := \frac{e^a - 1}{a} \in (0, 6.36)$. For larger a's this parameter α grows rapidly, which is numerically inconvenient . . .

Design example 3: stage 1



The characteristic polynomial,

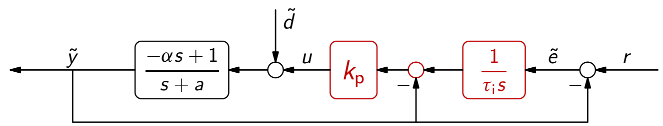
$$\begin{aligned} \tilde{\chi}_{cl}(s) &= k_p(-\alpha s + 1)(\tau_i s + 1) + \tau_i s(s+a) \\ &= \tau_i(1 - \alpha k_p)s^2 + (\tau_i(a + k_p) - \alpha k_p)s + k_p \end{aligned}$$

and the closed-loop transfer functions,

$$\begin{bmatrix} \tilde{T}_{yr}(s) \\ \tilde{T}_d(s) \end{bmatrix} = \begin{bmatrix} k_p \\ \tau_i s \end{bmatrix} \frac{-\alpha s + 1}{\tau_i(1 - \alpha k_p)s^2 + (\tau_i(a + k_p) - \alpha k_p)s + k_p}.$$

are still second-order. Mind that now $\tilde{T}_d(s)$ is bi-proper and both $\tilde{T}_{yr}(s)$ and $\tilde{T}_d(s)$ have a RHP zero, which might be misleading (the responses of the original system are inertial and should not normally exhibit undershoot).

Design example 3: stage 1 (contd)



Let's choose

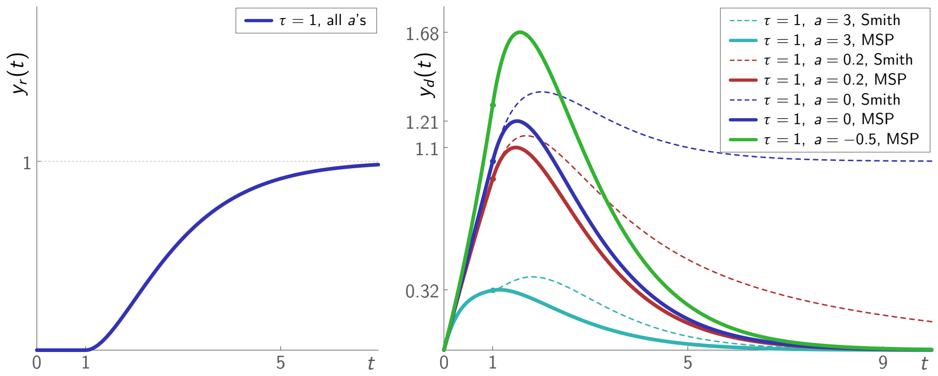
$$k_p = \frac{a(e^a - (a-1)^2)}{(e^a + a - 1)^2} > 0.099 \quad \text{and} \quad \tau_i = \frac{1 - e^{-a}(a-1)^2}{a} > 0.267,$$

(with $\lim_{a\to 0} k_p = \frac{3}{4}$ and $\lim_{a\to 0} \tau_i = 3$) for which

$$\tilde{T}_{yr}(s) = \frac{-\alpha s + 1}{(s+1)^2} \quad \text{and} \quad \tilde{T}_d(s) = \frac{s(-\alpha s + 1)}{(s+1)^2}.$$

Design example 3: stage 2

Closed-loop step responses now have no remainings of plant dynamics:



As a matter of fact, in this case $T_{yr}(s) = \frac{e^{-s}}{(s+1)^2}$ and

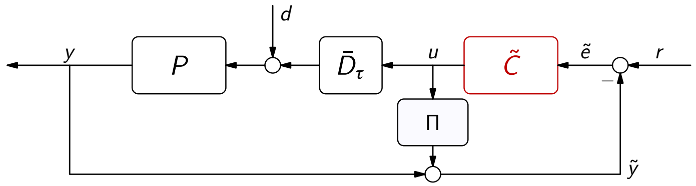
$$T_d(s) = \frac{e^{-a}}{a} s \left(\frac{1 - e^{-s}}{s} - \frac{e^a - e^{-s}}{s + a} + \frac{(e^a - 1)s + 2e^a + a - 2}{(s+1)^2} e^{-s} \right),$$

so both $s = 0$ and $s = -a$ are removable singularities (not poles) of $T_d(s)$.

MSP: loop transfer function

Loops can be analyzed in terms of their return difference transfer functions:

- at frequencies where $|1 + L(j\omega)| \gg 1$, the loop gain is high;
- at frequencies where $|1 + L(j\omega)| \approx 1$, the loop gain is low;
- at frequencies where $|1 + L(j\omega)| \ll 1$, it is close to the critical point.



The MSP return difference for the actual loop $L(s) = P(s)e^{-\tau s}C(s)$ is

$$1 + L(s) = 1 + \frac{P(s)e^{-\tau s}\tilde{C}(s)}{1 + \Pi(s)\tilde{C}(s)} = \frac{1 + \tilde{P}(s)\tilde{C}(s)}{1 + \Pi(s)\tilde{C}(s)} = \frac{1 + \tilde{L}(s)}{1 + \Pi(s)\tilde{C}(s)},$$

where $\tilde{L}(s) = \tilde{P}(s)\tilde{C}(s)$ is the designed loop in stage 1.

MSP: loop transfer function (contd)

Thus, the relation between the designed and actual loops, can be seen in

$$|1 + L(j\omega)| = |S_C(j\omega)|\,|1 + \tilde{L}(j\omega)|.$$

where $S_C(s) := \frac{1}{1+\Pi(s)\tilde{C}(s)}$ is the sensitivity function of the internal loop of the overall controller $C(s)$. Thus,

$|S_C(j\omega)| \approx 1 \implies$ transparent design

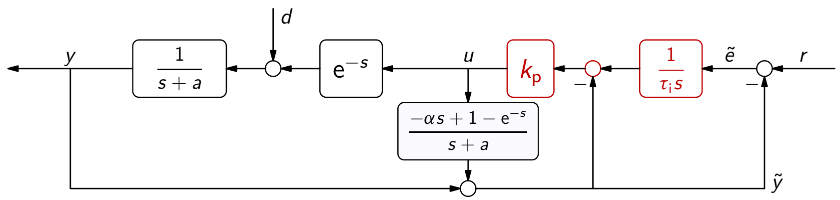
$|S_C(j\omega)| \ll 1 \implies$ poor $L(s)$, even from good $\tilde{L}(s)$

$|S_C(j\omega)| \gg 1 \implies$ possibly good L, even from poor $\tilde{L}$, but might yield
- high-gain L when low-gain $\tilde{L}$ is designed
- fragile implementation of the controller internal loop

Note that $|S_C(j\omega)| \approx 1 \iff |\Pi(j\omega)\tilde{C}(j\omega)| \ll 1$, so it may make sense to
- keep $|\Pi(j\omega)|$ small at frequencies of interest,

which actually implies that $\tilde{P}(j\omega)$ should approximate $P(j\omega)e^{-j\tau\omega}$ there...

Design example 3: loop transfer function



Here

$$S_C(s) = \frac{a(\alpha + 1)^2}{e^a a + e^a \frac{1 - e^{-s}}{s} - (a-1)^2 \frac{e^a - e^{-s}}{s+a}},$$

where $\alpha = \frac{e^a - 1}{a}$. It verifies $|S_C(j\omega)| > 1$ for all a and all ω:

some plots should be here, perhaps...