Control Systems I

Lecture 10: System Specifications

Readings: Guzzella, Chapter 10

Emilio Frazzoli

Institute for Dynamic Systems and Control
D-MAVT
ETH Zürich

November 24, 2017

Tentative schedule

#	Date	Topic
1	Sept. 22	Introduction, Signals and Systems
2	Sept. 29	Modeling, Linearization
3	Oct. 6	Analysis 1: Time response, Stability
4	Oct. 13	Analysis 2: Diagonalization, Modal coordinates.
5	Oct. 20	Transfer functions 1: Definition and properties
6	Oct. 27	Transfer functions 2: Poles and Zeros
7	Nov. 3	Analysis of feedback systems: internal stability, root locus
8	Nov. 10	Frequency response
9	Nov. 17	Analysis of feedback systems 2: the Nyquist condition
10	Nov. 24	Specifications for feedback systems
11	Dec. 1	Loop Shaping
12	Dec. 8	PID control
13	Dec. 15	Implementation issues
14	Dec. 22	Robustness

The Nyquist condition and robustness margins

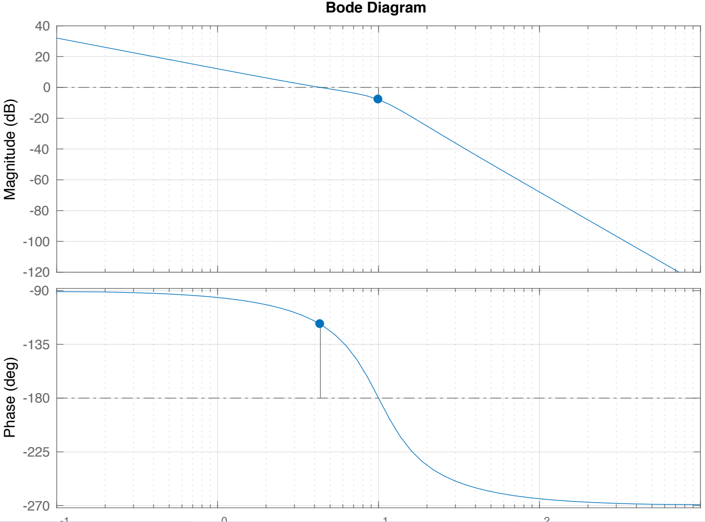
- If the open-loop is stable, then we know that in order for the closed-loop to be stable the Nyquist plot of $L(s)$ should **NOT** encircle the -1 point.
- The gain margin and phase margin measure how close the system is to closed-loop instability.

The Nyquist condition on Bode plots

- If the open-loop is stable, then we know that in order for the closed-loop to be stable the Nyquist plot of $L(s)$ should **NOT** encircle the -1 point.
- In other words, $|L(j\omega)| < 1$ whenever $\angle L(j\omega) = 180^\circ$.
- On the Bode plot, this means that the magnitude plot should be below the 0 dB line if/when the phase plot crosses the -180° line.
- Remember that this condition is valid only if the open loop is stable. In all other cases (including non-minimum phase zeros) it is strongly recommended to double check any conclusion on closed-loop stability using other methods (Nyquist, root locus).

Gain and Phase Margin

The “distance” from the Nyquist plot to the -1 point is a measure of robustness. On the bode plot, it is easy to measure this distance in terms of **gain** and **phase** margin.



Summary of the previous lecture

In the previous lecture, we learned:

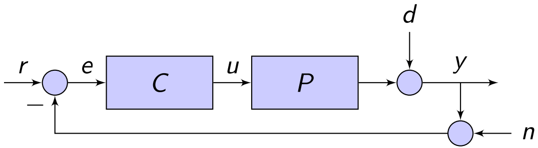
- How to sketch a polar plot (and hence a Nyquist plot), based on Bode plots
- The Nyquist condition to determine closed-loop stability using a Nyquist plot.
- How to check the Nyquist condition on a Bode plot.
- How to quickly assess the "robustness" of a feedback control system.
- Now we have three graphical methods to study closed-loop stability given the (open-)loop transfer function.
    1. **Root locus:** always correct if applicable (assumes finite-dimensional system)
    2. **Nyquist:** always correct, always appplicable;
    3. **Bode:** very useful for control system design, however may be misleading in determining closed-loop stability (e.g., for open-loop unstable systems).
- So far we have only looked at analysis issues, i.e., how to determine closed-loop stability; from now on we will concentrate on control synthesis, i.e., how to design a feedback control system that makes a system behave as desired.

Plan for this lecture

- We have seen how to establish whether the closed-loop system will be stable or not, based on the (open-)loop transfer function.
- The next step is to evaluate how well the closed-loop will behave, e.g.,
  - how quickly and how closely the closed-loop system can track commands,
  - how well it can reject disturbances, modeling errors, and noise.
- Specifications on the closed-loop behavior are typically given using two main paradigms, plus one that can be seen both ways:
  - Steady-state error
  - Time-domain specifications
  - Frequency-domain specifications
- In this lecture we will learn about both kinds of specifications, and what they mean in terms of the closed-loop system and/or the (open-)loop transfer function.

Proportional feedback

Steady-state error to step inputs



- Recall the following transfer function (and that $L(s) = P(s)C(s)$):

$$r \to e, d \to e : S(s) = \frac{1}{1 + L(s)}$$

- If the input is a unit step, i.e., $r(t) = 1 = e^{0t}$ for $t \geq 0$, the steady-state output will be

$$e_{ss}(t) = S(0)e^{0t} = \frac{1}{1 + L(0)}, \qquad \text{for } t \geq 0.$$

- So if $\lim_{s \to 0} L(s) = k_{Bode} = k_{DC}$ is finite then the steady-state error to a unit step is $1/(1 + k_{DC})$.
- If the limit is infinite (i.e., there is a pole at $s = 0$, i.e., $L(s)$ contains an integrator), the steady-state error is zero.

Steady-state error to higher-order ramps

- More in general, one may ask the closed-loop system to have a finite steady-state error to a unit ramp of order $m = \{0, 1, 2, \ldots\}$, i.e.,

$$r(t) = \frac{1}{m!} t^m, \qquad t \geq 0.$$

- The steady-state error can be computed as

$$e_{ss}(t) = \lim_{s \to 0} \left(\frac{1}{1 + L(s)} \cdot \frac{e^{st}}{s^m} \right)$$

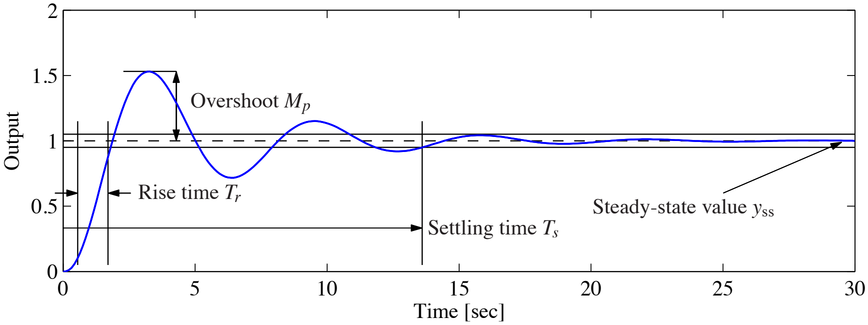
System type

- The result depends on the type of the system, i.e., the number of integrators in $L(s)$, i.e., the number of poles of $L(s)$ at $s = 0$, and is summarized as follows

e_{ss}	$m = 0$	$m = 1$	$m = 2$
Type 0	$\frac{1}{1 + k_{Bode}}$	∞	∞
Type 1	0	$\frac{1}{k_{Bode}}$	∞
Type 2	0	0	$\frac{1}{k_{Bode}}$

- In plain English:
  - The steady-state error is smaller as the (Bode) gain increases;
  - A requirement to have zero steady-state errors to ramps of order m, is the same as requiring to have at least $m + 1$ integrators on the path from the error e to the (reference/disturbanc) input.

Time domain: step response of a 2nd order system



- Time domain specifications are usually given in terms of the step response of a 2nd order system:

$$G(s) = \frac{\omega_n^2}{s^2 + 2\zeta\omega_n s + \omega_n^2} \Rightarrow y(t) = \cdot(1 - e^{-\sigma t}\cos(\omega t)).$$

- Recall that the poles are at $s = -\sigma \pm j\omega$, and that $\omega_n^2 = \sigma^2 + \omega^2$, $\zeta = \sigma/\omega_n$.

Second-order response: shape vs. time scale

Analytical approximations

- Rise time depends primarily on ω_n: note that $T_{100} = \frac{\pi}{2\omega}$ (assuming $\zeta < 1$). Other formulas available, e.g., $T_{90} \approx (0.14 + 0.4\zeta)\frac{2\pi}{\omega_n}$.
- Peak time: depends on the frequency ω, $T_p \approx \frac{\pi}{\omega}$;
- % Overshoot: depends on the damping ζ:

$$\ln M_p \approx -\frac{\sigma\pi}{\omega} = -\frac{\zeta\pi}{\sqrt{1-\zeta^2}}, \qquad \zeta^2 = \frac{(\ln M_p)^2}{\pi^2 + (\ln M_p)^2}.$$

- Settling time (e.g., to 2%): depends on the real part of the poles σ,

$$T_s = \frac{-\ln 2\%}{\sigma}, \qquad \sigma = \frac{-\ln 2\%}{T_s}.$$

- Note: the settling time is the only relevant specification (in addition to the steady-state error to a unit step) for a first-order system.

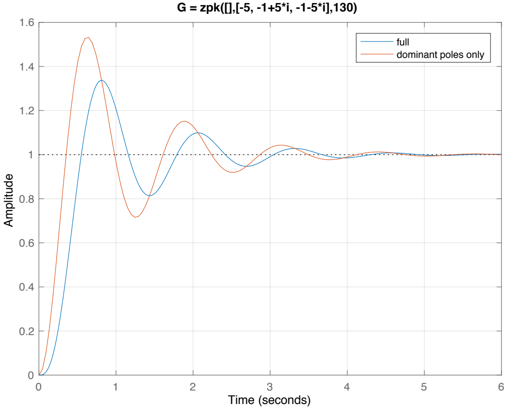
Dominant poles approximation

- What if the closed-loop system if of higher order? Often one can approximate it with a second-order (or even first-order) system, and apply the specifications to the approximation.
- The approximation is based on the concept of **dominant poles**.
  - Dominant poles are typically those with the largest real part (i.e., the slowest decay rate);
  - Exceptions are made when the poles with the largest real part also have very small residues (typically because of nearby zeros).

$$G(s) = \frac{r_1}{s - p_1} + \frac{r_2}{s - p_2} + \ldots \quad \Leftrightarrow \quad g(t) = r_1 e^{p_1 t} + r_2 e^{p_2 t} + \ldots$$

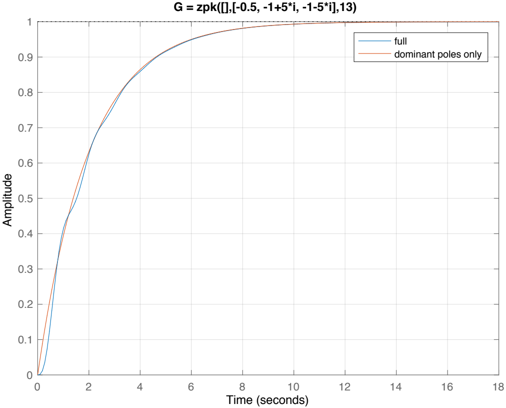
Dominant poles approximation examples

- Consider a third-order system with $G(s) = \dfrac{130}{(s+5)(s+1+5j)(s+1-5j)}$.
- The contribution to the response of the pole at $s = -5$ will decay as e^{-5t}, while that of the poles at $s = -1 \pm 5j$ will decay as e^{-t}.
- Dominant pole approximation: $G_{dom}(s) = \dfrac{26}{(s+1+5j)(s+1-5j)}$.



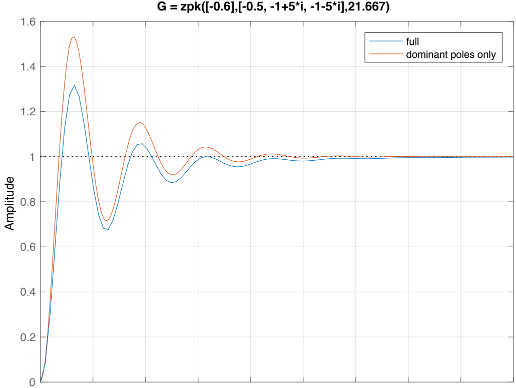
Dominant poles approximation examples

- Consider a third-order system with $G(s) = \dfrac{13}{(s+0.5)(s+1+5j)(s+1-5j)}$.
- The contribution to the response of the pole at $s = -0.5$ will decay as $e^{-0.5t}$, while that of the poles at $s = -1 \pm 5j$ will decay as e^{-t}.
- Dominant pole approximation: $G_{dom}(s) = \dfrac{0.5}{s+0.5}$.



Dominant poles approximation examples

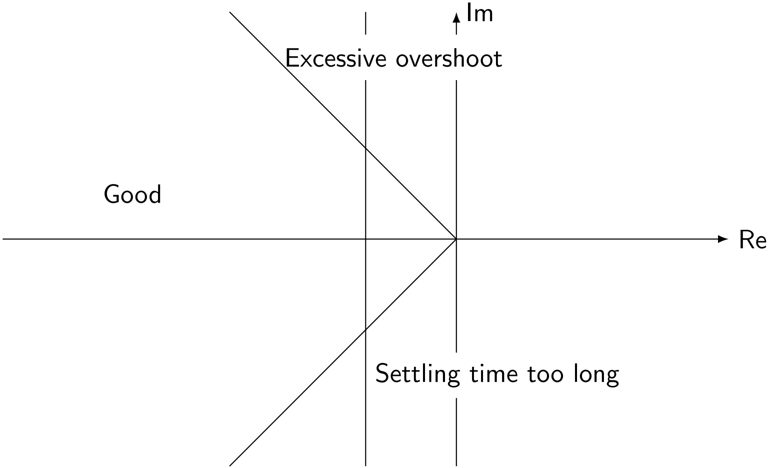
- Consider a third-order system with $G(s) = \dfrac{21.667(s+0.6)}{(s+0.5)(s+1+5j)(s+1-5j)}$.
- The zero at $s=-1$ makes the magnitude of the residue of the pole at $s=-0.5$ small wrt to the magnitudes of the residues of the other poles $\Rightarrow$ effectively we have a near-pole/zero cancellation.
- Dominant pole approximation: $G_{dom}(s) = \dfrac{26}{(s+1+5j)(s+1-5j)}$.



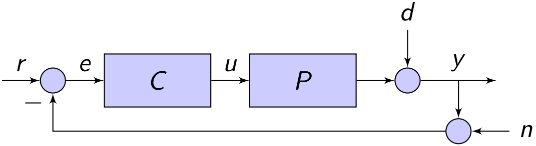
Using time-domain specifications

- Time-domain specifications impose constraints on the locations of the **dominant closed-loop poles.**
- In particular:
  - Settling time constraints require the real part of the dominant poles to be to the less than some maximum value.
  - Peak overshoot constraints require the damping ratio to be more than some minimum value—resulting in a sector of the left half plane.
  - Rise time constraints require that the distance of the dominant poles from the origin, and/or (the absolute value of) the imaginary part of the dominant poles to be more than some minimum value.
- When checking time-domain specifications, it is particularly convenient to use the root locus.

Time-domain specifications in the complex plane



Command tracking/disturbance rejection vs. noise rejection



- Recall: the sensitivity function $S(s) = \dfrac{1}{1 + L(s)}$ is the transfer function from the reference input r or the output disturbance d to the error e. Hence, if we want small errors to reference inputs, and want to reject disturbances, we need $S(s)$ to be "small."
- The transfer function from the noise input n to the output y is the complementary sensitivity $T(s) = \dfrac{L(s)}{1 + L(s)}$. If we do not want the effect of the noise to be observed at the output, then we need $T(s)$ to be "small."
- So if we want to reject both disturbances and noise, and want to follow commands well, we need both $T(s)$ and $S(s)$ to be "small". But $T(s) + S(s) = 1$, so they cannot be "small" at the same time!

Frequency-domain specifications

- The main point in using frequency domain specifications is exactly to handle such requirements.
- Typically commands and disturbances act at "low" frequency, e.g., no more than 10 Hz.
- Noise is typically a high-frequency phenomenon, e.g., more than 100 Hz.
- So we can reconcile both command tracking/disturbance rejection AND noise rejection by separating them frequency-wise!
  - Make $|S(j\omega)| << 1$ (hence $|T(j\omega)| \approx 1$ at low frequencies.
    e.g., "ensure that commands are tracked with max 10% error up to a frequency of 10Hz.
  - Make $|T(j\omega)| << 1$ at high frequencies.
    e.g., "ensure that noise is reduced by a factor of 10 at the output at frequencies higher than 100 Hz.

Frequency-domain specifications on the Bode plot

- Frequency-domain specifications are usually expressed in terms of closed-loop frequency response.
- Can we write them in terms of the open loop frequency response? Indeed we can.
- Remember that for good command tracking / disturbance rejection, we want $|S(j\omega)| = |1 + L(j\omega)|^{-1}$ to be small at low frequencies, i.e., we want $|L(j\omega)|$ to be large at low frequencies.
- Typically this is written as $|S(j\omega)| \cdot |W_1(j\omega)| < 1$ for some function $|W_1(j\omega)|$ that is large at low frequency. This translates to $|1 + L(j\omega)| > |W_1(j\omega)|$, which is approximated as

$$|L(j\omega)| > |W_1(j\omega)|.$$

- This can be seen as a "low frequency obstacle" on the magnitude Bode plot.

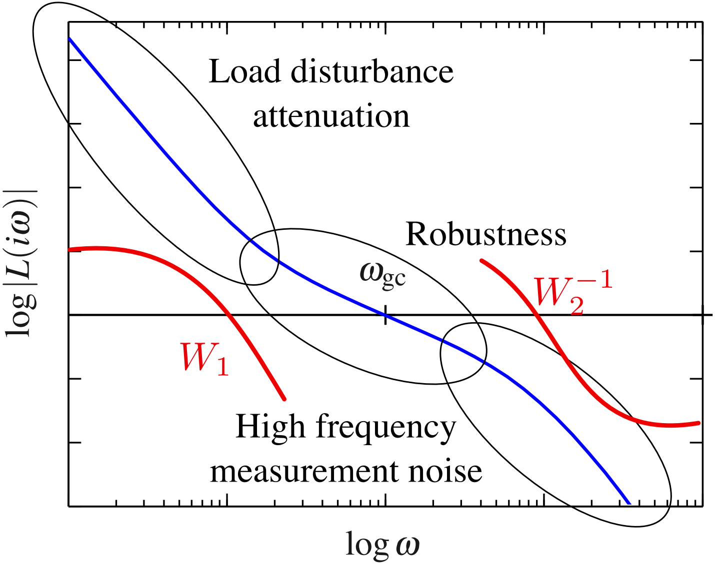
Frequency-domain specifications on the Bode plot

- For good noise rejection, we want $|T(j\omega)| = |L(j\omega)|/|1 + L(j\omega)|$ to be small at high frequencies.
- If $|T(j\omega)|$ is small, then $|L(j\omega)|$ has to be small, and $|T(j\omega)| \approx |L(j\omega)|$ (at high frequencies).
- Typically this is written as $|T(j\omega)| \cdot |W_2(j\omega)| < 1$ for some function $|W_2(j\omega)|$ that is large at high frequency. This translates to $|L(j\omega)| < |W_2(j\omega)|^{-1}$.
- This can be seen as a "high-frequency obstacle" on the Bode plot.

Closed-loop bandwidth and (open-loop) crossover

- The bandwidth of the closed-loop system is defined as the maximum frequency ω for which $|T(j\omega)| > 1/\sqrt{2}$, i.e., the output can track the commands to within a factor of ≈ 0.7.
- Let ω_c be the crossover frequency, such that $|L(j\omega_c)| = 1$. If we assume that the phase margin is about 90°, then $L(j\omega_c) = -j$, and $T(j\omega_c) = 1/\sqrt{2}$.
- In other words, the (open-loop) crossover frequency is approximately equal to the bandwidth of the closed-loop system.

Bode-plot "obstacle course"



Summary of the lecture

We discussed how we can write specifications for the closed-loop behavior in terms of:

- **Steady-state error**: this has consequences on the necessary DC gain, and/or the number of integrators in the loop transfer function (type of the system);
- **Time-domain specifications**: these describe how the closed-loop system, approximated as a second-order "dominant" system, should behave in time. These specifications are relatively easy to formulate and understand, but do not lend themselves directly to easy design—best interpreted on the complex plane/root locus. Matlab can help a lot in meeting these specifications.
- **Frequency-domain specifications**: these give a technically sophisticated, powerful set of specifications in terms of ability of the system to follow commands, and to reject disturbances and noise. These specifications apply directly to powerful methods for control design—the Bode plot "obstacle course".

In the next lecture we will finally start seeing how we can design a control system that satisfy all the requirements (closed-loop stability + all the specs we covered in this lecture).