DIGITAL CONTROLLER DESIGN

5.1: Direct digital design: Steady-state accuracy

- We have spent quite a bit of time discussing digital hybrid system analysis, and some time on controller design via emulation.
- We now look at "direct digital design."

Specifications:

- Steady-state accuracy,
- Transient response
- Absolute/ relative stability,
- Sensitivity,
- Disturbance rejection,
- Control effort.

Steady-state accuracy

- How well does a control system track step/ramp/... inputs?
- General formulation: $Y(z) = T(z)R(z)$.
- The error is: $e[k] = r[k] - y[k]$. In Laplace domain:

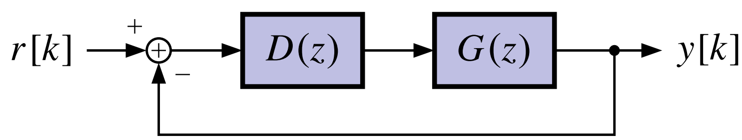
$$\begin{aligned} E(z) &= R(z) - T(z)R(z) \\ &= [1 - T(z)]R(z). \end{aligned}$$

- Use the final-value theorem to find e_{ss}.

$$e_{ss} = \lim_{z \to 1} (z-1)[1 - T(z)]R(z).$$

<u>*Unity-feedback systems:*</u>

- ***If*** the system is of the form (unity-feedback)



- Then,

$$T(z) = \frac{D(z)G(z)}{1 + D(z)G(z)}$$

$$1 - T(z) = \frac{1}{1 + D(z)G(z)}.$$

- Let

$$D(z)G(z) = \frac{K \prod_{i=1}^{m}(z - z_i)}{(z-1)^N \prod_{i=1}^{p}(z - p_i)}, \quad z_i \neq 1,\ p_i \neq 1.$$

- Also, define the "Bode Gain"

$$K_{dc} = K \left. \frac{\prod_{i=1}^{m}(z - z_i)}{\prod_{i=1}^{p}(z - p_i)} \right|_{z=1}$$

  which is the dc-gain with all poles at $z = 1$ removed.

- Consider a step input.

$$e_{ss} = \lim_{z \to 1} (z-1) \left[\frac{1}{1 + D(z)G(z)} \right] \frac{z}{z-1}$$

$$= \lim_{z \to 1} \frac{1}{1 + D(z)G(z)}.$$

  - If $K_p = \lim_{z \to 1} D(z)G(z)$ then $e_{ss} = \dfrac{1}{1 + K_p}$.

- If $N = 0$ then $e_{ss} = \dfrac{1}{1 + K_{dc}}$.
- If $N > 0$ then $e_{ss} = 0$.

■ Now, consider a ramp input.

$$\begin{aligned} e_{ss} &= \lim_{z\to 1}(z-1)\left[\frac{1}{1+D(z)G(z)}\right]\frac{Tz}{(z-1)^2} \\ &= \lim_{z\to 1}\frac{T}{(z-1)+(z-1)D(z)G(z)}. \end{aligned}$$

- If $K_v = \lim_{z\to 1}\dfrac{1}{T}(z-1)D(z)G(z)$ then $e_{ss} = \dfrac{1}{K_v} = \dfrac{T}{K_{dc}}$.
- If $N = 1$ then $e_{ss} = \dfrac{T}{K_{dc}}$.
- If $N > 1$ then $e_{ss} = 0$ (and so forth).

General unity-feedback result

■ Poles and large gains at $z = 1$ of $D(z)G(z)$ decrease e_{ss} but also decrease stability.

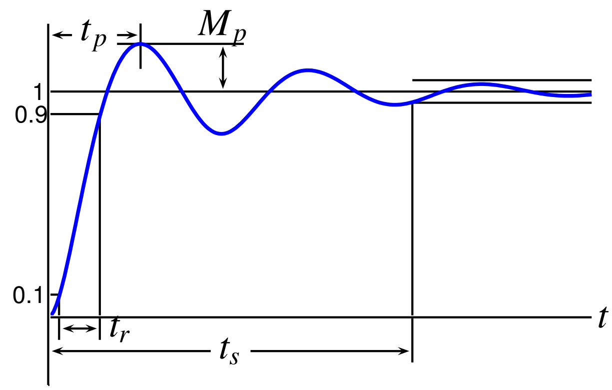
■ System design is a tradeoff between steady-state accuracy, relative stability, and complexity.

5.2: Direct digital design: Other requirements

Transient response

- If the system is dominated by second-order poles near $z = 1$, then we can determine pole locations for transient-response (step-response) specifications.

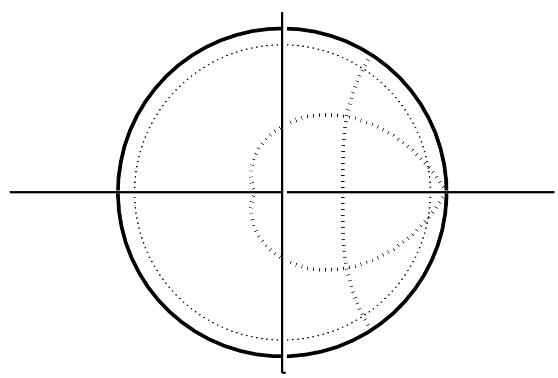
$$\omega_n \geq 1.8/t_r$$
$$\sigma \geq 4.6/t_s$$
$$M_p \approx e^{-\pi\zeta/\sqrt{1-\zeta^2}}.$$



- To convert these specifications to the z-plane,
  - $r = e^{-\sigma T}$.
  - Use `zgrid` to plot locus of constant ζ and ω_n.
  - Mark regions of acceptable poles on plot.

EXAMPLE: Plot the specifications for

- $M_p \leq 16\% \quad \longrightarrow \zeta \geq 0.5.$
- $t_s < 10$ s $\quad \longrightarrow \sigma \geq 0.5.$
- $t_r \approx 1.8$ s $\quad \longrightarrow \omega_n \approx 1.$
- $T = 0.2$ s.
- Then, $r = e^{-\sigma T} = e^{-0.1} = 0.9.$



- Our three boundaries are plotted. What is the region of acceptable closed-loop poles?

- *For dominant second-order systems*, transient-response design can also be done using frequency-response information.
- M_p (percent overshoot) and M_r (resonant peak) are related through ζ:

$$M_p = 1 + e^{-\zeta/\sqrt{1-\zeta^2}}$$
$$M_r = \frac{1}{2\zeta\sqrt{1-\zeta^2}}$$

  - Use $M_r < 2$ dB.
- Settling time,

$$t_s \approx \frac{4}{\omega_b \zeta}\sqrt{(1-2\zeta^2)+\sqrt{4\zeta^4-4\zeta^2+2}},$$

where ω_b is the closed-loop bandwidth.

- Also, $t_r\omega_b \approx 2$.

Relative stability

- Want $GM \gg 1$ and $PM \gg 0$.
- $PM \approx 100\zeta$ (especially for second-order systems).
- Only sure way to measure GM and PM is Nyquist plot.

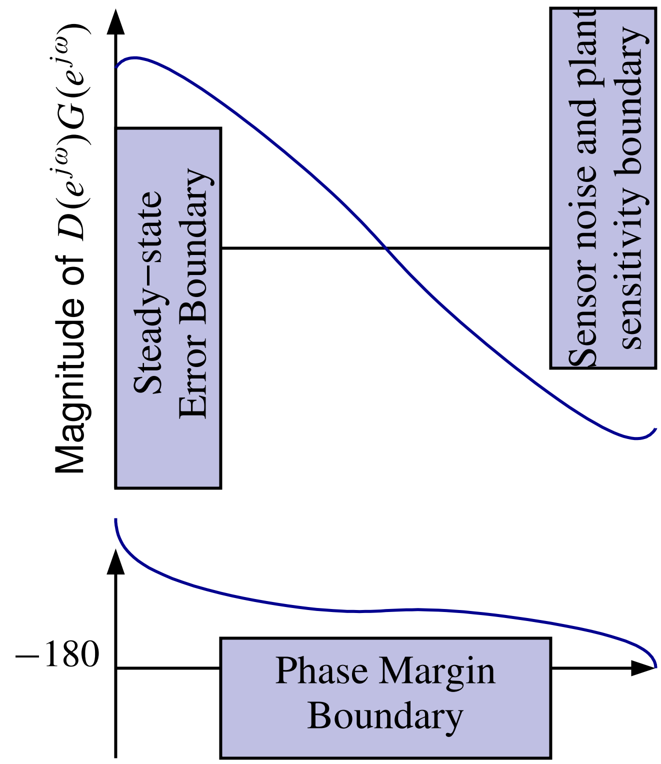
Sensitivity and disturbance rejection

- If we define $S(z) = 1 - T(z)$ then

$$\begin{aligned} Y(z) &= T(z)R(z) + S(z)W(z) + T(z)V(z) \\ &= [1 - S(z)]R(z) + S(z)W(z) + [1 - S(z)]V(z) \end{aligned}$$

for a unity-feedback system.

- Want sensitivity small for good disturbance rejection and tracking, but large for sensor-noise rejection and robustness.
- This typically places constraints on the “loop gain” $L(z) = D(z)\overline{GH}(z)$.



Control effort

- There are always physical constraints on control effort.

$$|u(t)| < \# \qquad \text{saturation limits}$$

$$\int_0^{t_f} |u(\tau)|\,\mathrm{d}\tau < \# \qquad \text{finite total resources}$$

$$|u^2(t)| < \# \qquad \text{finite power}$$

- Designing controllers with constrained control effort can be difficult. Subject of “optimal control.”

5.3: Phase-lag compensation

- Compensators/controllers come in many varieties.
  - Some common types are phase-lead, phase-lag, and PID.
- Root-locus and Bode techniques can be used to design compensators.
- We mostly consider compensation using a first-order system:

$$D(z) = K_d \frac{(z - z_0)}{(z - z_p)}.$$

- We first consider using Bode methods to design our compensator, so we need to convert $D(z) \to D(w)$.

$$D(w) = D(z)\big|_{z=\frac{1+(T/2)w}{1-(T/2)w}},$$

which is also first-order, and has transfer function

$$D(w) = a_0 \left[\frac{1 + w/\omega_{w_0}}{1 + w/\omega_{w_p}} \right],$$

where ω_{w_0} = zero location and ω_{w_p} = pole location in the w-plane; a_0 is the dc-gain.

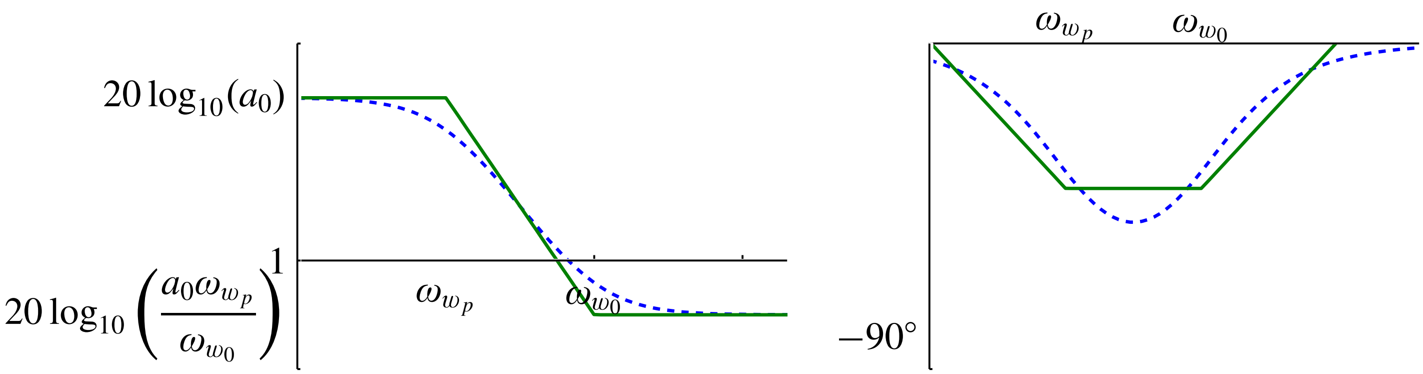
- We will eventually need to convert the compensator back to $D(z)$. The correspondences are

$$K_d = a_0 \left[\frac{\omega_{w_p}(\omega_{w_0} + 2/T)}{\omega_{w_0}(\omega_{w_p} + 2/T)} \right] \qquad z_0 = \frac{2/T - \omega_{w_0}}{2/T + \omega_{w_0}} \qquad z_p = \frac{2/T - \omega_{w_p}}{2/T + \omega_{w_p}}.$$

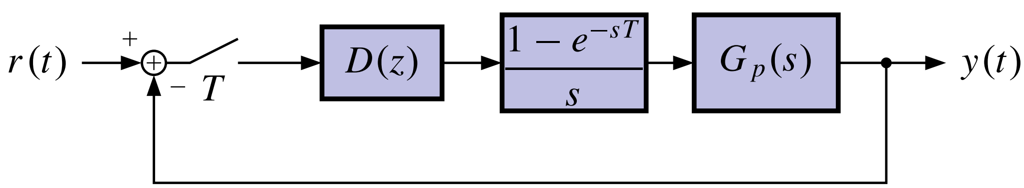
- If $|\omega_{w_0}| < |\omega_{w_p}|$ then compensator is a "lead" compensator.
- If $|\omega_{w_0}| > |\omega_{w_p}|$ then compensator is a "lag" compensator.

Phase-lag compensation

- A phase-lag compensator has its pole closer to the origin (of the w-plane) than its zero (closer to 1 in the z-plane).
- The Bode plot of a typical lag compensator is



- Low-frequency gain is a_0, high-frequency gain is $20\log_{10}\left(\dfrac{a_0\omega_{w_p}}{\omega_{w_0}}\right)$ dB.
- We consider designing a compensator for the system



- Let $G(s) = \left(\dfrac{1-e^{-sT}}{s}\right)G_p(s)$, $G(z) = \mathcal{Z}[G(s)]$, $G(w) = G(z)\big|_{z=\frac{1+(T/2)w}{1-(T/2)w}}$.
- Lag controller adds phase. Must be careful *NOT* to add phase near crossover of $G(j\omega_w)$.
- Therefore, keep both the pole and zero at low frequency.

Phase-lag design method (Bode)

- System e_{ss} specifications determine dc-gain a_0.
- Desired phase margin PM also specified.

1. Plot Bode plot of $G(w)$.
2. Determine frequency ω_{w_1} where phase of $G(w)$ is about $-180^\circ + PM + 5^\circ$. The crossover of the compensated system will occur at approximately this frequency.
3. Choose $\omega_{w_0} = 0.1\omega_{w_1}$. This ensures that little phase lag is introduced at new crossover (actually, about 5° … see above)
4. At ω_{w_1} we want $|D(\omega_{w_1})G(\omega_{w_1})| = 1$. The gain of the compensator at "high frequency" is $a_0\omega_{w_p}/\omega_{w_0}$.

$$\left|\frac{a_0\omega_{w_p}}{\omega_{w_0}}G(j\omega_{w_1})\right| = 1$$

$$\frac{a_0\omega_{w_p}}{\omega_{w_0}} = \frac{1}{|G(j\omega_{w_1})|}$$

   or

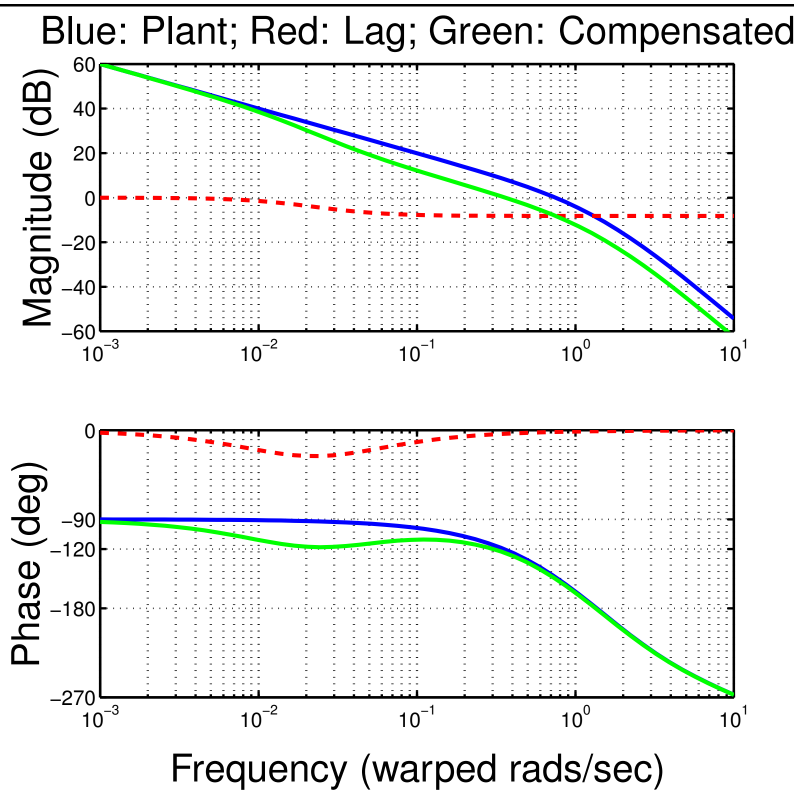
$$\omega_{w_p} = \frac{0.1\omega_{w_1}}{a_0|G(j\omega_{w_1})|}.$$

5. Design is complete since we know a_0, ω_{w_0}, and ω_{w_p}. Note that if $H(s) \neq 1$, then we replace $G(w)$ with $\overline{GH}(w)$.

EXAMPLE: Let $G_p(s) = \dfrac{1}{s(s+1)(0.5s+1)}$ and $T = 0.05$ s.

$$\begin{aligned} G(z) &= \frac{z-1}{z}\mathcal{Z}\left[\frac{1}{s^2(s+1)(0.5s+1)}\right] \\ &= \frac{z-1}{z}\left[\frac{0.005z}{(z-1)^2} - \frac{1.5z}{z-1} + \frac{2z}{z-0.9512} - \frac{0.5z}{z-0.9048}\right] \end{aligned}$$

- Plot Bode plot of $G(w)$



- Design spec gives $a_0 = 0$ dB, $PM = 55°$.
- From graph, we see that $\angle G(j\omega_w) = (-180° + 55° + 5°) = -120°$ at $\omega_{w_1} \approx 0.36$. Also, $|G(j\omega_{w_1})| \approx 2.57$.
- $\omega_{w_0} = 0.1\omega_{w_1} = 0.036$ and $\omega_{w_p} = \dfrac{0.1\omega_{w_1}}{a_0|G(j\omega_{w_1})|} = 0.014$.
- Combining the above, and converting from $D(w)$ to $D(z)$ we get

$$D(z) = \frac{0.3891z - 0.3884}{z - 0.9993}.$$

Finite-precision problem

- When implementing a phase-lag filter we may have difficulty.
- Coefficients of filter stored as binary fixed-point values.
- For example,

$$\text{Value} = \frac{b_7}{2^1} + \frac{b_6}{2^2} + \frac{b_5}{2^3} + \cdots + \frac{b_0}{2^8}$$

for an 8-bit fixed-point value.

EXAMPLE: $(0.11000001)_2 = \left(\frac{1}{2} + \frac{1}{4} + \frac{1}{256}\right)_{10} = (0.75390625)_{10}$.

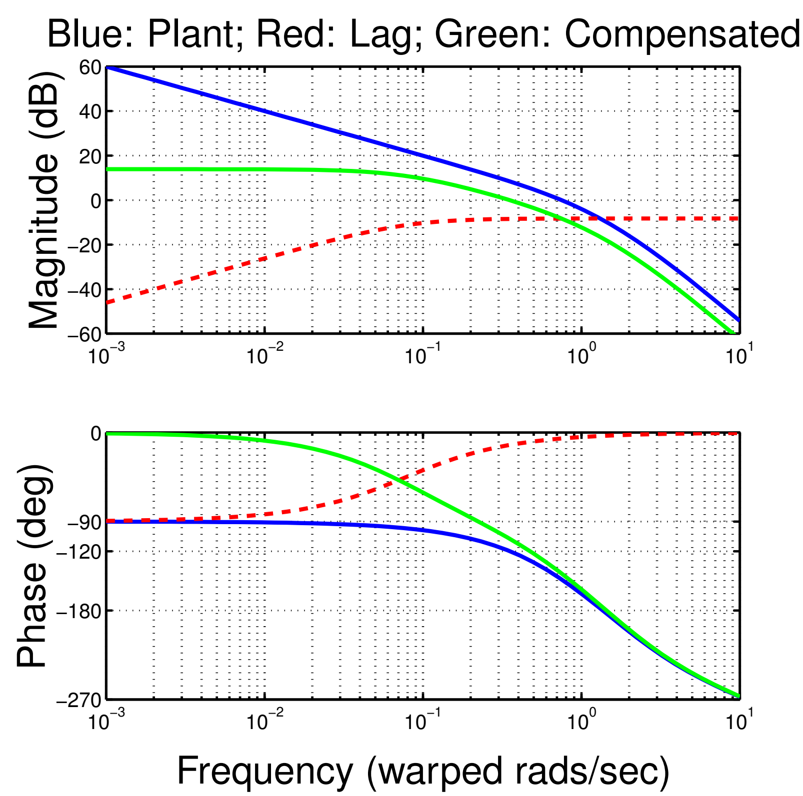
- Minimum value that can be represented is zero.
- Maximum value is $(0.11111111)_2 = (1 - 1/256)_{10} = (0.99609375)_{10}$.
- For previous example, need denominator coefficient of 0.9993 but implement 0.99609375. Need numerator coefficients

$$(0.3891)_{10} \Rightarrow (0.01100011)_2 = (0.38671875)_{10}$$

$$(0.38840)_{10} \Rightarrow (0.01100011)_2 = (0.38671875)_{10}$$

- Compensator zero is shifted to "1" causing a dc-gain of zero. We get

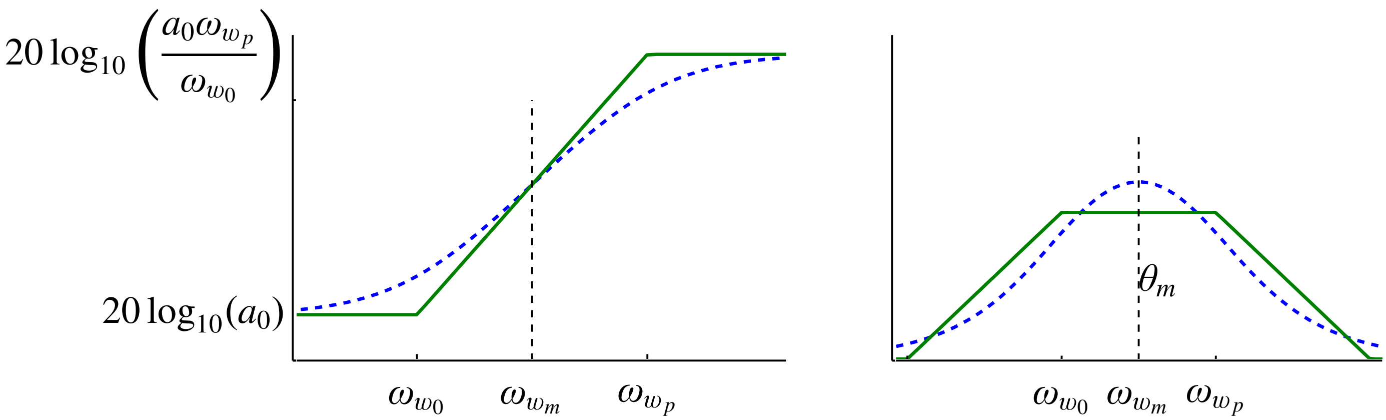
$$D(z)_{implemented} = \frac{0.38671875(z-1)}{z - 0.99609375}.$$



- Phase margin of 70° (good).
- System type reduced (bad).
- Need more bits in implementation or smarter implementation method.

5.4: Phase-lead compensation

- A phase-lead compensator has its zero closer to the origin in the w-plane than its pole (closer to 1 in the z-plane).
- Low-frequency gain of $20\log_{10} a_0$ dB.
- High-frequency gain of $20\log_{10}\dfrac{a_0\omega_{w_p}}{\omega_{w_0}}$dB. (same)



- Adds phase. Maximum phase shift occurs at

$$\omega_{w_m} = \sqrt{\omega_{w_p}\omega_{w_0}} \qquad \text{geometric mean}$$

The phase shift is

$$\theta_m = \tan^{-1}\left[\frac{1}{2}\left(\sqrt{\frac{\omega_{w_p}}{\omega_{w_0}}} - \sqrt{\frac{\omega_{w_0}}{\omega_{w_p}}}\right)\right].$$

At this location,

$$|D(j\omega_{w_m})| = a_0\sqrt{\frac{\omega_{w_p}}{\omega_{w_0}}}.$$

- Note that lead controller decreases phase near crossover.
  - Stabilizing effect, but
  - Increases high-frequency gain, ... destabilizing.
- Design tends to be trial-and-error.

Phase-lead design method (Bode)

- System e_{ss} specifications determine dc-gain a_0.
- Desired phase margin PM also specified.

1. Crossover frequency must be generated ω_{w_1} (more later). Then, spec is $D(j\omega_{w_1})G(\omega_{w_1}) = 1\angle(-180^\circ + PM)$. Gain margin not specified, but must be "adequate."
2. We see that
$$|D(j\omega_{w_1})| = \frac{1}{|G(j\omega_{w_1})|}$$
and
$$\angle D(j\omega_{w_1}) = -180^\circ + PM - \angle G(j\omega_{w_1}) \stackrel{\Delta}{=} \theta.$$
3. Can derive (Appendix I of Phillips/Nagle)
$$a_1 = \frac{1 - a_0|G(j\omega_{w_1})|\cos\theta}{\omega_{w_1}|G(j\omega_{w_1})|\sin\theta} \qquad \text{and} \qquad b_1 = \frac{\cos\theta - a_0|G(j\omega_{w_1})|}{\omega_{w_1}\sin\theta},$$
where $\omega_{w_0} = a_0/a_1$ and $\omega_{w_p} = 1/b_1$.

- Note that this design method only works when certain constraints on the value chosen for ω_{w_1} are met. First, $\theta > 0$ since this is a lead compensator. Also, system must be stable.

  I) $\theta > 0$ leads to $\angle G(j\omega_{w_1}) < -180^\circ + PM$.

  II) $|D(j\omega_{w_1})| > a_0$ for lead compensator, so $|G(j\omega_{w_1})| < 1/a_0$.

  III) b_1 must be positive for stability. $\cos\theta > a_0|G(j\omega_{w_1})|$.

- Note: If $H(s) \neq 1$ then replace $G(w)$ with $\overline{GH}(w)$ everywhere.

EXAMPLE: Revisit previous example.

- Unity dc-gain, $a_0 = 1$ and $PM \geq 55^\circ$.

- Choose ω_{w_1}: $\angle G(j\omega_{w_1}) < -125°$ by condition (I).
- $|G(j\omega_{w_1})| < 1$ by condition (II).
- Select (arbitrarily) $\omega_{w_1} = 1.2$. Then, $\angle G(j\omega_{w_1}) = -173°$, and $|G(j\omega_{w_1})| = 0.45$. Thus, conditions (I) and (II) are met.
- Verify: $\theta = -180 + 55 + 173 = 48°$. $\cos(\theta) = 0.67 > |G(j\omega_{w_1})| = 0.45$ so condition (III) is met as well.
- Therefore, our selection for ω_{w_1} is valid. Continue with design.

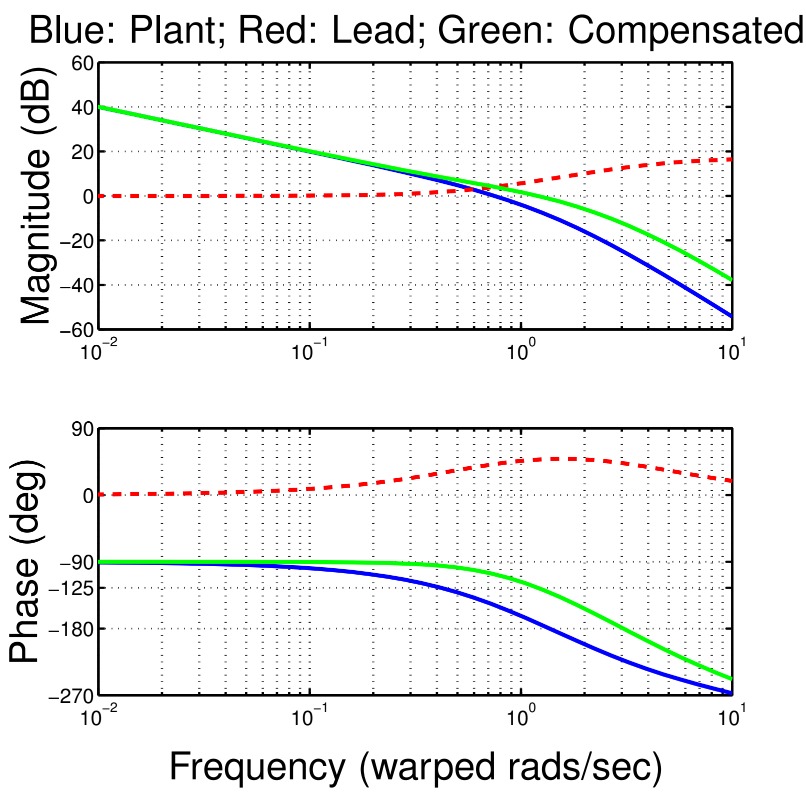
$$a_1 = \frac{1 - (1)(0.4576)(\cos(48°))}{(1.2)(0.4576)(\sin(48°))} = 1.701$$

$$b_1 = \frac{\cos(48°) - (1)(0.4576)}{(1.2)(\sin(48°))} = 0.2387.$$

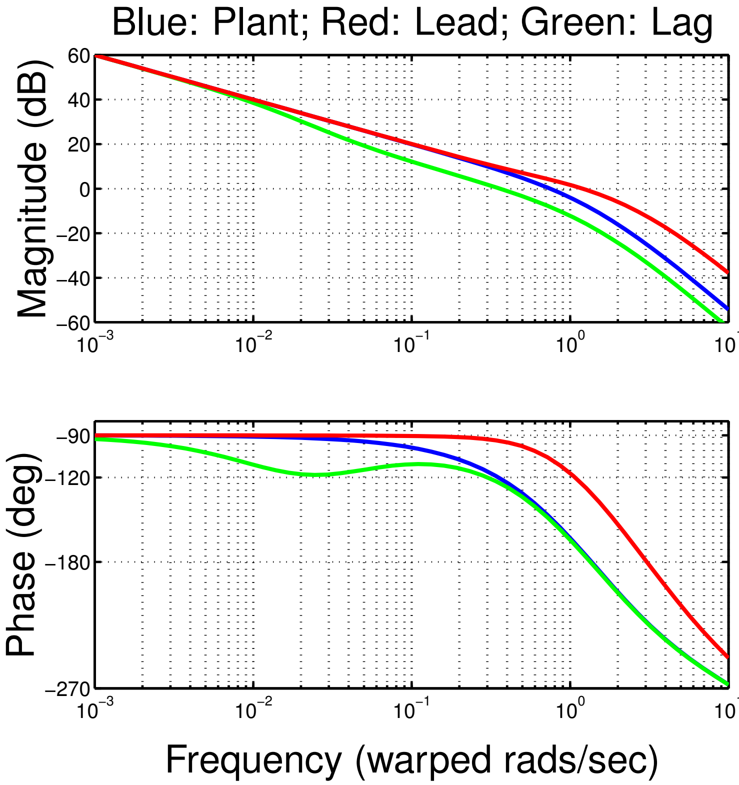
- So,

$$D(w) = \frac{a_1 w + a_0}{b_1 w + 1} = \frac{1.701w + 1}{0.2387w + 1}$$

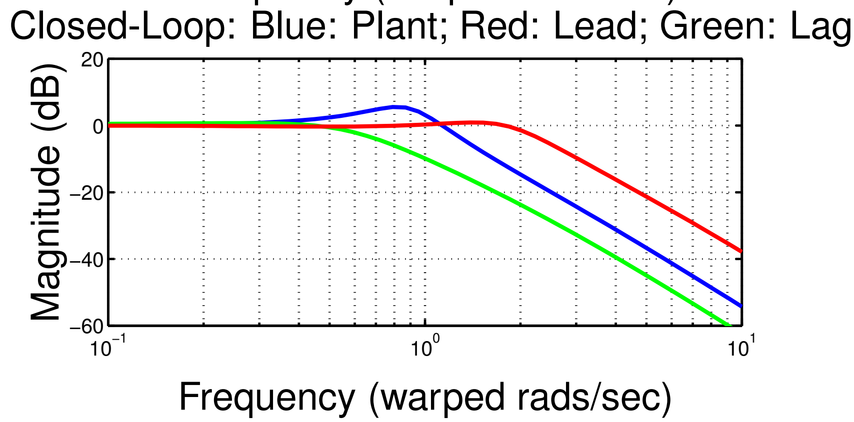
$$D(z) = \frac{6.539(z - 0.9710)}{z - 0.8106}.$$

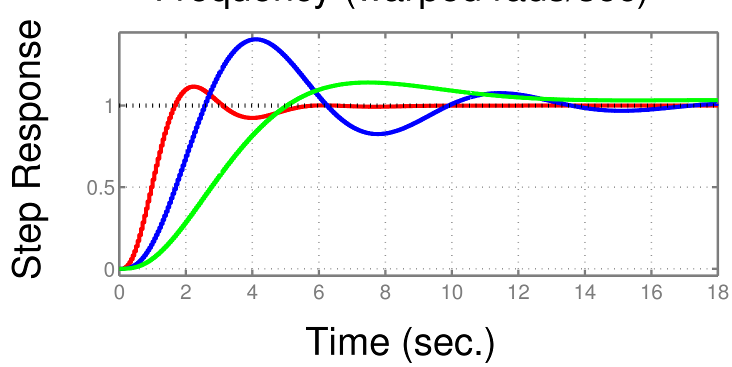


- The PM for this controller is $55°$ and the GM is 12.3 dB.
- A different choice of ω_{w_1} would give same PM but different GM.



- Compare the two examples from an open-loop perspective.
- Phase-lead has larger bandwidth.





- Compare from a closed-loop perspective.
- Again, phase-lead has large closed-loop bandwidth or large high-frequency gain.
- This may magnify effects due to sensor noise, and accentuate unmodeled high-frequency dynamics.
- One possible solution is to add a pole to $D(w)$ at high frequency (so not to change PM, but to reduce high-frequency gain).

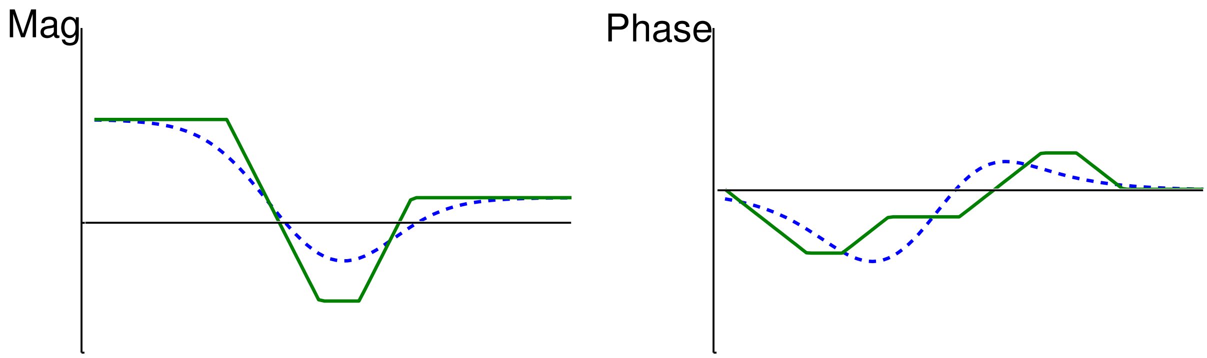
5.5: Lead/lag tradeoffs; other compensators

- In summary, some possible advantages of phase-lag compensation are:
  1. The low-frequency characteristics are maintained or improved.
  2. The stability margins are improved.
  3. The bandwidth is reduced, which is an advantage if high-frequency noise is a problem. Also, for other reasons, reduced bandwidth may be an advantage.
- Some possible disadvantages of phase-lag compensation are:
  1. The reduced bandwidth may be a problem in some systems.
  2. The system transient response will have one very slow term. This will become evident when root-locus design is covered.
  3. Numerical problems with filter coefficients may result.
- Some possible advantages of phase-lead compensation are:
  1. Stability margins are improved.
  2. High-frequency performance, such as speed-of-response, is improved.
  3. Phase-lead compensation is *required* to stabilize certain types of systems.
- Some possible disadvantages of phase-lead compensation are:
  1. Any high-frequency noise problems are accentuated.
  2. Large signals may be generated, which may damage the system or at least result in nonlinear operation of the system. Since the design assumed linearity, the results of the nonlinear operation will not be immediately evident.

Lag-Lead Compensation

- System specifications cannot always be achieved using a first-order (lead or lag) compensator.
- For example, low steady-state error may give very large bandwidth if a lead compensator is used.

- One option is to cascade lag and lead filters.



- Lag increases low-frequency gain.
- Lead increases bandwidth and stability margins.

Lead-lag design method (Bode)

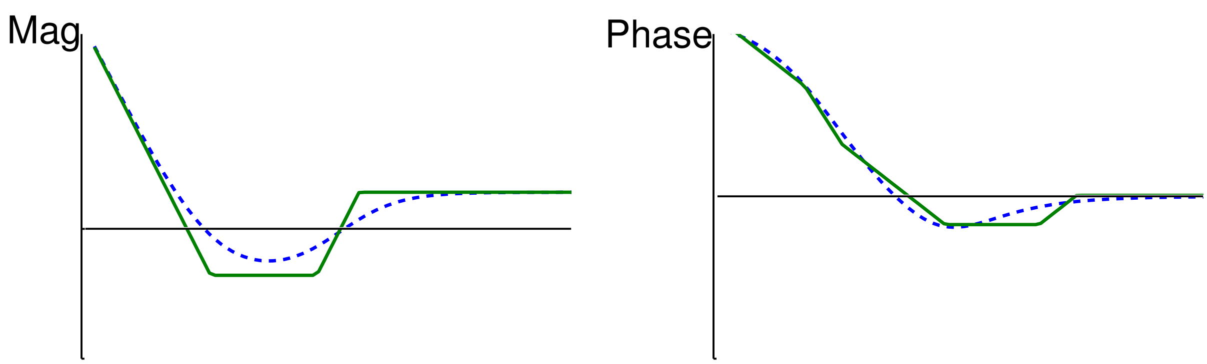
- Design lag first for acceptable Bode gain.
- Design lead for resulting system to give bandwidth and stability.

PID compensation

- A practical PID compensator has transfer function

$$D(z) = K\left[1 + \frac{T}{2T_I}\left(\frac{z+1}{z-1}\right) + T_D\left(\frac{z-1}{Tz}\right)\right].$$

Note that a bilinear integrator and a reverse-Euler derivative were used.



- Integrator term improves steady-state performance.
- Derivative term improves stability and PM.

PID design method (Bode)

- Very similar to lead design. Use $D(w) = K + \dfrac{K}{T_I w} + K T_D w$.
- Find $D(w)$ such that

$$D(j\omega_{w_1})G(j\omega_{w_1}) = 1\angle -180^\circ + PM$$

for a selected ω_{w_1}. Let $\theta = -180^\circ + PM - \angle G(j\omega_{w_1})$.

- Then, we can show that

$$K = \frac{\cos(\theta)}{|G(j\omega_{w_1})|}$$

and

$$T_D\omega_{w_1} + \frac{1}{T_I\omega_{w_1}} = \tan(\theta).$$

- Note that T_D and T_I are not uniquely specified. Choose one to meet some other specification. Increasing T_D increases bandwidth. Decreasing T_I decreases steady-state errors.

5.6: Root-locus design

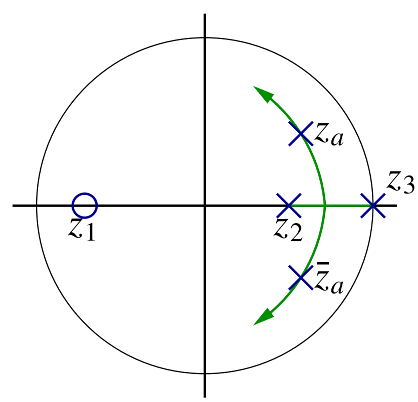
- An alternative design method is to use the root locus.
- Locus uses open-loop $D(z)\overline{GH}(z)$ information to plot locus of closed-loop poles.
- By adding dynamics to $D(z)$ we change the locus.

Phase-lag controller

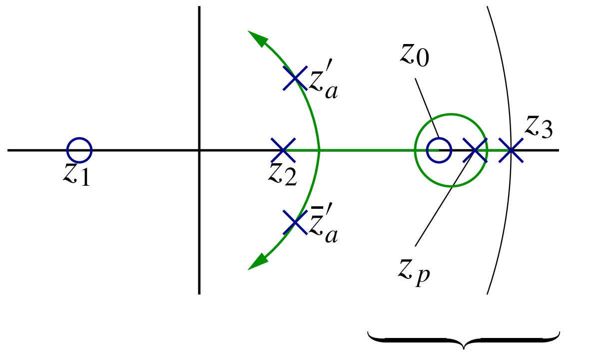
- To design a root-locus *lag* controller, we assume that the uncompensated system has good transient response but poor steady-state response.
- We set

$$D(z) = \underbrace{\left(\frac{1-z_p}{1-z_0}\right)}_{K_d}\left(\frac{z-z_0}{z-z_p}\right).$$

- Note $D(z)$ has unity dc-gain and $K_d < 1$. We place the pole near (very near) $z = 1$ and the zero a little to the left of the pole.
- Consider the uncompensated locus to the right. Pick a gain K_u to give pole locations z_a and $\bar{z}_a$. Assume that these pole locations are chosen to provide good transient response.



- Add the lag pole and zero very close to $z = 1$. The two poles and zero are so close together, they behave almost like a single pole. The locus is unchanged, except around $z = 1$.



- But, consider the gain K_c to get poles at z_a' and $\bar{z}_a'$. The compensator gain $K_d = \dfrac{1 - z_p}{1 - z_0} < 1$.
- The uncompensated gain is the gain to put poles at z_a and $\bar{z}_a$ without $D(z)$.
$$K_u = \frac{|z_a - z_2||z_a - z_3|}{|z_a - z_1|}.$$
- The compensated gain is
$$\begin{aligned} K_c &= \frac{|z_a' - z_p||z_a' - z_2||z_a' - z_3|}{K_d|z_a' - z_0||z_a' - z_1|} \\ &\approx \frac{|z_a - z_2||z_a - z_3|}{K_d|z_a - z_1|} \\ &= \frac{K_u}{K_d} > K_u. \end{aligned}$$
- Therefore, the lag compensator allows us to use larger gain for the same transient response and hence steady-state error is improved.

Phase-lag design procedure (root-locus method)

1. Plot locus of uncompensated system. Find acceptable pole locations for transient response on the locus. Set $K_u =$ gain to put poles there.
2. Determine from specifications the *required* gain K_c to meet steady-state error requirements.

3. Compute $K_d = K_u/K_c$.
4. Choose compensator pole location close to $z = 1$.
5. The compensator zero is

$$z_0 = 1 - \frac{1 - z_p}{K_d}.$$

6. Iterate (if necessary) to perfection (choosing slightly different pole location).

Phase-lead controller

- A phase-lead controller generally speeds up the transient response of a system.

$$D(z) = K_d \frac{(z - z_0)}{(z - z_p)}, \qquad K_d = \frac{1 - z_p}{1 - z_0} > 1.$$

- That is, the zero is closer than the pole to the unit circle.

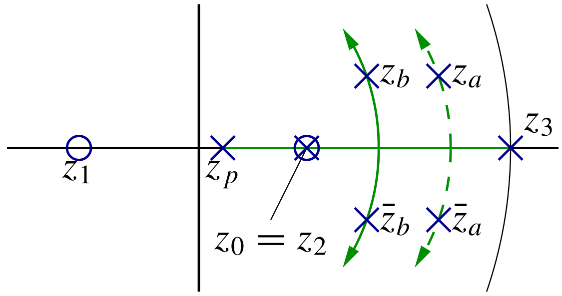
Phase-lead design procedure (root-locus method)

1. Choose *desired* pole locations z_b and $\bar{z}_b$.
2. Place the zero of the compensator to cancel a *stable* pole of $G(z)$.
3. Choose either the gain of the compensated system K_c or the compensator pole z_p such that $D(z)$ is phase-lead.
4. Then, for z_b to be on the locus, we must satisfy

$$K_c D(z) G(z)|_{z=z_b} = -1$$

   Solve for the unknown, which is either K_c or z_p.

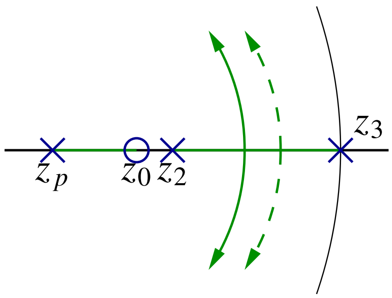
- Note that we only solve for one pole location. The other roots may not be satisfactory. May need trial-and-error approach to design.



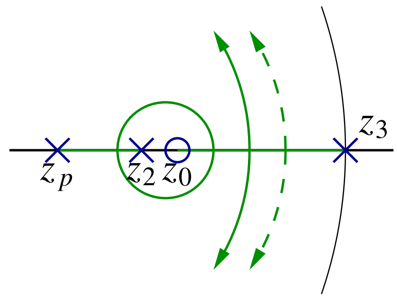
Pole-zero Cancellation

- Can it occur?

If our zero is too far left



If our zero is too far right



- Either way, the locus is still okay. (What if we tried to cancel an unstable pole?)

5.7: Numeric design of PID controllers

- A PID controller is implemented via difference equations: e.g.,

$$u[k] = u[k-1] + K\left[\left(1+\frac{T}{T_I}+\frac{T_D}{T}\right)e[k] - \left(1+\frac{2T_D}{T}\right)e[k-1] + \frac{T_D}{T}e[k-2]\right]$$

or,

$$u[k] = u[k-1] + q_0 e[k] + q_1 e[k-1] + q_2 e[k-2]$$

where q_0, q_1, and q_2 are constants.

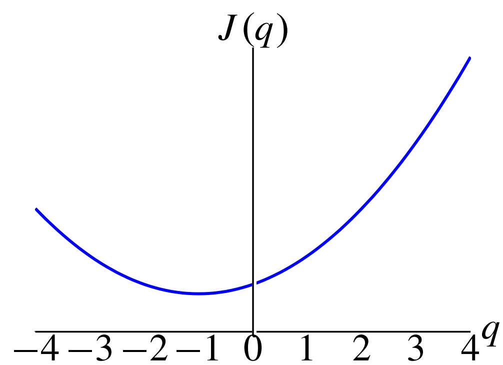
- Different sets of $\{q_0,\ q_1,\ q_2\}$ give different performance.
- Can select a set of parameters to satisfy some design specifications.
- Some types of design spec:

$$\text{IAE}: \quad J = \sum_k |e[k]| \qquad\qquad \text{ISE}: \quad J = \sum_k |e[k]|^2.$$

$$\text{ITAE}: \quad J = \sum_k k|e[k]|$$

- Over *large* number of samples, select $q = \{q_0,\ q_1,\ q_2\}$ to *minimize* J.

Function minimization in one dimension



- What value of q minimizes $J(q) = q^2 + 2q + 5$?
- Start with a guess, then refine until required accuracy is achieved.

1. Start with a guess. Say, $q = 3$.
2. Set the "step size" to a small value. Say $\gamma = 10^{-3}$, and set the gradient "test step size" δq to a value smaller than γ.

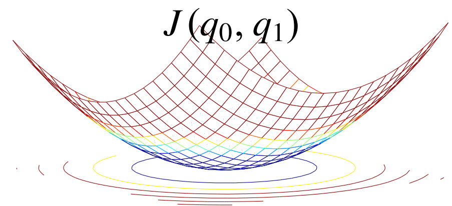
3. Estimate the slope: $\dfrac{\partial J}{\partial q} \approx \dfrac{J(q+\delta q) - J(q)}{\delta q}$.
4. Update parameters: $q \leftarrow q - \gamma \left(\dfrac{\partial J}{\partial q}\right)$.
5. Repeat from (3) until convergence.

- Note that update is in *negative* direction of gradient.

Function minimization in two dimensions

- Suppose we need to minimize a function of two variables, q_0 and q_1. For example,

$$J(q_0, q_1) = (q_0 - 2)^2 + (q_1 + 1)^2.$$



- Instead of a curve to track along, we have a three-dimensional surface.

1. Start with a guess. Say, $q_0 = 5$ and $q_1 = 3$.
2. Set the "step size" to a small value. Say $\gamma = 10^{-3}$, and set the gradient "test step sizes" $\delta q_0 = \delta q_1$ to a value smaller than γ .
3. Estimate the slope in the q_0-direction:
$$\frac{\partial J}{\partial q_0} \approx \frac{J(q_0 + \delta q_0, q_1) - J(q_0, q_1)}{\delta q_0}$$
4. Estimate the slope in the q_1-direction:
$$\frac{\partial J}{\partial q_1} \approx \frac{J(q_0, q_1 + \delta q_1) - J(q_0, q_1)}{\delta q_1}$$
5. Normalize the gradient for better performance
$$\Delta = \sqrt{\left(\frac{\partial J}{\partial q_0}\right)^2 + \left(\frac{\partial J}{\partial q_1}\right)^2}.$$

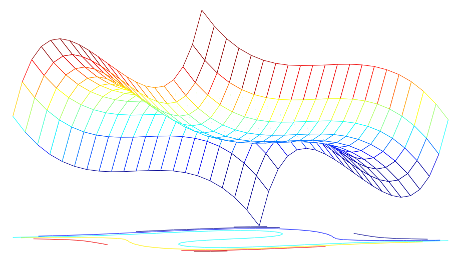
6. Update q_0 and q_1:

$$q_0 \leftarrow q_0 - \frac{\gamma}{\Delta}\left(\frac{\partial J}{\partial q_0}\right)$$

$$q_1 \leftarrow q_1 - \frac{\gamma}{\Delta}\left(\frac{\partial J}{\partial q_1}\right).$$

7. Repeat from (3) until convergence.

- Note: In many dimensions, performance optimization is complicated and "global" minimum is not guaranteed.



Function minimization in three(+) dimensions

- Our PID controller design problem is a problem in three variables.

1. Start with a guess. $q = \{q_0,\ q_1,\ q_2\}$.
2. Set the "step size" to a small value. Say $\gamma = 10^{-3}$, and set the gradient "test step sizes" $\delta q_0 = \delta q_1 = \delta q_2$ to a value smaller than γ.
3. Estimate the slope in the q_0-direction:

$$\frac{\partial J}{\partial q_0} \approx \frac{J(q_0 + \delta q_0, q_1, q_2) - J(q_0, q_1, q_2)}{\delta q_0}$$

4. Estimate the slope in the q_1-direction:

$$\frac{\partial J}{\partial q_1} \approx \frac{J(q_0, q_1 + \delta q_1, q_2) - J(q_0, q_1, q_2)}{\delta q_1}$$

5. Estimate the slope in the q_2-direction:

$$\frac{\partial J}{\partial q_2} \approx \frac{J(q_0, q_1, q_2 + \delta q_2) - J(q_0, q_1, q_2)}{\delta q_2}$$

6. Normalize the gradient for better performance

$$\Delta = \sqrt{\left(\frac{\partial J}{\partial q_0}\right)^2 + \left(\frac{\partial J}{\partial q_1}\right)^2 + \left(\frac{\partial J}{\partial q_2}\right)^2}.$$

7. Update q_0, q_1 and q_2:

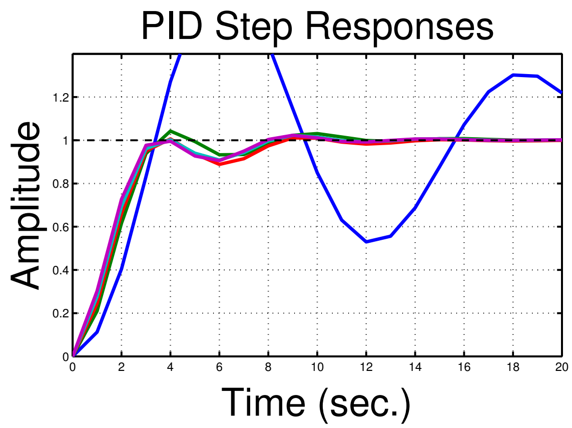
$$\begin{aligned} q_0 &\leftarrow q_0 - \frac{\gamma}{\Delta}\left(\frac{\partial J}{\partial q_0}\right) \\ q_1 &\leftarrow q_1 - \frac{\gamma}{\Delta}\left(\frac{\partial J}{\partial q_1}\right) \\ q_2 &\leftarrow q_2 - \frac{\gamma}{\Delta}\left(\frac{\partial J}{\partial q_2}\right). \end{aligned}$$

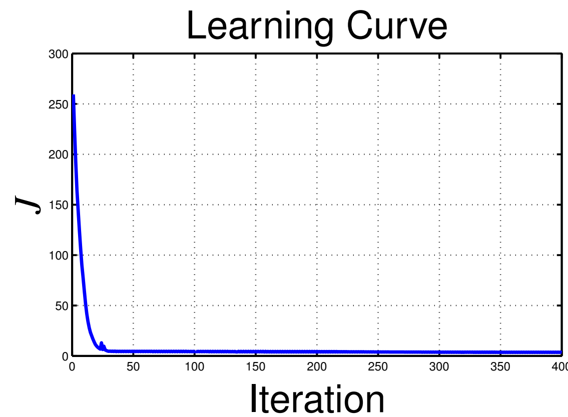
8. Repeat from (3) until convergence.

```
g=zpk([0.1+0.1*j 0.1-0.1*j],[0.3+0.5*j 0.3-0.5*j 0.8],1,-1);
gamma=0.01; dq=0.001; T=20;
wgtfn=[0:T]; % for ITAE

q0 = 0.1; q1 = 0.1; q2 = 0.1;

maxiter=410; J=zeros([1 maxiter]);

for i=1:maxiter,
  d=tf([q0+dq q1 q2],[1 -1 0],-1);
  t=feedback(d*g,1); s=step(t,T+1); J1=wgtfn*abs(1-s);

  d=tf([q0 q1+dq q2],[1 -1 0],-1);
  t=feedback(d*g,1); s=step(t,T+1); J2=wgtfn*abs(1-s);

  d=tf([q0 q1 q2+dq],[1 -1 0],-1);
  t=feedback(d*g,1); s=step(t,T+1); J3=wgtfn*abs(1-s);

  d=tf([q0 q1 q2],[1 -1 0],-1);
  t=feedback(d*g,1); s=step(t,T+1); J(i)=wgtfn*abs(1-s);
  if (mod(i-10,100)==0),
    J(i), plot(0:T,s); axis([0 T 0 1.4]); drawnow;
  end
```

```
  p1=(J1-J(i))/dq; p2=(J2-J(i))/dq; p3=(J3-J(i))/dq;
  Delta=sqrt(p1*p1+p2*p2+p3*p3);

  q0=q0-(gamma/Delta)*p1; q1=q1-(gamma/Delta)*p2; q2=q2-(gamma/Delta)*p3;
end
```

- Optimization progressing:



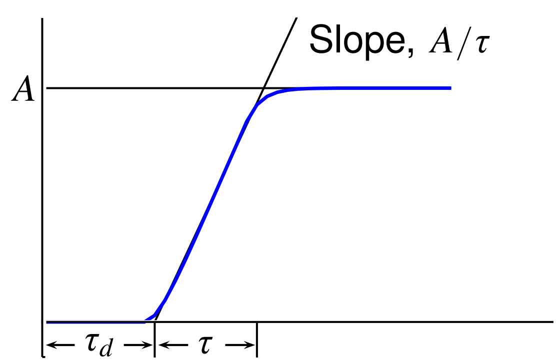


- Note that each time we adapt $\{q_0, q_1, q_2\}$ we need to simulate the system performance (*e.g.,* output to a step input) four times.

Initial guess for K, T_I and T_D:

- Ziegler–Nichols can give a good initial guess for PID parameters.
- "Rules of thumb" for selecting K, T_I, T_D.
- Not optimal in any sense—just provide "good" performance.

METHOD I: If system has step response like this,

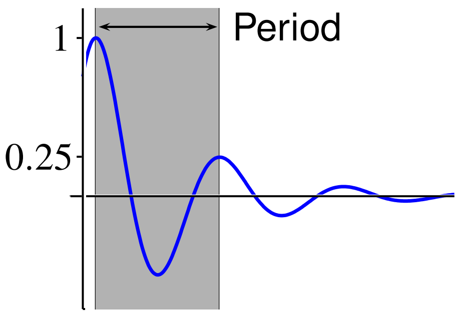


$$\frac{Y(s)}{U(s)} = \frac{Ae^{-\tau_d s}}{\tau s + 1},$$

(first-order system plus delay)

- We can easily identify A, τ_d, τ from this step response.

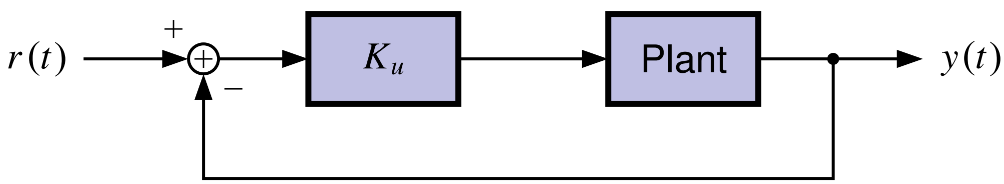
- Don't need complex model!
- Tuning criteria: Ripple in impulse response decays to 25% of its value in one period of ripple



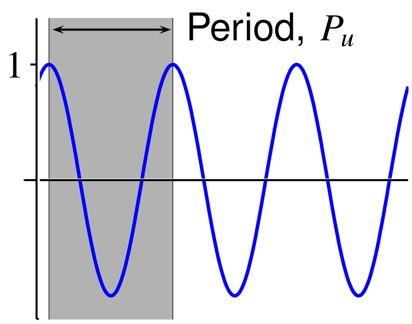
RESULTING TUNING RULES:

P	PI	PID
$K = \dfrac{\tau}{A\tau_d}$	$K = \dfrac{0.9\tau}{A\tau_d}$	$K = \dfrac{1.2\tau}{A\tau_d}$
$T_I = \infty$	$T_I = \dfrac{\tau_d}{0.3}$	$T_I = 2\tau_d$
$T_D = 0$	$T_D = 0$	$T_D = 0.5\tau_d$

METHOD II: Configure system as



- Turn up gain K_u until system produces oscillations (on stability boundary) K_u = "ultimate gain."



RESULTING TUNING RULES:

P	PI	PID
$K = 0.5K_u$	$K = 0.45K_u$	$K = 0.6K_u$
$T_I = \infty$	$T_I = \dfrac{1}{1.2}P_u$	$T_I = 0.5P_u$
$T_D = 0$	$T_D = 0$	$T_D = \dfrac{P_u}{8}$

5.8: Direct design method of Ragazzini

- An interesting design method *computes* $D(z)$ directly.
- Note: The closed-loop transfer function is

$$\begin{aligned} T(z) &= \frac{D(z)G(z)}{1+D(z)G(z)} \\ (1+D(z)G(z))T(z) &= D(z)G(z) \\ D(z)G(z)(T(z)-1) &= -T(z) \\ D(z) &= \frac{1}{G(z)}\frac{T(z)}{1-T(z)}. \end{aligned}$$

- Can control design be this simple?
- The problem is that this technique may ask for the impossible: (non-causal, unstable ...).

Causality:

- If $D(z)$ is causal, then it has no poles at ∞. $D(\infty)$ must be finite or zero. Therefore, $T(z)$ must have enough zeros at ∞ to cancel out poles at ∞ from $\dfrac{1}{G(z)}$.
  - $T(z)$ must have a zero at infinity of the same order as the order of the zero of $G(z)$ at infinity.
- Put another way, the delay in $T(z)$ must be at least as long as the delay in $G(z)$.

Stability:

- If $G(z)$ has unstable poles, they cannot be canceled directly by $D(z)$ or there will be trouble!

- The characteristic equation of the closed-loop system is

$$1 + D(z)G(z) = 0$$

Let $D(z) = \dfrac{c(z)}{d(z)}$ and $G(z) = \dfrac{b(z)}{a(z)}$. Then

$$1 + \frac{c(z)}{d(z)}\frac{b(z)}{a(z)} = 0.$$

- Let the unstable pole in $G(z)$ be at α, so $a(z) = (z-\alpha)\overline{a}(z)$. To cancel it, $c(z) = (z-\alpha)\overline{c}(z)$, and

$$(z-\alpha)\overline{a}(z)d(z) + (z-\alpha)\overline{c}(z)b(z) = 0$$

$$(z-\alpha)[\overline{a}(z)d(z) + \overline{c}(z)b(z)] = 0.$$

- The unstable root is still a factor of the characteristic equation! (oops).
- Unstable poles must be canceled via the feedback mechanism. This imposes constraints on $T(z)$.
  - $[1 - T(z)]$ must contain as zeros all the poles of $G(z)$ outside the unit circle.
  - $T(z)$ must contain as zeros all the zeros of $G(z)$ outside the unit circle.

Steady-state accuracy:

- Assume that the system is to be of type-I with velocity constant K_v.
- Note:

$$E(z) = [1 - T(z)]R(z).$$

- We must have zero steady-state error to a step. Therefore

$$\lim_{z\to 1}(z-1)[1-T(z)]\frac{z}{(z-1)} = 0$$

or $T(1) = 1$.

- Must have $1/K_v$ error to a unit ramp. Therefore

$$\lim_{z\to 1}(z-1)[1-T(z)]\frac{Tz}{(z-1)^2}=\frac{1}{K_v}.$$

- Use l'hôpital's rule to evaluate:

$$-T\left.\frac{\mathrm{d}T(z)}{\mathrm{d}z}\right|_{z=1}=\frac{1}{K_v}.$$

EXAMPLE: Consider $T=1$ s, $G(z)=0.0484\left[\dfrac{z+0.9672}{(z-1)(z-0.9048)}\right]$.

- Want $T(z)$ to approximate $s^2+s+1=0$, or, converting to z-plane,

$$z^2-0.7859z+0.3679=0.$$

- So,

$$T(z)=\frac{b_0+b_1z^{-1}+b_2z^{-2}+\cdots}{1-0.7859z^{-1}+0.3679z^{-2}}.$$

- Causality requires $T(z)|_{z=\infty}=0$ because $G(\infty)=0$, so $b_0=0$.
- Note that $G(z)$ has no poles *outside* the unit circle, so don't need to worry about that.
- Zero steady-state error to a step requires

$$T(1)=\frac{b_1+b_2+\cdots}{1-0.7859+0.3679}=1.$$

  Therefore,

$$b_1+b_2+\cdots=0.5820.$$

- Steady-state error of $1/K_v$ to a unit ramp (let $K_v=1$)

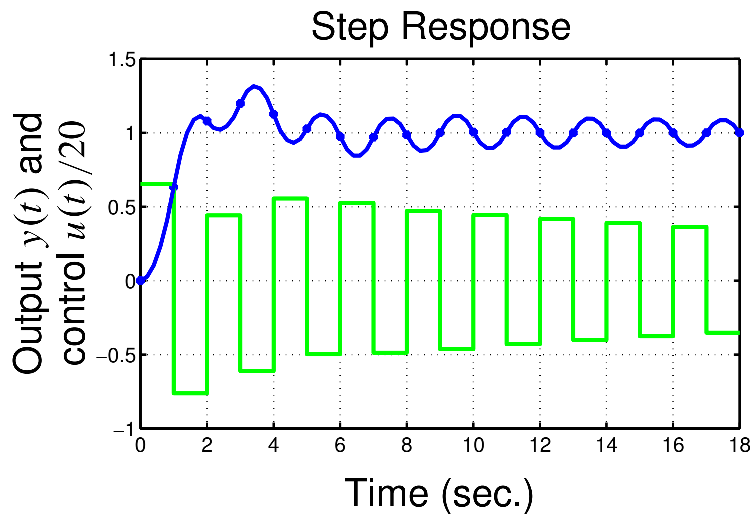
$$\frac{1}{K_v}=-\left.\frac{\mathrm{d}T(z)}{\mathrm{d}z}\right|_{z=1}$$

$$1=+\left.\frac{\mathrm{d}T(z)}{\mathrm{d}z^{-1}}\right|_{z=1}$$

$$= \frac{(0.5820)[b_1 + 2b_2 + 3b_3 + \cdots] - (0.5820)[-0.7859 + 0.3679(2)]}{(0.5820)^2},$$

or, $b_1 + 2b_2 + 3b_3 + \cdots = 0.5318$.

- So, we have two constraints. We can satisfy these constraints with just b_1 and b_2:



$$b_1 = 0.6321$$
$$b_2 = -0.0501.$$

- Then,

$$T(z) = \frac{0.6321z - 0.0501}{z^2 - 0.7859z + 0.3679}.$$

and

$$D(z) = \frac{1}{G(z)}\frac{T(z)}{1 - T(z)} = 13.07\frac{(z - 0.9048)}{(z + 0.9672)}\frac{(z - 0.07932)}{(z - 0.418)}.$$

- Note that $T(z)$ has two well damped poles. Why the oscillation?

$$\frac{U(z)}{R(z)} = \frac{D(z)}{1 + D(z)G(z)} = \frac{T(z)}{G(z)}$$
$$= 13.07\frac{(z - 0.0793)}{(z^2 - 0.7859z + 0.3679)}\frac{(z - 1)(z - 0.9048)}{(z + 0.9672)}.$$

- This has a poorly damped pole at -0.9672. Aha! This corresponds to the zero of $G(z)$ at -0.9672.

- A solution: Use a b_3 term in $T(z)$ and add the constraint that $T(z)|_{z=-0.9672} = 0$.