Solutions to Selected Problems

Contents

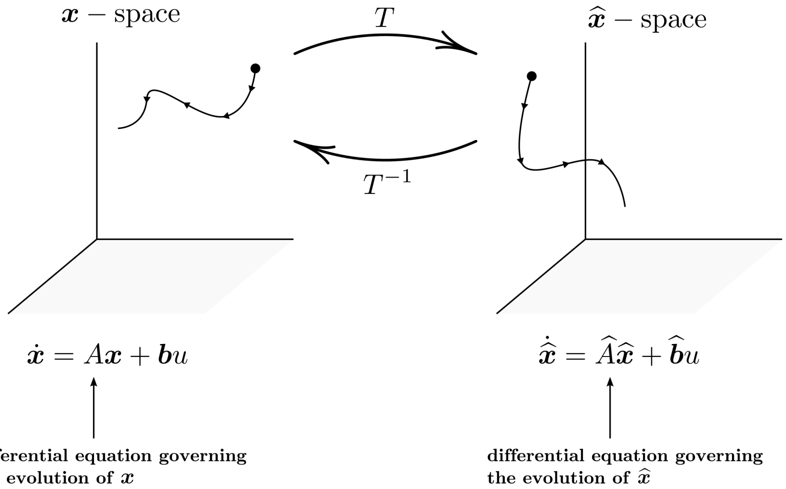
Notation

We will use the convention followed in the book; specifically:

- Random vectors are set in boldface sans serif type, as $\mathsf{\mathbf{x}}$ made up of scalar components x_i.
- Vectors (deterministic) are denoted by lower case letters in boldface type, as the vector $\boldsymbol{x}$ made up of scalar components x_i.
- $\mathbf{col}(\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots \boldsymbol{x}_N)$ is a column vector formed by vertically stacking column vectors $\boldsymbol{x}_i$'s, $1 \leq i \leq N$.

Chapter 1

A Brief Note On Controllable Canonical Form



They represent the same ***system*** (I/O behavior, internal stability property, identical modes, etc.) but each shows a different representation of the system.

Why is Controllable Canonical Form Useful?

Consider the state-space form of a linear dynamical system $\dot{\boldsymbol{x}} = A\boldsymbol{x} + \boldsymbol{b}u$ (SISO model). Transform it into the controllable canonical form using T by defining the state $\widehat{\boldsymbol{x}} = T\boldsymbol{x}$: $\dot{\widehat{\boldsymbol{x}}} = \widehat{A}\widehat{\boldsymbol{x}} + \widehat{\boldsymbol{b}}u$.

Using the input $u = -\boldsymbol{k}^T\widehat{\boldsymbol{x}}$, we can easily design the feedback gain $\boldsymbol{k}^T$ such that the eigenvalues of the closed-loop system $\dot{\widehat{\boldsymbol{x}}} = (\widehat{A} - \widehat{\boldsymbol{b}}\boldsymbol{k}^T)\widehat{\boldsymbol{x}}$ are placed at desired locations on the complex plane.

Applying this input to the system represented by $\dot{\boldsymbol{x}} = A\boldsymbol{x} + \boldsymbol{b}u$ yields:

$$\dot{\boldsymbol{x}} = A\boldsymbol{x} + \boldsymbol{b}(-\boldsymbol{k}^T\widehat{\boldsymbol{x}}) = (A - \boldsymbol{b}\boldsymbol{k}^T T)\boldsymbol{x} = A_{cl}\boldsymbol{x} \quad \text{where} \begin{cases} A_{cl} = A - \boldsymbol{b}\boldsymbol{k}^T T \\ \widehat{A}_{cl} = \widehat{A} - \widehat{\boldsymbol{b}}\boldsymbol{k}^T \end{cases}$$

But eigenvalues of A_{cl} and $\widehat{A}_{cl}$ are identical!

$$\widehat{A}_{cl} = TAT^{-1} - T\boldsymbol{b}\boldsymbol{k}^T = TAT^{-1} - T\boldsymbol{b}\boldsymbol{k}^T TT^{-1} = T\{A - \boldsymbol{b}\boldsymbol{k}^T T\}T^{-1} = TA_{cl}T^{-1}$$

Therefore, $\widehat{A}_{cl}$ and A_{cl} are similar matrices $\Longleftrightarrow \sigma(\widehat{A}_{cl}) = \sigma(A_{cl})$. As a result, applying the control $u = -\boldsymbol{k}^T T\boldsymbol{x}$ places the eigenvalues of A_{cl} at desired locations.

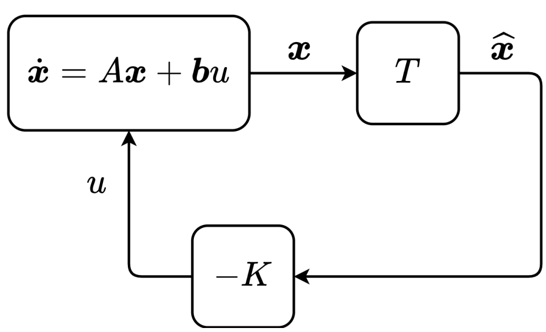


Figure 1: Example

Problem 2.1

$$G(s) = \frac{\overset{c_3}{\downarrow}1s^3 + \overset{c_2}{\downarrow}3s^2 + \overset{c_1}{\downarrow}5s + \overset{c_0}{\downarrow}8}{\underset{a_3}{\uparrow}s^4 + 7s^3 + \underset{a_2}{\uparrow}14s^2 + \underset{a_1}{\uparrow}8s + \underset{a_0}{\uparrow}0}$$

$$\boldsymbol{a} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 8 \\ 14 \\ 7 \end{bmatrix} \qquad A_{cc} = \left[\begin{array}{c|c} \mathbf{0} & I \\ \hline \multicolumn{2}{c}{-\boldsymbol{a}^T} \end{array}\right] = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -8 & -17 & -7 \end{bmatrix}$$

$$\boldsymbol{c} = \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 8 \\ 5 \\ 3 \\ 1 \end{bmatrix} \qquad \boldsymbol{b}_{cc} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \qquad \boldsymbol{c}_{cc}^T = \begin{bmatrix} 8 & 5 & 3 & 1 \end{bmatrix}$$

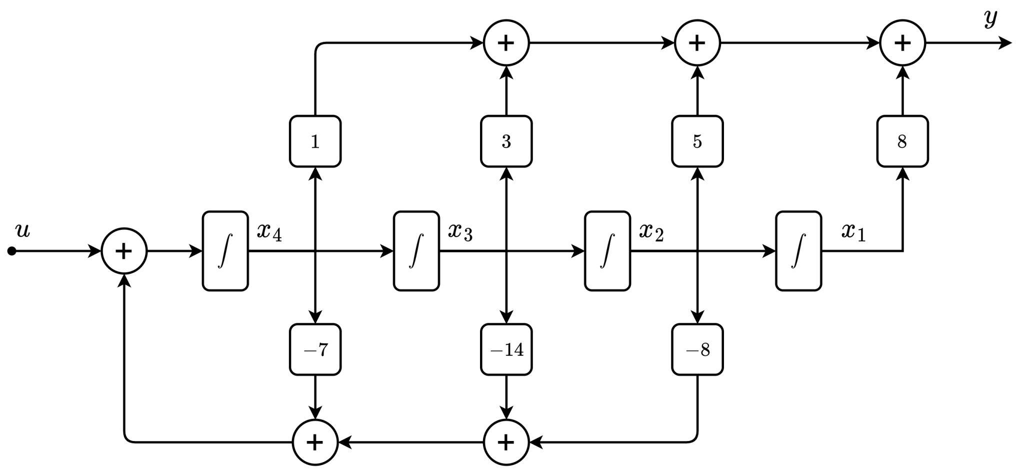


Figure 2: Example

Observable <u>Canonical</u> Form:

$$A_{oc} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & -8 \\ 0 & 1 & 0 & -14 \\ 0 & 0 & 1 & -7 \end{bmatrix} \qquad \boldsymbol{b}_{oc} = \begin{bmatrix} 8 \\ 5 \\ 3 \\ 1 \end{bmatrix}$$

$$\boldsymbol{c}_{oc}^T = \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix}$$

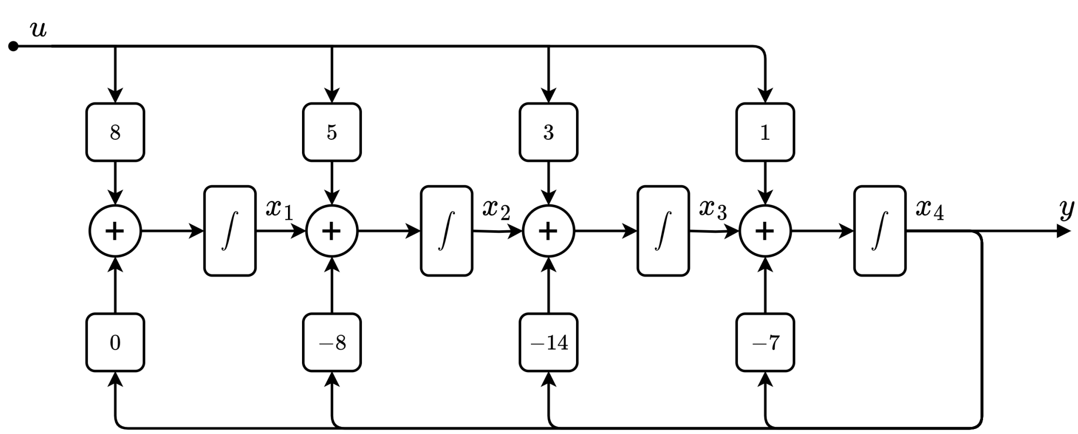


Figure 3: Example

Observable Companion Form:

$$A_{ocmp} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -8 & -14 & -7 \end{bmatrix} \qquad \boldsymbol{b}_{ocmp} = \begin{bmatrix} 1 \\ -4 \\ 19 \\ -77 \end{bmatrix}$$

$$\boldsymbol{c}_{ocmp}^T = \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix}$$

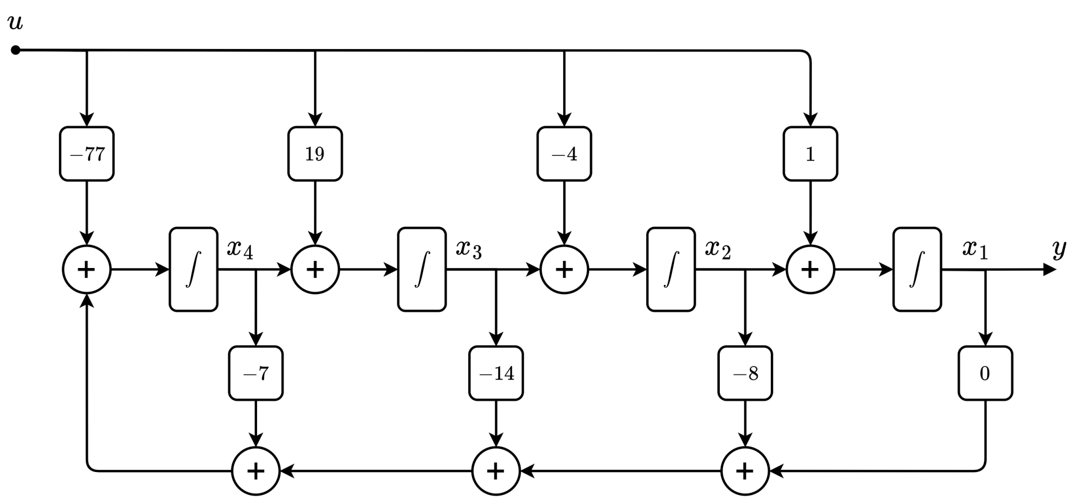
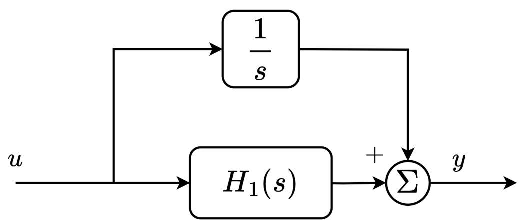


Figure 4: Example

Part a) How to interpret the singularity of A?

- A has a zero eigenvalue: a pole at the origin
- A has an eigenvector corresponding to this zero eigenvalue: $\boldsymbol{v} = Null(A)$
- Since Real$\{\lambda(A)\} < 0$ for all other eigenvalues λ, system states are absorbed to $\boldsymbol{v} = Null(A)$
- After modal decomposition, the ***component*** lying on $\boldsymbol{v}$ remains constant.
- Since there is only 1 pole at the origin:
  1. System is stable in the sense of Lyapunov.
  2. System is **NOT** BIBO stable.

     System can be decomposed in the way shown in the above figure. It is now evident that applying a unit step signal at the input $u(t)$ = unit step (BI) results in unbounded output.
  3. System is type 1: the closed loop system is able to track constant references inputs.

Part b)

$$G(s) = \frac{1}{s} + \frac{-5/3}{s+1} + \frac{1/2}{s+2} + \frac{7/6}{s+4}$$

$$A = \begin{bmatrix} 0 & & & \\ & -1 & & \\ & & -2 & \\ & & & -4 \end{bmatrix} \qquad \boldsymbol{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\boldsymbol{c}^T = \begin{bmatrix} 1 & \frac{-5}{3} & \frac{1}{2} & \frac{7}{6} \end{bmatrix}$$

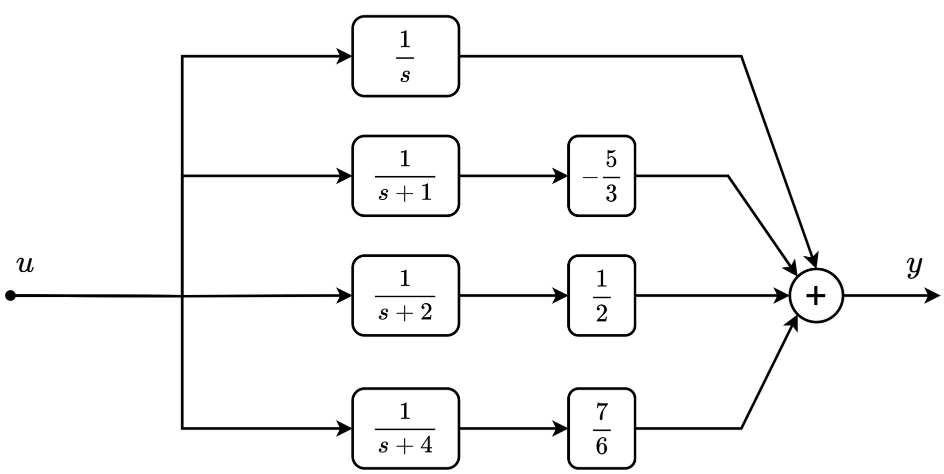


Figure 5: Example

Part c) Yes!

Problem 2.2

$$G(s) = \frac{\overset{c_1}{\downarrow}\hspace{-0.3em}10s + \overset{c_0}{\downarrow}\hspace{-0.3em}40}{\underset{a_2}{\uparrow}\hspace{-0.3em}s^3 + 3s^2 + \underset{a_1}{\uparrow}\hspace{-0.3em}2s + \underset{a_0}{\uparrow}\hspace{-0.3em}0}$$

$$\boldsymbol{a} = \begin{bmatrix} a_0 \\ a_1 \\ a_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 3 \end{bmatrix} \qquad A_{cc} = \left[\begin{array}{c|c} \boldsymbol{0} & I \\ \hline \multicolumn{2}{c}{-\boldsymbol{a}^T} \end{array}\right] = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -2 & -3 \end{bmatrix}$$

$$\boldsymbol{c} = \begin{bmatrix} c_0 \\ c_1 \\ c_2 \end{bmatrix} = \begin{bmatrix} 40 \\ 10 \\ 0 \end{bmatrix} \qquad \boldsymbol{b}_{cc} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \qquad \boldsymbol{c}_{cc}^T = \begin{bmatrix} 40 & 10 & 0 \end{bmatrix}$$

$$\rho(s) = s^3 + 3s^2 + 2s = s(s^2 + 3s + 2) = s(s+2)(s+1) \Rightarrow \begin{cases} s_1 = 0 \\ s_2 = -1 \\ s_3 = -2 \end{cases} \longrightarrow T = \left[\ \boldsymbol{v}_1 \,\middle|\, \boldsymbol{v}_2 \,\middle|\, \boldsymbol{v}_3\ \right]$$

$$A^{(0)}\boldsymbol{v}_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -2 & -3 \end{bmatrix} \boldsymbol{v}_1 = 0 \qquad \Rightarrow \boldsymbol{v}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$A^{(-1)}\boldsymbol{v}_2 = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & -2 & -2 \end{bmatrix} \boldsymbol{v}_2 = 0 \qquad \Rightarrow \boldsymbol{v}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

$$A^{(-2)}\boldsymbol{v}_3 = \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & -2 & -1 \end{bmatrix} \boldsymbol{v}_3 = 0 \qquad \Rightarrow \boldsymbol{v}_3 = \begin{bmatrix} 1 \\ -2 \\ 4 \end{bmatrix}$$

$$\Rightarrow T = \left[\ \boldsymbol{v}_1 \,\middle|\, \boldsymbol{v}_2 \,\middle|\, \boldsymbol{v}_3\ \right] = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -1 & -2 \\ 0 & 1 & 4 \end{bmatrix}$$ Which is the same as Vandermonde Matrix

$\widehat{A} = PAP^{-1} \quad \widehat{\boldsymbol{b}} = P\boldsymbol{b}$: set $P \triangleq T^{-1}$

Using $\widehat{\boldsymbol{x}} = P\boldsymbol{x}$:

$$\dot{\widehat{\boldsymbol{x}}} = \begin{bmatrix} 0 & & \\ & -1 & \\ & & -2 \end{bmatrix} \widehat{\boldsymbol{x}} + \begin{bmatrix} \frac{1}{2} \\ -1 \\ \frac{1}{2} \end{bmatrix} u$$

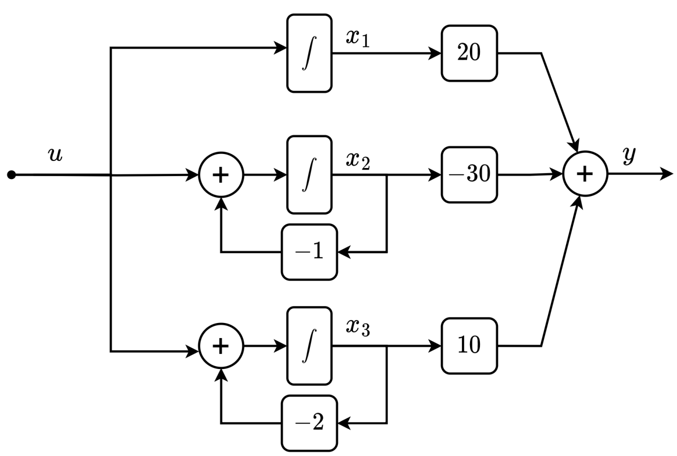
$$y = \begin{bmatrix} 40 & 30 & 20 \end{bmatrix} \widehat{\boldsymbol{x}}$$

To make $\widehat{\boldsymbol{b}} \leftarrow \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$:

$$\widetilde{\boldsymbol{x}} = \underbrace{\begin{bmatrix} 2 & & \\ & -1 & \\ & & 2 \end{bmatrix}}_{\widetilde{P}} \widehat{\boldsymbol{x}}$$

$$\dot{\widetilde{\boldsymbol{x}}} = \begin{bmatrix} 0 & & \\ & -1 & \\ & & -2 \end{bmatrix} \widetilde{\boldsymbol{x}} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} u$$

$$y = \begin{bmatrix} 20 & -30 & 10 \end{bmatrix} \widetilde{\boldsymbol{x}}$$



Using partial fraction expansion, the factorized transfer function can be expressed as

$$G(s) = \frac{20}{s} + \frac{-30}{s+1} + \frac{10}{s+2}$$

Problem 2.3

Decompose the given transfer function into its factors:

$$G(s) = \frac{\overset{\overset{c_0}{\downarrow}}{1}}{\underset{\underset{a_2}{\uparrow}}{s^3 + 7s^2} + \underset{\underset{a_1}{\uparrow}}{31s} + \underset{\underset{a_0}{\uparrow}}{25}} = \frac{1/20}{s+1} + \frac{\frac{1}{-32-16j}}{s-(-3+4j)} + \frac{\frac{1}{-32+16j}}{s-(-3-4j)}$$

We can now write the modal canonical form as:

$$A = \begin{bmatrix} -1 & & \\ & -3+4j & \\ & & -3-4j \end{bmatrix} \quad \boldsymbol{b} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$\boldsymbol{c}^T = \begin{bmatrix} \frac{1}{20} & \frac{1}{-32-16j} & \frac{1}{-32+16j} \end{bmatrix}$$

To eliminate complex entries, use the similarity transformation $T = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & j & -j \end{bmatrix}$. Let $\widehat{\boldsymbol{x}} = T\boldsymbol{x}$, then

$$\widehat{A} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -3 & 4 \\ 0 & -4 & -3 \end{bmatrix} \quad \widehat{\boldsymbol{b}} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$$

$$\widehat{\boldsymbol{c}}^T = \begin{bmatrix} \frac{1}{20} & \frac{-1}{40} & \frac{1}{80} \end{bmatrix}$$

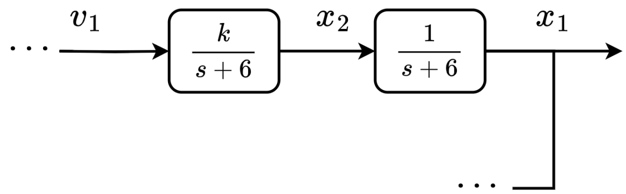
Controllable Canonical From:

$$A_{cc} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -25 & -31 & -7 \end{bmatrix} \quad \boldsymbol{b}_{cc} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

$$\boldsymbol{c}_{cc}^T = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$$

Problem 2.4

Part a)



$$\begin{cases} \dot{x}_2 = -6x_2 + kv_1 \\ \dot{x}_1 = -6x_1 + x_2 \\ v_1 = u - 3x_1 - x_2 \end{cases}$$

$$\Longrightarrow \begin{cases} \dot{x}_1 = -6x_1 + x_2 \\ \dot{x}_2 = -6x_2 + ku - 3kx_1 - kx_2 \end{cases} \qquad \Longrightarrow \begin{cases} \dot{x}_1 = -6x_1 + x_2 \\ \dot{x}_2 = -3kx_1 + (-6-k)x_2 + ku \end{cases}$$

$$\Rightarrow \dot{\boldsymbol{x}} = \begin{bmatrix} -6 & -1 \\ -3k & -k-6 \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} 0 \\ k \end{bmatrix} u \qquad \frac{Y(s)}{U(s)} = \begin{bmatrix} 1 & 0 \end{bmatrix} (sI - A)^{-1} \begin{bmatrix} 0 \\ k \end{bmatrix}$$

$$y = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \boldsymbol{x}$$

Part b)

Simplify the block diagram:

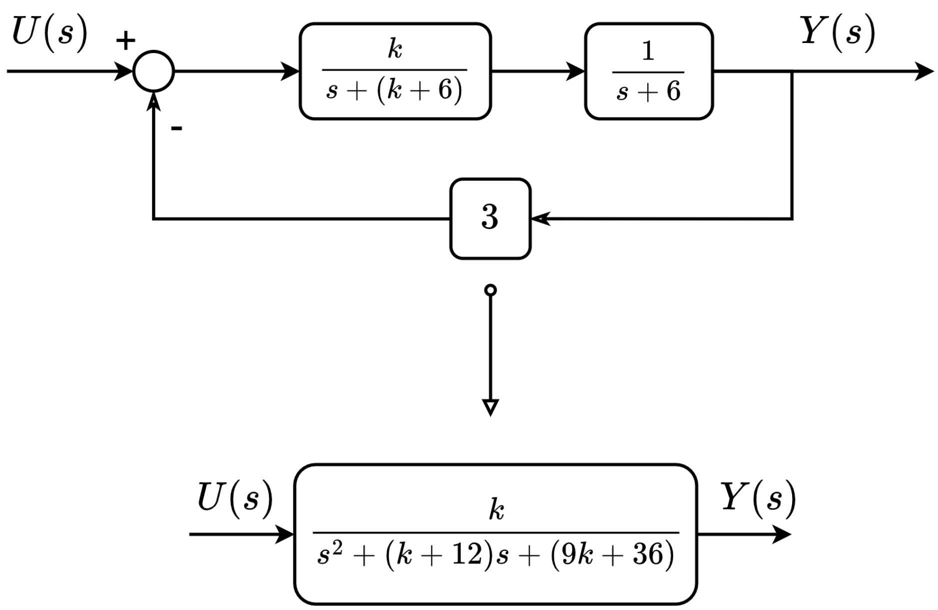


Figure 6: Example

Part c)

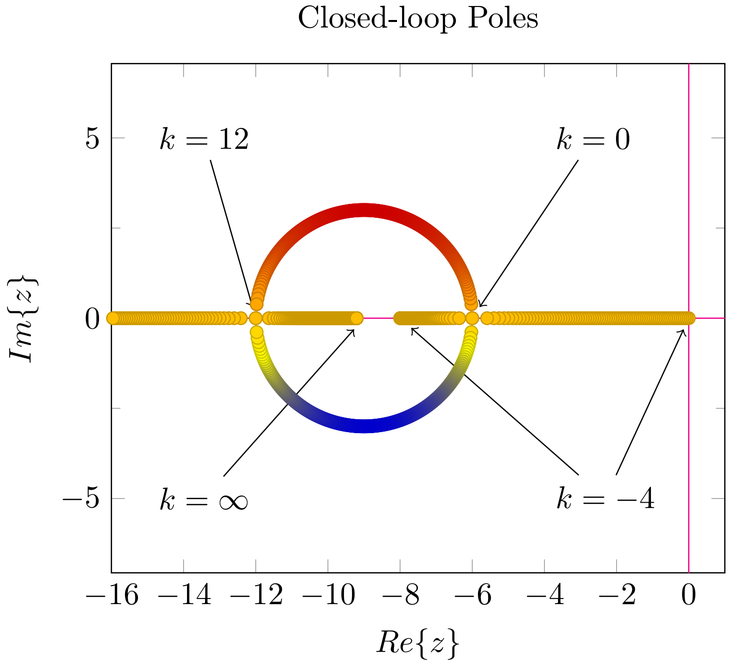
This is a 2nd order system whose poles are affected by the change in the gain value k. We can plot the locus of the poles of the closed loop system as function of k.
For the closed-loop system to be stable, according to Routh-Hurwitz criterion, the following conditions must hold:

$$\left.\begin{array}{ll} c > 0 \Rightarrow 9k + 36 > 0 & \Rightarrow k > -4 \\ b > 0 \Rightarrow k + 12 > 0 & \Rightarrow k > -12 \end{array}\right\} \Rightarrow k > -4$$

The roots can be obtained as

$$s_{1,2} = (-k/2 - 6) \pm \sqrt{k(k-12)}$$

For $0 < k < 12$, the roots appear in complex conjugate pairs.


Closed-loop Poles
$k = 12$
$k = 0$
$k = \infty$
$k = -4$
$Im\{z\}$
$Re\{z\}$
5
0
−5
−16
−14
−12
−10
−8
−6
−4
−2
0

Problem 2.5

State-space formulation of the dynamical system is given by

$$\dot{\boldsymbol{x}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -a_1 & -a_2 & -a_3 \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} u$$

$$\Rightarrow \Phi(s) = (sI - A)^{-1} = \frac{1}{s^3 + a_3 s^2 + a_2 s + a_1} \begin{bmatrix} s^2 + a_3 s + a_2 & s + a_3 & 1 \\ -a_1 & s^2 + a_3 s & s \\ -a_1 s & -a_2 s - a_1 & s^2 \end{bmatrix}$$

Appending the output equation $y = \begin{bmatrix} c_1 & c_2 & c_3 \end{bmatrix} \boldsymbol{x}$ and matching the resulting state-space model to the controllable canonical form yields

$$G(s) = \frac{Y(s)}{U(s)} = \frac{c_1 s^2 + c_2 s + c_3}{s^3 + a_3 s^2 + a_2 s + a_1}$$

The corresponding differential equation is then readily obtained as

$$\dddot{y} + a_3 \ddot{y} + a_2 \dot{y} + a_1 y = c_1 \ddot{u} + c_2 \dot{u} + c_3 u$$

Studying the steady-state response of the system to unit impulse, unit step inputes:

(I) $U(s) = 1 \Rightarrow sY(s) = \dfrac{s(c_1 s^2 + c_2 s + c_3)}{s^3 + a_3 s^2 + a_2 s + a_1}$

(a) if $G(s)$ has a single pole at the origin and all other poles in the *open* left-half complex plane, then: $a_1 = 0 \Rightarrow sY(s) = \dfrac{c_1 s^2 + c_2 s + c_3}{s^3 + a_3 s^2 + a_2 s} \Rightarrow y(\infty) = \lim_{s \to 0} sY(s) \Rightarrow \boxed{y(\infty) = \dfrac{c_3}{a_2}}$

(b) if $G(s)$ is BIBO stable (all poles lie the *open* left-half complex plane), then: $y(\infty) = \lim_{s \to 0} sY(s) \Rightarrow \boxed{y(\infty) = 0}$

(II) $U(s) = \dfrac{1}{s} \Rightarrow sY(s) = \dfrac{c_1 s^2 + c_2 s + c_3}{s^3 + a_3 s^2 + a_2 s + a_1} = G(s)$

if $G(s)$ is BIBO stable $\Rightarrow y(\infty) = \lim_{s \to 0} sY(s) \Rightarrow \boxed{y(\infty) = \dfrac{c_3}{a_1}}$

The model characterized by the given A and $\boldsymbol{b}$ is in controllable canonical form and thus is automatically controllable:

$$M_c = \left[\ \boldsymbol{b} \mid A\boldsymbol{b} \mid A^2\boldsymbol{b}\ \right] = \left[\begin{array}{c|c|c} 0 & 0 & 1 \\ 0 & 1 & -a_3 \\ 1 & -a_3 & -a_2 - a_3^2 \end{array}\right] \Rightarrow det(M_c) = -1 \Rightarrow \text{Controllable} \checkmark$$

Let $\boldsymbol{c}^T = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$ and form the observability matrix:

$$M_o = \begin{bmatrix} \boldsymbol{c}^T \\ \hline \boldsymbol{c}^T A \\ \hline \boldsymbol{c}^T A^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow det(M_o) = 1 \Rightarrow \text{Observable} \checkmark$$

As you can see, $det(M_c)$ and $det(M_o)$ are both non-zero and **independent** of the values of a_1, a_2, a_3.

Problem 2.6

let $\boldsymbol{x} \triangleq \begin{bmatrix} v \\ f \end{bmatrix}$, then the state-space equations are:

$$\begin{cases} \dot{v} = -\dfrac{b}{m}v - \dfrac{1}{m}f + \dfrac{1}{m}u & \text{(I)} \\ \dot{f} = 2kv & \text{(II)} \end{cases}$$

Reformulating the dynamical equations in terms of the new state $\widehat{\boldsymbol{x}} \triangleq \begin{bmatrix} f \\ \dot{f} \end{bmatrix}$:

$$\begin{aligned} \text{from (II):} \quad \ddot{f} &= 2k\left(\frac{-b}{m}v - \frac{1}{m}f + \frac{1}{m}u\right) \\ &= (2k)\left(\frac{-b}{m}\right)\left(\frac{\dot{f}}{2k}\right) - \frac{2k}{m}f + \frac{2k}{m}u \\ &= -\frac{b}{m}\dot{f} - \frac{2k}{m}f + \frac{2k}{m}u \end{aligned}$$

$$\Rightarrow \begin{cases} \dot{\widehat{x}}_1 = \dfrac{d}{dt}f = \dot{f} = \widehat{x}_2 \\ \dot{\widehat{x}}_2 = \dfrac{d}{dt}\dot{f} = \ddot{f} = -\dfrac{2k}{m}\widehat{x}_1 - \dfrac{b}{m}\widehat{x}_2 + \dfrac{2k}{m}u \end{cases} \xrightarrow{\text{Output Equations}} \begin{cases} z_1 = v = \dfrac{1}{2k}\dot{f} = \dfrac{1}{2k}\widehat{x}_2 \\ z_2 = f = \widehat{x}_1 \end{cases}$$

The new state-space equations are:

$$\dot{\widehat{\boldsymbol{x}}} = \begin{bmatrix} 0 & 1 \\ \dfrac{-2k}{m} & \dfrac{-b}{m} \end{bmatrix} \widehat{\boldsymbol{x}} + \begin{bmatrix} 0 \\ \dfrac{2k}{m} \end{bmatrix} u$$

$$\boldsymbol{z} = \begin{bmatrix} 0 & \dfrac{1}{2k} \\ 1 & 0 \end{bmatrix} \widehat{\boldsymbol{x}}$$

To obtain the transformation T relating these two state vectors, $\widehat{\boldsymbol{x}} = T\boldsymbol{x}$, note that $\begin{bmatrix} f \\ \dot{f} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 2k & 0 \end{bmatrix} \begin{bmatrix} v \\ f \end{bmatrix}$, therefore $T \triangleq \begin{bmatrix} 0 & 1 \\ 2k & 0 \end{bmatrix}$.

Using the Laplace transform, we get

$$\frac{\widehat{X}_2(s)}{U(s)} = \frac{2ks}{ms^2 + bs + 2k}$$

The form of time response depends on the parameters (whether it is underdamped, overdamped, or critically damped).

Problem 2.7

$$\begin{cases} I_{cm}\dot{\omega}_{im} &= M_m + M_{int} \\ \dot{e}_{sig} &= S_{sg}\left(\omega_{cmd} - \omega_{im} + \omega_d\right) \\ M_m &= \left(\dfrac{S_c}{S_{sg}}\right) F_c(D) e_{sig} \end{cases}$$

Part a)

Using the Laplace Transform:

$$I_{cm}s^2\omega_{im}(s) = sM_i(s) + \left(\frac{S_c}{S_{sg}}\right) F_c(s) S_{sg}\{\omega_{cmd}(s) - \omega_{im}(s) + \omega_d(s)\}$$
$$\left(Is^2 + S_cF_c(s)\right)\omega_{im}(s) = S_cF_c(s)\left(\omega_{cmd}(s) + \omega_d(s)\right) + sM_i(s)$$

Part b)

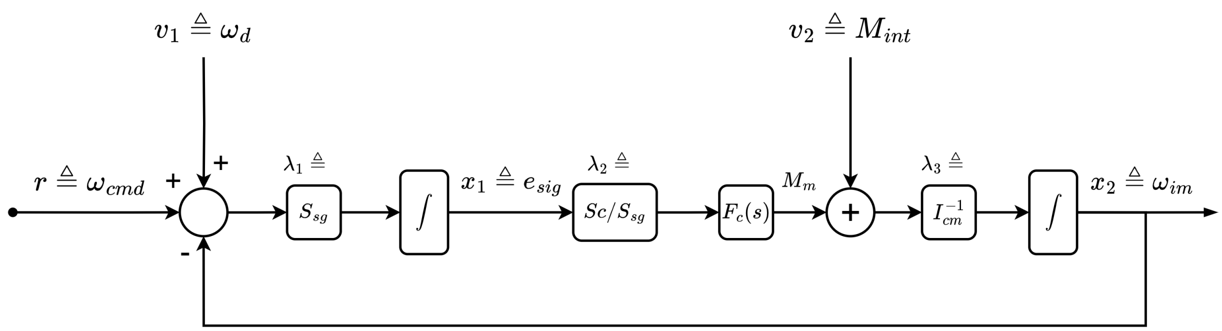


Figure 7: Example

Part c)

$$\begin{cases} \dot{x}_1 = \lambda_1 x_2 + \lambda_1 v_1 + \lambda_1 r \\ \dot{x}_2 = \lambda_2\lambda_3 x_1 + \lambda_3 v_2 \end{cases}$$

states: $\boldsymbol{x} = \begin{bmatrix} e_{sig} & \omega_{im} \end{bmatrix}^T$

disturbances: $\boldsymbol{v} = \begin{bmatrix} v_1 & v_2 \end{bmatrix}^T$

reference input: $r = \omega_{cmd}$

The new state-space equations are:

$$\dot{\boldsymbol{x}} = \underbrace{\begin{bmatrix} 0 & -\lambda_1 \\ \lambda_2\lambda_3 & 0 \end{bmatrix}}_{A} \boldsymbol{x} + \underbrace{\begin{bmatrix} \lambda_1 \\ 0 \end{bmatrix}}_{\boldsymbol{b}} r + \underbrace{\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_3 \end{bmatrix}}_{G} \underbrace{\begin{bmatrix} v_1 \\ v_2 \end{bmatrix}}_{\boldsymbol{v}}$$

Part d)

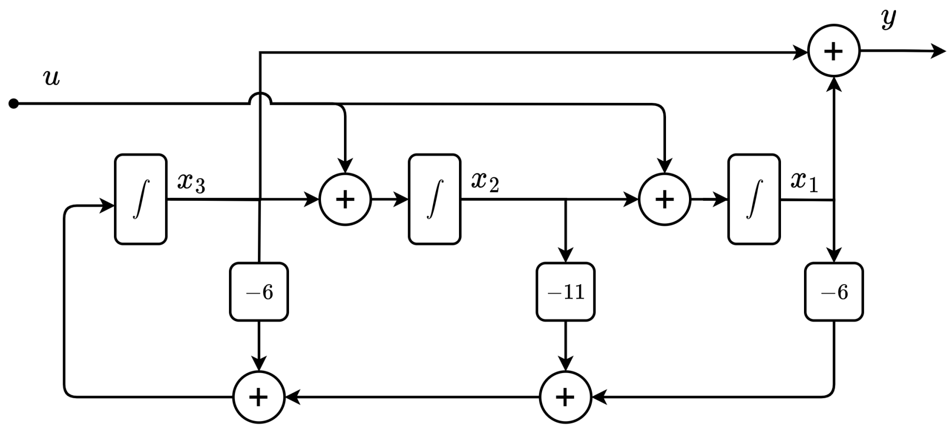
set $r = 0(t) \Rightarrow \dot{\boldsymbol{x}} = A\boldsymbol{x} + G\boldsymbol{v}$. We use Laplace transform to obtain $\frac{X_2(s)}{V_1(s)}$ and $\frac{X_2(s)}{V_2(s)}$ and then to analyze the behavior of the system under the action of disturbances.

$$\left.\begin{aligned} \frac{X_2(s)}{V_1(s)} &= \frac{\lambda_1\lambda_2\lambda_3}{s^2 + \lambda_1\lambda_2\lambda_3} = \frac{\alpha^2}{s^2+\alpha^2} \\ \frac{X_2(s)}{V_2(s)} &= \frac{\lambda_3 s}{s^2 + \lambda_1\lambda_2\lambda_3} \end{aligned}\right\} \Rightarrow \begin{array}{l}\text{Not BIBO stable:} \\ \text{Any disturbance would render the system unstable}\end{array}$$

Problem 2.9

From the block diagram, variables are related as

$$\begin{cases} \dot{x}_1 = x_2 + u \\ \dot{x}_2 = x_3 + u \\ \dot{x}_3 = -6x_1 - 11x_2 - 6x_3 \\ \hline y = x_1 + x_3 \end{cases}$$



In state-space form:

$$\dot{\boldsymbol{x}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -6 & -11 & -6 \end{bmatrix} \boldsymbol{x} + \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} u$$
$$y = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} \boldsymbol{x}$$

Compute controllability and observability matrices:

$$M_c = \left[\ \boldsymbol{b} \mid A\boldsymbol{b} \mid A^2\boldsymbol{b}\ \right] = \left[\begin{array}{c|c|c} 1 & 1 & 0 \\ 1 & 0 & -17 \\ 0 & -17 & 96 \end{array}\right] \Rightarrow \text{Full Rank} \Rightarrow \text{Controllable} \checkmark$$

$$M_o = \left[\begin{array}{c} \boldsymbol{c}^T \\ \hline \boldsymbol{c}^T A \\ \hline \boldsymbol{c}^T A^2 \end{array}\right] = \begin{bmatrix} 1 & 0 & 1 \\ -6 & -10 & -6 \\ 36 & 60 & 26 \end{bmatrix} \Rightarrow \text{Full Rank} \Rightarrow \text{Observable} \checkmark$$

Obtain the transfer function:

$$G(s) = \boldsymbol{c}^T(sI - A)^{-1}\boldsymbol{b} = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} \frac{1}{s^3 + 6s^2 + 11s + 6} \begin{bmatrix} s^2 + 6s + 11 & s + 6 & 1 \\ -6 & s^2 + 6s & s \\ -6s & -11s - 6 & s^2 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$$

$$= \frac{s^2 - 10s + 11}{s^3 + 6s^2 + 11s + 6} \longrightarrow \begin{cases} \text{Poles: } -1,\ -2,\ -3 & \Rightarrow \text{Stable} \\ \text{Zeros: } 8.74,\ 1.25 & \Rightarrow \text{Non-minimum Phase} \end{cases}$$

Find the eigen-decomposition of A: $V = \begin{array}{c} \begin{matrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 \end{matrix} \\ \begin{bmatrix} 1 & 1 & 1 \\ -1 & -2 & -3 \\ 1 & 4 & 9 \end{bmatrix} \end{array}$

let $T \triangleq V^{-1} = \begin{bmatrix} 3 & 5/2 & 1/2 \\ -3 & -4 & -1 \\ 1 & 3/2 & 1/2 \end{bmatrix}$. Define the new state $\widehat{\boldsymbol{x}} \triangleq T\boldsymbol{x}$:

$$\begin{aligned} \dot{\widehat{\boldsymbol{x}}} &= TAT^{-1}\widehat{\boldsymbol{x}} + T\boldsymbol{b}u \\ &= \widehat{A}\widehat{\boldsymbol{x}} + \widehat{\boldsymbol{b}}u \\ &= \begin{bmatrix} -1 & & \\ & -2 & \\ & & -3 \end{bmatrix} \widehat{\boldsymbol{x}} + \begin{bmatrix} 11/2 \\ -7 \\ 5/2 \end{bmatrix} u \end{aligned}$$

To make $\boldsymbol{b} \leftarrow \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$, define another transformation $Q \triangleq \begin{bmatrix} 2/11 & & \\ & -1/7 & \\ & & 2/5 \end{bmatrix}$ and let $\widetilde{\boldsymbol{x}} \triangleq Q\widehat{\boldsymbol{x}}$:

$$\begin{aligned} \dot{\widetilde{\boldsymbol{x}}} &= Q\widehat{A}Q^{-1}\widetilde{\boldsymbol{x}} + Q\widehat{\boldsymbol{b}}u \\ &= \begin{bmatrix} -1 & & \\ & -2 & \\ & & -3 \end{bmatrix} \widetilde{\boldsymbol{x}} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} u \end{aligned}$$

Combine the two transformations:

$$\widetilde{\boldsymbol{x}} = Q\widehat{\boldsymbol{x}} = QT\boldsymbol{x} = P\boldsymbol{x} \quad : P \triangleq QT$$

$$\begin{cases} \dot{\widetilde{\boldsymbol{x}}} = \widetilde{A}\widetilde{\boldsymbol{x}} + \widetilde{\boldsymbol{x}}u \\ y = \boldsymbol{c}^T P^{-1}\widetilde{\boldsymbol{x}} \end{cases} \implies \begin{aligned} \dot{\widetilde{\boldsymbol{x}}} &= \begin{bmatrix} -1 & & \\ & -2 & \\ & & -3 \end{bmatrix} \widetilde{\boldsymbol{x}} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} u \\ y &= \begin{bmatrix} 11 & -35 & 25 \end{bmatrix} \widetilde{\boldsymbol{x}} \end{aligned}$$

System response to impulse and unit step:

$$\begin{aligned} Y(s) = U(s)G(s) = &\xrightarrow[\text{impulse}]{U(s)=1} Y(s) = \frac{s^2 - 10s + 11}{s^3 + 6s^2 + 11s + 6} \\ &\xrightarrow{\mathcal{L}^{-1}\{\cdot\}} \boxed{y(t) = 11e^{-t} - 35e^{-2t} + 25e^{-3t}} \\ Y(s) = U(s)G(s) = &\xrightarrow[\text{step}]{U(s)=\frac{1}{s}} Y(s) = \frac{s^2 - 10s + 11}{s(s^3 + 6s^2 + 11s + 6)} \\ &\xrightarrow{\mathcal{L}^{-1}\{\cdot\}} \boxed{y(t) = -11e^{-t} + \frac{35}{2}e^{-2t} - \frac{25}{3}e^{-3t} + \frac{11}{6}} \end{aligned}$$

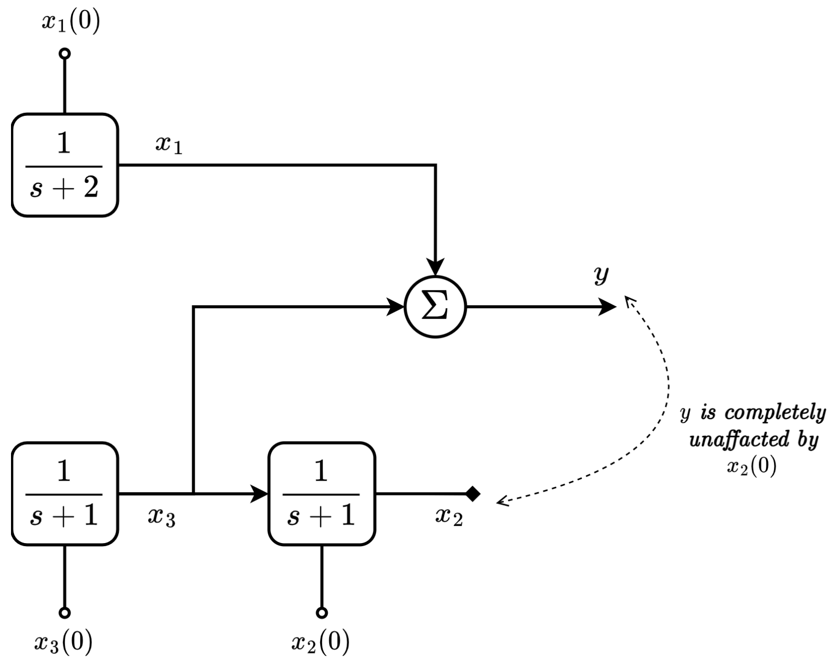
Problem 2.10

Part a)

Even if we set $u_1 = 0$, the presence of u_2 casts the system into controllable canonical form: $\boxed{\text{Controllable}}$. Observability can be investigated through modal decomposition:

$$\hat{A} = \left[\begin{array}{c|cc} -2 & & \\ \hline & -1 & 1 \\ & & -1 \end{array}\right] \longleftrightarrow$$

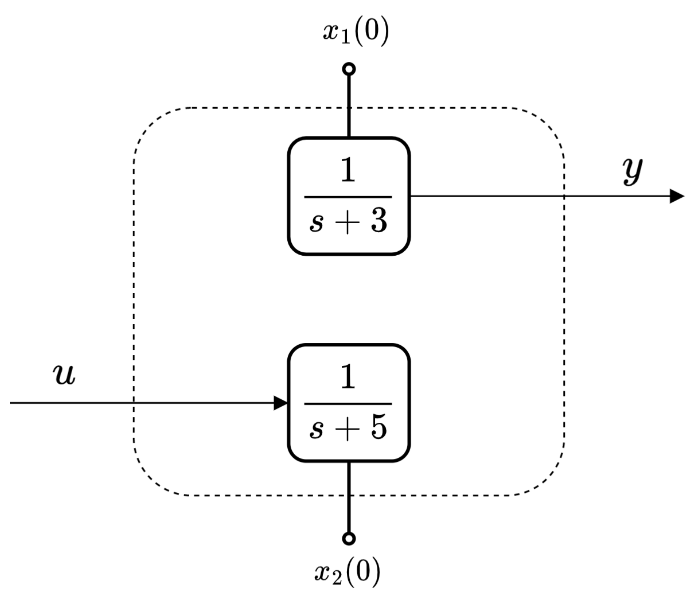
$$\hat{\mathbf{c}}^T = \left[\begin{array}{ccc} 1 & 0 & 1 \end{array}\right]$$



It is evident now that $y(t)$ does not in any way sense the change in $x_2(0)$: $\boxed{\text{Unobservable}}$.

Part b)

Since $x_1(t)$ is unaffected by the input, the system is $\boxed{\text{Uncontrollable}}$. Also, $x_2(t)$ is not observed in the output: $\boxed{\text{Unobservable}}$.



Part c)

The system is both $\boxed{\text{Controllable}}$ and $\boxed{\text{Observable}}$.

Problem 2.11

$$W(t_0,t) = \int_{t_0}^{t} e^{A(t_0-\tau)}B(\tau)B(\tau)^T e^{A^T(t_0-\tau)}d\tau \;\Rightarrow\; W(0,t) = \int_0^t e^{-A\tau}B(\tau)B(\tau)^T e^{-A^T\tau}d\tau$$

$A = \begin{bmatrix} 0 & \omega \\ -\omega & 0 \end{bmatrix}$. Since $\mathcal{L}\{e^{At}\} = (sI-A)^{-1} \;\Rightarrow\; \mathcal{L}\{e^{-At}\} = (sI+A)^{-1} \Rightarrow e^{-At} = \mathcal{L}^{-1}\{(sI+A)^{-1}\}$

$$sI + A = \begin{bmatrix} s & \omega \\ -\omega & s \end{bmatrix} \Rightarrow (sI+A)^{-1} = \frac{1}{s^2+\omega^2}\begin{bmatrix} s & -\omega \\ \omega & s \end{bmatrix} \Rightarrow \boxed{e^{-At} = \begin{bmatrix} \cos(\omega\tau) & -\sin(\omega\tau) \\ \sin(\omega\tau) & \cos(\omega\tau) \end{bmatrix}}$$

$$\Rightarrow W(0,t) = \int_0^t \begin{bmatrix} \cos(\omega\tau) & -\sin(\omega\tau) \\ \sin(\omega\tau) & \cos(\omega\tau) \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \cos(\omega\tau) & \sin(\omega\tau) \\ -\sin(\omega\tau) & \cos(\omega\tau) \end{bmatrix} d\tau$$

$$= \int_0^t \begin{bmatrix} \sin^2(\omega\tau) & -\sin(\omega\tau)\cos(\omega\tau) \\ -\sin(\omega\tau)\cos(\omega\tau) & \cos^2(\omega\tau) \end{bmatrix} d\tau$$

$$= \begin{bmatrix} \frac{t}{2} - \frac{\sin(2\omega t)}{4\omega} & -\frac{1}{2\omega}\sin^2(\omega t) \\ -\frac{1}{2\omega}\sin^2(\omega t) & \frac{t}{2} + \frac{\sin(2\omega t)}{4\omega} \end{bmatrix}$$

$\det(W(0,t)) = \frac{1}{4\omega^2}\left((t\omega)^2 - \sin^2(t\omega)\right) > 0 \quad \forall t > 0 \;\Rightarrow$ Controllable ✓

$M_c = \left[\; \boldsymbol{b} \;\middle|\; A\boldsymbol{b} \;\right] = \left[\begin{array}{c|c} 0 & \omega \\ 1 & 0 \end{array}\right] \Rightarrow \det(M_c) = -\omega \neq 0 \Rightarrow$ Controllable ✓

Problem 2.15

Part a)

Let $\Lambda \triangleq \Phi(t_1, t_0)$ and define $\Psi(t, t_1, t_0) \triangleq \Phi(t, t_1)\Phi(t_1, t_0)$ viewed as matrix-valued function of t. $\Phi(t, t_0)$ satsifies the following IVP:

$$\begin{cases} \dfrac{d}{dt}\Phi(t, t_0) = A(t)\Phi(t, t_0) \\ \Phi(t_1, t_0) = \Lambda \end{cases}$$

This is a linear matrix differential equation with specified initial condition which has a unique solution. Realize that

$$\Psi(t_1, t_1, t_0) = \Phi(t_1, t_1)\Phi(t_1, t_0) = \mathbb{I}\ \Phi(t_1, t_0) = \Lambda$$

Differentiating $\Psi(t, t_1, t_0)$ with respect to t yields:

$$\frac{d}{dt}\Psi(t, t_1, t_0) = A(t)\Phi(t, t_1)\Phi(t_1, t_0) = A(t)\Psi(t, t_1, t_0)$$

As a result, $\Psi(t, t_1, t_0)$ satisfies the same IVP:

$$\begin{cases} \dfrac{d}{dt}\Psi(t, t_1, t_0) = A(t)\Psi(t, t_1, t_0) \\ \Psi(t_1, t_1, t_0) = \Lambda \end{cases}$$

And by uniqueness of the solution, we conclude that $\Phi(t, t_0) = \Psi(t, t_1, t_0) \Longrightarrow \boxed{\Phi(t, t_0) = \Phi(t, t_1)\Phi(t_1, t_0)}$ for any t, t_0, and t_1.

Part b)

Substitute $t = t_0$ in $\Phi(t, t_0) = \Phi(t, t_1)\Phi(t_1, t_0)$ to obtain $\Phi(t_0, t_0) = \mathbb{I} = \Phi(t_0, t_1)\Phi(t_1, t_0) \Longrightarrow \Phi(t_0, t_1) = \Phi^{-1}(t_1, t_0)$. Since this is valid for any t_0 and t_1, $\Phi(t_0, t) = \Phi^{-1}(t, t_0)$.

Problem 2.16

The solution to the state differential equations $\dot{\boldsymbol{x}}(t) = A(t)\boldsymbol{x}(t) + B(t)\boldsymbol{u}(t)$ is given by $\boldsymbol{x}(t) = \Phi(t,t_0)\boldsymbol{x}(t_0) + \boldsymbol{\alpha}(t)$ where $\Phi(t,t_0)$ is unique and satisfies

$$\begin{cases} \dot{\Phi}(t,t_0) = \dfrac{d}{dt}\Phi(t,t_0) = A(t)\Phi(t,t_0) \\ \Phi(t_0,t_0) = I \end{cases} \tag{$\star$}$$

Consider the state $\widehat{\boldsymbol{x}}(t)$ which is related to $\boldsymbol{x}(t)$ as $\widehat{\boldsymbol{x}}(t) = T\boldsymbol{x}(t)$. The differential equations governing the evolution of $\widehat{\boldsymbol{x}}(t)$ are obtained as: (dropping dependencies on t to simplify notation)

$$\begin{aligned} \dot{\widehat{\boldsymbol{x}}} &= \dot{T}\boldsymbol{x} + T(A\boldsymbol{x} + B\boldsymbol{u}) \\ \Rightarrow \dot{\widehat{\boldsymbol{x}}} &= \left(\dot{T} + TA\right)T^{-1}\widehat{\boldsymbol{x}} + TB\boldsymbol{u} \\ &= \widehat{A}\widehat{\boldsymbol{x}} + \widehat{B}\boldsymbol{u} \end{aligned}$$

The solution to this system of differential equations is expressed as $\widehat{\boldsymbol{x}}(t) = \widehat{\Phi}(t,t_0)\widehat{\boldsymbol{x}}(t_0) + \widehat{\boldsymbol{\alpha}}(t)$. Since $\boldsymbol{x}(t) = \Phi(t,t_0)\boldsymbol{x}(t_0) + \boldsymbol{\alpha}(t)$:

$$\xRightarrow{\widehat{\boldsymbol{x}}=T\boldsymbol{x}} T^{-1}(t)\widehat{\boldsymbol{x}}(t) = \Phi(t,t_0)T^{-1}(t_0)\widehat{\boldsymbol{x}}(t_0) + \boldsymbol{\alpha}(t) \Rightarrow \widehat{\boldsymbol{x}}(t) = \underbrace{\{T(t)\Phi(t,t_0)T^{-1}(t_0)\}}_{\widehat{\Phi}(t,t_0)}\widehat{\boldsymbol{x}}(t_0) + T(t)\boldsymbol{\alpha}(t) = \widehat{\Phi}(t,t_0)\widehat{\boldsymbol{x}}(t_0) + \widehat{\boldsymbol{\alpha}}(t)$$

The last equality results from the uniqueness of the solution of differential equations. As such, the two matrices are related as

$$\boxed{\widehat{\Phi}(t,t_0) = T(t)\Phi(t,t_0)T^{-1}(t_0)}$$

Also, note that $\widehat{\Phi}(t_0,t_0) = T(t_0)\underbrace{\Phi(t_0,t_0)}_{I}T^{-1}(t_0) = I$ which is consistent with the properties of the state transition matrix.

To obtain the differential equation satisfied by $\widehat{\Phi}(t_0,t_0)$, first, note that

$$\frac{d}{dt}\left(\Phi(t,t_0)\right) = \dot{\Phi}(t,t_0) \overset{\text{from}(\star)}{=\!=\!=} A(t)T^{-1}(t)\widehat{\Phi}(t,t_0)T(t_0) \tag{$\star\star$}$$

Next, differentiating the boxed equation above (after isolating $\Phi(t,t_0)$):

$$\begin{aligned} \Phi(t,t_0) = T^{-1}(t)\widehat{\Phi}(t,t_0)T(t_0) \xRightarrow{\frac{d}{dt}(\cdot)} \dot{\Phi}(t,t_0) &= \dot{T}^{-1}(t)\widehat{\Phi}(t,t_0)T(t_0) + T^{-1}(t)\dot{\widehat{\Phi}}(t,t_0)T(t_0) \\ &= A(t)T^{-1}(t)\widehat{\Phi}(t,t_0)T(t_0) && \text{(from } (\star\star)) \end{aligned}$$

$$\begin{aligned} &\Rightarrow T^{-1}\dot{\widehat{\Phi}}(t,t_0) = T^{-1}\dot{T}T^{-1}\widehat{\Phi}(t,t_0) + A(t)T^{-t}\widehat{\Phi}(t,t_0) \\ &\Rightarrow \boxed{\dot{\widehat{\Phi}}(t,t_0) = \left(TAT^{-1} + \dot{T}T^{-1}\right)\widehat{\Phi}(t,t_0)} \end{aligned}$$

Problem 2.18

First, we find the differential equation satisfied by $\boldsymbol{p}(t)$:

$$\boldsymbol{x}^T\boldsymbol{p} = c \xrightarrow{\frac{d}{dt}(\cdot)} \dot{\boldsymbol{x}}^T\boldsymbol{p} + \boldsymbol{x}^T\dot{\boldsymbol{p}} = 0 \Rightarrow \boldsymbol{x}^T A^T\boldsymbol{p} + \boldsymbol{x}^T\dot{\boldsymbol{p}} = 0 \Rightarrow \boldsymbol{x}^T\left(A^T\boldsymbol{p} + \dot{\boldsymbol{p}}\right) = 0$$

Since this must hold <u>for all</u> $\boldsymbol{x}(t)$, then we must have

$$\boxed{\dot{\boldsymbol{p}} = -A^T\boldsymbol{p}}$$

Suppose $\boldsymbol{x}(t)$ is the solution of the IVP $\{\dot{\boldsymbol{x}}(t) = A(t)\boldsymbol{x}(t); \boldsymbol{x}(t_0) = \boldsymbol{x}_0\}$, then define $\boldsymbol{p}(t)$ as a time function such that the relation $\boldsymbol{x}^T(t)\boldsymbol{p}(t) = c$ holds for all t. In particular, $\boldsymbol{x}^T(t)\boldsymbol{p}(t) = \boldsymbol{x}^T(t_0)\boldsymbol{p}(t_0) = c$

$$\begin{cases} \dot{\boldsymbol{x}}(t) = A(t)\boldsymbol{x}(t) \\ \boldsymbol{x}(t) = \Phi(t,t_0)\boldsymbol{x}(t_0) \\ \text{where } \dot{\Phi}(t,t_0) = A(t)\Phi(t,t_0) \end{cases} \qquad \begin{cases} \dot{\boldsymbol{p}}(t) = -A^T(t)\boldsymbol{p}(t) \\ \boldsymbol{p}(t) = \widehat{\Phi}(t,t_0)\boldsymbol{p}(t_0) \\ \text{where } \dot{\widehat{\Phi}}(t,t_0) = -A^T(t)\widehat{\Phi}(t,t_0) \end{cases}$$

Proceed as follows:

$$\begin{aligned} \boldsymbol{x}^T(t)\boldsymbol{p}(t) = \boldsymbol{x}^T(t_0)\boldsymbol{p}(t_0) &\Rightarrow \boldsymbol{x}^T(t_0)\Phi^T(t,t_0)\widehat{\Phi}(t,t_0)\boldsymbol{p}(t_0) = \boldsymbol{x}^T(t_0)\boldsymbol{p}(t_0) \\ &\Rightarrow \boldsymbol{x}^T(t_0)\{\Phi^T(t,t_0)\widehat{\Phi}(t,t_0) - I\}\boldsymbol{p}(t_0) = 0 \end{aligned}$$

This must be true for <u>any</u> choice of $(\boldsymbol{x}(t_0), \boldsymbol{p}(t_0))$, therefore, we can conclude that $\Phi^T(t,t_0)\widehat{\Phi}(t,t_0) - I = 0$ and thus

$$\boxed{\widehat{\Phi}(t,t_0) = \Phi^{-T}(t,t_0)} \qquad (\star)$$

There is an alternative way to show this result. Consider the linear differential matrix equation with the specified initial condition: $\begin{cases} \dot{X} = 0 \\ X(t_0) = I \end{cases}$

We claim that $X(t) \triangleq I$ satisfies this system of equations and thus is *the* solution for this IVP. Indeed this function passes through the specified initial condition at time t_0: $X(t_0) = I$ and also $\dfrac{d}{dt}(X(t)) = \dfrac{d}{dt}(I) = 0$.

We claim that $Y(t) \triangleq \Phi^T(t,t_0)\widehat{\Phi}(t,t_0)$ also satisfies this differential equation:

$$\begin{aligned} &(1) \quad Y(t_0) = \Phi^T(t_0,t_0)\widehat{\Phi}(t_0,t_0) = (I)(I) = I \checkmark \\ &(2) \quad \frac{d}{dt}Y(t) = \frac{d}{dt}\left(\Phi^T(t,t_0)\widehat{\Phi}(t,t_0)\right) = \dot{\Phi}^T\widehat{\Phi} + \Phi^T\dot{\widehat{\Phi}} = \Phi^T A^T\widehat{\Phi} + \Phi^T(-A^T)\widehat{\Phi} = 0 \checkmark \end{aligned}$$

Since the solution must be unique, it must be the case that $Y(t) \equiv X(t)$:

$$\Phi^T(t,t_0)\widehat{\Phi}(t,t_0) = I \Rightarrow \ \widehat{\Phi}(t,t_0) = \Phi^{-T}(t,t_0)$$

Using the identity $(\star)$ (note the arrangement of input arguments):

$$\widehat{\Phi}^T(t,\tau) = \Phi^{-1}(t,\tau) \Rightarrow \widehat{\Phi}^T(t,\tau) = \Phi(\tau,t) \xrightarrow[\text{variables}]{\text{rename}} \widehat{\Phi}^T(\tau,t) = \Phi(t,\tau)$$

$$\frac{d}{d\tau}\Phi(t,\tau) = \left(\frac{d}{d\tau}\widehat{\Phi}(\tau,t)\right)^T = \left(-A^T(\tau)\widehat{\Phi}(\tau,t)\right)^T = -\widehat{\Phi}^T(\tau,t)A(\tau) = -\Phi(t,\tau)A(\tau)$$

$$\boxed{\frac{d}{d\tau}\Phi(t,\tau) = -\Phi(t,\tau)A(\tau)}$$

Problem 2.19

Let $I_1 = I_2$ and define the constants $\alpha \triangleq \frac{I_2 - I_3}{I_1}$, $\beta \triangleq \frac{1}{I_1}$ and $\gamma \triangleq \frac{1}{I_3}$. Define the state vector $\boldsymbol{x} \triangleq [w_1\ w_2\ w_3]^T$, and the input vector $\boldsymbol{u} = [u_1\ u_2\ u_3]^T$. The nonlinear state differential equations are compactly expressed as $\dot{\boldsymbol{x}} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u})$:

$$\dot{\boldsymbol{x}} = \begin{bmatrix} \alpha x_2 x_3 + \beta u_1 \\ -\alpha x_1 x_3 + \beta u_2 \\ \gamma u_3 \end{bmatrix} = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u})$$

The following pair of time functions $(\boldsymbol{x}^\star(t), \boldsymbol{u}^\star(t))$ is a solution for this system of differential equation: (**the nominal solution presented in the book appears to be erroneous, as it fails to satisfy the differential equations**)

$$\boldsymbol{x}^\star(t) = \begin{bmatrix} \cos(\alpha\omega_0 t) \\ -\sin(\alpha\omega_0 t) \\ \omega_0 \end{bmatrix}, \quad \boldsymbol{u}^\star(t) = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

This can be verified by checking the validity of $\frac{d}{dt}\boldsymbol{x}^\star(t) = \boldsymbol{f}(\boldsymbol{x}^\star(t), \boldsymbol{u}^\star(t))$.
To find the dynamics of the deviation from this nominal solution, we need to determine the dynamics of the time function $\boldsymbol{e}(t) \triangleq \boldsymbol{x}(t) - \boldsymbol{x}^\star(t)$:

$$\dot{\boldsymbol{e}}(t) = \dot{\boldsymbol{x}}(t) - \dot{\boldsymbol{x}}^\star(t) = \boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}) - \boldsymbol{f}(\boldsymbol{x}^\star, \boldsymbol{u}^\star) = \boldsymbol{f}(\boldsymbol{e}(t) + \boldsymbol{x}^\star(t), \boldsymbol{u}(t)) - \boldsymbol{f}(\boldsymbol{x}^\star(t), \boldsymbol{u}^\star(t))$$

For small deviations of $\boldsymbol{x}$ from $\boldsymbol{x}^\star$ and $\boldsymbol{u}$ from $\boldsymbol{u}^\star$:

$$\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{u}) - \boldsymbol{f}(\boldsymbol{x}^\star, \boldsymbol{u}^\star) \approx \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right|_{\substack{\boldsymbol{x}^\star \\ \boldsymbol{u}^\star}} (\boldsymbol{x} - \boldsymbol{x}^\star) + \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}\right|_{\substack{\boldsymbol{x}^\star \\ \boldsymbol{u}^\star}} (\boldsymbol{u} - \boldsymbol{u}^\star) = A(t)\boldsymbol{e}(t) + B(t)\boldsymbol{u}(t)$$

$$A(t) \triangleq \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right|_{\substack{\boldsymbol{x}^\star \\ \boldsymbol{u}^\star}} = \left.\begin{bmatrix} 0 & \alpha x_3 & \alpha x_2 \\ -\alpha x_3 & 0 & -\alpha x_1 \\ 0 & 0 & 0 \end{bmatrix}\right|_{\substack{\boldsymbol{x}^\star \\ \boldsymbol{u}^\star}} = \begin{bmatrix} 0 & \alpha\omega_0 & -\alpha\sin(\alpha\omega_0 t) \\ -\alpha\omega_0 & 0 & -\alpha\cos(\alpha\omega_0 t) \\ 0 & 0 & 0 \end{bmatrix}$$

$$B(t) \triangleq \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{u}}\right|_{\substack{\boldsymbol{x}^\star \\ \boldsymbol{u}^\star}} = \begin{bmatrix} \beta & & \\ & \beta & \\ & & \gamma \end{bmatrix}$$

The obtained linearized perturbation equation is a valid approximation for the dynamics of $\boldsymbol{e}$ for small $\boldsymbol{u}$ (around zero) and small $\boldsymbol{e}$:

$$\boxed{\dot{\boldsymbol{e}} = A(t)\boldsymbol{e} + B\boldsymbol{u}}$$

Problem 2.20

Define the state variables $x_1 \triangleq c$ and $x_2 \triangleq \dot{c}$. The dynamical system is described as

$$\dot{\boldsymbol{x}} = \begin{bmatrix} x_2 \\ t^2 x_1 - \sin(x_1) - x_1^3 x_2^2 + u \end{bmatrix}$$

$$\dot{\boldsymbol{x}} = \boldsymbol{f}(t, \boldsymbol{x}, u)$$

The pair $(\boldsymbol{x}^\star(t) = \mathbf{0}, u^\star(t) = 0)$ is a solution for this system of differential equations. As before,

$$\dot{\boldsymbol{e}} \approx \left.\frac{\partial \boldsymbol{f}}{\partial \boldsymbol{x}}\right|_{\substack{\boldsymbol{x}^\star \\ u^\star}} (\boldsymbol{x} - \boldsymbol{x}^\star) + \left.\frac{\partial \boldsymbol{f}}{\partial u}\right|_{\substack{\boldsymbol{x}^\star \\ u^\star}} (u - u^\star) = \left.\begin{bmatrix} 0 & 1 \\ t^2 - \cos(x_1) - 3x_1^2 x_2^2 & -2x_1^3 x_2 \end{bmatrix}\right|_{\substack{\boldsymbol{x}^\star \\ u^\star}} \boldsymbol{e} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u$$

$$\dot{\boldsymbol{e}} = \begin{bmatrix} 0 & 1 \\ t^2 - 1 & 0 \end{bmatrix} \boldsymbol{e} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u = A(t)\boldsymbol{e} + \boldsymbol{b}u$$

The pair $\left(\boldsymbol{x}^\star(t) = \begin{bmatrix} t \\ 1 \end{bmatrix}, u^\star(t) = \sin(t)\right)$ is also a nominal solution for this dynamical system:

$$\dot{\boldsymbol{e}} = \begin{bmatrix} 0 & 1 \\ -2t^2 - \cos(t) & -2t^3 \end{bmatrix} \boldsymbol{e} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} (u - u^\star)$$

We can set the input $u = v + u^\star$ to get $\dot{\boldsymbol{e}} = A(t)\boldsymbol{e} + \boldsymbol{b}v$.
In the equivalent discrete-time model, the input signal to the physical continuous-time process is in the staircase form (generated by the digital controller and transmitted through the hold device), that is, $u(t) = u(t_i)$ for $t \in [t_i, t_{i+1})$. Now the equivalent discrete-time model is given by

$$\begin{cases} \boldsymbol{e}(t_{k+1}) = \Phi(t_{k+1}, t_k)\boldsymbol{e}(t_k) + \left(\int_{t_k}^{t_{k+1}} \Phi(t_{k+1}, \tau)B(\tau)d\tau\right) u(t_k) \\ y(t_k) = [1 \quad 0]\boldsymbol{e}(t_k) \end{cases}$$

Observe that $y = [\ 1 \quad 0\ ]\boldsymbol{e} = \Delta c$ so that the perturbed output Δc is obtained in the measurement. But, how to compute $\Phi(t, t_0)$ for this time-varying dynamics matrix? In general, it is not possible to derive an explicit form for $\Phi(t, t_0)$. However, we can obtain the value of the matrix $\Phi(t, t_0)$ using numerical integration techniques to any desired accuracy by numerically solving the following IVP in each sampling interval:

$$\begin{cases} \dot{\Phi}(t, t_k) = A(t)\Phi(t, t_k) \\ \Phi(t_k, t_k) = I \end{cases} \qquad t \in [t_k, t_{k+1}]$$

Problem 2.21

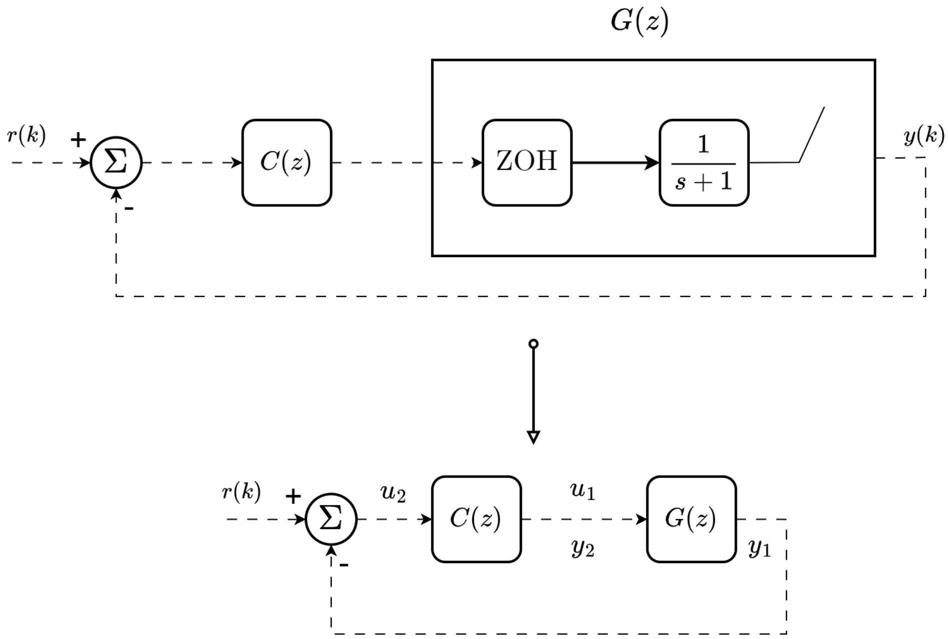


Figure 8: Example

Plant is assumed to be a continuous-time process whose dynamics is described by $\dot{x} = -x + u$. When placed into the setting depicted above (preceded by a hold device and followed by a sampler), the controller *sees* a discrete-time process.

$\xrightarrow{①}$ Form equivalent discrete-time model: $x_1(k+1) = \widehat{A}x_1(k) + \widehat{b}u_1(k)$ where $\widehat{A} = e^{AT} = e^{-T}$ and $\widehat{b} = \int_0^T e^{-\tau} d\tau = (1 - e^{-T})$ where T is the sampling period.

$$\begin{cases} x_1(k+1) = e^{-T}x_1(k) + (1 - e^{-T})u_1(k) \\ y_1(k) = x_1(k) \end{cases}$$

$\xrightarrow{②}$ State-space model of the controller:

$$\begin{cases} x_2(k+1) = u_2(k) \\ y_2(k) = -\frac{1}{T}x_2(k) + \frac{1}{T}u_2(k) \end{cases}$$

$\xrightarrow{③}$ Signal connections in the feedback loop:

$$\begin{cases} u_2(k) = r(k) - y_1(k) \\ u_1(k) = y_2(k) \end{cases}$$

Assemble the equations to generate the state-space model for the closed-loop system:

$$\begin{aligned} x_1(k+1) &= e^{-T}x_1(k) + (1 - e^{-T})\{-\frac{1}{T}x_2(k) + \frac{1}{T}\left(r(k) - x_1(k)\right)\} \\ &= \left(e^{-T} - \frac{(1 - e^{-T})}{T}\right)x_1(k) - \frac{1 - e^{-T}}{T}x_2(k) + \frac{1 - e^{-T}}{T}r(k) \end{aligned}$$

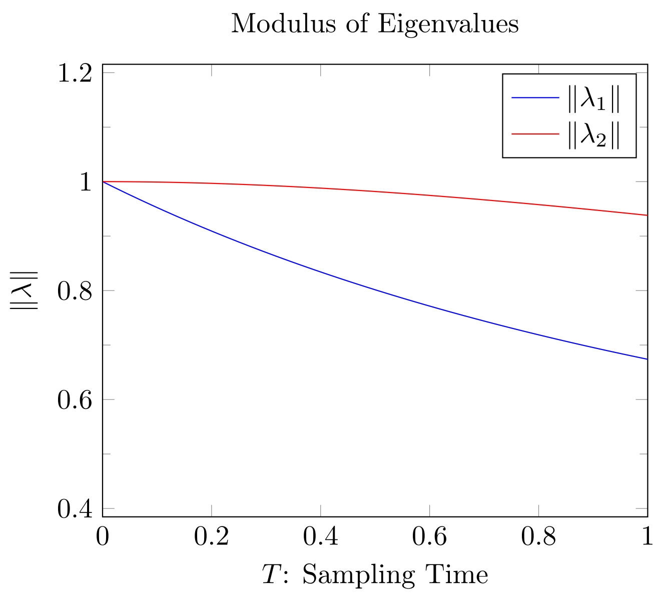
$$x_2(k+1) = -x_1(k) + r(k)$$

Let $a \triangleq e^{-T}$ and $b \triangleq 1 - e^{-T}$, then

$$\boldsymbol{x}(k+1) = \begin{bmatrix} a - {}^{b}\!/\!_{T} & -{}^{b}\!/\!_{T} \\ -1 & 0 \end{bmatrix} \boldsymbol{x}(k) + \begin{bmatrix} {}^{b}\!/\!_{T} \\ 1 \end{bmatrix} r(k)$$

$$z(k) = \begin{bmatrix} 1 & 0 \end{bmatrix} \boldsymbol{x}(k)$$

Eigenvalues of $A_d \triangleq \begin{bmatrix} a - {}^{b}\!/\!_{T} & -{}^{b}\!/\!_{T} \\ -1 & 0 \end{bmatrix}$ are plotted as function of the sampling period T in the following graph. Note that $|\lambda| < 1$ for any chosen T; The discrete-time closed loop system is stable for any choice of the T:



Problem 2.22

There is a stable state-space realization for the integration operation ($\frac{1}{s}$): $\dot{\boldsymbol{x}} = \boldsymbol{u}$. To implement this operator into the computer we need to construct a discrete-time dynamical system (a recursive algorithm) that approximates this differential equation. This falls into the domain of numerical discretization of ODEs and there are lot of techniques to achieve this: forward Euler, backward Euler, trapezoidal method, Runge-Kutta methods, etc. In the following, we use the backward Euler method.

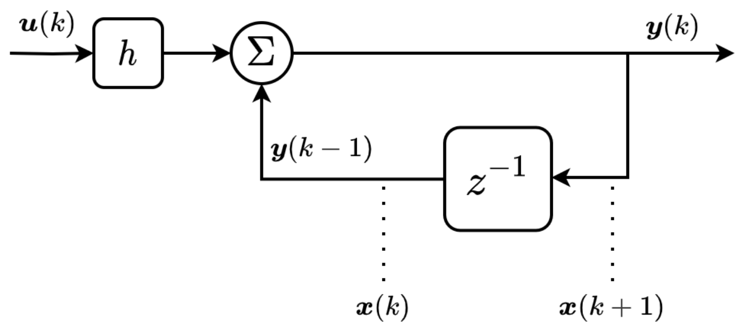
$$\dot{\boldsymbol{x}} = \boldsymbol{f}(t, \boldsymbol{x}) \xRightarrow{discretize} \boldsymbol{x}(n+1) = \boldsymbol{x}(n) + h\boldsymbol{f}(t(n+1), \boldsymbol{x}(n+1))$$

$$\dot{\boldsymbol{x}} = \boldsymbol{u} \xRightarrow{discretize} \boxed{\boldsymbol{x}(n+1) = \boldsymbol{x}(n) + h\boldsymbol{u}(n+1)}$$

h is the time-step (sampling period). To obtain state-space formulation of $\boldsymbol{x}(n+1) = \boldsymbol{x}(n) + h\boldsymbol{u}(n+1)$, block diagrams would help. Note that $\boldsymbol{x}(n)$ and $\boldsymbol{u}(n)$ are the inputs/outputs of this system, therefore we use the following symbols to be consistent with the conventional notations used in control theory.

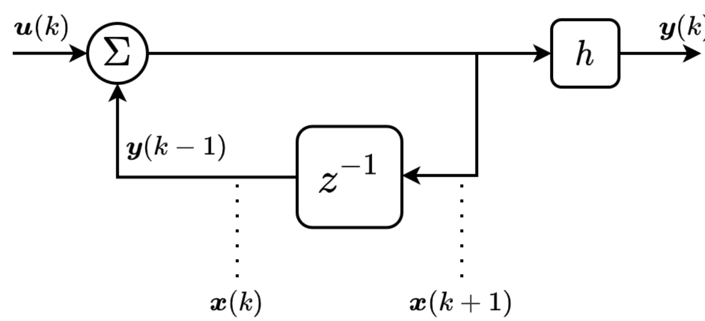
$$\boldsymbol{y}(n+1) = \boldsymbol{y}(n) + h\boldsymbol{u}(n+1) \leftrightarrow \boldsymbol{y}(k) = \boldsymbol{y}(k-1) + h\boldsymbol{u}(k)$$

$$\begin{cases} \boldsymbol{x}(k+1) = \boldsymbol{x}(k) + h\boldsymbol{u}(k) \\ \boldsymbol{y}(k) = \boldsymbol{x}(k) + h\boldsymbol{u}(k) \end{cases} \longleftrightarrow$$



Alternatively, we could use the following equivalent block diagram to obtain another equivalent state-space representation of this system:

$$\begin{cases} \boldsymbol{x}(k+1) = \boldsymbol{x}(k) + \boldsymbol{u}(k) \\ \boldsymbol{y}(k) = h\boldsymbol{x}(k) + h\boldsymbol{u}(k) \end{cases} \longleftrightarrow$$



In general, for SISO systems, replacing s in the transfer function by $\dfrac{1 - z^{-1}}{h}$ results in the backward Euler discretized model.

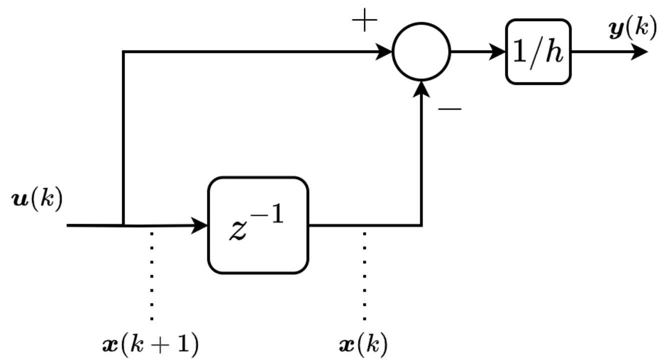
$$\frac{Y(s)}{U(s)} = \frac{1}{s} \Rightarrow \frac{Y(z)}{U(z)} = \frac{h}{1 - z^{-1}}$$

$$\Rightarrow Y(z) - z^{-1}Y(z) = hU(z)$$

$$\xRightarrow{\mathcal{Z}^{-1}\{\cdot\}} y(k) = y(k-1) + hu(k)$$

The derivative operator $G(s) = s$ has no state-space realization. Therefore, we need to use our *brain* to come up with an algorithm that performs similar operation on sequences. One such algorithm reads $y(k) = \dfrac{1}{h}(u(k) - u(k-1))$. Block-diagram helps in drawing state-space formulation.

$$\begin{cases} \boldsymbol{x}(k+1) = \boldsymbol{u}(k) \\ \boldsymbol{y}(k) = -\dfrac{1}{h}\boldsymbol{x}(k) + \dfrac{1}{h}\boldsymbol{u}(k) \end{cases} \longleftrightarrow$$



Problem 2.23

First, we obtain the state-space model of the continuous-time plant:

$$\frac{s+3}{(s+1)(s+2)} = \frac{2}{s+1} + \frac{-1}{s+2} \quad \Longrightarrow \quad \begin{aligned} \dot{\boldsymbol{x}}_p &= \begin{bmatrix} -1 & \\ & -2 \end{bmatrix} \boldsymbol{x}_p + \begin{bmatrix} 1 \\ 1 \end{bmatrix} u_p \\ y_p &= \begin{bmatrix} 2 & -1 \end{bmatrix} \boldsymbol{x}_p \end{aligned}$$

Discretize the plant:

$$\left.\begin{aligned} A_p &\triangleq e^{Ah} = \begin{bmatrix} e^{-h} & \\ & e^{-2h} \end{bmatrix} \\ \boldsymbol{b}_p &\triangleq \int_0^h e^{A\tau}\boldsymbol{b}d\tau = \int_0^h \begin{bmatrix} e^{-\tau} \\ e^{-2\tau} \end{bmatrix} d\tau = \begin{bmatrix} 1-e^{-h} \\ \frac{1}{2}(1-e^{-2h}) \end{bmatrix} \\ \boldsymbol{c}_p^T &= \begin{bmatrix} 2 & -1 \end{bmatrix} \end{aligned}\right\} \Longrightarrow \begin{cases} \boldsymbol{x}_p(k+1) = A_p\boldsymbol{x}_p(k) + \boldsymbol{b}_p u_p(k) \\ y_p(k) = \boldsymbol{c}_p^T \boldsymbol{x}_p(k) \end{cases}$$

Digital PID controller:

$$\left.\begin{aligned} I &: \begin{cases} x_1(k+1) = x_1(k) + u_c(k) \\ y_1(k) = hx_1(k) + hu_c(k) \end{cases} \\ D &: \begin{cases} x_2(k+1) = u_c(k) \\ y_2(k) = -\frac{1}{h}x_2(k) + \frac{1}{h}u_c(k) \end{cases} \end{aligned}\right\} \xrightarrow{\boldsymbol{x}_c = \begin{bmatrix} x_2 \\ x_1 \end{bmatrix}} \begin{cases} \boldsymbol{x}_c(k+1) = \overbrace{\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}}^{A_c \triangleq} \boldsymbol{x}_c(k) + \overbrace{\begin{bmatrix} 1 \\ 1 \end{bmatrix}}^{\boldsymbol{b}_c \triangleq} u_c(k) \\ y_c(k) = \underbrace{\begin{bmatrix} \frac{-k_d}{h} & hk_i \end{bmatrix}}_{\boldsymbol{c}_c^T \triangleq} \boldsymbol{x}_c(k) + \underbrace{(\frac{k_d}{h} + hk_i + k_p)}_{d_c \triangleq} u_c(k) \end{cases}$$

Form the closed-loop system by connecting signals according to the block-diagram:

$$\begin{cases} \boldsymbol{x}_c(k+1) = A_c\boldsymbol{x}_c(k) + \boldsymbol{b}_c(r - \boldsymbol{c}_p^T\boldsymbol{x}_p(k)) \\ \boldsymbol{x}_p(k+1) = A_p\boldsymbol{x}_p(k) + \boldsymbol{b}_p\{\boldsymbol{c}_c^T\boldsymbol{x}_c(k) + d_c(r - \boldsymbol{c}_p^T\boldsymbol{x}_p(k))\} \end{cases}$$

Define the augmented state vector $\boldsymbol{x}(k) \triangleq \begin{bmatrix} \boldsymbol{x}_c(k) \\ \boldsymbol{x}_p(k) \end{bmatrix}$. The dynamics of the closed-loop system is then given by:

$$\boldsymbol{x}(k+1) = \begin{bmatrix} A_c & -\boldsymbol{b}_c\boldsymbol{c}_p^T \\ \boldsymbol{b}_p\boldsymbol{c}_c^T & A_p - \boldsymbol{b}_p d_c \boldsymbol{c}_p^T \end{bmatrix} \boldsymbol{x}(k) + \begin{bmatrix} \boldsymbol{b}_c \\ \boldsymbol{b}_p d_c \end{bmatrix} r(k)$$

$$y(k) = \begin{bmatrix} \boldsymbol{0}^T & \boldsymbol{c}_p^T \end{bmatrix} \boldsymbol{x}(k)$$

Suppose that the state $\boldsymbol{x}(k)$ of the system is known at $t = t_k$: $\boldsymbol{x}(k)$ is known. We need to find $\boldsymbol{x}_p(t_k+\tau)$ for all $\tau \in [0, h)$, where h is the sampling period. Then, we can use it to compute $y_p(t_k+\tau)$: the value of the output between sampling time instants. Recall that the plant input $u_p(t)$ remains constant during the sampling interval: $u_p(t) = u_p(t_k)$ for $t \in [t_k, t_{k+1})$:

$$\begin{aligned} \boldsymbol{x}_p(t_k+\tau) &= e^{A_p\tau}\boldsymbol{x}_p(t_k) + \left(\int_0^\tau e^{A_p s}\boldsymbol{b}_p ds\right) u_p(t_k) \\ \widehat{A}(\tau) &\triangleq e^{A_p\tau} \\ \widehat{\boldsymbol{b}}(\tau) &\triangleq \int_0^\tau e^{A_p s}\boldsymbol{b}_p ds \end{aligned}$$

$$\begin{aligned} \Longrightarrow \boldsymbol{x}_p(t_k+\tau) &= \widehat{A}(\tau)\boldsymbol{x}_p(t_k) + \widehat{\boldsymbol{b}}(\tau)u_p(t_k) \\ &= \widehat{A}(\tau)\boldsymbol{x}_p(t_k) + \widehat{\boldsymbol{b}}(\tau)\left\{\boldsymbol{c}_c^T\boldsymbol{x}_c(t_k) + d_c\left(r(t_k) - \boldsymbol{c}_p^T\boldsymbol{x}_p^T(t_k)\right)\right\} \\ &= \left(\widehat{A}(\tau) - \widehat{\boldsymbol{b}}(\tau)d_c\boldsymbol{c}_p^T\right)\boldsymbol{x}_p(t_k) + \left(\widehat{\boldsymbol{b}}(\tau)\boldsymbol{c}_c^T\right)\boldsymbol{x}_c(t_k) + \left(\widehat{\boldsymbol{b}}(\tau)d_c\right) r(t_k) \end{aligned}$$

$$y_p(t_k+\tau) = \boldsymbol{c}_p^T\left(\widehat{A}(\tau) - \widehat{\boldsymbol{b}}(\tau)d_c\boldsymbol{c}_p^T\right)\boldsymbol{x}_p(t_k) + \boldsymbol{c}_p^T\left(\widehat{\boldsymbol{b}}(\tau)\boldsymbol{c}_c^T\right)\boldsymbol{x}_c(t_k) + \boldsymbol{c}_p^T\left(\widehat{\boldsymbol{b}}(\tau)d_c\right) r(t_k)$$

that the state $\boldsymbol{x}(t_k) = \begin{bmatrix} \boldsymbol{x}_c(t_k) \\ \boldsymbol{x}_p(t_k) \end{bmatrix}$ is known, the right hand side of the last relation is fully available and therefore $y_p(t_k + \tau)$ is computable for all $\tau \in [0, h)$

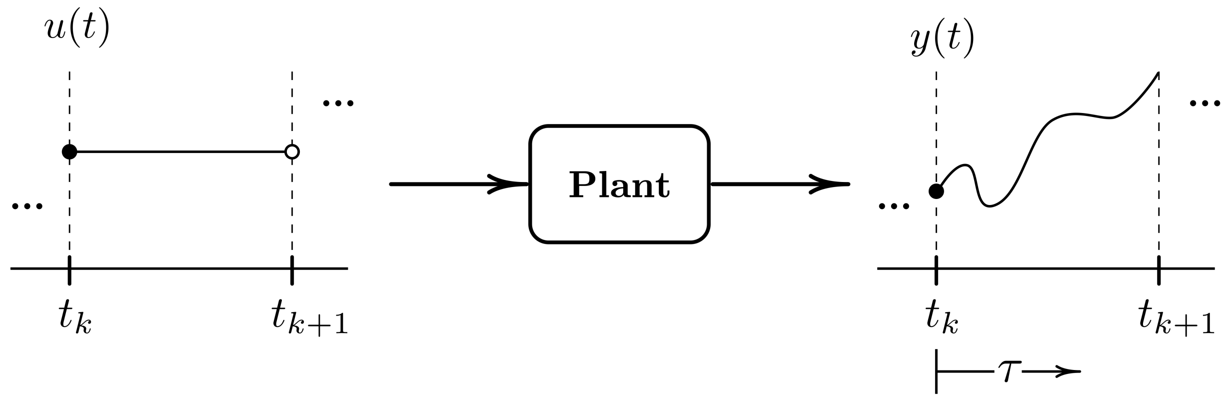


Figure 9: Example

Chapter 5

Problem 5.1

The distribution function of $\mathbf{x}$ matches that of a Gaussian random vector, therefore $\mathbf{x}$ is a Gaussian random variable; $\mathbf{x} \sim \mathcal{N}(\mathbf{0}, P_{\mathbf{xx}})$, where $P_{\mathbf{xx}} = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix}$. Let $\mathsf{z} \triangleq \mathsf{x}_1 = H\mathbf{x} = \begin{bmatrix} 1 & 0 \end{bmatrix} \mathbf{x}$. Since z is obtained by performing a linear transformation on $\mathbf{x}$, it is also a Gaussian random variable; $z \sim \mathcal{N}(0, P_{\mathsf{zz}})$ where $P_{\mathsf{zz}} = HP_{\mathbf{xx}}H^T = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} = 1$. In fact, $\mathbf{x}$ and z are jointly Gaussian, in other words, $\begin{bmatrix} \mathbf{x} \\ \mathsf{z} \end{bmatrix}$ is a Gaussian RV because it is a linear transformation of $\mathbf{x}$:

$$\begin{bmatrix} \mathbf{x} \\ \mathsf{z} \end{bmatrix} = \begin{bmatrix} I \\ H \end{bmatrix} \mathbf{x} = A\mathbf{x}$$

We can now obtain the covariance matrix of $\begin{bmatrix} \mathbf{x} \\ \mathsf{z} \end{bmatrix}$ as

$$\begin{bmatrix} P_{\mathbf{xx}} & P_{\mathbf{xz}} \\ P_{\mathbf{zx}} & P_{\mathsf{zz}} \end{bmatrix} = AP_{\mathbf{xx}}A^T = \begin{bmatrix} I \\ H \end{bmatrix} P_{\mathbf{xx}} \begin{bmatrix} I & H^T \end{bmatrix} = \begin{bmatrix} P_{\mathbf{xx}} & HP_{\mathbf{xx}} \\ P_{\mathbf{xx}}H^T & HP_{\mathbf{xx}}H^T \end{bmatrix}$$

Notice that $\mathbf{x}|\mathsf{z} = \rho$ is also a Gaussian RV (see Section 3.10) whose mean and covariance matrix are given by (3-112a) and (3-112b):

$$\boldsymbol{m}_{\mathbf{x}|\mathsf{z}}(\rho) = \boldsymbol{m}_{\mathbf{x}} + P_{\mathbf{xz}}P_{\mathsf{zz}}^{-1}(\rho - m_{\mathsf{z}}) = 0 + \begin{bmatrix} 1 \\ 1/2 \end{bmatrix} (1)(\rho - 0) = \rho \begin{bmatrix} 1 \\ 1/2 \end{bmatrix}$$

$$P_{\mathbf{x}|\mathsf{z}} = P_{\mathbf{xx}} - P_{\mathbf{xz}}P_{\mathsf{zz}}^{-1}P_{\mathbf{zx}} = \begin{bmatrix} 1 & 1/2 \\ 1/2 & 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1/2 \end{bmatrix} (1) \begin{bmatrix} 1 & 1/2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 3/4 \end{bmatrix}$$

As such, $\mathbf{x}|\mathsf{z} = \rho \sim \mathcal{N}\left(\begin{bmatrix} \rho \\ \rho/2 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 3/4 \end{bmatrix}\right)$ and for the particular measured value of $\rho = 1$, $\mathbf{x}|\mathsf{z} = 1 \sim \mathcal{N}\left(\begin{bmatrix} 1 \\ 1/2 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 3/4 \end{bmatrix}\right)$ which is consistent with intuition: Since $\mathsf{x}_1 = \mathsf{z}$ is known, there are no uncertainties in x_1 direction and thus $P_{\mathbf{x}|\mathsf{z}}$ becomes positive semidefinite and $\mathbf{x}|\mathsf{z} = \rho$ becomes a degenerate Gaussian RV (the joint distribution function collapses in x_1 direction).

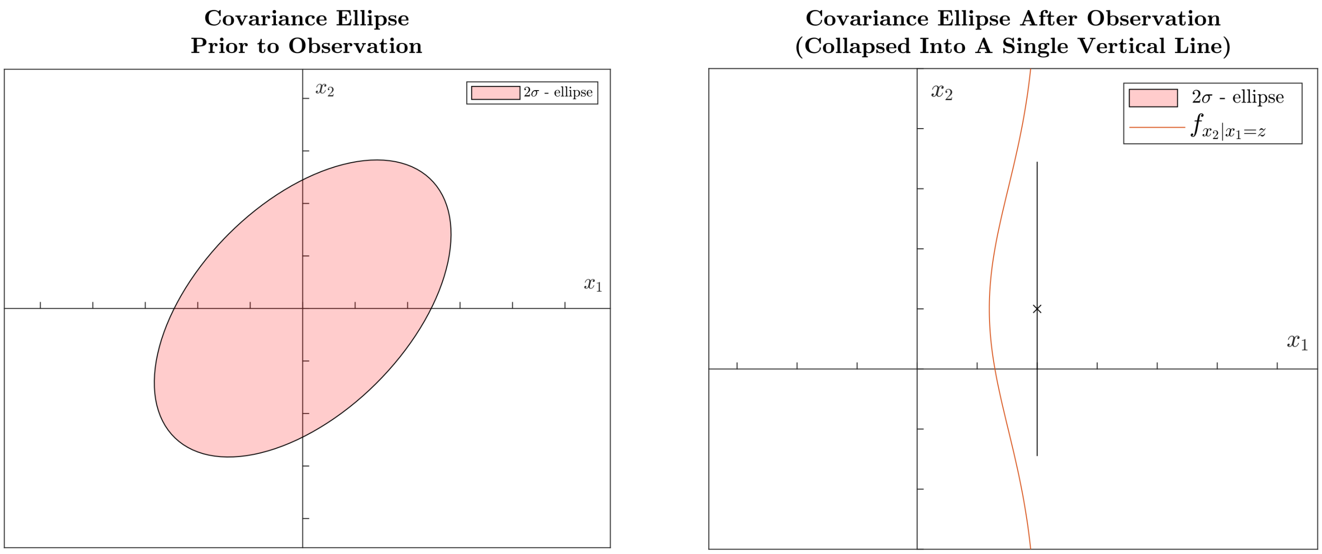


Figure 10: Covariance ellipse of $\mathbf{x}$ prior and after the measurement

Problem 5.2

Since A is symmetric, A^{-1} is also a symmetric matrix. As a result, $F^T = G^T$.

$$AA^{-1} = I \Rightarrow \begin{bmatrix} P^{-1} & H^T \\ H & -R \end{bmatrix} \begin{bmatrix} D & F \\ G^T & E \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$$

$$P^{-1}D + H^T G^T = I \tag{1}$$

$$P^{-1}F + H^T E = 0 \implies F = -PH^T E \qquad \implies F = PH^T(HPH^T + R)^{-1} \tag{2}$$

$$HD - RG^T = 0 \implies G^T = R^{-1}HD \tag{3}$$

$$HF - RE = I \tag{4}$$

(4) : using(2), $-HPH^T E - RE = I \implies E = -(HPH^T + R)^{-1}$
(1) : using(3), $P^{-1}D + H^T R^{-1} HD = I \implies D = (P^{-1} + H^T R^{-1} H)^{-1}$

To obtain the expanded form of D, we use (1), (2), (3) and the fact that $F^T = G^T$:

$$P^{-1}D = I - H^T G^T \implies D = P - PH^T G^T \implies D = P - PH^T(HPH^T + R)^{-1}HP$$

$$\implies \boxed{(P^{-1} + H^T R^{-1} H)^{-1} = P - PH^T(HPH^T + R)^{-1}HP}$$

Problem 5.3

First, we prove $(5-29)$: Post-multiply $(5-28)$ by $H^T R^{-1}$ to obtain

$$\begin{aligned}[P^{-1}+H^TR^{-1}H]^{-1}H^TR^{-1} &= PH^TR^{-1} - PH^T[HPH^T+R]^{-1}HPH^TR^{-1} \\ &= PH^T\left\{R^{-1} - \left[HPH^T+R\right]^{-1}HPH^TR^{-1}\right\} \qquad (*)\end{aligned}$$

We now show that the term is the curly braces equals $[HPH^T+R]^{-1}$:

$$\begin{aligned} I = [HPH^T+R]^{-1}[HPH^T+R] &\implies R^{-1} = [HPH^T+R]^{-1}[HPH^TR^{-1}+I] \\ &\implies R^{-1} = [HPH^T+R]^{-1} + [HPH^T+R]^{-1}HPH^TR^{-1} \\ &\implies R^{-1} - [HPH^T+R]^{-1}HPH^TR^{-1} = [HPH^T+R]^{-1} \qquad (**) \\ &\xrightarrow[(**)]{(*)} \boxed{[P^{-1}+H^TR^{-1}H]^{-1}H^TR^{-1} = PH^T[HPH^T+R]^{-1}} \end{aligned}$$

To prove $(5-30)$, Squeeze $(5-28)$ by $H(\cdot)H^T$ (to simply notation, let $\triangle \triangleq [P^{-1}+H^TR^{-1}H]$, and $\Box \triangleq [HPH^T+R]$):

$$\begin{aligned} H\triangle^{-1}H^T &= HPH^T - HPH^T\Box^{-1}HPH^T \\ &= HPH^T\{I - \Box^{-1}HPH^T\} \\ &= HPH^T\Box^{-1}R \qquad \text{(use identity(**))} \end{aligned}$$

We show that $HPH^T\Box^{-1}R = R - R\Box^{-1}R$:

$$\begin{aligned} \Box\Box^{-1} = I &\implies \Box\Box^{-1}R = R \implies (HPH^T+R)\Box^{-1}R = R \\ &\implies HPH^T\Box^{-1}R + R\Box^{-1}R = R \\ &\implies HPH^T\Box^{-1}R = R - R\Box^{-1}R. \qquad (***) \end{aligned}$$

We have now established $H\triangle^{-1}H^T = R - R\Box^{-1}R$:

$$\implies \boxed{H[P^{-1}+H^TR^{-1}H]^{-1}H^T = R - R[HPH^T+R]^{-1}R}$$

Problem 5.4

Ⓐ First, an explicit formula for K is obtained. Using the notation introduced in the previous problem and preforming the matrix multiplication to obtain the $(2,1)$ entry of P^*:

$$HP^- = \Box K^T \implies \boxed{K^T = \Box^{-1} HP^-}$$

We are only required to establish $P^- = P^+ + P^- H^T K^T$ which is the equation that corresponds to the (1,1) entry of P^+. Substituting K:

$$P^+ = P^- - P^- H^T \Box^{-1} HP^- \overset{\text{(5-28)}}{=\!=} \triangle^{-1}$$

which does indeed match (5-32). As a consequence

$$|P^*| = |P^+||HP^- H^T + R| \tag{$\star$}$$

Ⓑ

$$\begin{aligned}
&\begin{bmatrix} X_1 & X_2 \\ X_2^T & X_3 \end{bmatrix} \begin{bmatrix} I & -X_1^{-1} X_2 \\ & I \end{bmatrix} \begin{bmatrix} X_1^{-1} & \\ & O^{-1} \end{bmatrix} \begin{bmatrix} I & \\ -X_2^T X_1^{-1} & I \end{bmatrix} \\
=&\begin{bmatrix} X_1 & X_2 \\ X_2^T & X_3 \end{bmatrix} \begin{bmatrix} X_1^{-1} & -X_1^{-1} X_2 O^{-1} \\ & O^{-1} \end{bmatrix} \begin{bmatrix} I & \\ -X_2^T X_1^{-1} & I \end{bmatrix} \\
=&\begin{bmatrix} X_1 & X_2 \\ X_2^T & X_3 \end{bmatrix} \begin{bmatrix} X_1^{-1} + X_1^{-1} X_2 O^{-1} X_2^T X_1^{-1} & -X_1^{-1} X_2 O^{-1} \\ -O^{-1} X_2^T X_1^{-1} & O^{-1} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}
\end{aligned}$$

Let $A \triangleq \begin{bmatrix} X_1 & X_2 \\ X_2^T & X_3 \end{bmatrix}$, then

$$A^{-1} = \begin{bmatrix} I & -X_1^{-1} X_2 \\ & I \end{bmatrix} \begin{bmatrix} X_1^{-1} & \\ & O^{-1} \end{bmatrix} \begin{bmatrix} I & \\ -X_2^T X_1^{-1} & I \end{bmatrix}$$

This implies that

$$|A^{-1}| = |X_1^{-1}||O^{-1}| \implies \boxed{|A| = |X_1||X_3 - X_2^T X_1^{-1} X_2|}$$

Now, let $P^* = A$:

$$|P^*| = |P^-| \left| HP^- H^T + R - HP^- \underbrace{(P^-)^{-1} P^-}_{=I} H^T \right| = |P^-||R| \tag{$\star\star$}$$

Ⓒ

$$\xrightarrow[(**)]{(*)} |P^*| = |P^-||R| = |P^+| \left| HP^- H^T + R \right| \implies \frac{|HP^- H^T + R|^{1/2}}{|P^-|^{1/2} |R|^{1/2}} = \frac{1}{|P^+|^{1/2}}$$

Problem 5.6

I am pretty sure this has a simpler solution, but here we are. We adopt the notation $\mathbf{y} = \langle \mathbf{x}_1, \mathbf{x}_2, \cdots \mathbf{x}_N \rangle$ to indicate that $\mathbf{y}$ is an affine transformation of RVs $\mathbf{x}_1, \mathbf{x}_2, \cdots \mathbf{x}_N$: $\mathbf{y} = A_1\mathbf{x}_1 + A_2\mathbf{x}_2 + \cdots A_N\mathbf{x}_N + \boldsymbol{b}$ where A_is are deterministic real-valued matrices and $\boldsymbol{b}$ is a deterministic real-valued vector. If $\{\mathbf{x}_1, \mathbf{x}_2, \cdots \mathbf{x}_N\}$ are jointly Gaussian and $\mathbf{y} = \langle \mathbf{x}_1, \mathbf{x}_2, \cdots \mathbf{x}_N \rangle$, then $\{\mathbf{y}, \mathbf{x}_{j_1}, \mathbf{x}_{j_2}, \cdots \mathbf{x}_{j_n}\}$ are also jointly Gaussian for any choices of $j_i \in \{1, 2, \cdots, N\}$ and any $n \leq N$. In other words,

$$\mathbf{r} = \begin{bmatrix} \mathbf{y} \\ \mathbf{x}_{j_1} \\ \vdots \\ \mathbf{x}_{j_n} \end{bmatrix}$$

is a Gaussian RV. This is because $\mathbf{r}$ can be expressed as an affine transformation of $\mathbf{col}(\mathbf{x}_1, \mathbf{x}_2, \cdots \mathbf{x}_N)$ which is assumed to be Gaussian. As an example, let $\mathbf{y} = \langle \mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3 \rangle$ where $\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{bmatrix}$ is Gaussian. Then $\mathbf{r} = \begin{bmatrix} \mathbf{y} \\ \mathbf{x}_1 \\ \mathbf{x}_3 \end{bmatrix}$ is also Gaussian because

$$\begin{bmatrix} \mathbf{y} \\ \mathbf{x}_1 \\ \mathbf{x}_3 \end{bmatrix} = \begin{bmatrix} A_1 & A_2 & A_3 \\ I & 0 & 0 \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \\ \mathbf{x}_3 \end{bmatrix} + \begin{bmatrix} \boldsymbol{b} \\ 0 \\ 0 \end{bmatrix}$$

We now return to the problem. Proof is by induction. We first show that $\{\widehat{\mathbf{x}}(t_0^+), \mathbf{x}(t_0)\}$ are jointly Gaussian.

$$\begin{aligned} \widehat{\mathbf{x}}(t_0^+) &= \widehat{\boldsymbol{x}}_0 + K(t_0)\{\mathbf{z}(t_0) - H(t_0)\widehat{\boldsymbol{x}}_0\} \\ &= \widehat{\boldsymbol{x}}_0 + K(t_0)\{H(t_0)\mathbf{x}(t_0) + \mathbf{v}(t_0) - H(t_0)\widehat{\boldsymbol{x}}_0\} \\ &= \{K(t_0)H(t_0)\}\ \mathbf{x}(t_0) \quad + K(t_0)\ \mathbf{v}(t_0) + \{\widehat{\boldsymbol{x}}_0 - K(t_0)H(t_0)\widehat{\boldsymbol{x}}_0\} \\ &= \qquad A_1 \qquad \mathbf{x}(t_0) \quad + \quad A_2 \quad \mathbf{v}(t_0) + \qquad\qquad \boldsymbol{b} \end{aligned}$$

As a result, $\widehat{\mathbf{x}}(t_0^+) = \langle \mathbf{x}(t_0), \mathbf{v}(t_0) \rangle$. Recall that $\widehat{\boldsymbol{x}}_0$ is a real-valued deterministic vector used as the initial estimate of the state process at t_0 which is available at the beginning of the filtering operation. By assumption, the measurement noise process $\mathbf{v}(\cdot, \cdot)$ is independent of $\mathbf{x}(t_0)$ which implies that $\{\mathbf{x}(t_0), \mathbf{v}(t_0)\}$ are independent RVs and therefore are jointly Gaussian. We can now conclude $\{\widehat{\mathbf{x}}(t_0^+), \mathbf{x}(t_0)\}$ are jointly Gaussian RVs.

Let us now assume that $\{\widehat{\mathbf{x}}(t_{i-1}^+), \mathbf{x}(t_{i-1})\}$ are jointly Gaussian. We will show that $\{\widehat{\mathbf{x}}(t_i^+), \mathbf{x}(t_i)\}$ must be jointly Gaussian. By similar expansions used previously it is readily shown that $\widehat{\mathbf{x}}(t_i^+) = \langle \widehat{\mathbf{x}}(t_i^-), \mathbf{x}(t_i), \mathbf{v}(t_i) \rangle$. Realize that $\widehat{\mathbf{x}}(t_i^-) = \Phi(t_i, t_{i-1})\widehat{\mathbf{x}}(t_{i-1}^+) + \boldsymbol{u}_d(t_{i-1}) \Rightarrow \widehat{\mathbf{x}}(t_i^-) = \langle \widehat{\mathbf{x}}(t_{i-1}^+) \rangle$ ($\boldsymbol{u}_d(\cdot)$ is a deterministic signal). Also, $\mathbf{x}(t_i) = \Phi(t_i, t_{i-1})\mathbf{x}(t_{i-1}) + \mathbf{w}_d(t_{i-1}) + \boldsymbol{u}_d(t_{i-1}) \Rightarrow \mathbf{x}(t_i) = \langle \mathbf{x}(t_{i-1}), \mathbf{w}_d(t_{i-1}) \rangle$ (see 4-122). Combining these results yields

$$\widehat{\mathbf{x}}(t_i^+) = \langle \widehat{\mathbf{x}}(t_{i-1}^+), \mathbf{x}(t_{i-1}), \mathbf{w}_d(t_{i-1}), \mathbf{v}(t_i) \rangle$$

Recall that $\widehat{\mathbf{x}}(t_{i-1}^+) = f(\mathbf{Z}(t_{i-1}))$; once the value of RV $\mathbf{Z}(t_{i-1})$ is realized (the entire measurement history up to time t_{i-1} becomes available), we can proceed with the filtering algorithm to sequentially obtain state estimates up until $t = t_{i-1}$. In other words, the value of $\widehat{\mathbf{x}}(t_{i-1}^+)$ also becomes available as a real-valued vector $\widehat{\boldsymbol{x}}(t_{i-1})$. Furthermore,

1. $\mathbf{v}(t_i)$ is independent of $\mathbf{w}_d(t_{i-1})$: $\mathbf{v}(\cdot, \cdot)$ and $\mathbf{w}(\cdot, \cdot)$ are independent processes

2. $\mathbf{v}(t_i)$ is independent of $\mathbf{x}(t_{i-1})$: recall that (see problem 4-4)

$$\begin{aligned} \mathbf{x}(t_{i-1}) &= \Phi(t_{i-1}, t_0)\mathbf{x}(t_0) \\ &\quad + \sum_{k=1}^{i-1} \Phi(t_{i-1}, t_k)G_d(t_{k-1})\mathbf{w}_d(t_{k-1}) \\ &\quad + \sum_{k=1}^{i-1} \Phi(t_{i-1}, t_k)B_d(t_{k-1})\boldsymbol{u}_d(t_{k-1}) \\ &\Longrightarrow \mathbf{x}(t_{i-1}) = f\left(\mathbf{x}(t_0), \mathbf{w}_d(t_0), \cdots \mathbf{w}_d(t_{i-2})\right) \end{aligned} \qquad (\star)$$

   This shows that $\mathbf{x}(t_{i-1})$ is completely determined by $\mathbf{x}(t_0)$ and $\mathbf{w}_d(t_0), \cdots, \mathbf{w}_d(t_{i-2})$, which are independent of $\mathbf{v}(t_i)$. As a result, $\mathbf{v}(t_i)$ must be independent of any functions of $\mathbf{x}(t_0), \mathbf{w}_d(t_0), \cdots, \mathbf{w}_d(t_{i-2})$. In particular, it must independent of $\mathbf{x}(t_{i-1})$.

3. $\mathbf{w}_d(t_{i-1})$ is independent of $\mathbf{x}(t_{i-1})$:

   Equation ($\star$) shows that $\mathbf{x}(t_{i-1})$ is determined by $\mathbf{x}(t_0)$ and $\mathbf{w}_d(t_0), \cdots, \mathbf{w}_d(t_{i-2})$, which are independent of $\mathbf{w}_d(t_{i-1})$. As a result, $\mathbf{w}_d(t_{i-1})$ must be independent of any functions of $\mathbf{x}(t_0), \mathbf{w}_d(t_0), \cdots, \mathbf{w}_d(t_{i-2})$. In particular, it must independent of $\mathbf{x}(t_{i-1})$. This is also intuitively true: $\mathbf{w}_d(t_{i-1}) \triangleq \int_{t_{i-1}}^{t_i} \Phi(t_i, \tau) G(\tau) d\boldsymbol{\beta}(\tau)$ is the resultant effect of the noise process exerted on the system during the time interval $[t_{i-1}, t_i]$; it is happening after t_{i-1}.

4. Both $\mathbf{w}_d(t_{i-1})$ and $\mathbf{v}(t_i)$ are independent of $\mathbf{Z}(t_{i-1})$ (and therefore any function of it):

$$\mathbf{Z}(t_{i-1}) = \begin{bmatrix} \mathbf{z}(t_0) \\ \vdots \\ \mathbf{z}(t_{i-1}) \end{bmatrix} = \begin{bmatrix} H(t_0)\mathbf{x}(t_0) + \mathbf{v}(t_0) \\ \vdots \\ H(t_{i-1})\mathbf{x}(t_{i-1}) + \mathbf{v}(t_{i-1}) \end{bmatrix} = F\left(\mathbf{x}(t_0), \cdots, \mathbf{x}(t_{i-1}); \mathbf{v}(t_0), \cdots, \mathbf{v}(t_{i-1})\right)$$

   We have already shown that $\mathbf{v}(t_i)$ is independent of $\mathbf{x}(t_k)$ for $0 \le k \le i-1$ (you can simply extend the argument made in part 2). Coupling this with the fact that $\mathbf{v}(t_i)$ is a white process implies that $\mathbf{v}(t_i)$ is independent of $\mathbf{Z}(t_{i-1})$. Independence of $\mathbf{w}_d(t_{i-1})$ and $\mathbf{x}(t_k)$ for $0 \le k \le i-1$ was previously proved. This, along with the fact that $\mathbf{v}(\cdot,\cdot)$ and $\mathbf{w}(\cdot,\cdot)$ are independent processes implies that $\mathbf{w}_d(t_{i-1})$ is independent of $\mathbf{Z}(t_{i-1})$.

These, plus the induction assumption ($\{\hat{\mathbf{x}}(t_{i-1}^+), \mathbf{x}(t_{i-1})\}$ being jointly Gaussian) are enough to conclude that $\{\underbrace{\hat{\mathbf{x}}(t_{i-1}^+)}_{f(\mathbf{Z}(t_{i-1}))}, \mathbf{x}(t_{i-1}), \mathbf{w}_d(t_{i-1}), \mathbf{v}(t_i)\}$ are jointly Gaussian.

A Quick Note to Prove the Previous Statement

$\{\mathbf{a}, \mathbf{b}\}$ are jointly Gaussian. $\mathbf{c}$ is Gaussian and independent of both $\mathbf{a}$ and $\mathbf{b}$. Show that $\begin{bmatrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{c} \end{bmatrix}$ is Gaussian.

We show that $\mathbf{c}$ is independent of $\mathbf{m} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \end{bmatrix}$.

$$f_{\mathbf{m}|\mathbf{c}}(\boldsymbol{\zeta}|\boldsymbol{\xi}) = f_{\mathbf{a},\mathbf{b}|\mathbf{c}}(\zeta_1, \zeta_2|\boldsymbol{\xi}) = f_{\mathbf{a}|\mathbf{b},\mathbf{c}}(\zeta_1|\zeta_2, \boldsymbol{\xi}) \cdot f_{\mathbf{b}|\mathbf{c}}(\zeta_2|\boldsymbol{\xi}) = f_{\mathbf{a}|\mathbf{b}}(\zeta_1, \zeta_2) \cdot f_{\mathbf{b}}(\zeta_2) = f_{\mathbf{a},\mathbf{b}}(\zeta_1, \zeta_2) = f_{\mathbf{m}}(\boldsymbol{\zeta})$$

Therefore, $\mathbf{c}$ and $\mathbf{m}$ are independent Gaussian RVs. Thus $\begin{bmatrix} \mathbf{m} \\ \mathbf{c} \end{bmatrix} = \begin{bmatrix} \mathbf{a} \\ \mathbf{b} \\ \mathbf{c} \end{bmatrix}$ is Gaussian.

Now, consider $\begin{bmatrix} \hat{\mathbf{x}}(t_i^+) \\ \mathbf{x}(t_i) \end{bmatrix}$:

$$\begin{bmatrix} \hat{\mathbf{x}}(t_i^+) \\ \mathbf{x}(t_i) \end{bmatrix} = \begin{bmatrix} \hat{\mathbf{x}}(t_i^+) \\ \Phi(t_i, t_{i-1})\mathbf{x}(t_{i-1}) + \mathbf{w}_d(t_{i-1}) + \boldsymbol{u}_d(t_{i-1}) \end{bmatrix} = \begin{bmatrix} I & 0 & 0 \\ 0 & \Phi(t_i, t_{i-1}) & I \end{bmatrix} \begin{bmatrix} \hat{\mathbf{x}}(t_i^+) \\ \mathbf{x}(t_{i-1}) \\ \mathbf{w}_d(t_{i-1}) \end{bmatrix} + \begin{bmatrix} 0 \\ \boldsymbol{u}_d(t_{i-1}) \end{bmatrix}$$

But $\begin{bmatrix} \hat{\mathbf{x}}(t_i^+) \\ \mathbf{x}(t_{i-1}) \\ \mathbf{w}_d(t_{i-1}) \end{bmatrix}$ is Gaussian; this is because $\hat{\mathbf{x}}(t_i^+) = \langle \hat{\mathbf{x}}(t_{i-1}^+), \mathbf{x}(t_{i-1}), \mathbf{w}_d(t_{i-1}), \mathbf{v}(t_i) \rangle$ and it was previously shown that $\{\hat{\mathbf{x}}(t_{i-1}^+), \mathbf{x}(t_{i-1}), \mathbf{w}_d(t_{i-1}), \mathbf{v}(t_i)\}$ are jointly Gaussian. This implies that $\begin{bmatrix} \hat{\mathbf{x}}(t_i^+) \\ \mathbf{x}(t_i) \end{bmatrix} = T \begin{bmatrix} \hat{\mathbf{x}}(t_i^+) \\ \mathbf{x}(t_{i-1}) \\ \mathbf{w}_d(t_{i-1}) \end{bmatrix} + \boldsymbol{b}$ is a Gaussian RV and the proof is complete.

The next part of the problem is easy to show. Note that $\hat{\mathbf{x}}(t_i^+) = f(\mathbf{Z}(t_i))$. As such, $\hat{\mathbf{x}}(t_i^+)|\mathbf{Z}(t_i) = Z_i$ is the deterministic real-valued vector $\hat{\boldsymbol{x}}(t_i^+)$:

$$\mathbb{E}\left[\, \mathbf{x}(t_i)\hat{\mathbf{x}}(t_i^+)^T \;\middle|\; \mathbf{Z}(t_i) = Z_i \,\right] = \underbrace{\mathbb{E}[\, \mathbf{x}(t_i) \mid \mathbf{Z}(t_i) = Z_i \,]}_{\hat{\boldsymbol{x}}(t_i^+) \text{ by definition}} \hat{\boldsymbol{x}}(t_i^+)^T = \hat{\boldsymbol{x}}(t_i^+)\hat{\boldsymbol{x}}(t_i^+)^T$$

Problem 5.8

Define the voltage of capacitors as state variables: $x_1 \triangleq V_A$, $x_2 \triangleq V_B$:

$$\dot{x}_1 = \frac{1}{c_1} i_2$$
$$\dot{x}_2 = \frac{1}{c_2}(i_1 - i_2)$$

The measurement equation reads $y = x_2$.

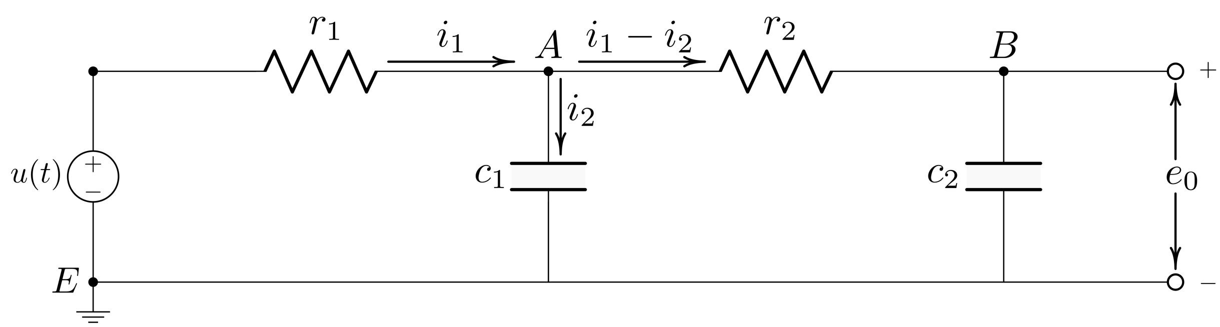


Figure 11: Circuit diagram

Circuit equations are:

$$V_E + u - r_1 i_1 = V_A = x_1 \Rightarrow i_1 = \frac{1}{r_1} u - \frac{1}{r_1} x_1$$
$$V_A - r_2(i_1 - i_2) = V_B \Rightarrow x_1 - r_2(i_1 - i_2) = x_2 \Rightarrow i_1 - i_2 = \frac{1}{r_2}(x_1 - x_2)$$

The state-space model of the system is obtained as:

$$\dot{x}_1 = -\frac{1}{c_1}\left(\frac{1}{r_1} + \frac{1}{r_2}\right) x_1 + \frac{1}{c_1 r_2} x_2 + \frac{1}{c_1 r_1} u$$
$$\dot{x}_2 = \frac{1}{c_2 r_2} x_1 - \frac{1}{c_2 r_2} x_2$$
$$y = x_2$$

Substituting given values for the parameters ($r_1 = r_2 = c_1 = c_2 = 1$) and assuming a WGN at the input and also in the measurement:

$$\dot{\mathsf{x}} = \begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix} \mathsf{x} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} \mathsf{w}$$

$$\mathsf{y} = \begin{bmatrix} 0 & 1 \end{bmatrix} \mathsf{x} + \mathsf{v}$$

As specified in the problem description, the system starts up with no charge in the capacitors and thus the voltages across both capacitors are initially zero. We can now setup the filtering algorithm by identifying $F(t) = \begin{bmatrix} -2 & 1 \\ 1 & -1 \end{bmatrix}$, $B(t) \equiv 0$, $G(t) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, $H(t_i) = \begin{bmatrix} 0 & 1 \end{bmatrix}$, $Q(t) = 2$, $\hat{\boldsymbol{x}}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, $P_0 = \epsilon I$.

Since the system is time-invariant and the sampling is uniform ($T_s = \Delta t = \frac{1}{2}$), we can readily compute $F_d \triangleq \Phi(t_{i+1}, t_i)$ as

$$\Phi(t_{i+1}, t_i) = e^{F(t_{i+1} - t_i)} = e^{FT_s} = \begin{bmatrix} 0.4238 & 0.2487 \\ 0.2487 & 0.6725 \end{bmatrix}$$

Furthermore, since noise processes are stationary,

$$Q_d \triangleq \int_{t_{i-1}}^{t_i} \Phi(t_i, \tau) G(\tau) Q(\tau) G^T(\tau) \Phi^T(t_i, \tau) d\tau = \int_0^{Ts} e^{F\sigma} G Q G^T e^{F^T\sigma} d\sigma = \begin{bmatrix} 0.4561 & 0.0917 \\ 0.0917 & 0.0299 \end{bmatrix}$$

Part a)

In the case of perfect measurements ($R(t_i) \equiv 0$), for the filtering algorithm to run smoothly without any numerical difficulties, it is sufficient (as suggested in section 5.10) that $P_0 > 0$ and $Q_d > 0$. We may chose any (small; since we *know* that the system is initially at rest) positive value for ϵ in $P_0 = \epsilon I$, for $P_0 > 0$ to be satisfied.

$$P(t_i^-) = F_d P(t_{i-1}^+) F_d^T + Q_d \qquad \text{(Time Update Equations)}$$

$$P(t_i^+) = P(t_i^-) - P(t_i^-) H^T \left[H P(t_i^-) H^T \right]^{-1} H P(t_i^-) \qquad \text{(Measurement Update Equations)}$$

Running these recursive equations for two seconds over the uniform time-grid $\{0, 0.5, 1, 1.5, 2\}$ with initial conditions $P(t_0^-) = P(0^-) = P_0 = I$ yields

$$P(0^+) = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},\ P(0.5^+) = \begin{bmatrix} 0.2121 & 0 \\ 0 & 0 \end{bmatrix},\ P(1^+) = \begin{bmatrix} 0.1915 & 0 \\ 0 & 0 \end{bmatrix},\ P(1.5^+) = \begin{bmatrix} 0.1904 & 0 \\ 0 & 0 \end{bmatrix},\ P(2^+) = \begin{bmatrix} 0.1903 & 0 \\ 0 & 0 \end{bmatrix}$$

From this point onward ($t \geq 2$) steady-state condition is essentially reached. Since the measurement of x_2 is perfect, the uncertainty in the estimate of x_2 is identically zero for all times. Also, the steady-state value for the Kalman gain is obtained as $K_\infty^T = \begin{bmatrix} 2.683 & 1 \end{bmatrix}$. The following is the filtering of a sample process:

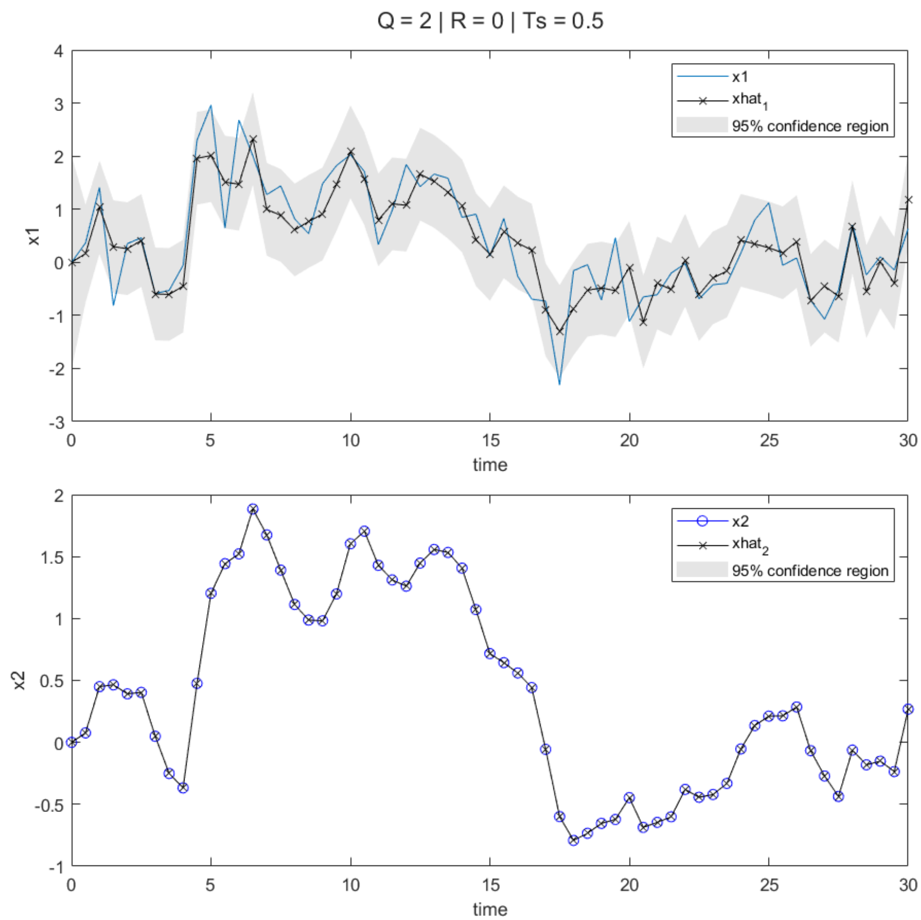


Figure 12: Estimation of the voltage across c_1 ($\hat{x}_1$) by the perfect measurement of the c_2 voltage

$(1,1)$-entries of $P(t_i^+)$ are the variance of the error in the estimate of x_1.

Part b)

We now assume that the measurement is noisy ($R(t_i) = 0.2$) and set up covariance update equations as:

$$P(t_i^-) = F_d P(t_{i-1}^+) F_d^T + Q_d \qquad \text{(Time Update Equations)}$$

$$P(t_i^+) = P(t_i^-) - P(t_i^-) H^T \left[H P(t_i^-) H^T + R \right]^{-1} H P(t_i^-) \qquad \text{(Measurement Update Equations)}$$

Covariance matrices are obtained as:

$$P(0.0^+) = \begin{bmatrix} 1.0000 & 0.0000 \\ 0.0000 & 0.1667 \end{bmatrix} \quad P(0.5^+) = \begin{bmatrix} 0.5081 & 0.1226 \\ 0.1226 & 0.0910 \end{bmatrix} \quad P(1.0^+) = \begin{bmatrix} 0.4588 & 0.1182 \\ 0.1182 & 0.0835 \end{bmatrix}$$

$$P(1.5^+) = \begin{bmatrix} 0.4552 & 0.1162 \\ 0.1162 & 0.0808 \end{bmatrix} \quad P(2.0^+) = \begin{bmatrix} 0.4550 & 0.1161 \\ 0.1161 & 0.0800 \end{bmatrix}$$

Beyond this point ($t \geq 2$) steady-state condition is essentially reached:

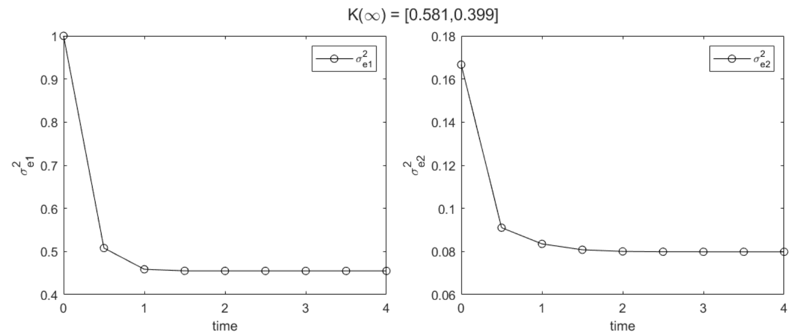


Figure 13: Error variances in the estimates for x_1 and x_2

As expected, the variance of the error in the estimate of x_1 is higher in this case (0.455 compared to 0.1903). Also, the steady-state value for the Kalman gain is obtained as $K_\infty^T = \begin{bmatrix} 0.58 & 0.4 \end{bmatrix}$ which is lower than the previous case (not counting too much on the observed values due to noisy measurements).

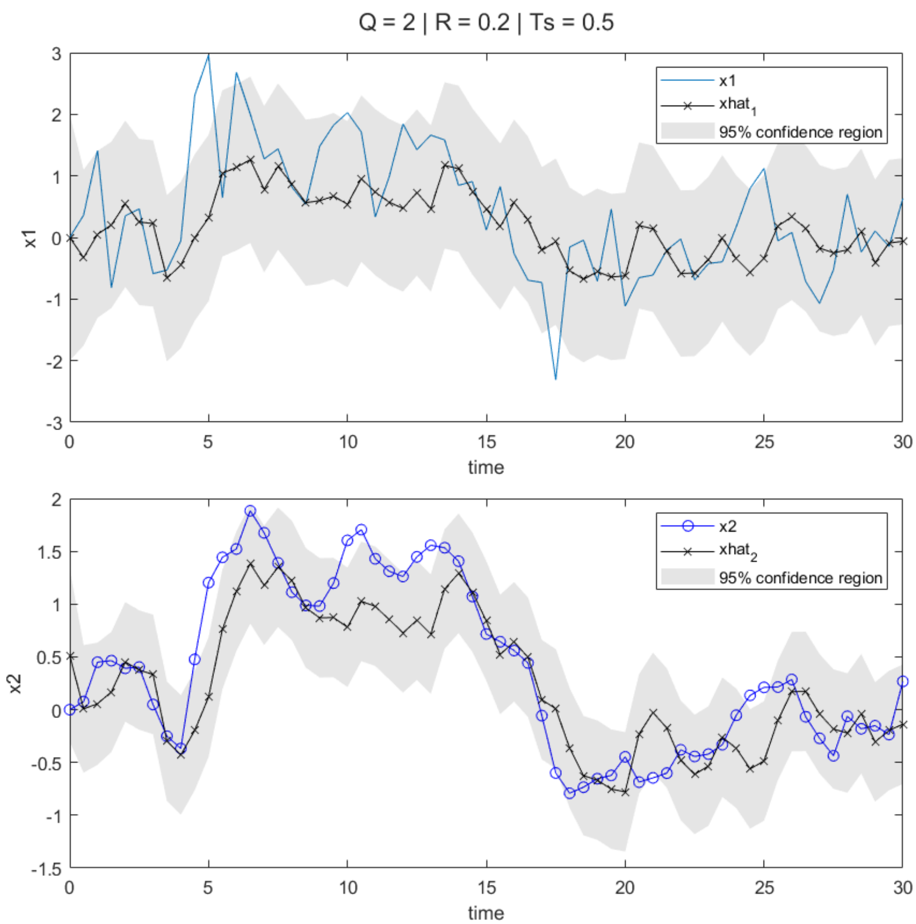


Figure 14: Estimation of the voltage across c_1 ($\hat{x}_1$) by the noisy measurement of the c_2 voltage

Problem 5.9

State-space equations of the given dynamical system are obtained by defining the state variables $\mathsf{x}_1 \triangleq \mathsf{y}$ and $\mathsf{x}_2 \triangleq \dot{\mathsf{y}}$, and the state vector $\mathbf{x} \triangleq [\mathsf{x}_1 \quad \mathsf{x}_2]^T$. Then,

$$\dot{\mathsf{x}}_1 = \dot{\mathsf{y}} = \mathsf{x}_2$$
$$\dot{\mathsf{x}}_2 = \ddot{\mathsf{y}} = -\mathsf{y} = -\mathsf{x}_1$$

Thus, the state equation is:

$$\dot{\mathbf{x}} = \begin{bmatrix} \dot{\mathsf{x}}_1 \\ \dot{\mathsf{x}}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \mathbf{x}$$

The initial state is given as $\mathbf{x}(t_0) \sim \mathcal{N}(\widehat{\mathbf{x}}_0, P_0)$, where:

$$\widehat{\mathbf{x}}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad P_0 = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix}$$

Appending the measurement equation, the state-space model is obtained as:

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \mathbf{x}(t)$$
$$\mathsf{z}(t_i) = \begin{bmatrix} 1 & 0 \end{bmatrix} \mathbf{x}(t_i) + \mathsf{v}(t_i)$$

The optimal filtering algorithms can now be set up by identifying: $F(t) = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, $B(t) \equiv 0$, $G(t) \equiv 0$, $H(t_i) = \begin{bmatrix} 1 & 0 \end{bmatrix}$, $Q(t) \equiv 0$ (process noise covariance), $R(t_i) \equiv 1$ (measurement noise variance). Since F is constant, we can compute the state transition matrix $\Phi(t_i, t_{i-1})$:

$$\Phi(t_i, t_{i-1}) = e^{F(t_i - t_{i-1})}$$

Let $\Delta t_i = t_i - t_{i-1}$. Then,

$$\Phi(\Delta t_i) = e^{F\Delta t_i} = \begin{bmatrix} \cos(\Delta t_i) & \sin(\Delta t_i) \\ -\sin(\Delta t_i) & \cos(\Delta t_i) \end{bmatrix}$$

Let us for simplicity assume a uniform sampling of the process, so $\Delta t_i = T_s$ (a constant sampling period). Then $\Phi \triangleq e^{FT_s}$ becomes a constant matrix. The time update equations are:

$$\widehat{\mathbf{x}}(t_i^-) = \Phi \widehat{\mathbf{x}}(t_{i-1}^+)$$
$$P(t_i^-) = \Phi P(t_{i-1}^+) \Phi^T$$

The measurement update equations are:

$$K(t_i) = P(t_i^-) H^T (H P(t_i^-) H^T + R)^{-1}$$
$$\widehat{\boldsymbol{x}}(t_i^+) = \widehat{\boldsymbol{x}}(t_i^-) + K(t_i)[\boldsymbol{z}(t_i) - H\widehat{\boldsymbol{x}}(t_i^-)]$$
$$P(t_i^+) = (I - K(t_i)H) P(t_i^-)$$

Specifically, since $H = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and $R = 1$:

$$H P(t_i^-) H^T + R = P_{11}(t_i^-) + 1$$

$$K(t_i) = \frac{1}{P_{11}(t_i^-) + 1} \begin{bmatrix} P_{11}(t_i^-) \\ P_{21}(t_i^-) \end{bmatrix}$$

To determine an explicit relation for the $\dot{\mathsf{y}}(t_i)$ estimator (which is $\widehat{\mathsf{x}}_2(t_i^+)$), we can write out each component of the previous vector relations:

$$\widehat{x}_2(t_i^+) = \widehat{x}_2(t_i^-) + \frac{P_{21}(t_i^-)}{P_{11}(t_i^-) + 1} [z(t_i) - \widehat{x}_1(t_i^-)]$$

Some remarks are in order:

- The given dynamical system represents an undamped oscillator. The state vector rotates around the origin with unit speed (angular frequency $\omega = 1$).

- We are making noisy measurements of the position of the oscillator (x_1) in the hope that we can estimate its *velocity* (x_2) in the long run.
- We can interpret this optimality (among other interpretations) as the most likely value of the velocity of the system given the history of noisy position measurements.
- Notice that Φ is in fact a rotation matrix; $\Phi\widehat{\mathbf{x}}(t_{i-1}^+)$ rotates the estimate $\widehat{\mathbf{x}}(t_{i-1}^+)$ by T_s radians in the clockwise direction (due to the $-sin(\Delta t_i)$ term in the Φ_{21} position, this corresponds to the system dynamics $\dot{\mathsf{x}}_2 = -\mathsf{x}_1$).
- Furthermore, $\Phi P(t_{i-1}^+)\Phi^T$ is a similarity transformation (recall that rotation matrices are orthonormal: $\Phi^T = \Phi^{-1}$). This also implies that in periods without measurement, the uncertainty in the state of the system remains constant (recall that eigenvalues are similarity-invariant).

Now, if sampling is performed every $T_s = 2k\pi$ seconds (for integer k), we are essentially making observations of the x_1 component of a "fixed" state vector in the state space. In this case, $T_s = 2k\pi \Rightarrow \Phi = I$ (identity matrix, meaning a complete rotation and back to the same spot). Then, the time update for covariance becomes:

$$P(t_i^-) = \Phi P(t_{i-1}^+)\Phi^T = IP(t_{i-1}^+)I^T = P(t_{i-1}^+)$$

The measurement update for the inverse covariance (Information Filter form) is:

$$P^{-1}(t_i^+) = P^{-1}(t_i^-) + H^T R^{-1} H$$

With $H = \begin{bmatrix}1 & 0\end{bmatrix}$ and $R = 1$, $H^T R^{-1} H = \begin{bmatrix}1\\0\end{bmatrix}(1)\begin{bmatrix}1 & 0\end{bmatrix} = \begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}$. So,

$$P^{-1}(t_i^+) = P^{-1}(t_{i-1}^+) + \begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}$$

Solving this recursive relation, yields $P^{-1}(t_k^+)$:

$$P^{-1}(t_k^+) = P^{-1}(t_0) + k\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}$$

As $k \to +\infty$:

$$P_{11}(t_k^+) \to 0$$

$$P_{22}(t_k^+) \text{ remains intact (i.e., } P_{22}(t_0))$$

This means that we can estimate the position x_1 with increasing accuracy, but our knowledge of the velocity x_2 does not improve under these specific sampling conditions ($T_s = 2k\pi$).

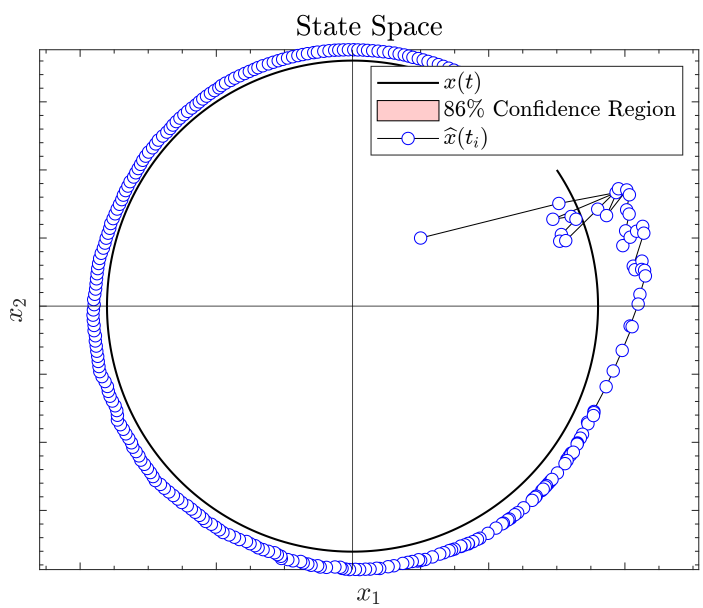


Figure 15: Estimating the states of the harmonic oscillator: State-Space view

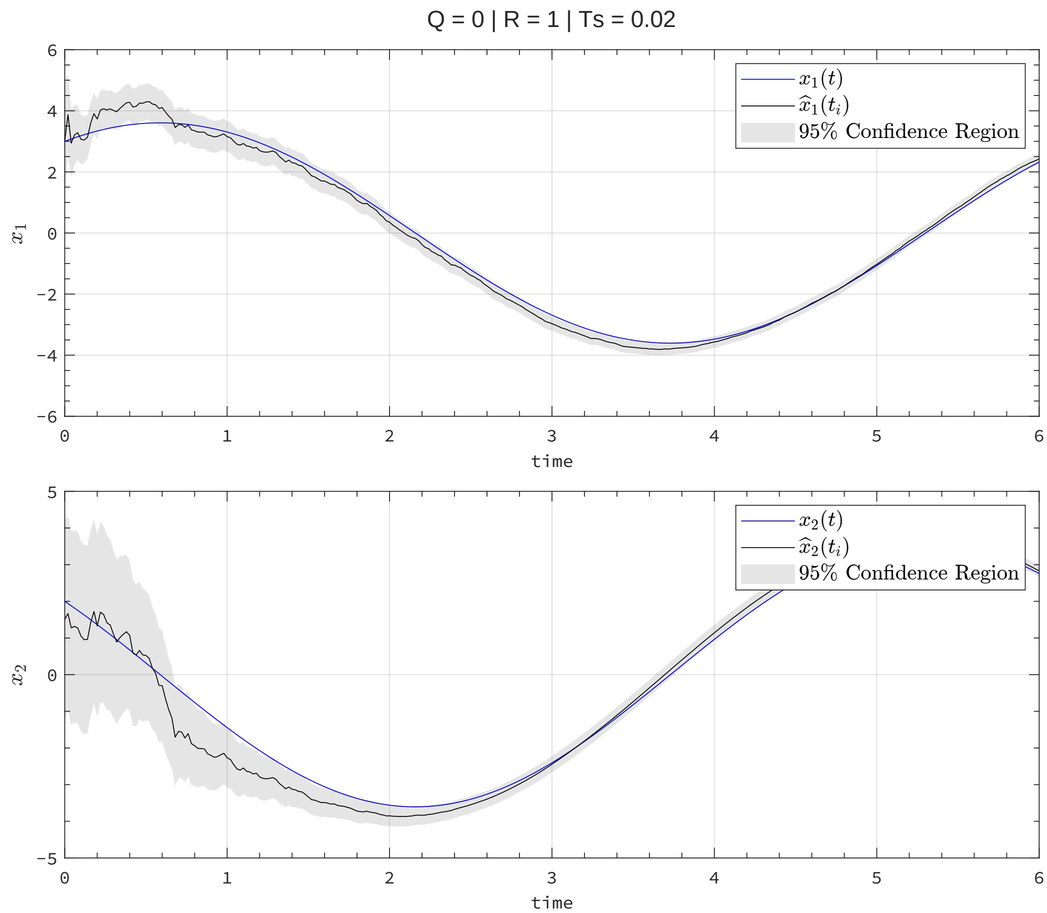


Figure 16: Estimating the states of the harmonic oscillator

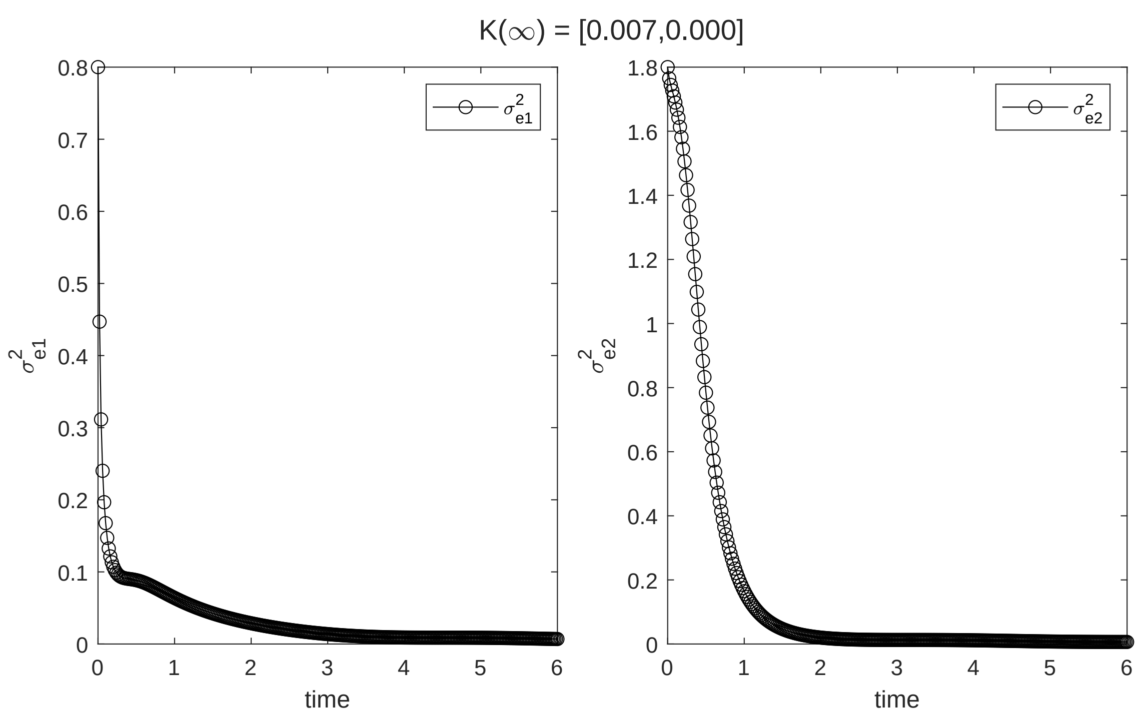


Figure 17: Variance of the error in the estimates of position and velocity

Problem 5.12

State space model of the given system in controllable canonical form is given by:

$$G(s) = \frac{\overset{c_1}{\downarrow}\;\;\; \overset{c_0}{\downarrow}}{} \frac{0.3s + 0.003}{\underset{\uparrow\, a_1}{s^2 + 0.006s} + \underset{\uparrow\, a_0}{0.003}} \qquad \boldsymbol{a} = \begin{bmatrix} a_0 \\ a_1 \end{bmatrix} = \begin{bmatrix} 0.003 \\ 0.006 \end{bmatrix} \qquad A = \left[\begin{array}{c|c} \mathbf{0} & \mathbf{I} \\ \hline \multicolumn{2}{c}{-\boldsymbol{a}^T} \end{array}\right] = \begin{bmatrix} 0 & 1 \\ -0.003 & -0.006 \end{bmatrix}$$

$$\boldsymbol{c} = \begin{bmatrix} c_0 \\ c_1 \end{bmatrix} = \begin{bmatrix} 0.003 \\ 0.3 \end{bmatrix} \qquad \boldsymbol{b} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad \boldsymbol{c}^T = \begin{bmatrix} 0.003 & 0.3 \end{bmatrix}$$

Let $\widetilde{\boldsymbol{x}} = \begin{bmatrix} \tilde{x}_1 & \tilde{x}_2 \end{bmatrix}^T$ be the state vector for the process dynamics, then the system model is given by

$$\begin{aligned} \dot{\widetilde{\boldsymbol{x}}}(t) &= A\widetilde{\boldsymbol{x}}(t) + \boldsymbol{b}h_c(t) \\ h(t) &= \boldsymbol{c}^T\widetilde{\boldsymbol{x}}(t) \end{aligned}$$

In response to a step reference command ($h_c(s) = \frac{\alpha}{s}$), asymptotic tracking is achieved ($h(\infty) = \lim_{s\to 0} s\frac{\alpha}{s}G(s) = \alpha$), however, since the poles are located at $\{-0.003\pm 0.0547j\}$ the response signal exhibits a slowly damped ($\zeta \approx 0.055 \ll 1$) and oscillatory behavior with a long settling time ($T_s \approx 1870s$).

The input signal $h_c(t)$ is a stochastic process $\mathsf{h_c}(t)$ comprised of two components: a random bias $\mathsf{h_{c0}}$, and a zero-mean white Gaussian noise process $\delta\mathsf{h_c}(t)$. The random bias $\mathsf{h_{c0}}$ is modeled as the output of an integrator with no inputs:

$$\begin{cases} \dot{\mathsf{h}}_{\mathsf{c0}}(t) = 0 \\ \mathsf{h_{c0}}(t_0) \sim \mathcal{N}(\mu, P) \end{cases}$$

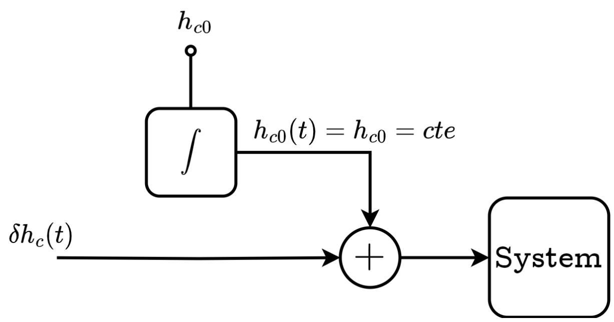


Figure 18: Input signal

With this input, the stochastic model is given by

$$\begin{aligned} \dot{\widetilde{\mathbf{x}}}(t) &= A\widetilde{\mathbf{x}}(t) + \boldsymbol{b}\mathsf{h_{c0}}(t) + \boldsymbol{b}\delta\mathsf{h_c}(t) \\ \mathsf{h}(t) &= \boldsymbol{c}^T\widetilde{\mathbf{x}}(t) \end{aligned}$$

We define the augmented state vector $\mathbf{x} \triangleq \begin{bmatrix} \widetilde{\mathbf{x}} \\ \mathsf{h_{c0}} \end{bmatrix}$. $\mathbf{x}$ satisfies the following stochastic differential equation:

$$\dot{\mathbf{x}}(t) = \begin{bmatrix} \dot{\widetilde{\mathbf{x}}} \\ \dot{\mathsf{h}}_{\mathsf{c0}} \end{bmatrix} = \begin{bmatrix} A\widetilde{\mathbf{x}} + \boldsymbol{b}\mathsf{h_{c0}} + \boldsymbol{b}\delta\mathsf{h_c} \\ 0 \end{bmatrix} = \begin{bmatrix} A & \boldsymbol{b} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \widetilde{\mathbf{x}} \\ \mathsf{h_{c0}} \end{bmatrix} + \begin{bmatrix} \boldsymbol{b} \\ 0 \end{bmatrix} \delta\mathsf{h_c}(t) = F\mathbf{x}(t) + G\mathsf{w}(t)$$

The measurement equation is

$$\mathsf{z}(t) = \mathsf{h_m}(t) = \begin{bmatrix} \boldsymbol{c}^T & 0 \end{bmatrix} \begin{bmatrix} \widetilde{\mathbf{x}} \\ \mathsf{h_{c0}} \end{bmatrix} + \delta_{\mathsf{m}}(t) = H\mathbf{x}(t) + \mathsf{v}(t)$$

As such, the augmented system model can be compactly expressed as

$$\dot{\mathsf{x}}(t) = F\mathsf{x}(t) + G\mathsf{w}(t)$$
$$\mathsf{z}(t) = H\mathsf{x}(t) + \mathsf{v}(t)$$

With $F = \begin{bmatrix} 0 & 1 & 0 \\ -0.003 & -0.006 & 1 \\ 0 & 0 & 0 \end{bmatrix}$, $G = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$, $H = \begin{bmatrix} 0.003 & 0.3 & 0 \end{bmatrix}$, $Q(t) \equiv 400$, $R_c(t) \equiv 900$, the continuous-time filtering equations can now be set up:

$$\begin{aligned}
\dot{P}(t) &= FP(t) + P(t)F^T + GQG^T - P(t)H^T R_c^{-1} HP(t) \\
\dot{\widehat{\boldsymbol{x}}}(t) &= F\widehat{\boldsymbol{x}}(t) + P(t)H^T R_c^{-1} \left[z(t) - H\widehat{\boldsymbol{x}}(t)\right] \\
\widehat{h}(t) &= H\widehat{\boldsymbol{x}}(t)
\end{aligned}$$

These equations are solved using the initial conditions:

$$\widehat{\boldsymbol{x}}(t_0) = \begin{bmatrix} \widetilde{\boldsymbol{x}}_0 \\ 10\,000 \end{bmatrix}$$
$$P(t_0) = \begin{bmatrix} \widetilde{P}_0 & 0 \\ 0 & 250\,000 \end{bmatrix}$$

where $\widetilde{\boldsymbol{x}}_0$ is an initial estimate of the state process with its corresponding confidence level expressed through $\widetilde{P}_0$ (it can be related to your initial estimate of the altitude). $\widehat{\boldsymbol{x}}(t)$ is the optimal estimate of the state process $\mathsf{x}(t)$ at time t. The optimal (also minimum variance) estimate of the altitude $\mathsf{h}(t)$ can be obtained by evaluating $\widehat{h}(t) = H\widehat{\boldsymbol{x}}(t)$ at each desired point in time.

With discrete-time measurements, the state model is

$$\dot{\mathsf{x}}(t) = F\mathsf{x}(t) + G\mathsf{w}(t)$$
$$\mathsf{z}(t_i) = H\mathsf{x}(t_i) + \mathsf{v}(t_i)$$

Since the system is time-invariant and the sampling is uniform ($T_s = \Delta t = 1$), we can readily compute $F_d \triangleq \Phi(t_{i+1}, t_i)$ as

$$\Phi(t_{i+1}, t_i) = e^{F(t_{i+1}-t_i)} = e^{FT_s} = \begin{bmatrix} 0.9985 & 0.9965 & 0.4989 \\ -0.0030 & 0.9925 & 0.9965 \\ 0 & 0 & 1.0000 \end{bmatrix}$$

Furthermore, since noise processes are stationary,

$$Q_d \triangleq \int_{t_{i-1}}^{t_i} \Phi(t_i, \tau)G(\tau)Q(\tau)G^T(\tau)\Phi^T(t_i, \tau)d\tau = \int_0^{Ts} e^{F\sigma}GQG^T e^{F^T\sigma} d\sigma = \begin{bmatrix} 132.6554 & 198.6055 & 0 \\ 198.6055 & 397.2128 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

With that, the filtering algorithm can be set up:[1]

$$P(t_i^-) = F_d P(t_{i-1}^+) F_d^T + Q_d \qquad \text{(Time Update Equations)}$$
$$K(t_i) = P(t_i^-)H^T \left[HP(t_i^-)H^T + R(t_i)\right]^{-1}$$
$$P(t_i^+) = P(t_i^-) - K(t_i)HP(t_i^-) \qquad \text{(Measurement Update Equations)}$$

$$\widehat{\boldsymbol{x}}(t_i^-) = F_d\ \widehat{\boldsymbol{x}}(t_{i-1}^+) \qquad \text{(Time Update Equations)}$$
$$\widehat{\boldsymbol{x}}(t_i^+) = \widehat{\boldsymbol{x}}(t_i^-) + K(t_i)\left[\boldsymbol{z}_i - H\widehat{\boldsymbol{x}}(t_i^-)\right] \qquad \text{(Measurement Update Equations)}$$
$$\widehat{h}(t_i) = H\widehat{\boldsymbol{x}}(t_i) \qquad \text{(Output Estimate)}$$

Below is a simulation of the system with the following parameters:

[1] Unlike the convention used in the book, we assume that the first measurement is taken at $t = t_0$. This is also the time point at which an estimate of the state process ($\widehat{\boldsymbol{x}}_0, P_0$) is available. As such, the filtering algorithm begins by evaluation of *measurement update equations*.

- True value of the altitude at time t_0 (unknown to the observer) is $h(t_0) = 6\,000$ ft
- Initial Estimate (at time t_0) of the altitude is $\widehat{h}(t_0) = 2\,000$ ft with standard deviation of $1\,000$ ft equivalent to a variance of $1\,000\,000$ ft^2
- True value of the commanded altitude (unknown to observer) is $h_c(t_0) = 12\,000$ ft
- Initial Estimate (at time t_0) of the commanded altitude is $\widehat{h}_c(t_0) = 10\,000$ ft with standard deviation of 500 ft equivalent to a variance of $250\,000$ ft^2

The filtering algorithm begins with the following initial estimates over the uniform time-grid $\{0, 1, 2, \cdots, 1\,000\}$.

$$\widehat{\boldsymbol{x}}(t_0) = \begin{bmatrix} 0 \\ 6\,666.7 \\ 10\,000 \end{bmatrix} \Longrightarrow \widehat{h}(t_0) = H\widehat{\boldsymbol{x}}(t_0) = 2\,000$$

$$P(t_0) = \begin{bmatrix} 1\,000\,000 & 0 & 0 \\ 0 & 11\,110\,000 & 0 \\ 0 & 0 & 250\,000 \end{bmatrix} \Longrightarrow P_h(t_0) = HP(t_0)H^T = 1\,000\,000$$

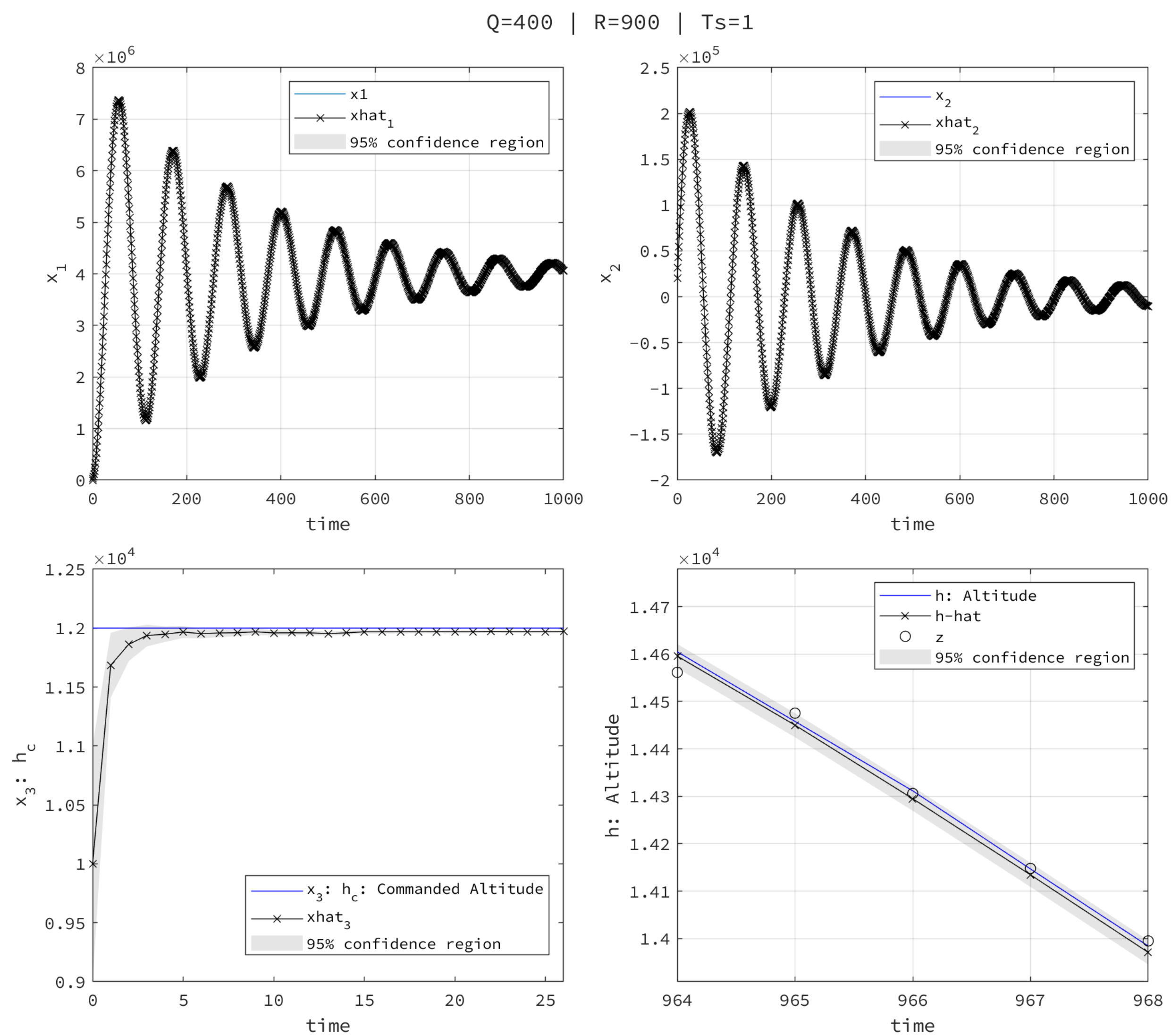


Figure 19: System Simulation: State Estimates

Error variance in the *commanded altitude* estimate reduces to a value of $\sigma^2_{\widehat{h}_c} = 1.2$ ft^2 at the end of simulation time. It means that it is very likely that the true value of the commanded altitude is within just ± 4 feet of the estimate. You can also observe the rapid convergence of $\widehat{h}_c(t)$ to h_{c0} in the figure. Likewise, the terminal value of the error variance in the *altitude* estimate is $\sigma_h^2 = 160$ ft^2. It means that we are very confident that the true value of the altitude lies within an error bound of ± 40 feet of the estimate.

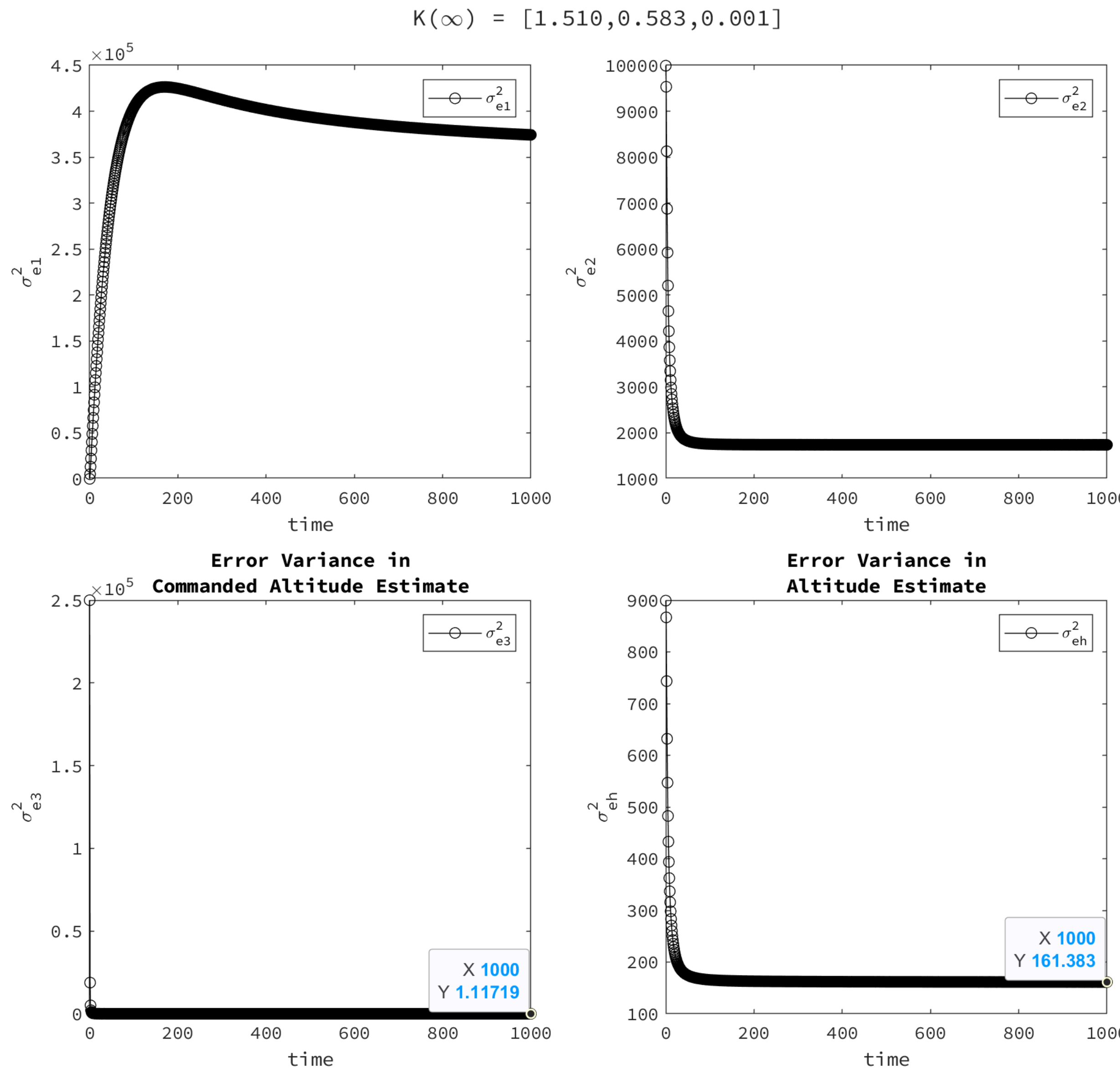


Figure 20: System Simulation: Error Variance

Problem 5.13

The given dynamics model for the position variable is assumed to be valid over the *uniform*[2] time-grid $\{0, 1, 2, \cdots, 10\}$. Let us suppose that the first measurement is taken at t_1 $(0 \leq t_1 \leq 8)$ and the second measurement is taken at t_2 $(t_1 < t_2 \leq 9)$ where $t_1, t_2 \in \mathbb{Z}_+$. Valid values for (t_1, t_2) satisfying these constraints are shown in the following figure:

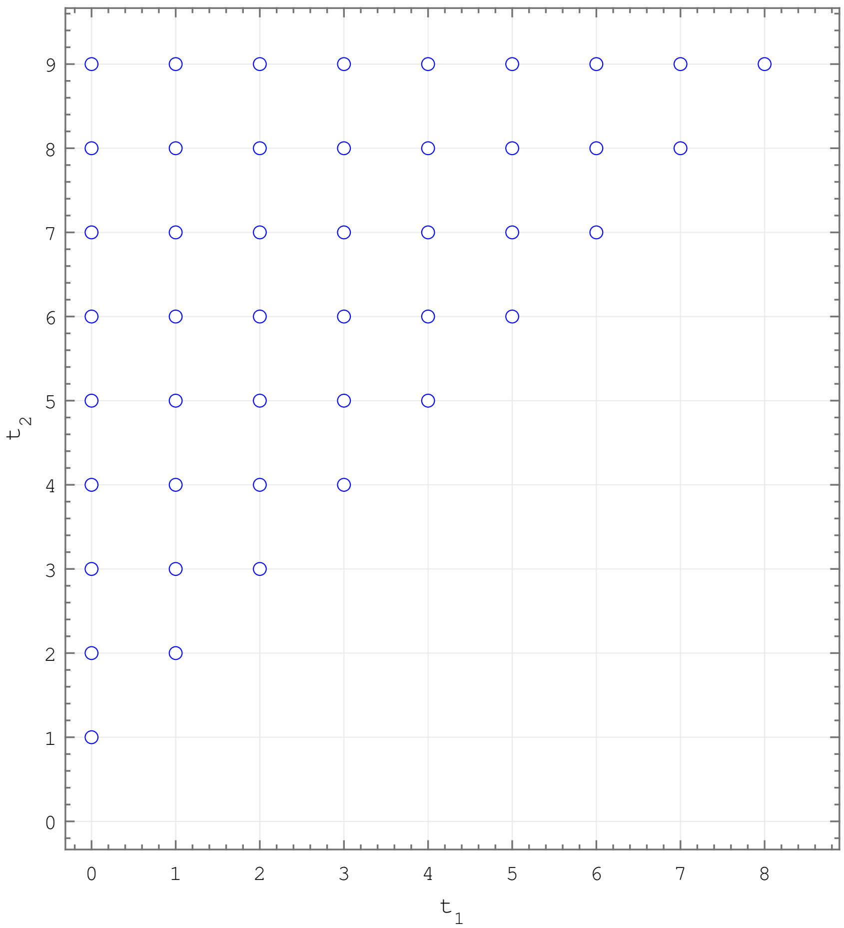


Figure 21: Valid time points for the two measurements

The problem is well-suited for applying Kalman filter as the optimal estimator. Since the state variable is scalar, the error covariance (variance) time update equation is given by:

$$P(t_i^-) = F_d P(t_{i-1}^+) F_d^T + Q_d$$
$$\implies p(t_i^-) = p(t_{i-1}^+) + q \qquad \text{(Time Update Equations)}$$

Likewise, measurement update equation for error covariance (variance) is

$$P(t_i^+) = P(t_i^-) - P(t_i^-)H^T \left[HP(t_i^-)H^T + R\right]^{-1} HP(t_i^-)$$
$$\implies p(t_i^+) = p(t_i^-) - \frac{p(t_i^-)^2}{p(t_i^-) + r}$$
$$\implies p(t_i^+) = \frac{p(t_i^-)r}{p(t_i^-) + r} \qquad \text{(Measurement Update Equations)}$$

Error covariance of the optimal estimate at $t_f = 10$ is obtained as follows:

1. Propagate error covariance from $t_0 = 0$ to t_1:

$$p(t_1^-) = p_0 + t_1 q$$

[2] For non-uniform time-grid, $Q_d(t_i) \triangleq \mathbb{E}\left[w_d^2(t_i)\right]$ is expected to be time-varying to account for the accumulated effect of the continuous-time noise process on the system over $[t_i,\ t_{i+1}]$ (see Eq. 4-127b). Constancy of $Q_d(t_i)$ implies (logically) that the given discrete-time model is only valid over a uniform mesh; $t_{i+1} - t_i = T = \text{cte}$.

2. Make a measurement at t_1 and update the covariance:

$$p(t_1^+) = \frac{p(t_1^-)r}{p(t_1^-) + r}$$

3. Propagate error covariance from t_1 to t_2:

$$p(t_2^-) = p(t_1^+) + (t_2 - t_1)q$$

4. Make a measurement at t_2 and update the covariance:

$$p(t_2^+) = \frac{p(t_2^-)r}{p(t_2^-) + r}$$

5. Finally, propagate error covariance from t_2 to $t_f = 10$:

$$p(t_f) = p(t_2^+) + (t_f - t_2)q$$

By combining these steps, we can obtain a closed form formula for $p(t_f)$ as a function of t_1 and t_2:

$$p(t_f) = \frac{(p_0 + t_1 q)r^2 + (t_2 - t_1)qr(p_0 + t_1 q + r)}{(p_0 + t_1 q)r + ((t_2 - t_1)q + r)(p_0 + t_1 q + r)} + (t_f - t_2)q \qquad (\star)$$

We can now vary (t_1, t_2) over the grid shown in figure 21 and compute the resulting values for the position error variance at time $t_f = 10$.

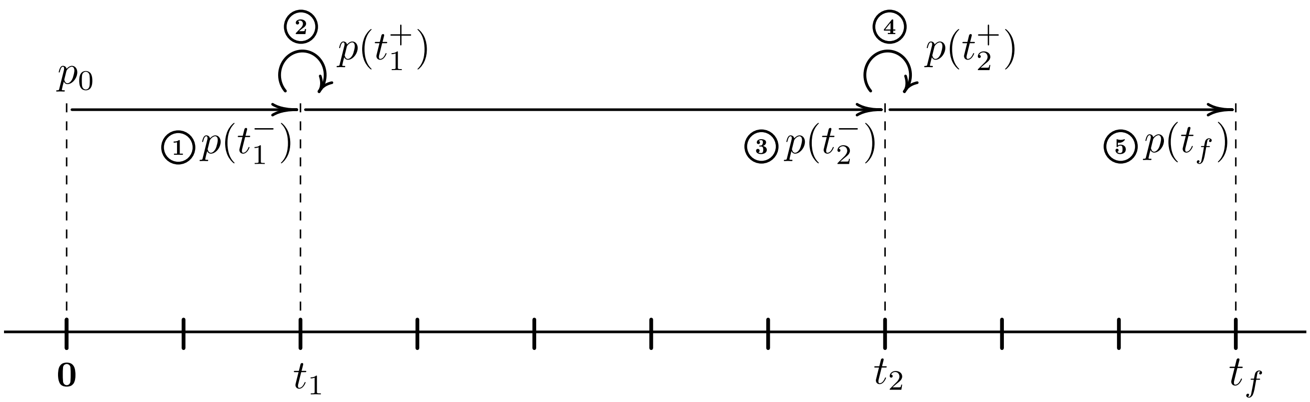


Figure 22: Steps 1-5 illustrated

Intuitively, making measurements closer to $t_f = 10$ yields more precise results for the position estimate at $t = t_f$. If there is a time gap between t_2 and t_f, position estimate becomes more and more unreliable due to the constant injection of noise into the system as we move from t_2 to t_f. This is also evident in covariance time update equation as q is a non-zero forcing term in $p(t_i^-) = p(t_{i-1}^+) + q$ causing $p(t)$ to increase over time.
We can rewrite Eq.($\star$) as

$$p(t_f) = r\left\{1 - \frac{r(p_0 + t_1 q + r)}{(p_0 + t_1 q)r + ((t_2 - t_1)q + r)(p_0 + t_1 q + r)}\right\} + (t_f - t_2)q$$

From this equation, it is now readily observed that reducing $(t_2 - t_1)$ and/or $(t_f - t_2)$ lowers $p(t_f)$.

Part a)

It is evident from Eq.($\star$) that scaling p_0, q, and r equally, simply scales $p(t_f)$ by the same factor. As such, we use $p_0 = q = r = 1$ ft^2 for subsequent computations. Figure 23 is obtained by computing Eq.($\star$) over the time points shown in figure 21. As expected, the least error variance is obtained by taking measurements at $(t_1, t_2) = (8, 9)$. As t_2 becomes closer to the target time t_f (fixing t_1 and increasing t_2), smaller values for $p(t_f)$ are obtained. Also, as the time difference between t_1 and t_2 is reduced (fixing t_2 and increasing t_1), more accurate estimate is obtained.

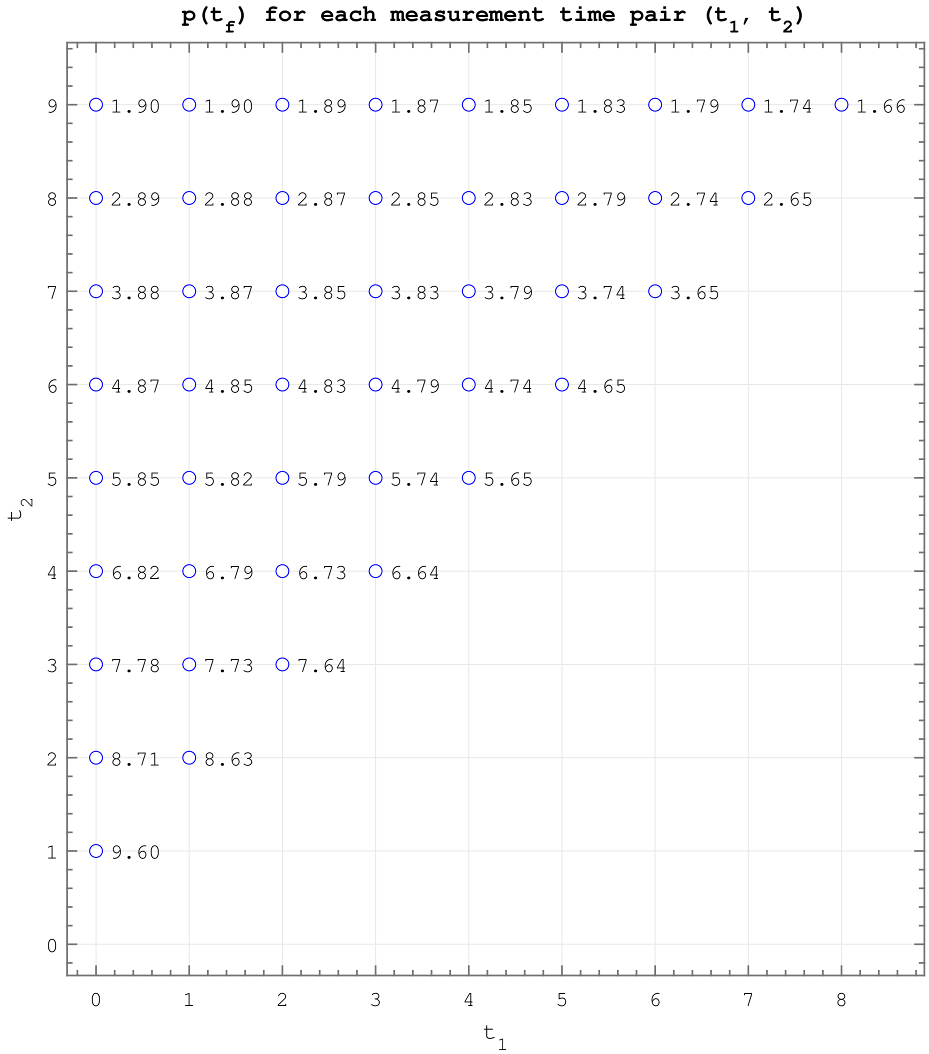


Figure 23: $p(t_f)$ over valid measurement time points (t_1, t_2)

Thus, the optimal rms terminal position error is $\sqrt{p(t_f)} = \sqrt{1.65E4 \text{ ft}^2} \approx 128.6$ ft.

Part (b) - (e)

We perform the same calculations under the condition that $p(t^-) \leq (2.5 \text{ ft})^2 = 6.25 \text{ ft}^2$ must be satisfied prior to each measurement.

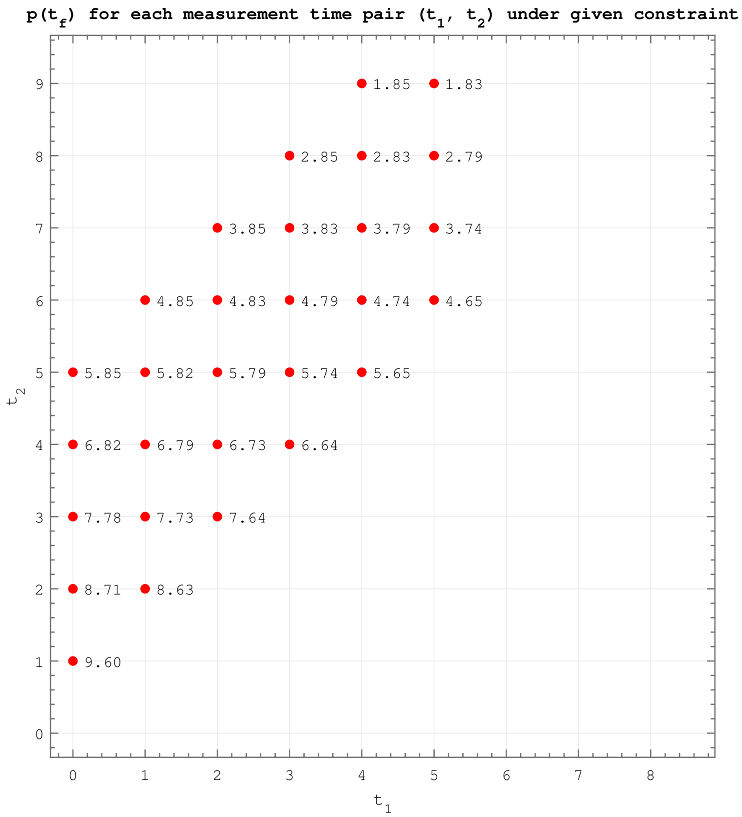


Figure 24: $p(t_f)$ over valid measurement time points (t_1, t_2) under the given constraint

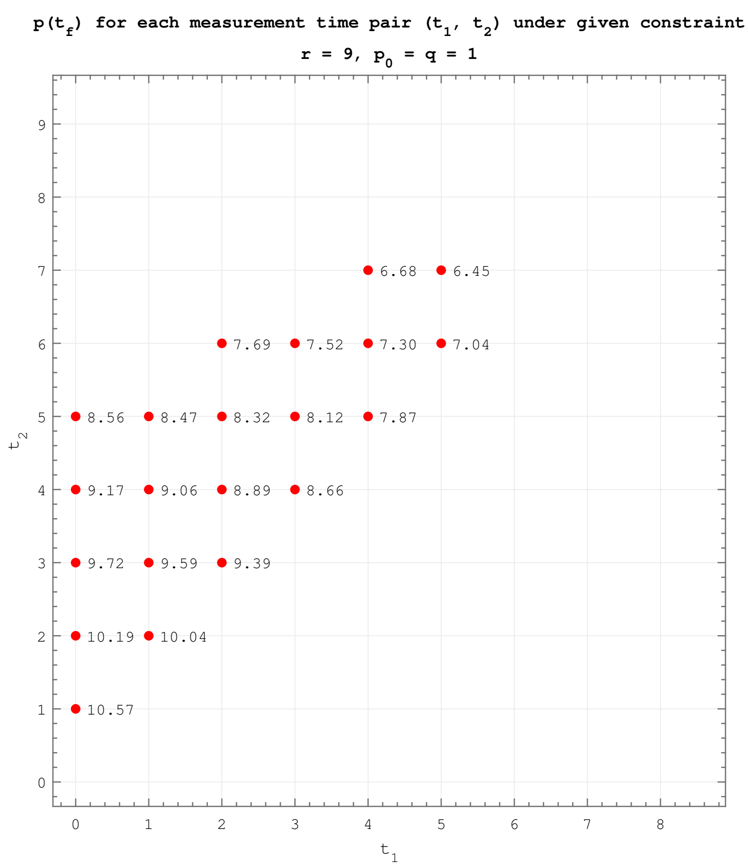


Figure 25: $p(t_f)$ over valid measurement time points (t_1, t_2) under the given constraint

Problem 5.20

The small orbital deviations we wish to observe from the ground station are the state variables of a linear dynamical system described by

$$\dot{\mathbf{x}}(t) = F\mathbf{x}(t) + B\boldsymbol{u}(t)$$

where $F = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 2\omega^2 & 0 & 0 & 2r_0\omega^2 \\ 0 & 0 & 0 & 1 \\ 0 & -{}^{2\omega}/_{r_0} & 0 & 0 \end{bmatrix}$, and $B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & {}^{1}/_{r_0} \end{bmatrix}$.

The state vector $\mathbf{x}(t)$ is the deviation signal $\mathbf{x}(t) = \begin{bmatrix} \delta r(t) & \delta\dot{r}(t) & \delta\theta(t) & \delta\dot{\theta}(t) \end{bmatrix}^T$ from the nominal trajectory of the object in a circular orbit: $r^*(t) \equiv r_0$, $\dot{r}^*(t) \equiv 0$, $\theta^*(t) = \omega t$, $\dot{\theta}^*(t) = \omega$, $u_t^*(t) = u_r^*(t) \equiv 0$. We assume that the nominal trajectory is known to the observer, allowing a measurement of an orbital variable to be converted into a measurement of its deviation. For example, by obtaining a measurement of angle through $\theta(t) + \mathsf{v}(t)$, we can simply define and evaluate $\mathsf{z}(t) \triangleq \{\theta(t) + \mathsf{v}(t)\} - \theta^*(t) = \{\theta(t) - \theta^*(t)\} + \mathsf{v}(t) = \delta\theta(t) + \mathsf{v}(t) = \mathsf{x}_3(t) + \mathsf{v}(t)$ which is a measurement of the state variable $\mathsf{x}_3(t)$.

Choosing $H = \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix}$ renders the system observable, meaning that, in the absence of dynamic driving noise ($Q \equiv 0$), the estimated states will converge to the true states over time. This is demonstrated in the following simulation.

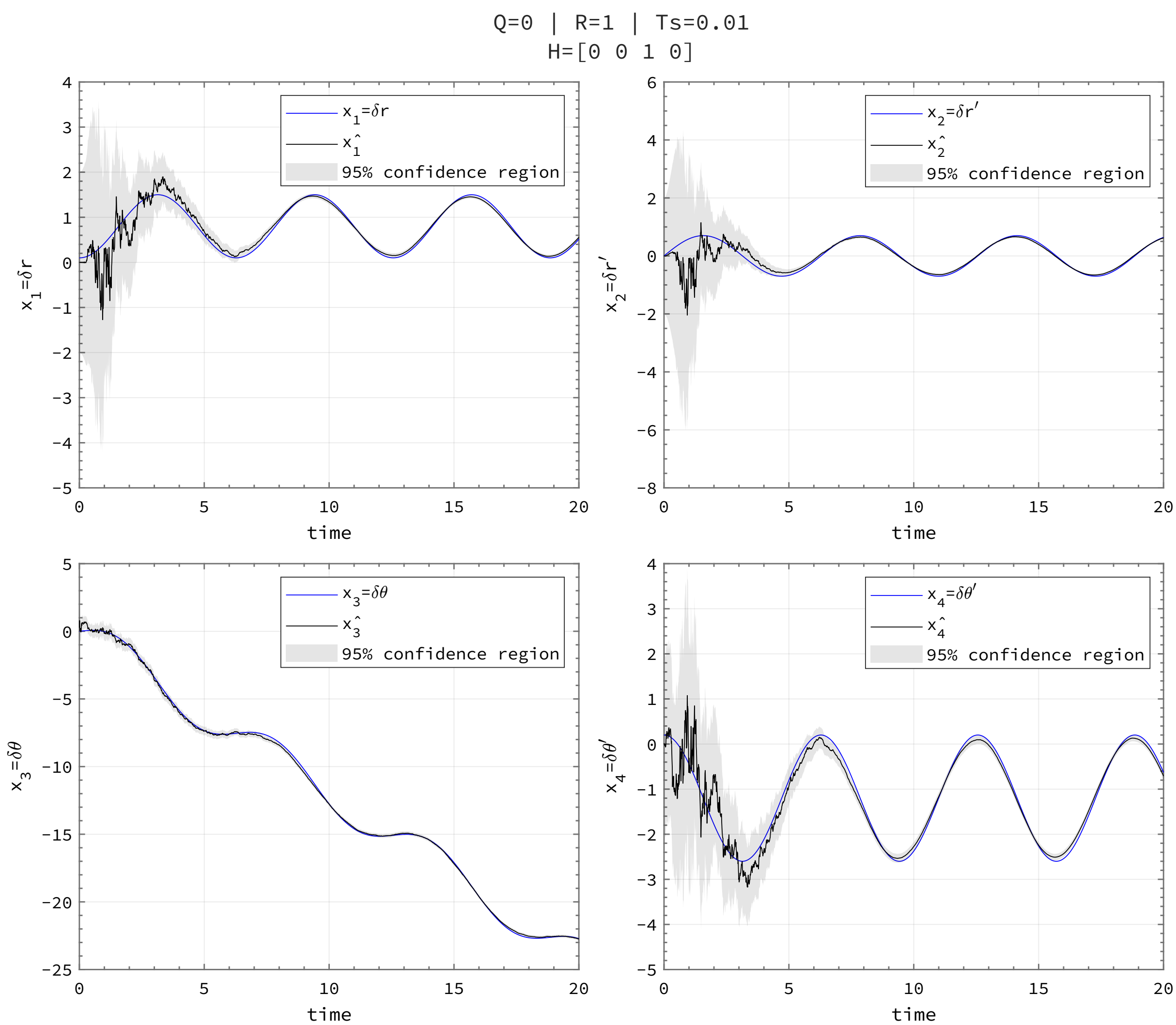


Figure 26: Estimating deviation signals by observing $\theta(t)$.

Choosing $H = \begin{bmatrix}1 & 0 & 0 & 0\end{bmatrix}$ results in a unobservable subsystem. Decomposing the system into observable/unobservable components by applying a linear transformation to the state vector using

$$\widetilde{\mathsf{x}} = T\mathsf{x}, \qquad T = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

produces the observability staircase form, which for the present systems takes a simple form. Realize that T is merely a permutation matrix indicating that x_3 is not observed in the output. Since x_3 is not channeled into the output signal, we cannot improve our estimate of it through observations.

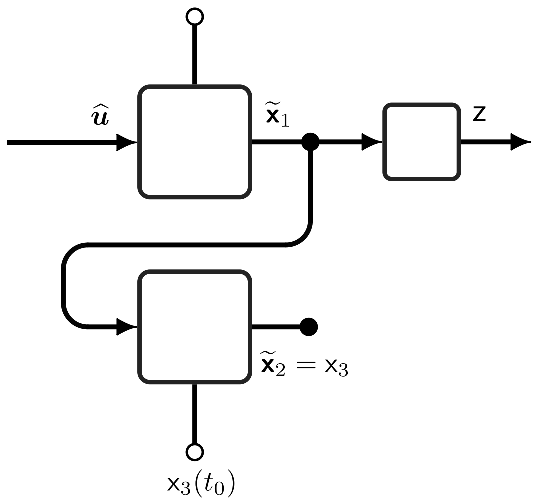


Figure 27: The state variable x_3 is not sensed by the output signal z.

It is therefore expected that the variance of the error in the estimate of x_3 remains constant or grows over time. This is also confirmed by the simulation results that follow. As shown in Fig. 29, x_1, x_2, and x_4 are in the observable part of the system and therefore are successfully recovered by the filtering process. However the uncertainty in the value of x_3 is not improved by the measurement process.
Choosing $H = \begin{bmatrix}1 & 0 & 0 & 0; 0 & 0 & 1 & 0\end{bmatrix}$ seems to accelerate the recovery process. Simulations indicate that the magnitude of variances decays more rapidly (especially in the initial phase of the filtering) when both measurements are incorporated into the filtering operation.

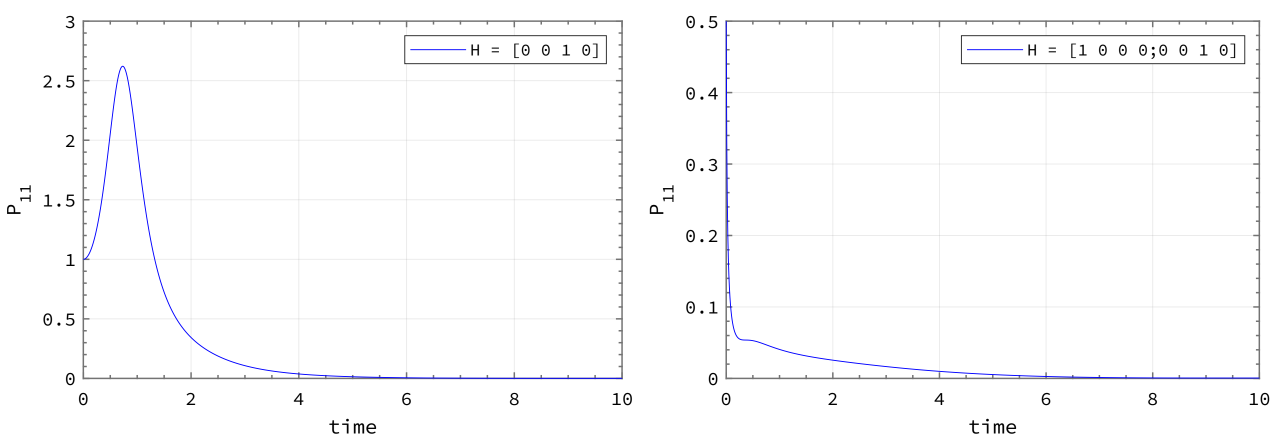


Figure 28: Variance of the error in the estimate of x_1 for two different observation models.

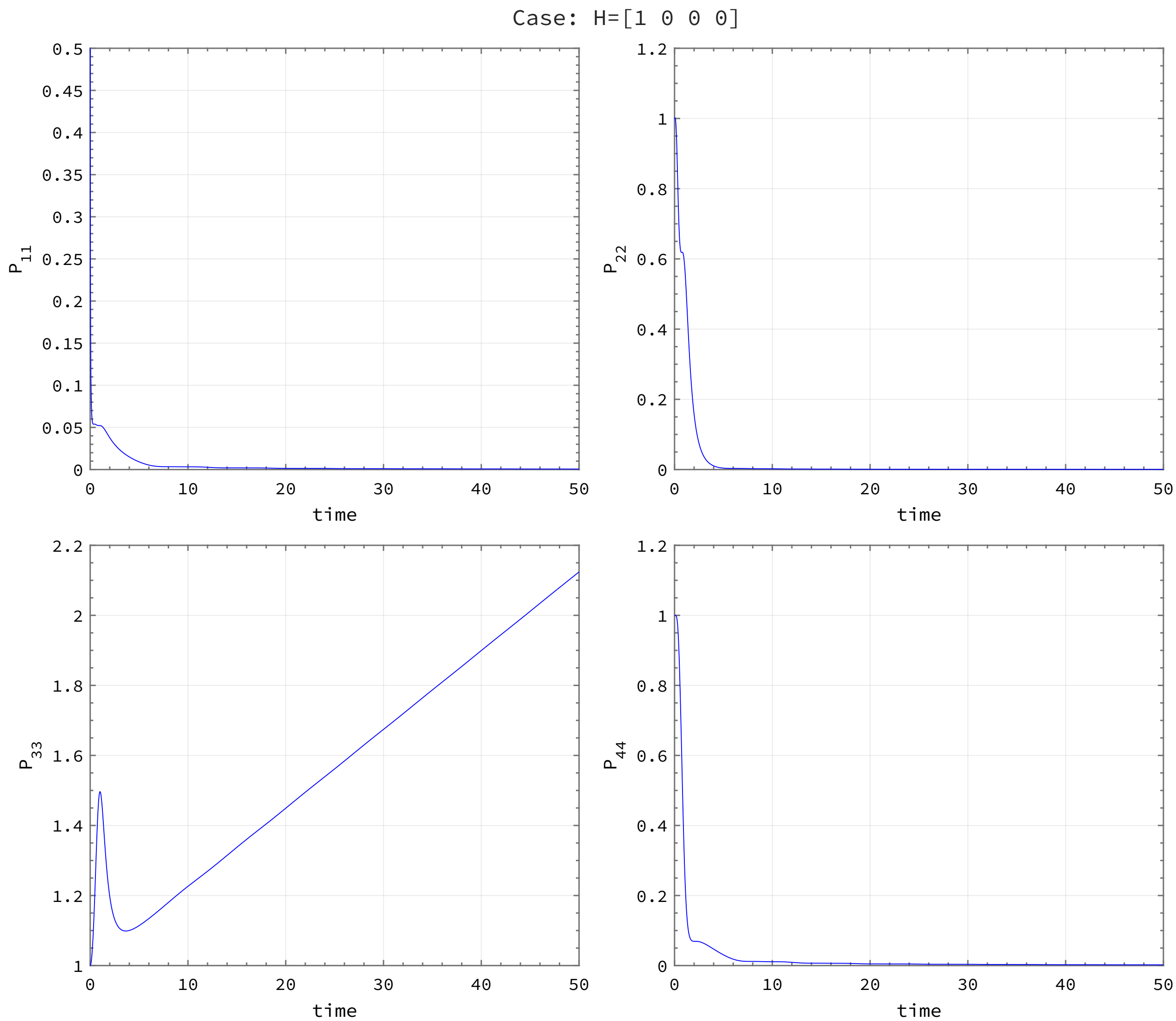


Figure 29: Variance of the error in the estimate of the system state variables when x_1 is used for the measurement process.