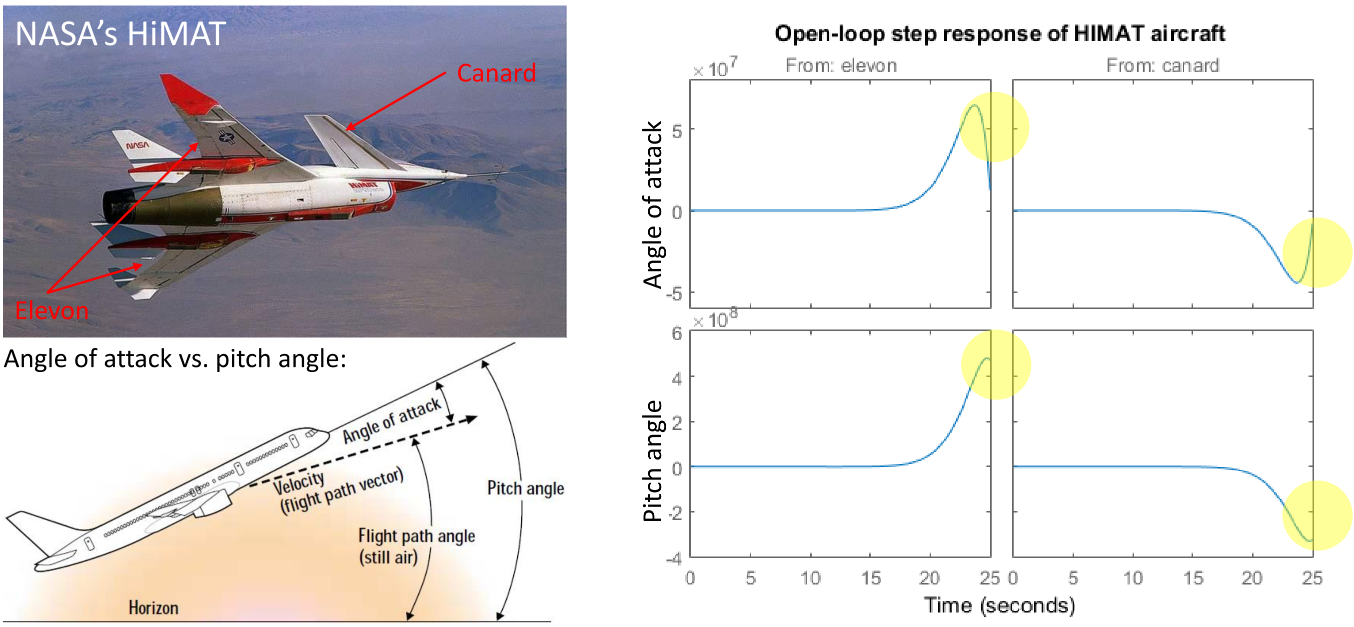
EEET 3046 Control Systems (2020)

Lecture 3: Stability and steady-state error

Dr. Yee Wei Law ⟨yeewei.law@unisa.edu.au⟩

Contents



URLs: http://www.boeing.com/commercial/aeromagazine/aero_12/whatisaoa.pdf
http://au.mathworks.com/help/robust/examples/loop-shaping-of-himat-pitch-axis-controller.html

Figure 1: NASA's HiMAT is a remotely piloted tailless aircraft that uses *elevons* and *canards* to control its *angle of attack* and *pitch angle*. Elevons are flight control surfaces that combine the functions of the elevators (pitch) and the ailerons (roll). The *open-loop step responses* of these angles are *unstable* — the yellow spots show they grow indefinitely with a step change in the deflection angles of the elevons and canards. However, the yellow spots also show the responses reverse their growing trend at some point, but this is due to model inaccuracy.

1 Introduction

Consider the HiMAT (Highly Maneuverable Aircraft Technology) aircraft, developed by Rockwell for NASA in the 1970s, to enhance the transonic maneuverability of American fighter jets at that time. To change its angle of attack or pitch angle, the HiMAT changes the deflection angles of its *elevons* and *canards* (see Figure 1). In open loop, the *angle of attack* and *pitch angle* grow indefinitely with a step change in the deflection angles; we say the *open-loop step responses* of the angle of attack and pitch angle are unstable. This is an example of an open-loop unstable system — we design/synthesize controllers to stabilize such systems.

Physically, an unstable system whose response grows without bound could be self-damaging, unsafe, if not plainly useless. Therefore, when designing a control system, we always ensure it is stable, before we worry about its performance, which is often achieved in a trade-off between transient response and steady-state error.

In this lecture, we will learn

- the definitions of stability;
- how to determine the stability of a system;
- how to determine the steady-state error of a stable system;
- the effect of disturbances, and the meaning of the sensitivity functions — this paves way for the classical controller design technique of *loop shaping*.

This lecture follows the directions of [Nis15, Chs. 6 and 7] and [DB11, Ch. 5].

2 Definitions of stability

In Lecture 2, we derived the output response for the RC circuit in Figure 2 as

$$Y(s) = \underbrace{\frac{1}{RCs+1}X(s)}_{\text{forced response}} + \underbrace{\frac{RC}{RCs+1}y(0)}_{\text{natural response}}. \tag{1}$$

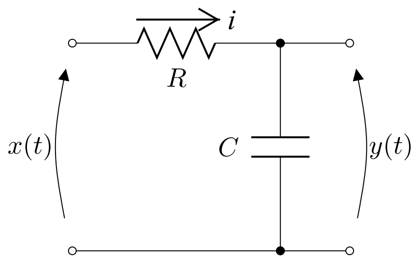


Figure 2: An RC circuit.

- When the initial state is zero, $Y(s)$ reduces to the first term, representing the *forced response* (aka *zero-state response*), which depends only on the input. The stability of the forced response determines the system's *external stability*.
- When the input is zero, $Y(s)$ reduces to the second term, representing the *natural response* (aka *free response* or *zero-input response*), which depends only on the initial state. The stability of the natural response determines the system's *internal stability*.

The stability of $Y(s)$ necessitates both external and internal stability, but we shall see internal stability implies external stability. Let us discuss external stability and internal stability in turn.

2.1 External stability

The external stability of an LTI system is equivalent to the stability of its forced response. Since only the input and output of the system are considered, external stability is synonymous with *input-output stability*, and the more expressive term *bounded-input bounded-output stability* (BIBO stability).

> **Definition: BIBO stability [Kai80, Sect. 2.6.1], [Che99, Ch. 5, p. 122]**
>
> A system is BIBO stable if every bounded input produces a bounded output.

- A signal is bounded if the signal is less than a finite value for all time.
- Note that if an input is unbounded, the system's BIBO stability cannot be determined.
- Figure 3 shows physical analogies for external stability.

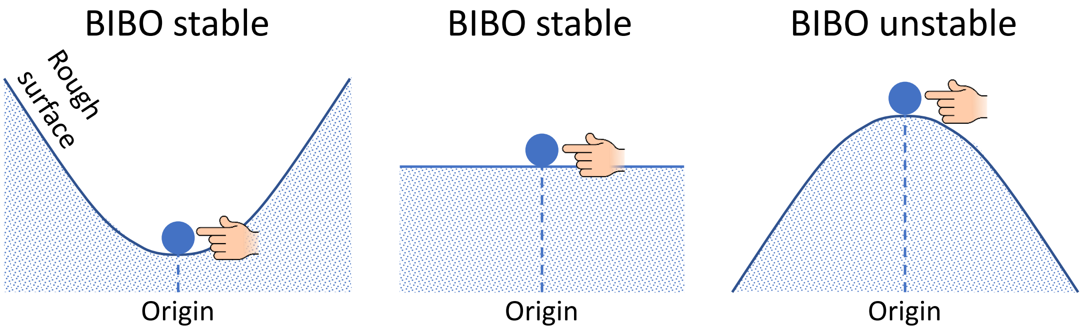


Figure 3: Physical analogies for external stability. Say the ball is at the origin, and we give the ball a poke of finite strength. Think of the distance between the ball and the origin as the *forced response*. If the ball eventually stops at a certain spot, the system is externally stable (BIBO stable). If the ball keeps moving away, the system is externally unstable (BIBO unstable).

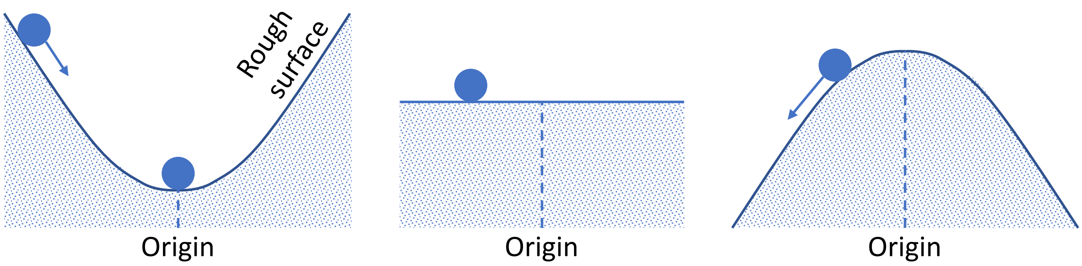


Figure 4: Physical analogies for internal stability. At $t = 0$, we place a ball at a random spot other than the origin, and see if the ball will settle at the origin (stable), or settle at some point other than the origin (marginally stable), or keep moving away from the origin (unstable). Think of the the distance between the ball and the origin as the *natural response*. Illustration inspired by [WL07, Figure 6.1].

2.2 Internal stability

The internal stability of an LTI system is equivalent to the stability of its natural response.

> **Definition: Internal stability [Nis15, Sect. 6.1]**
>
> An LTI system is
> - *stable* if its natural response approaches zero with time;
> - *marginally stable* if its natural response neither decays nor grows but remains constant or oscillates as time approaches infinity;

- *unstable* if its natural response grows without bound with time.

Figure 4 shows physical analogies for internal stability.

The example in Eq. (1) shows that the natural response shares the same characteristic polynomial with the transfer function. Whether this natural response is stable depends on the roots of this characteristic polynomial, or equivalently, the poles of the transfer function. Let us study how the location of the poles in the s plane affects the stability of the natural response (see Figure 5):

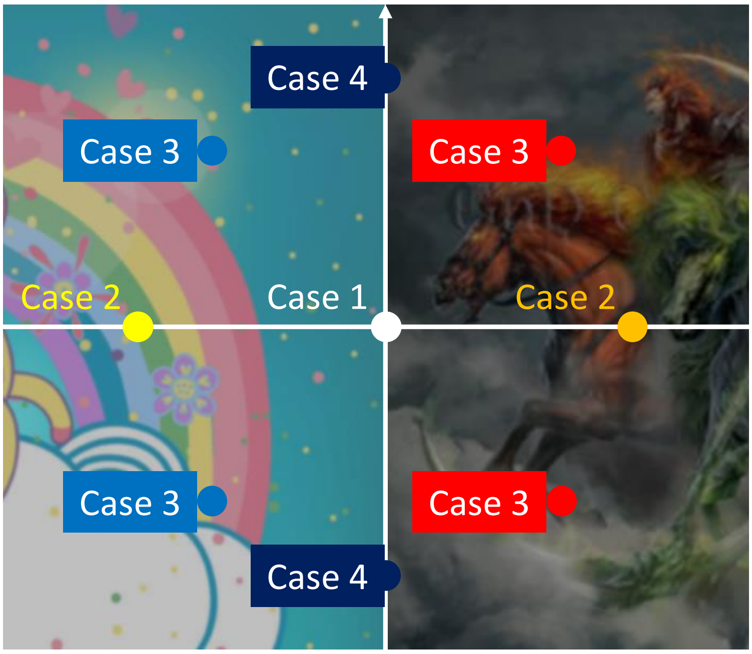


Figure 5: Pole locations determine stability: see text for a discussion of the different cases.

Case 1 One or more poles at the origin:

- One pole at the origin is associated with a step response in the time domain (because $\mathcal{L}\{1\} = 1/s$), which is marginally stable.
- $n > 1$ poles (i.e., poles with *multiplicity* $n > 1$) at the origin are associated with a time response of the form t^{n-1} (because $\mathcal{L}\left\{t^{n-1}\right\} = (n-1)!/s^n$), which is unstable.

Case 2 One or more poles on the real axis excluding the origin:

- A pole at $-\sigma \neq 0$ is associated with the exponential response $\exp(-\sigma t)$, which is stable if $\sigma > 0$, but unstable if $\sigma < 0$.
- A pole with multiplicity $n > 1$ at $-\sigma \neq 0$ is associated with the exponential response $t^{n-1}\exp(-\sigma t)$, which is stable if $\sigma > 0$, but unstable if $\sigma < 0$.

Quiz 1

$$\mathcal{L}^{-1}\left\{\frac{1}{(s+\sigma)^n}\right\} = ?$$

Case 3 One or more strictly complex poles:

- For physically realizable systems, these poles come in conjugate pairs and have the form $-\sigma \pm j\omega$, where $\sigma, \omega \neq 0$. A conjugate pair is associated with a time response of the form $e^{-\sigma t}\sin(\omega t + \phi)$, which is stable if $\sigma > 0$, but unstable if $\sigma < 0$.
- A conjugate pair with multiplicity $n > 1$ is associated with a linear combination of responses of the form $e^{-\sigma t}\sin(\omega t + \phi_0)$, $te^{-\sigma t}\sin(\omega t + \phi_1)$, ..., $t^{n-1}e^{-\sigma t}\sin(\omega t + \phi_{n-1})$. The combined response is stable if $\sigma > 0$, but unstable if $\sigma < 0$.

Quiz 2

Using the formula $\mathcal{L}\left\{t^n f(t)\right\} = (-1)^n F^{(n)}(s)$, determine $\mathcal{L}\left\{te^{-\sigma t}\sin(\omega t)\right\}$.

Case 4 One or more poles on the imaginary axis excluding the origin:

- For physically realizable systems, these poles come in conjugate pairs and have the form $\pm j\omega$, where $\omega \neq 0$. A conjugate pair is associated with a sinusoidal response, which is marginally stable.
- A conjugate pair with multiplicity $n > 1$ is associated with a linear combination of responses of the form $\sin(\omega t + \phi_0)$, $t\sin(\omega t + \phi_1)$, ..., $t^{n-1}\sin(\omega t + \phi_{n-1})$. The combined response is unstable.

Quiz 3

Using the formula $\mathcal{L}\left\{t^n f(t)\right\} = (-1)^n F^{(n)}(s)$, determine $\mathcal{L}\left\{t\sin(\omega t)\right\}$.

In summary,

- A *stable* system has only poles in the open left half plane.
- A *marginally stable* system has poles on the imaginary axis with multiplicity 1, and no poles in the open right half plane.
- An *unstable* system has poles in the open right half plane, and/or poles on the imaginary axis with multiplicity greater than 1.

Example 1

Consider the system $G(s) = \dfrac{1}{(s+10)(s^2+\omega^2)}$. $G(s)$ has a pole at -10, and a pair of poles at $\pm j\omega$, so it is marginally stable. As for BIBO stability, we have not really covered the formal method for assessing BIBO stability, but consider the following types of input:

- Step input: The output response will consist of a step response, a decaying exponential, and a sinusoid. So this output response is bounded.
- Sinusoidal input of angular frequency ω: The output response will consist of a decaying exponential, a sinusoid, and a growing sinusoid. So this output response is unbounded.

In conclusion, $G(s)$ is marginally stable and BIBO unstable.

Example 1 is an example of how marginal stability does not ensure BIBO stability. Unless we are designing oscillators, stability is the only desirable condition. Theorem 1 says that internal stability implies BIBO stability; and BIBO stability implies internal stability if the transfer function is rational and proper.

Theorem 1: [Che99, Theorem 5.3]

A SISO LTI system with proper rational transfer function $G(s)$ is BIBO stable iff every pole of $G(s)$ has a negative real part, or equivalently, lies inside the left half s-plane.

Example 2

This example is adapted from [Che99, Example 5.2]. Consider the system $G(s) = 1$. In terms of external stability, $G(s)$ passes the input unchanged to the output, so it is clearly BIBO stable. However, in terms of internal stability, $G(s)$ does not have any pole, so it is not (internally) stable. In this example, $G(s)$ is not a rational transfer function since it is not a fraction of polynomials of s (either the numerator or the denominator can be a degenerate polynomial, but not both), so it does not satisfy the condition for Theorem 1.

Example 2 reminds us that internal stability implies BIBO stability, but the converse is not necessarily true. This is why when analyzing the stability of a SISO LTI system, we only consider its internal stability.

When a characteristic polynomial has only roots with a negative real part, it is so-called *Hurwitz*, and the corresponding system is stable. Now, the question is: is it possible to determine whether a characteristic polynomial is Hurwitz without factoring it? The answer to this question is the *Routh-Hurwitz stability criterion*, which is discussed in Supplementary Lecture B. While the Routh-Hurwitz criterion only applies to rational transfer functions, the *Nyquist stability criterion* applies to both rational and nonrational transfer functions — this is to be covered in a later lecture where the concept of frequency response is introduced.

3 Steady-state error

For a stable system, we are interested in its *steady-state error*, i.e., error in the steady state. To motivate our discussion, let us use the antenna azimuth position control system in Figure 6 as an example. The system is used to position a radio telescope antenna. The input is the desired azimuth angle, which is converted into a voltage value by a potentiometer. The output is the actual azimuth angle, which is also converted into a voltage value by another potentiometer. The difference between the two voltage values, called the *error*, is amplified. The larger the error, the more the motor will turn. The gain is how much the error is amplified. The higher the gain, the harder the motor will be driven. Typically, the higher the gain, the smaller the steady-state error. The goal of the control system is to minimize this steady-state error without destabilizing the system and worsening the transient response significantly.

A system gives different steady-state errors depending on the input. We are concerned with three kinds of input:

Input	$r(t)$	$R(s)$	Sample application
Step	1	$1/s$	Position control
Ramp	t	$1/s^2$	Tracking of constant-velocity targets
Parabola	$t^2/2$	$1/s^3$	Tracking of constant-acceleration targets

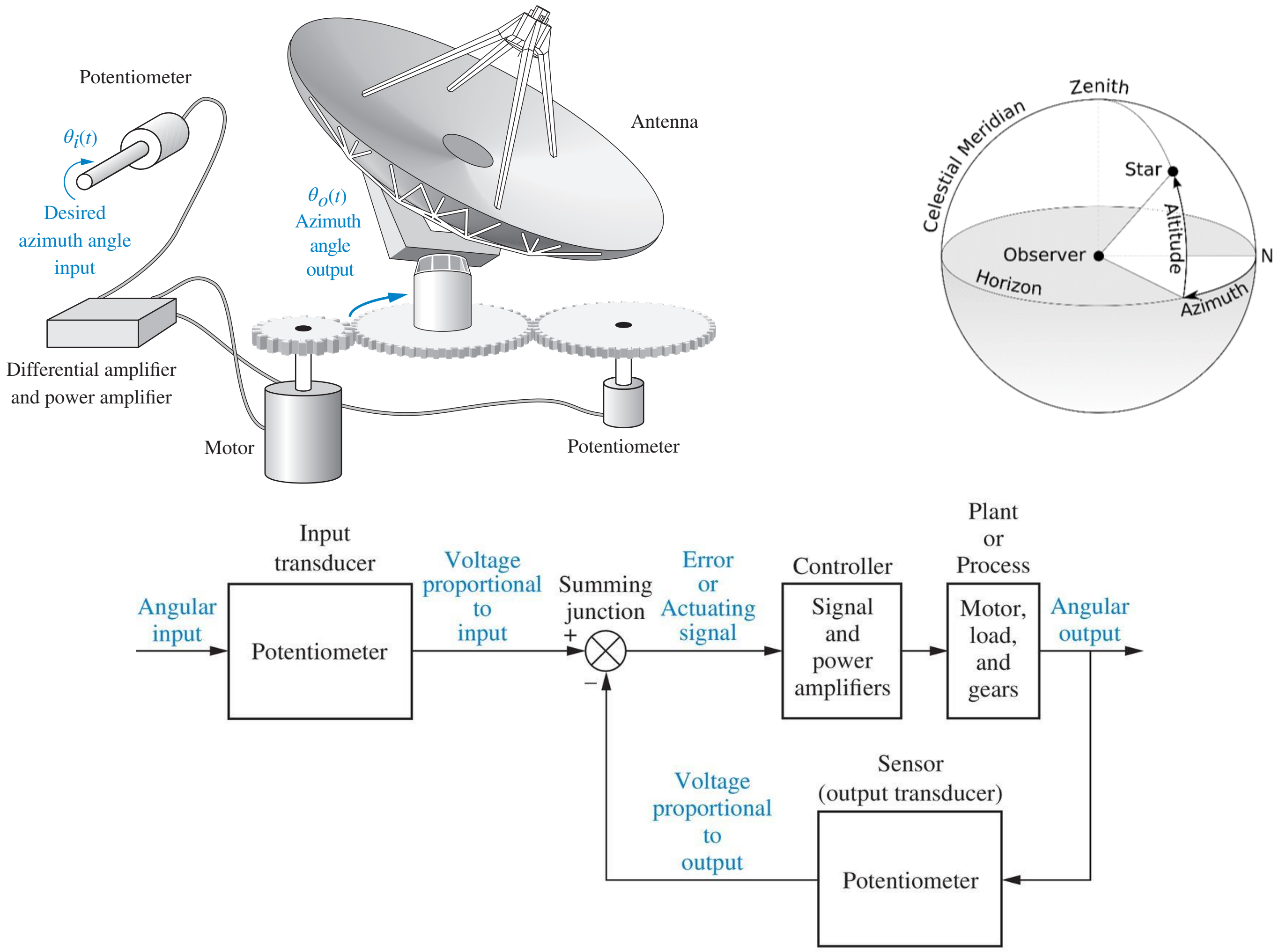


Figure 6: An antenna azimuth position control system [Nis15, FIGURE 1.8].

We shall call the steady-state error associated with each input type x as "x-response steady-state error", e.g., step-response steady-state error.

For each type of input, how much steady-state error a system produces depends on its configuration. In the next subsection, we shall derive the formulas for calculating the three types of steady-state errors for the *unity feedback configuration* (see Figure 7(a)), and then extend the results to more general configurations.

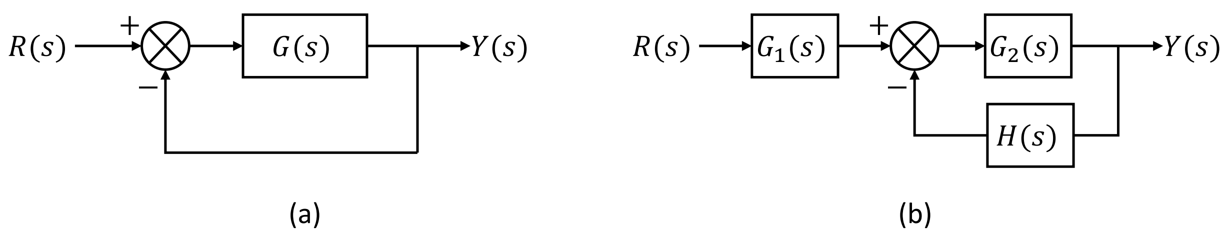


Figure 7: Unity vs nonunity feedback configurations, where $H(s)$ is nonzero and nonunity.

3.1 Unity feedback configuration

By "unity feedback configuration", we mean the configuration in Figure 7(a). Based on the block diagram,

$$Y(s) = G(s)[R(s) - Y(s)] = Y(s) = \frac{G(s)}{1+G(s)}R(s),$$

and the error in the Laplace domain is

$$E(s) \stackrel{\text{def}}{=} R(s) - Y(s) = \frac{1}{1+G(s)}R(s).$$

By the final value theorem, the steady-state error is then

$$e(\infty) \stackrel{\text{def}}{=} \lim_{t\to\infty} e(t) = \lim_{s\to 0} sE(s) = \lim_{s\to 0} \frac{1}{1+G(s)} sR(s). \tag{2}$$

- Step input: $R(s) = 1/s$,

$$e_{\text{step}}(\infty) = \frac{1}{1+\lim_{s\to 0} G(s)} = \frac{1}{1+K_p}, \tag{3}$$

  where K_p is called the *position error constant* (or simply *position constant*):

$$K_p \stackrel{\text{def}}{=} \lim_{s\to 0} G(s). \tag{4}$$

  Observe that if $G(s)$ has one or more uncanceled poles at the origin, $K_p \to \infty$, and $e_{\text{step}}(\infty) \to 0$.

- Ramp input: $R(s) = 1/s^2$,

$$e_{\text{ramp}}(\infty) = \frac{1}{\lim_{s\to 0} sG(s)} = \frac{1}{K_v}, \tag{5}$$

  where K_v is called the *velocity error constant* (or simply *velocity constant*):

$$K_v \stackrel{\text{def}}{=} \lim_{s\to 0} sG(s). \tag{6}$$

  Observe that if $G(s)$ has two or more uncanceled poles at the origin, $K_v \to \infty$, and $e_{\text{ramp}}(\infty) \to 0$.

- Parabolic input: $R = 1/s^3$,

$$e_{\text{para}}(\infty) = \frac{1}{\lim_{s\to 0} s^2G(s)} = \frac{1}{K_a}, \tag{7}$$

  where K_a is called the *acceleration error constant* (or simply *acceleration constant*):

$$K_a \stackrel{\text{def}}{=} \lim_{s\to 0} s^2 G(s). \tag{8}$$

Observe that if $G(s)$ has three or more uncanceled poles at the origin, $K_a \to \infty$, and $e_{\text{para}}(\infty) \to 0$.

Note:

- K_p, K_v and K_a are collectively called the *static error constants*.
- The discussion above suggests the *number of uncanceled poles at the origin* in the forward-path transfer function plays an important role in the minimization of the steady-state error — this number of poles is called the *system type* [Nis15, p. 347-348]. For example, if the unity feedback system in Figure 7(a) has $G(s) = 1/s$, it is a type 1 system. A type 1 system can track any step input with zero steady-state error. You can similarly infer the tracking ability of type 2, 3, . . . systems.

Example 3

Consider the unity feedback system in Figure 7(a) with

$$G(s) = \frac{K\prod_{j=1}^{n}(s+b_j)}{s\prod_{i=1}^{m}(s+a_i)},$$

where $K, a_i, b_j \in \mathbb{R}\setminus\{0\}$. What is the type number of the system? Furthermore, supposing the system is stable, calculate the static error constants, and the corresponding steady-state errors.

Solution: Since $G(s)$ has only one pole at the origin, the system is a type 1 system.

$$\begin{aligned} K_p &= \lim_{s\to 0} G(s) = \infty \implies e_{\text{step}}(\infty) = \frac{1}{1+K_p} = 0. \\ K_v &= \lim_{s\to 0} sG(s) = \frac{K\prod_{j=1}^{n} b_j}{\prod_{i=1}^{m} a_i} \implies e_{\text{ramp}}(\infty) = \frac{1}{K_v} = \frac{\prod_{i=1}^{m} a_i}{K\prod_{j=1}^{n} b_j}. \\ K_a &= \lim_{s\to 0} s^2G(s) = 0 \implies e_{\text{para}}(\infty) = \frac{1}{K_a} = \infty. \end{aligned}$$

3.2 Nonunity feedback configuration

The Eqs. (3)–(8) for calculating steady-state errors are only applicable to unity feedback systems. However, any nonunity feedback configuration, such as the one in Figure 7(b), can be converted into a unity feedback configuration — all it takes are just skills for reading block diagrams and doing high-school algebra.

The closed-loop transfer function (CLTF) for the unity feedback system in Figure 7(a) is

$$\frac{Y(s)}{R(s)} = \frac{G(s)}{1+G(s)}. \tag{9}$$

The CLTF for the nonunity feedback system in Figure 7(b) can be derived as follows:

$$\begin{aligned} Y(s) = G_2(s)[G_1(s)R(s) - H(s)Y(s)] &\implies Y(s)[1+G_2(s)H(s)] = G_2(s)G_1(s)R(s) \\ &\implies \frac{Y(s)}{R(s)} = \frac{G_2(s)G_1(s)}{1+G_2(s)H(s)}. \end{aligned} \tag{10}$$

Converting a nonunity feedback configuration into a unity feedback configuration is equivalent to finding an expression for $G(s)$ in terms of $G_1(s)$, $G_2(s)$ and $H(s)$, such that the right-hand side of Eq. (9) equals the right-hand side of Eq. (10), i.e.,

$$\frac{G}{1+G} = \frac{G_2 G_1}{1+G_2 H} \implies G + G G_2 H = G_1 G_2 + G G_1 G_2,$$

$$\therefore G = \frac{G_1 G_2}{1 - G_1 G_2 + G_2 H}. \tag{11}$$

Instead of the algebraic approach above, a graphical approach for deriving Eq. (11) can be found in [Nis15, pp. 358-359].

Example 4

Consider the nonunity feedback system in Figure 7(b) with

$$G_1(s) = 1, \qquad G_2(s) = \frac{K}{s(s+a_1)}, \qquad H(s) = \frac{1}{s+a_2},$$

where $K, a_1, a_2 \in \mathbb{R} \setminus \{0\}$. Assuming the system is stable, determine the value of a_2 such that the step-response steady-state error is 0.

Solution: Suppose the equivalent unity feedback system has the forward-path transfer function $G(s)$, then applying Eq. (11),

$$\begin{aligned} G = \frac{G_1 G_2}{1 - G_1 G_2 + G_2 H} &= \frac{\dfrac{K}{s(s+a_1)}}{1 - \dfrac{K}{s(s+a_1)} + \dfrac{K}{s(s+a_1)(s+a_2)}} \\ &= \frac{K(s+a_2)}{s(s+a_1)(s+a_2) - K(s+a_2) + K}. \end{aligned}$$

Assuming the system is stable,

$$K_p = \lim_{s \to 0} G = \frac{a_2}{1-a_2} \implies e_{\text{step}}(\infty) = \frac{1}{1+K_p} = 1 - a_2.$$

So that $e_{\text{step}}(\infty) = 0$, we need $a_2 = 1$.

3.3 Disturbances and sensitivity functions

No study on stability and steady-state error is complete without considering disturbances. Examples of disturbances include wind gusts to an aircraft, waves rocking a ship, varying road surfaces and gradients acting on a car's suspension or cruise controller [rMss, Ch. 12]. A disturbance is a stochastic process, and is typically modeled either as a *load disturbance* (aka *input disturbance*) or as an *output disturbance* (see Fig 8).

To account for the effect of disturbances on the steady-state error, Eqs. (3)–(8) no longer suffice. However, the calculation involves no more than reading the block diagram and applying the final

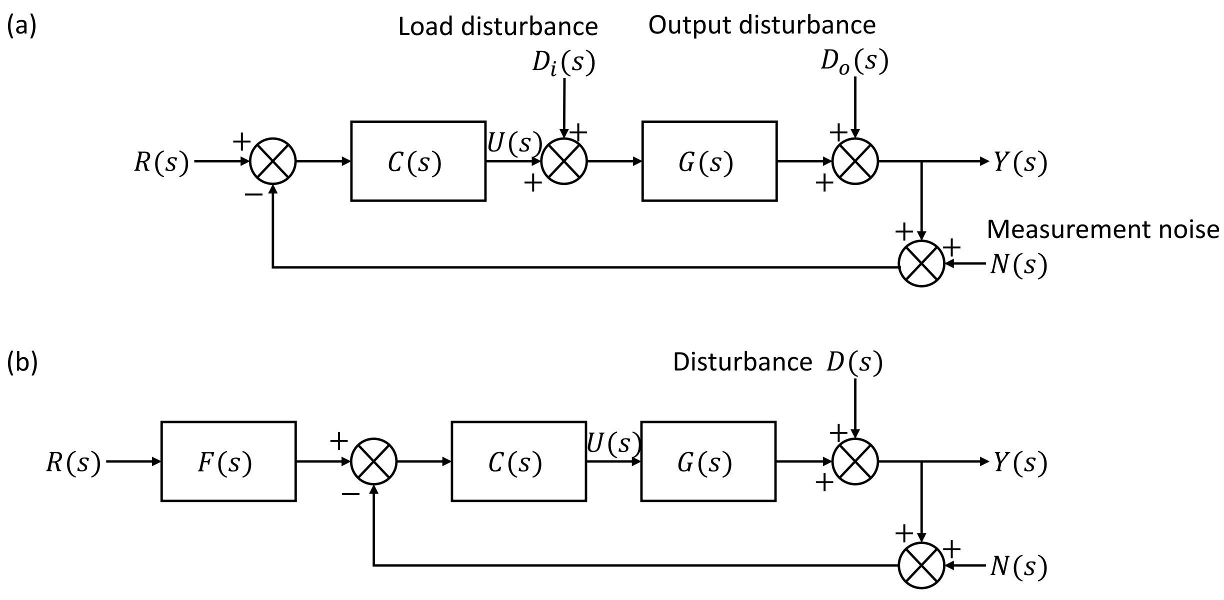


Figure 8: Block diagrams of (a) a 1DOF system; and (b) a 2DOF system.

value theorem. For example, consider the system in Figure 8(a). The output response is

$$\begin{aligned} Y &= D_o + G[C(R - Y - N) + D_i] = D_o + GCR - GCY - GCN + GD_i \\ \implies Y &= \frac{1}{1+GC}[GCR + GD_i + D_o - GCN]. \end{aligned} \tag{12}$$

$$\begin{aligned} \therefore E &= R - Y = R - \frac{1}{1+GC}[GCR + GD_i + D_o - GCN] \\ \implies E &= \frac{1}{1+GC}[R - GD_i - D_o + GCN]. \end{aligned} \tag{13}$$

Assuming the system is closed-loop stable, the steady-state error can then be calculated by applying the final value theorem to Eq. (13).

Example 5

Consider the unity feedback system in Figure 8(a) with $G(s) = \dfrac{1}{(s+a)^2}$, where $a \in \mathbb{R} \setminus \{0\}$. Suppose the system is stable.

(a) Assuming $C(s) = K$, where $K \in \mathbb{R}^+$, $D_i(s) = \dfrac{1}{s}$, $D_o(s) = 0$ and $N(s) = 0$, calculate the steady-state error due to the load disturbance.

(b) Assuming $C(s) = \dfrac{K}{s}$, where $K \in \mathbb{R}^+$, $D_i(s) = \dfrac{1}{s}$, $D_o(s) = 0$ and $N(s) = 0$, calculate the steady-state error due to the load disturbance.

Solution:

(a) Applying Eq. (13) to the problem, we have

$$E_d(s) = -\frac{G(s)}{1+G(s)C(s)} D_i(s) = -\frac{1}{(s+a)^2+K} \cdot \frac{1}{s},$$
$$\therefore e_{d,\text{step}}(\infty) = \lim_{s\to 0} sE_d(s) = -\lim_{s\to 0} s \cdot \frac{1}{(s+a)^2+K} \cdot \frac{1}{s} = -\frac{1}{a^2+K}.$$

Notice the larger we make K, the more we can reduce $e_{d,\text{step}}(\infty)$.

(b) Applying Eq. (13) again, we have

$$E_d(s) = -\frac{G(s)}{1+G(s)C(s)} D_i(s) = -\frac{s}{s(s+a)^2+K} \cdot \frac{1}{s},$$
$$\therefore e_{d,\text{step}}(\infty) = \lim_{s\to 0} sE_d(s) = -\lim_{s\to 0} s \cdot \frac{s}{s(s+a)^2+K} \cdot \frac{1}{s} = 0.$$

Notice injecting a pole at the origin into the controller $C(s)$ has the effect of reducing $e_{\text{step}}(\infty)$ to zero, even in the presence of a step disturbance.

The preceding example highlights the *disturbance rejection* property of integral control. Essentially, the integral controller contributes a pole that increases the system type by 1.

The unity feedback configuration in Figure 8(a), with a cascade controller, has so-called *one degree of freedom* (1DOF). This is because only one transfer function is available for shaping both reference response and disturbance response, i.e., once $C(s)$ is fixed, all $\frac{E(s)}{R(s)}$, $\frac{E(s)}{D_i(s)}$, $\frac{E(s)}{D_0(s)}$ and $\frac{E(s)}{N(s)}$ are fixed. In comparison, the configuration in Figure 8(b) has *two degrees of freedom* (2DOF) due to the additional $F(s)$ block, which is called a *set-point filter* or *reference filter* [GGS00, Sect. 5.2]. For the 2DOF architecture, we can revise Eqs. (12)–(13) as follows:

$$Y = \frac{1}{1+GC}[GCFR + D - GCN], \tag{14}$$

$$E = \frac{1}{1+GC}[(1+GC-GCF)R - D + GCN], \tag{15}$$

where $D(s) \overset{\text{def}}{=} G(s)D_i(s) + D_o(s)$. In the equation above, while $C(s)$ can be tuned to shape disturbance response, $F(s)$ can be tuned to shape reference response, to achieve both the desired disturbance rejection performance and the desired reference tracking performance. We will encounter the 2DOF architecture again in the context of the 2DOF PID controller in Lecture 7.

For now, we shall investigate the transfer functions $\frac{Y(s)}{D(s)}$, $\frac{Y(s)}{N(s)}$, $\frac{E(s)}{D(s)}$, $\frac{E(s)}{N(s)}$ in Eq. (14)–(15) in more detail. These are *sensitivity functions*, and understanding them allows us to design controllers that are *robust* to model uncertainty and disturbances.

Advanced: Robustness

The concept of robustness is not easy to grasp on an introductory level, but it can be viewed in two parts:

1. **Robust stability**: The system is stable for all perturbed plants about the nominal model up

to the worst-case model uncertainty [SP05, Sect. 1.2].

2. **Robust performance**: The system satisfies the performance specifications for all perturbed plants about the nominal model up to the worst-case model uncertainty [SP05, Sect. 1.2].

Standard measures of robustness are the *gain margin* and *phase margin* (see Lecture 8), which are related to the *maximum sensitivity*:

$$M_s \stackrel{\text{def}}{=} \max_{\omega} |S(j\omega)|. \tag{16}$$

For robustness, we need as per [AV16, Eq. (3.12)]

$$\text{gain margin} > \frac{M_s}{M_s - 1}, \qquad \text{phase margin} > 2\arcsin\left(\frac{1}{2M_s}\right).$$

Ultimately, controller design is about striking a balance between performance and robustness.

Revisiting the 2DOF architecture in Figure 8(b) and Eqs. (14)–(15), we have

$$Y(s) = \frac{G(s)C(s)F(s)}{1+G(s)C(s)}R(s) + \underbrace{\frac{1}{1+G(s)C(s)}}_{\text{Sensitivity function}}D(s) - \underbrace{\frac{G(s)C(s)}{1+G(s)C(s)}}_{\text{Complementary sensitivity function}}N(s), \tag{17}$$

$$E(s) = \left[1 - \frac{G(s)C(s)F(s)}{1+G(s)C(s)}\right]R(s) - \underbrace{\frac{1}{1+G(s)C(s)}}_{\text{Sensitivity function}}D(s) + \underbrace{\frac{G(s)C(s)}{1+G(s)C(s)}}_{\text{Complementary sensitivity function}}N(s), \tag{18}$$

$$U(s) = \frac{C(s)F(s)}{1+G(s)C(s)}R(s) - \underbrace{\frac{C(s)}{1+G(s)C(s)}}_{\text{Noise sensitivity function}}[D(s)+N(s)], \tag{19}$$

$$\underbrace{\frac{1}{1+G(s)C(s)}}_{\text{Sensitivity function}}D(s) = \underbrace{\frac{G(s)}{1+G(s)C(s)}}_{\text{Load sensitivity function}}D_i(s) + \underbrace{\frac{1}{1+G(s)C(s)}}_{\text{Sensitivity function}}D_o(s). \tag{20}$$

Collecting the definitions of the sensitivity functions in one place, we have:

$\frac{1}{1+G(s)C(s)}$	Sensitivity function, denoted $S(s)$
$\frac{G(s)C(s)}{1+G(s)C(s)}$	Complementary sensitivity function, denoted $T(s)$
$\frac{G(s)}{1+G(s)C(s)}$	Load sensitivity function / input disturbance sensitivity function
$\frac{C(s)}{1+G(s)C(s)}$	Noise sensitivity function / control sensitivity function

- Some authors affectionately call the four functions above the "*Gang of Four*" [rMss, Ch. 12].

- Denote by $S(s)$ the sensitivity function, then in Eq. (18), $S(s)$ measures the sensitivity of the tracking error to disturbance. Thus, good disturbance rejection requires $S(s)$ to be small, which in turn requires $C(s)$ to be large. The effect is the load sensitivity function will become small as well, but the the noise/control sensitivity function will become large, causing the control action to be sensitive to disturbance and noise, as per Eq. (19).
- Denote by $T(s)$ the complementary sensitivity function, then in Eq. (18), $T(s)$ measures the sensitivity of the tracking error to noise. Thus, good noise rejection requires $T(s)$ to be small, which in turn requires $C(s)$ to be small. Clearly, $S(s)$ and $T(s)$ cannot be made small for the same s (think "frequency").
- In fact, $S(s)$ and $T(s)$ are related through the identity:

$$S(s) + T(s) = 1, \tag{21}$$

Eq. (21) is of fundamental importance because it reflects a fundamental trade-off in control system design. Nevertheless, disturbances are typically concentrated in the low end of the frequency spectrum, whereas noise is typically concentrated in the high end of the frequency spectrum. This observation motivates the *loop shaping* technique for designing controllers, where the closed-loop frequency response of a system is shaped by making $S(s)$ small at low frequencies but $T(s)$ small at high frequencies.

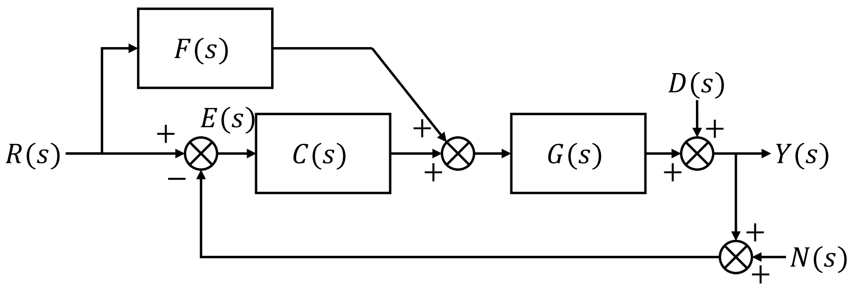


Figure 9: A unity feedback system with feedforward control.

Example 6

Consider the unity feedback system in Figure 9 with integrated feedforward control provided by the $F(s)$ block. Determine the sensitivity and complementary sensitivity functions. Furthermore, provide an expression for $F(s)$ that eliminates the steady-state error when $D(s) = N(s) = 0$.

Solution: Based on the block diagram,

$$Y = G\left[C(R - Y - N) + FR\right] + D \implies Y = \frac{G(C+F)}{1+GC}R + \frac{1}{1+GC}D - \frac{GC}{1+GC}N.$$

Therefore, in the presence of feedforward control, the sensitivity and complementary sensitivity function remain as $S = 1/(1 + GC)$ and $T = GC/(1 + GC)$ respectively.

Assuming $G(s) = G_N(s)/G_D(s)$, $C(s) = C_N(s)/C_D(s)$ and $F(s) = F_N(s)/F_D(s)$ are rational, then

$$Y = \frac{G_N(C_N F_D + C_D F_N)}{F_D(G_D C_D + G_N C_N)} R + \frac{G_D C_D}{G_D C_D + G_N C_N} D - \frac{G_N C_N}{G_D C_D + G_N C_N} N. \tag{22}$$

We can make the following observations:
- For Y to be stable, both F_D and $G_D C_D + G_N C_N$ must be Hurwitz.
- When $D = N = 0$, if we choose $F = G^{-1}$, then

$$E = R - Y = R - \frac{G(C+F)}{1+GC} R = R - \frac{GC+1}{1+GC} R = 0.$$

  This zero steady-state error is achieved with, in theory, instantaneous response.

Example 6 is meant to introduce the topic of feedforward control. In a 2DOF architecture like that for *internal model control*, feedforward is meant to provide good performance, while feedback is meant to provide robustness and good disturbance rejection [ÅH06, Sect. 5.2]. Applications of feedforward control include process control [Cor04, Sect. 6.6], vehicle control and robotics [SHV06, Cor11]. Nevertheless, there are some caveats for using feedforward control [ÅH06, Sect. 5.2]:

- As shown in Example 6, both F_D and $G_D C_D + G_N C_N$ must be Hurwitz.
- Since $F(s) = G(s)^{-1}$, a reasonably accurate plant model $G(s)$ is required, and furthermore the plant $G(s)$ must be *rational* and *minimum-phase*. A transfer function with dead time is not rational, and its inverse is noncausal (i.e., output starting before input).
- For a strictly proper $G(s)$ with relative degree 1, $F(s)$ involves differentiation. However, differentiation amplifies noise, and pure differentiation is unrealizable in software or hardware.
- For a strictly proper $G(s)$ with higher relative degree than 1, $F(s)$ involves impractical higher-order differentiation.
- In view of the difficulties above, $F(s)$ is often implemented as just an approximation of $G(s)^{-1}$ [ÅH06, Sect. 5.2].

4 Summary

- The stability of an LTI system can be viewed in terms of its internal stability and external stability, but since internal stability implies external stability (see Theorem 1), we are only concerned with internal stability.
- A system is (internally) stable if any of the following equivalent statements is true:
  - All its poles have a negative real part.
  - All its poles are located in the open left half s-plane.
  - Its characteristic polynomial is Hurwitz.
- Static error constants are defined to facilitate the calculation of step-response, ramp-response and parabolic-response steady-state errors for the unity feedback configuration (see Figure 7(a)). Nonunity feedback configurations can be converted into unity feedback configurations using Eq. (11).

- A type n system has n poles at the origin in the forward path, and can track a $1/s^n$ reference signal with zero steady-state error. Additionally, it can reject a $1/s^n$ disturbance.
- The sensitivity function, $S(s)$, should be small for good disturbance rejection, whereas the complementary sensitivity function, $T(s)$, should be small for noise rejection. Since $S(s) + T(s) = 1$, the standard practice is to make $S(s)$ small for low frequencies, but $T(s)$ small for high frequencies.
- In a two-degree-of-freedom architecture (e.g., Figure 9), feedforward control is designed to provide good performance, while feedback control is designed to provide good disturbance rejection.

References

[ÅH06] Karl Johan Åström and Tore Hägglund. *Advanced PID control*. ISA-Instrumentation, Systems, and Automation Society, 2006.

[ARMAH14] Haitham Abu-Rub, Mariusz Malinowski, and Kamal Al-Haddad, editors. *Power Electronics for Renewable Energy Systems, Transportation and Industrial Applications*. John Wiley & Sons Ltd, 2014.

[AV16] Victor M. Alfaro and Ramon Vilanova. *Model-Reference Robust Tuning of PID Controllers*. Springer International Publishing, 2016.

[Che99] Chi-Tsong Chen. *Linear System Theory and Design*. Oxford University Press, 3rd edition, 1999.

[Coo07] Michale V. Cook. *Flight Dynamics Principles: A Linear Systems Approach to Aircraft Stability and Control*. Elsevier Ltd., 2nd edition, 2007.

[Cor04] J.-P. Corriou. *Process Control: Theory and Applications*. Springer-Verlag London Ltd., 2004.

[Cor11] Peter Corke. *Robotics, Vision and Control: Fundamental Algorithms in MATLAB®*. Springer-Verlag Berlin Heidelberg, 2011.

[Cur14] Howard D. Curtis. *Orbital mechanics for engineering students*. Elsevier Ltd., 3rd edition, 2014.

[DB11] R.C. Dorf and R.H. Bishop. *Modern Control Systems*. Pearson, 12th edition, 2011.

[GGS00] Graham C. Goodwin, Stefan F. Graebe, and Mario E. Salgado. *Control System Design*. Prentice Hall, 2000.

[Kai80] Thomas Kailath. *Linear Systems*. Prentice-Hall, Inc., 1980.

[Nis15] Norman S. Nise. *Control Systems Engineering*. Wiley, 7th edition, 2015.

[rMss] K.J. Åström and R.M. Murray. *Feedback Systems: An Introduction for Scientists and Engineers*. Princeton University Press, 2nd edition, in progress. `http://www.cds.caltech.edu/~murray/amwiki/index.php/Second_Edition`.

[SHV06] Mark W Spong, Seth Hutchinson, and Mathukumalli Vidyasagar. *Robot modeling and control*. John Wiley & Sons, 2006.

[SP05] S. Skogestad and I. Postlethwaite. *Multivariable Feedback Control: Analysis and Design*. John Wiley & Sons, 2nd edition, 2005.

[WL07] Robert L. Williams II and Douglas A. Lawrence. *Linear State-Space Control Systems*. John Wiley & Sons, 2007.