Lecture 07: Understanding Poles and Zeros

System Poles and Zeros

The transfer function provides a basis for determining important system response characteristics without solving the complete differential equation. As defined, the transfer function is a rational function in the complex variable $s = \sigma + j\omega$, that is

$$H(s) = \frac{b_m s^m + b_{m-1} s^{m-1} + \ldots + b_1 s + b_0}{a_n s^n + a_{n-1} s^{n-1} + \ldots + a_1 s + a_0} \quad (1)$$

It is often convenient to factor the polynomials in the numerator and denominator, and to write the transfer function in terms of those factors

$$H(s) = \frac{N(s)}{D(s)} = K \frac{(s - z_1)(s - z_2)\cdots(s - z_{m-1})(s - z_m)}{(s - p_1)(s - p_2)\cdots(s - p_{n-1})(s - p_n)} \quad (2)$$

$$H(s) = \frac{N(s)}{D(s)} = K \frac{\prod_{j=1}^{m}(s - z_j)}{\prod_{k=1}^{n}(s - p_j)} \quad (3)$$

where the numerator and denominator polynomials, $N(s)$ and $D(s)$, have real coefficients defined by the system's differential equation and $K = b_m/a_n$. As written in Eq. (2) the z_i's are the roots of the equation.

$$N(s) = 0, \quad (3)$$

and are defined to be the system *zeros*, and the p_i's are the roots of the equation

$$D(s) = 0, \quad (4)$$

and are defined to be the system *poles*. In Eq. (2) the factors in the numerator and denominator are written so that when $s = z_i$ the numerator $N(s) = 0$ and the transfer function vanishes, that is

$$\lim_{s \to z_i} H(s) = 0.$$

and similarly when $s = p_i$ the denominator polynomial $D(s) = 0$ and the value of the transfer function becomes unbounded,

$$\lim_{s \to p_i} H(s) = \infty.$$

All of the coefficients of polynomials $N(s)$ and $D(s)$ are real, therefore the poles and zeros must be either purely real or appear in complex conjugate pairs. In general for the poles, either $p_i = \sigma_i$, or else $p_i, p_{i+1} = \sigma_i \pm j\omega_i$. The existence of a single complex pole without a corresponding conjugate pole would generate complex coefficients in the polynomial $D(s)$. Similarly, the system zeros are either real or appear in complex conjugate pairs.

■ *Example*: A linear system is described by the differential equation. Find the system poles and zeros.

$$\frac{d^2y}{dt^2} + 5\frac{dy}{dt} + 6y = 2\frac{du}{dt} + 1$$

Solution: From the differential equation the transfer function is

$$H(s) = \frac{2s+1}{s^2+5s+6}. \tag{5}$$

which may be written in factored form

$$\begin{aligned} H(s) &= \frac{1}{2}\frac{s+1/2}{(s+3)(s+2)} \\ &= \frac{1}{2}\frac{s-(-1/2)}{(s-(-3))(s-(-2))}. \end{aligned} \tag{6}$$

The system therefore has a single real zero at $s = -1/2$, and a pair of real poles at $s = -3$ and $s = -2$.

-- -----------------------------

The poles and zeros are properties of the transfer function, and therefore of the differential equation describing the input-output system dynamics. Together with the gain constant K they completely characterize the differential equation and provide a complete description of the system.

■ Example: A system has a pair of complex conjugate poles p_1, p_2 = -1 ± $j2$, a single real zero z_1 = -4, and a gain factor $K=3$. Find the differential equation representing the system.

Solution: The transfer function is

$$\begin{aligned} H(s) &= K\frac{s-z}{(s-p_1)(s-p_2)} \\ &= 3\frac{s-(-4)}{(s-(-1+j2))(s-(-1-j2))} \\ &= 3\frac{(s+4)}{s^2+2s+5} \end{aligned} \tag{7}$$

and the differential equation is

$$\frac{d^2y}{dt^2} + 2\frac{dy}{dt} + 5y = 3\frac{du}{dt} + 12u \tag{8}$$

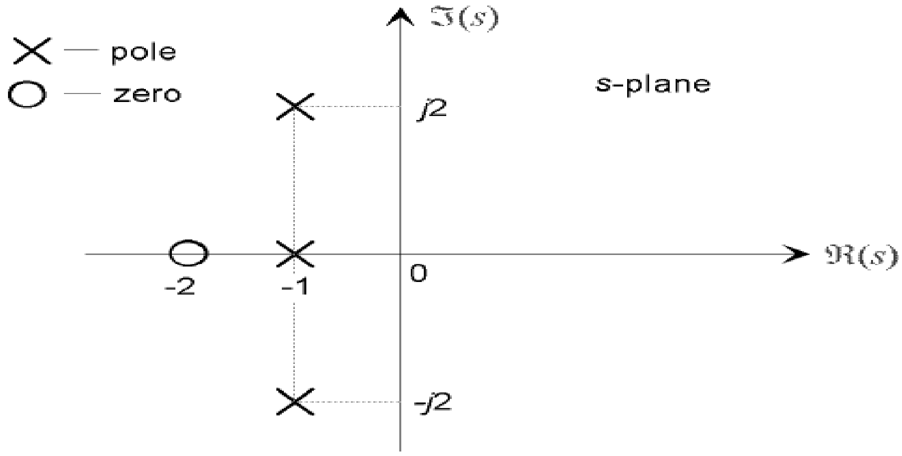


Figure 1: The pole-zero plot for a typical third-order system with one real pole and a complex conjugate pole pair, and a single real zero.

1.1 Pole-zero plots:

A system is characterized by its poles and zeros in the sense that they allow reconstruction of the input/output differential equation. In general, the poles and zeros of a transfer function may be complex, and the system dynamics may be represented graphically by plotting their locations on the complex *s*-plane, whose axes represent the real and imaginary parts of the complex variable *s*. Such plots are known as *pole-zero plots*. It is usual to mark a zero location by a circle (◦) and a pole location a cross (×). The location of the poles and zeros provide qualitative insights into the response characteristics of a system. Figure 1 is an example of a pole-zero plot for a third-order system with a single real zero, a real pole and a complex conjugate pole pair, that is;

$$H(s) = \frac{(3s+6)}{(s^3+3s^2+7s+5)} = 3\frac{(s-(-2))}{(s-(-1))(s-(-1-2j))(s-(-1+2j))}$$

System poles and the homogeneous response:

Because the transfer function completely represents a system differential equation, its poles and zeros effectively define the system response. In particular, the system poles directly define the components in the homogeneous response. The unforced response of a linear SISO system to a set of initial conditions is

$$y_h(t) = \sum_{i=1}^{n} C_i e^{\lambda_i t}$$

where the constants C_i are determined from the given set of initial conditions and the exponents λ_i are the roots of the *characteristic equation* or the system *eigenvalues*. The characteristic equation is

$$D(s) = s^n + a_{n-1}s^{n-1} + ... + a_1 s + a_0 = 0$$

and its roots are the system poles, that is $\lambda_i = p_i$, leading to the following important relationship:

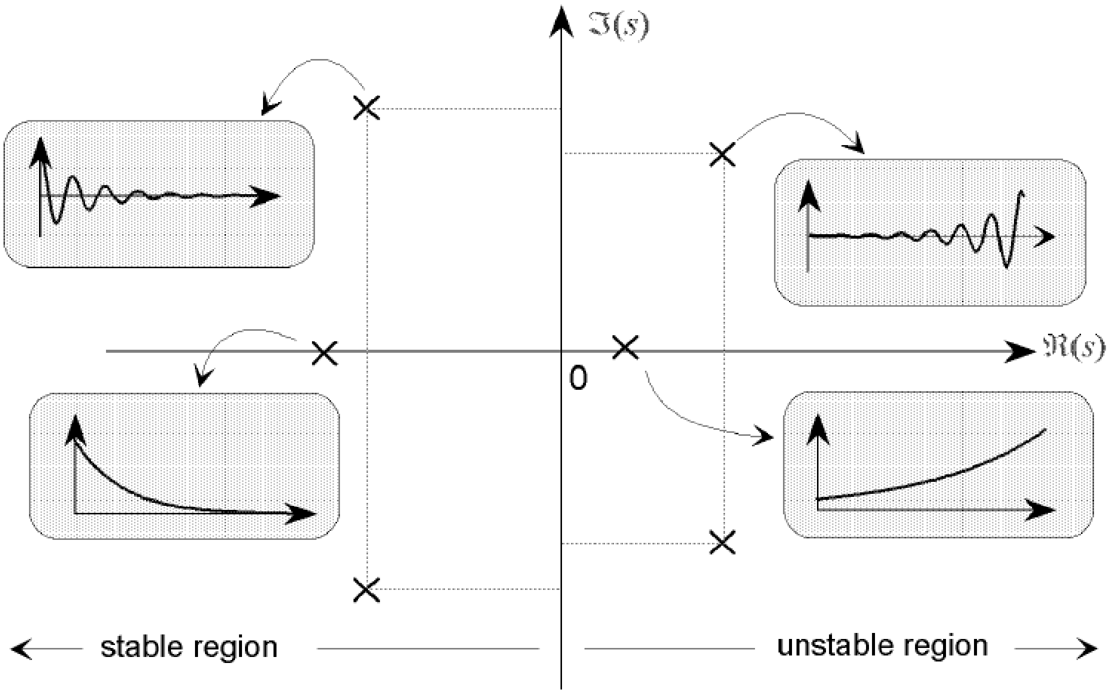


Figure 2: The specification of the form of components of the homogeneous response from the system pole locations on the pole-zero plot.

The transfer function poles are the roots of the characteristic equation, and also the eigenvalues of the system A matrix. The homogeneous response may, therefore, be written

$$y_h(t) = \sum_{i=1}^{n} C_i e^{p_i t}$$

The location of the poles in the *s*-plane, therefore, define the *n* components in the homogeneous response as described below:

1. A real pole $p_i = -\sigma$ in the left half of the *s*-plane defines an exponentially decaying component, $Ce^{-\sigma t}$, in the homogeneous response. The rate of the decay is determined by the pole location; poles far from the origin in the left-half plane correspond to components that decay rapidly, while poles near the origin correspond to slowly decaying components.

2. A pole at the origin $p_i = 0$ defines a component that is constant in amplitude and defined by the initial conditions.

3. A real pole in the right-half plane corresponds to an exponentially increasing component $Ce^{\sigma t}$ in the homogeneous response; thus defining the system to be unstable.

4. A complex conjugate pole pair $\sigma \pm j\omega$ in the left half of the *s*-plane combine to generate a response component that is a decaying sinusoid of the form $Ae^{-\sigma t} \sin(\omega t + \varphi)$ where A and φ are determined by the initial conditions. The rate of decay is specified by σ; the frequency of oscillation is determined by ω.

5. An imaginary pole pair, that is a pole pair lying on the imaginary axis, $\pm j\omega$ generates an oscillatory component with a constant amplitude determined by the initial conditions.
6. A complex pole pair in the right half-plane generates an exponentially increasing component.

These results are summarized in Fig. 2.

■ Example: Comment on the expected form of the response of a system with a pole-zero plot shown in Fig. 3 to an arbitrary set of initial conditions.

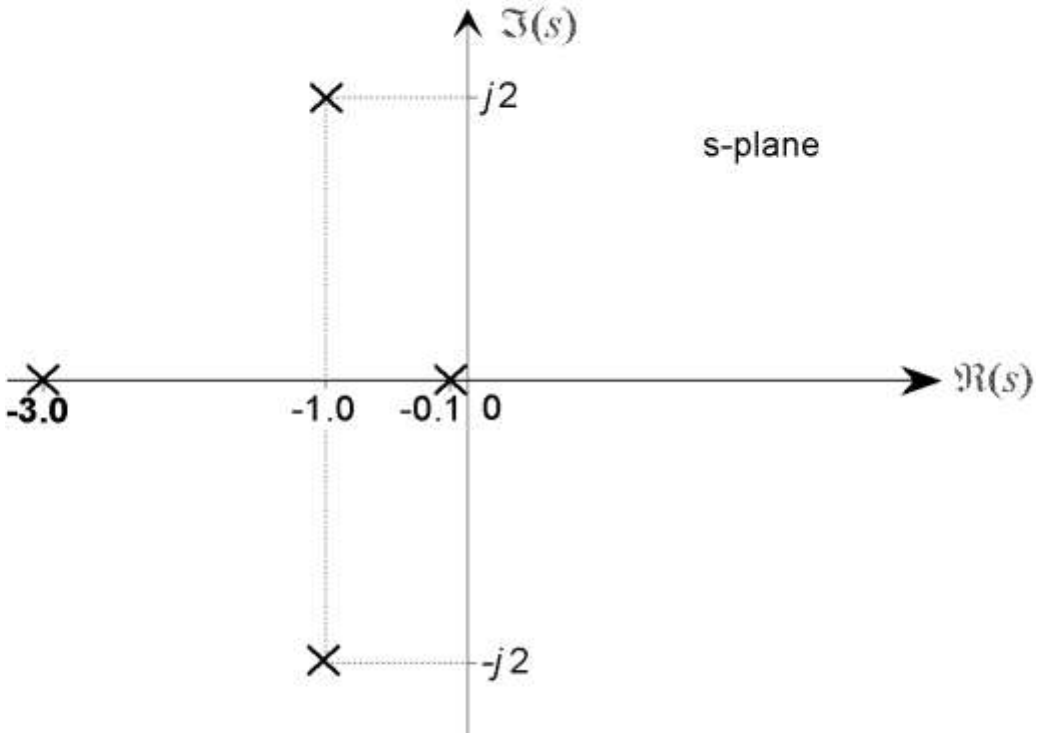


Figure 3: Pole-zero plot of a fourth-order system with two real and two complex conjugate poles.

Solution: The system has four poles and no zeros. The two real poles correspond to decaying exponential terms C_1e^{-3t} and $C_2e^{-0.1t}$, and the complex conjugate pole pair introduce an oscillatory component $Ae^{-t}\sin(2t+\varphi)$ so that the total homogeneous response is

$$y_h(t) = C_1e^{-3t} + C_2e^{-0.1t} + Ae^{-t}\sin(2t+\varphi)$$

Although the relative strengths of these components in any given situation is determined by the set of initial conditions, the following general observations may be made:

1. The term e^{-3t}, with a time-constant τ of 0.33 seconds, decays rapidly and is significant only for approximately 4τ or 1.33seconds.
2. The response has an oscillatory component $Ae^{-t}\sin(2t+\varphi)$ defined by the complex conjugate pair and exhibits some overshoot. The oscillation will decay in approximately four seconds because of the e^{-t} damping term.
3. The term $e^{-0.1t}$, with a time-constant τ = 10 seconds, persists for approximately 40 seconds. It is, therefore, the *dominant* long term response component in the overall homogeneous response.

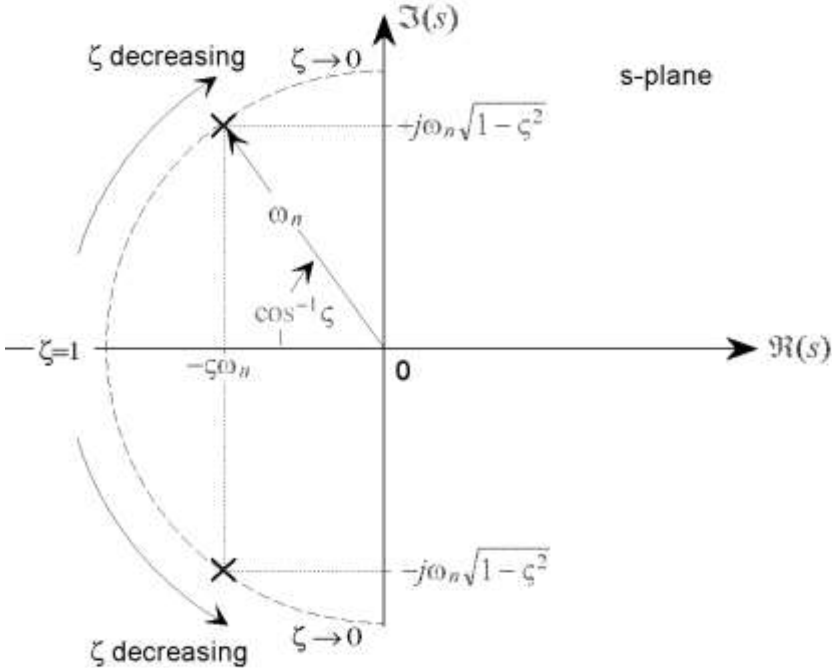


Figure 4: Definition of the parameters ω_n and ζ for an underdamped, second-order system from the complex conjugate pole locations.

The pole locations of the classical second-order homogeneous system

$$\frac{d^2y}{dt^2} + 2\zeta\omega_n \frac{dy}{dt} + \omega_n^2 y = 0$$

Roots are: $p_{1,2} = -\zeta\omega_n \pm \omega_n\sqrt{\zeta^2 - 1}$

If $\zeta \geq 1$, corresponding to an overdamped system, the two poles are real and lie in the left-half plane. For an underdamped system, $0 \leq \zeta < 1$, the poles form a complex conjugate pair,

$$p_{1,2} = -\zeta\omega_n \pm j\omega_n\sqrt{1-\zeta^2}$$

and are located in the left-half plane, as shown in **Fig. 4**. From this figure, it can be seen that the poles lie at a distance ω_n from the origin, and at an angle $\pm \cos^{-1}(\zeta)$ from the negative real axis. The poles for an underdamped second-order system, therefore, lie on a semi-circle with a radius defined by ω_n, at an angle defined by the value of the damping ratio ζ.

System stability

The stability of a linear system may be determined directly from its transfer function. An nth order linear system is asymptotically stable only if all of the components in the homogeneous response from a finite set of initial conditions decay to zero as time increases, or

$$\lim_{t\to\infty}\sum_{i=1}^{n} C_i e^{p_i t} = 0$$

where the p_i are the system poles. In a stable system, all components of the homogeneous response must decay to zero as time increases. If any pole has a positive real part there is a component in the output that increases without bound, causing the system to be unstable.

In order for a linear system to be stable, all of its poles must have negative real parts, that is they must all lie within the left-half of the s-plane. An "unstable" pole, lying in the right half of the s-plane, generates a component in the system homogeneous response that increases without bound from any finite initial conditions. A system having one or more poles lying on the imaginary axis of the s-plane has non-decaying oscillatory components in its homogeneous response, and is defined to be *marginally* stable.

2 Geometric Evaluation of the Transfer Function

The transfer function may be evaluated for any value of $s = \sigma + j\omega$, and in general, when s is complex the function $H(s)$ itself is complex. It is common to express the complex value of the transfer function in polar form as a magnitude and an angle:

$$H(s) = |H(s)|\, e^{j\phi(s)}, \tag{17}$$

with a magnitude $|H(s)|$ and an angle $\phi(s)$ given by

$$|H(s)| = \sqrt{\Re\{H(s)\}^2 + \Im\{H(s)\}^2}, \tag{18}$$

$$\phi(s) = \tan^{-1}\left(\frac{\Im\{H(s)\}}{\Re\{H(s)\}}\right) \tag{19}$$

where $\Re\{\}$ is the real operator, and $\Im\{\}$ is the imaginary operator. If the numerator and denominator polynomials are factored into terms $(s - p_i)$ and $(s - z_i)$ as in Eq. (2),

$$H(s) = K\frac{(s-z_1)(s-z_2)\dots(s-z_{m-1})(s-z_m)}{(s-p_1)(s-p_2)\dots(s-p_{n-1})(s-p_n)} \tag{20}$$

each of the factors in the numerator and denominator is a complex quantity, and may be interpreted as a *vector* in the s-plane, originating from the point z_i or p_i and directed to the point s at which the function is to be evaluated. Each of these vectors may be written in polar form in terms of a magnitude and an angle, for example for a pole $p_i = \sigma_i + \omega_i$, the magnitude and angle of the vector to the point $s = \sigma + \omega$ are

$$|s - p_i| = \sqrt{(\sigma - \sigma_i)^2 + (\omega - \omega_i)^2}, \tag{21}$$

$$\angle(s - p_i) = \tan^{-1}\left(\frac{\omega - \omega_i}{\sigma - \sigma_i}\right) \tag{22}$$

as shown in Fig. 5a. Because the magnitude of the product of two complex quantities is the product of the individual magnitudes, and the angle of the product is the sum of the component angles (Appendix B), the magnitude and angle of the complete transfer function may then be written

$$|H(s)| = K\frac{\prod_{i=1}^{m}|(s-z_i)|}{\prod_{i=1}^{n}|(s-p_i)|} \tag{23}$$

$$\angle H(s) = \sum_{i=1}^{m}\angle(s-z_i) - \sum_{i=1}^{n}\angle(s-p_i). \tag{24}$$

The magnitude of each of the component vectors in the numerator and denominator is the distance of the point s from the pole or zero on the s-plane. Therefore if the vector from the pole p_i to the point s on a pole-zero plot has a length q_i and an angle θ_i from the horizontal, and the vector from

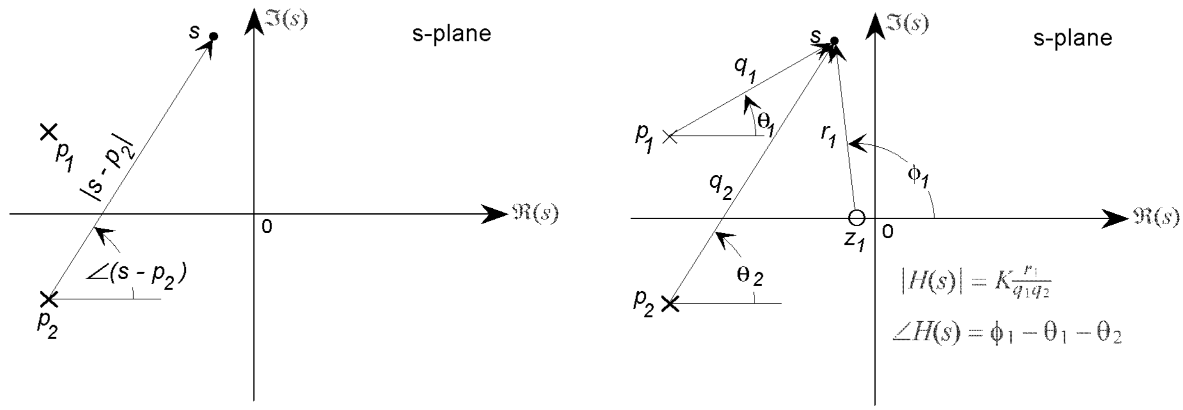


Figure 5: (a) Definition of s-plane geometric relationships in polar form, (b) Geometric evaluation of the transfer function from the pole-zero plot.

the zero z_i to the point s has a length r_i and an angle ϕ_i, as shown in Fig. 5b, the value of the transfer function at the point s is

$$|H(s)| = K\frac{r_1 \dots r_m}{q_1 \dots q_n} \tag{25}$$

$$\angle H(s) = (\phi_1 + \dots + \phi_m) - (\theta_1 + \dots + \theta_n) \tag{26}$$

The transfer function at any value of s may therefore be determined geometrically from the pole-zero plot, except for the overall "gain" factor K. The magnitude of the transfer function is proportional to the product of the geometric distances on the s-plane from each zero to the point s divided by the product of the distances from each pole to the point. The angle of the transfer function is the sum of the angles of the vectors associated with the zeros minus the sum of the angles of the vectors associated with the poles.

■ Example

A second-order system has a pair of complex conjugate poles a $s = -2 \pm j3$ and a single zero at the origin of the s-plane. Find the transfer function and use the pole-zero plot to evaluate the transfer function at $s = 0 + j5$.

Solution: From the problem description

$$\begin{aligned} H(s) &= K\frac{s}{(s-(-2+j3))(s-(-2-j3))} \\ &= K\frac{s}{s^2+4s+13} \end{aligned} \tag{27}$$

The pole-zero plot is shown in Fig. 6. From the figure the transfer function is

$$\begin{aligned} |H(s)| &= K\frac{\sqrt{(0-5)^2}}{\sqrt{(0-(-2))^2+(5-3)^2}\sqrt{(0-(-2))^2+(5-(-3))^2}} \\ &= K\frac{5}{4\sqrt{34}} \end{aligned} \tag{28}$$

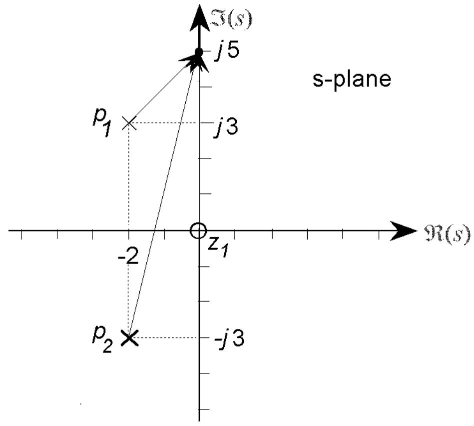


Figure 6: The pole-zero plot for a second order system with a zero at the origin.

and

$$\begin{aligned}\angle H(s) &= \tan^{-1}(5/0) - \tan^{-1}(2/2) - \tan^{-1}(8/2) \\ &= -31^0 \end{aligned} \tag{29}$$

---

3 Frequency Response and the Pole-Zero Plot

The frequency response may be written in terms of the system poles and zeros by substituting $j\omega$ for s directly into the factored form of the transfer function:

$$H(j\omega) = K\frac{(j\omega - z_1)(j\omega - z_2)\dots(j\omega - z_{m-1})(j\omega - z_m)}{(j\omega - p_1)(j\omega - p_2)\dots(j\omega - p_{n-1})(j\omega - p_n)}. \tag{30}$$

Because the frequency response is the transfer function evaluated on the imaginary axis of the s-plane, that is when $s = j\omega$, the graphical method for evaluating the transfer function described above may be applied directly to the frequency response. Each of the vectors from the n system poles to a test point $s = j\omega$ has a magnitude and an angle:

$$|j\omega - p_i| = \sqrt{\sigma_i^2 + (\omega - \omega_i)^2}, \tag{31}$$

$$\angle(s - p_i) = \tan^{-1}\left(\frac{\omega - \omega_i}{-\sigma_i}\right), \tag{32}$$

as shown in Fig. 7a, with similar expressions for the vectors from the m zeros. The magnitude and phase angle of the complete frequency response may then be written in terms of the magnitudes and angles of these component vectors

$$|H(j\omega)| = K\frac{\prod_{i=1}^{m}|(j\omega - z_i)|}{\prod_{i=1}^{n}|(j\omega - p_i)|} \tag{33}$$

$$\angle H(j\omega) = \sum_{i=1}^{m}\angle(j\omega - z_i) - \sum_{i=1}^{n}\angle(j\omega - p_i). \tag{34}$$

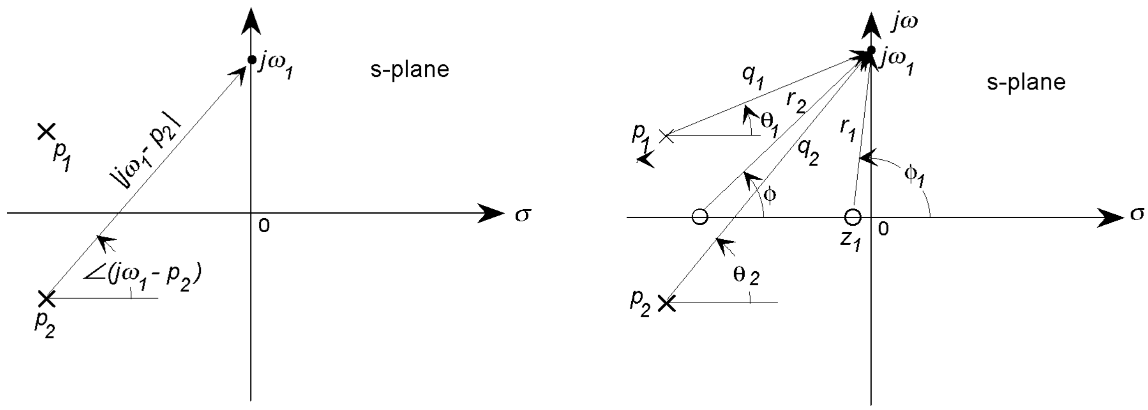


Figure 7: Definition of the vector quantities used in defining the frequency response function from the pole-zero plot. In (a) the vector from a pole (or zero) is defined, in (b) the vectors from all poles and zeros in a typical system are shown.

As defined above, if the vector from the pole p_i to the point $s = j\omega$ has length q_i and an angle θ_i from the horizontal, and the vector from the zero z_i to the point $j\omega$ has a length r_i and an angle ϕ_i, as shown in Fig. 7b, the value of the frequency response at the point $j\omega$ is

$$|H(j\omega)| = K\frac{r_1 \dots r_m}{q_1 \dots q_n} \tag{35}$$

$$\angle H(j\omega) = (\phi_1 + \dots + \phi_m) - (\theta_1 + \dots + \theta_n) \tag{36}$$

The graphical method can be very useful for deriving a qualitative picture of a system frequency response. For example, consider the sinusoidal response of a first-order system with a pole on the real axis at $s = -1/\tau$ as shown in Fig. 8a, and its Bode plots in Fig. 8b. Even though the gain constant K cannot be determined from the pole-zero plot, the following observations may be made directly by noting the behavior of the magnitude and angle of the vector from the pole to the imaginary axis as the input frequency is varied:

1. At low frequencies the gain approaches a finite value, and the phase angle has a small but finite lag.

2. As the input frequency is increased the gain decreases (because the length of the vector increases), and the phase lag also increases (the angle of the vector becomes larger).

3. At very high input frequencies the gain approaches zero, and the phase angle approaches $\pi/2$.

As a second example consider a second-order system, with the damping ratio chosen so that the pair of complex conjugate poles are located close to the imaginary axis as shown in Fig. 9a. In this case there are a pair of vectors connecting the two poles to the imaginary axis, and the following conclusions may be drawn by noting how the lengths and angles of the vectors change as the test frequency moves up the imaginary axis:

1. At low frequencies there is a finite (but undetermined) gain and a small but finite phase lag associated with the system.

2. As the input frequency is increased and the test point on the imaginary axis approaches the pole, one of the vectors (associated with the pole in the second quadrant) decreases in length

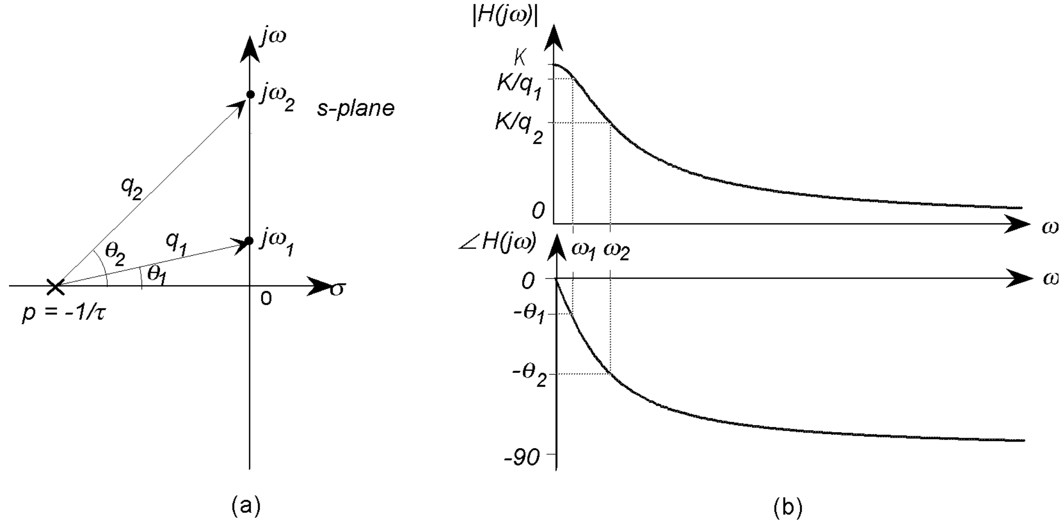


Figure 8: The pole-zero plot of a first-order system and its frequency response functions.

and at some point reaches a minimum. There is an increase in the value of the magnitude function over a range of frequencies close to the pole.

3. At very high frequencies, the lengths of both vectors tend to infinity, and the magnitude of the frequency response tends to zero, while the phase approaches an angle of π radians because the angle of each vector approaches $\pi/2$.

The following generalizations may be made about the sinusoidal frequency response of a linear system, based upon the geometric interpretation of the pole-zero plot:

1. If a system has an excess of poles over the number of zeros the magnitude of the frequency response tends to zero as the frequency becomes large. Similarly, if a system has an excess of zeros the gain increases without bound as the frequency of the input increases. This cannot happen in physical energetic systems because it implies an infinite power gain through the system.

2. If a system has a pair of complex conjugate poles close to the imaginary axis, the magnitude of the frequency response has a "peak", or resonance at frequencies in the proximity of the pole. If the pole pair lies directly upon the imaginary axis, the system exhibits an infinite gain at that frequency.

3. If a system has a pair of complex conjugate zeros close to the imaginary axis, the frequency response has a "dip" or "notch" in its magnitude function at frequencies in the vicinity of the zero. Should the pair of zeros lie directly upon the imaginary axis, the response is identically zero at the frequency of the zero, and the system does not respond at all to sinusoidal excitation at that frequency.

4. A pole at the origin of the s-plane (corresponding to a pure integration term in the transfer function) implies an infinite gain at zero frequency.

5. Similarly a zero at the origin of the s-plane (corresponding to a pure differentiation) implies a zero gain for the system at zero frequency.

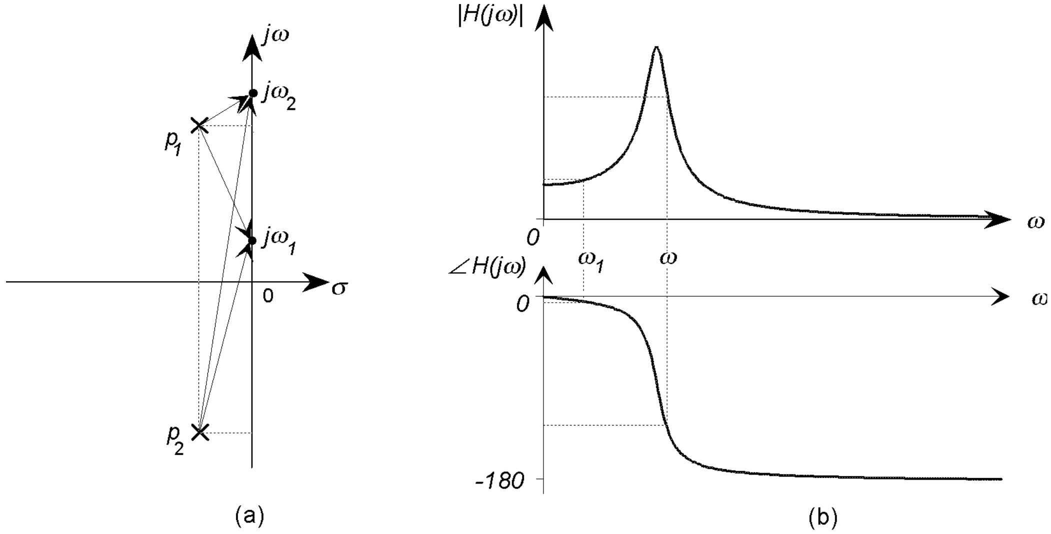


Figure 9: The pole-zero plot for a second-order system and its its frequency response functions.

3.1 A Simple Method for constructing the Magnitude Bode Plot directly from the Pole-Zero Plot

The pole-zero plot of a system contains sufficient information to define the frequency response except for an arbitrary gain constant. It is often sufficient to know the shape of the magnitude Bode plot without knowing the absolute gain. The method described here allows the magnitude plot to be sketched by inspection, without drawing the individual component curves. The method is based on the fact that the overall magnitude curve undergoes a *change* in slope at each break frequency.

The first step is to identify the break frequencies, either by factoring the transfer function or directly from the pole-zero plot. Consider a typical pole-zero plot of a linear system as shown in Fig. 10a. The break frequencies for the four first and second-order blocks are all at a frequency equal to the radial distance of the poles or zeros from the origin of the s-plane, that is $\omega_b = \sqrt{\sigma^2 + \omega^2}$. Therefore all break frequencies may be found by taking a compass and drawing an arc from each pole or zero to the positive imaginary axis. These break frequencies may be transferred directly to the logarithmic frequency axis of the Bode plot.

Because all low frequency asymptotes are horizontal lines with a gain of 0dB, a pole or zero does not contribute to the magnitude Bode plot below its break frequency. Each pole or zero contributes a change in the *slope* of the asymptotic plot of ± 20 dB/decade above its break frequency. A complex conjugate pole or zero pair defines *two* coincident breaks of ± 20 dB/decade (one from each member of the pair), giving a total change in the slope of ± 40 dB/decade. Therefore, at any frequency ω, the slope of the asymptotic magnitude function depends only on the number of break points at frequencies less than ω, or to the left on the Bode plot. If there are Z breakpoints due to zeros to the left, and P breakpoints due to poles, the slope of the curve at that frequency is $20 \times (Z - P)$ dB/decade.

Any poles or zeros at the origin cannot be plotted on the Bode plot, because they are effectively to the left of all finite break frequencies. However, they define the initial slope. If an arbitrary starting frequency and an assumed gain (for example 0dB) at that frequency are chosen, the shape of the magnitude plot may be easily constructed by noting the initial slope, and constructing the

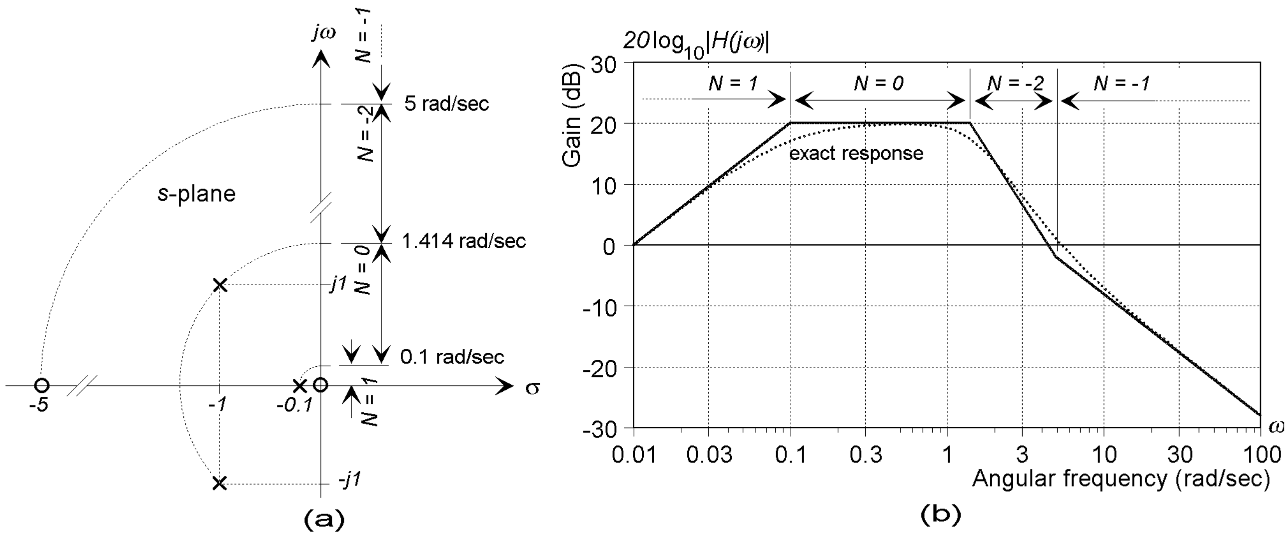
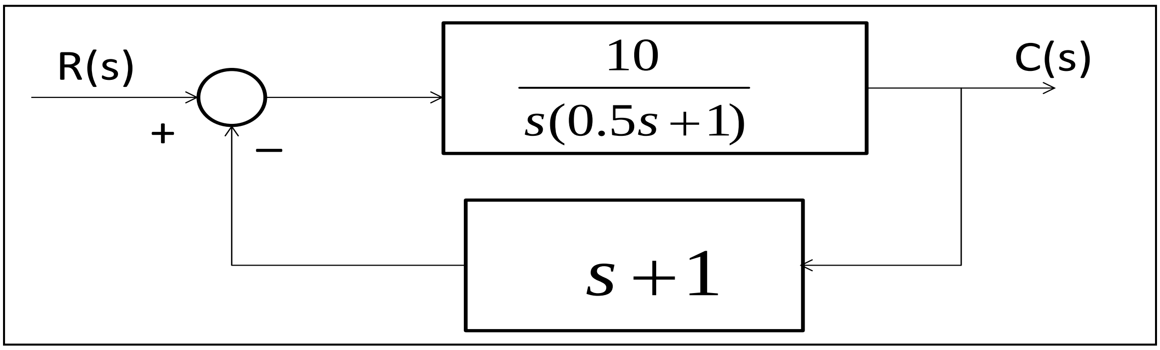


Figure 10: Construction of the magnitude Bode plot from the pole-zero diagram: (a) shows a typical third-order system, and the definition of the break frequencies, (b) shows the Bode plot based on changes in slope at the break frequencies

curve from straight line segments that change in slope by units of ± 20 dB/decade at the breakpoints. The arbitrary choice of the reference gain results in a vertical displacement of the curve.

Figure 10b shows the straight line magnitude plot for the system shown in Fig. 10a constructed using this method. A frequency range of 0.01 to 100 radians/sec was arbitrarily selected, and a gain of 0dB at 0.01 radians/sec was assigned as the reference level. The break frequencies at 0, 0.1, 1.414, and 5 radians/sec were transferred to the frequency axis from the pole-zero plot. The value of N at any frequency is $Z - P$, where Z is the number of zeros to the left, and P is the number of poles to the left. The curve was simply drawn by assigning the value of the slope in each of the frequency intervals and drawing connected lines.

Q1. For the following system, find the poles and zeros. Also, establish the differential equation for the system.



Summary: Differences between Poles and Zeros of a transfer function

Let us have a look at the differences between Poles and Zeros and their effects for a given function:

$$H(s) = \frac{N(s)}{D(s)}$$

1. Definition:
   - Poles are the roots of the denominator of a transfer function.
   - Zeros are the roots of the nominator of a transfer function.

2. Determination:
   - Poles are determined by equating D(s) with 0 and solving for s.
   - *Zeros are determined by equating N(s) with 0 and solving for s.*

3. Amount:
   - The number of poles is always greater or equal to the Zeros.
   - The numbers of Zeros are lesser or equal to Poles.

4. Determination of output:
   - Poles in a transfer function explain that the output has reached to infinity.
   - Whereas, the zeros in a transfer function indicate that the output has reached zero.

5. Effect of Additional Poles and Zeros In first-order systems:
   - Additional Poles delay the response of a system.
   - Left half-plane zeros speed up the response of a system and the right half-plane cause the response to go in the opposite direction.

6. Effect of Additional Poles and Zeros in Second-order systems:
   - Additional Poles in a dominantly second-order system decrease the number of oscillations.
   - Additional Zeros in a dominantly second-order system increases the number of oscillations.

Conclusion

The frequencies that turn nominator or denominator zero are called zeros and poles of a transfer function respectively. They determine the stability and working of a system.