Contents

Articles

References

Article Licenses

LTI system theory

Linear time-invariant theory, commonly known as **LTI system theory,** comes from applied mathematics and has direct applications in NMR spectroscopy, seismology, circuits, signal processing, control theory, and other technical areas. It investigates the response of a linear and time-invariant system to an arbitrary input signal. Trajectories of these systems are commonly measured and tracked as they move through time (e.g., an acoustic waveform), but in applications like image processing and field theory, the LTI systems also have trajectories in spatial dimensions. Thus, these systems are also called *linear translation-invariant* to give the theory the most general reach. In the case of generic discrete-time (i.e., sampled) systems, *linear shift-invariant* is the corresponding term. A good example of LTI systems are electrical circuits that can be made up of resistors, capacitors, and inductors.[1]

Overview

The defining properties of any LTI system are *linearity* and *time invariance*.

- *Linearity* means that the relationship between the input and the output of the system is a linear map: If input $x_1(t)$ produces response $y_1(t),$ and input $x_2(t)$ produces response $y_2(t),$ then the *scaled* and *summed* input $a_1x_1(t)+a_2x_2(t)$ produces the scaled and summed response $a_1y_1(t)+a_2y_2(t)$ where a_1 and a_2 are real scalars. It follows that this can be extended to an arbitrary number of terms, and so for real numbers $c_1, c_2, \ldots, c_k$,

  $$\text{Input} \quad \sum_k c_k\, x_k(t) \quad \text{produces output} \quad \sum_k c_k\, y_k(t).$$

  In particular,

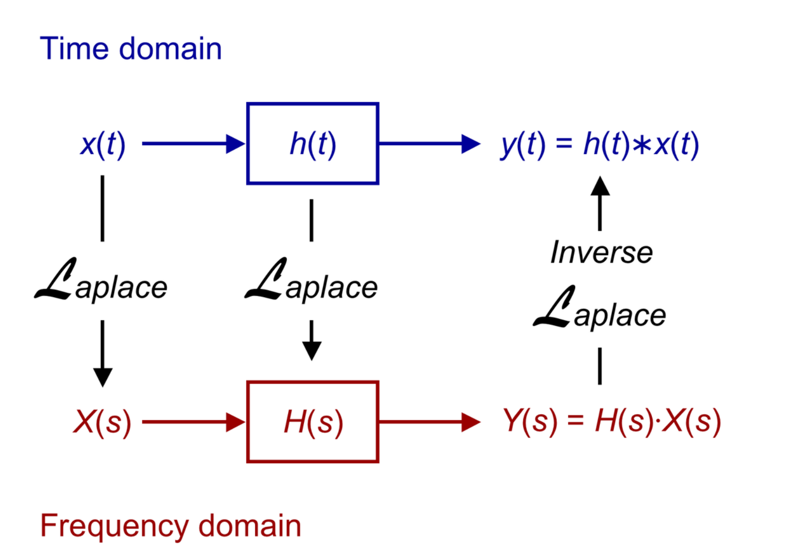
  $$\text{Input} \quad \int_{-\infty}^{\infty} c_\omega\, x_\omega(t)\, \mathrm{d}\omega \quad \text{produces output} \quad \int_{-\infty}^{\infty} c_\omega\, y_\omega(t)\, \mathrm{d}\omega \qquad \textbf{\textit{(Eq.1)}}$$

  where c_ω and x_ω are scalars and inputs that vary over a continuum indexed by ω. Thus if an input function can be represented by a continuum of input functions, combined "linearly", as shown, then the corresponding output function can be represented by the corresponding continuum of output functions, *scaled* and *summed* in the same way.

- *Time invariance* means that whether we apply an input to the system now or *T* seconds from now, the output will be identical except for a time delay of the *T* seconds. That is, if the output due to input $x(t)$ is $y(t)$, then the output due to input $x(t-T)$ is $y(t-T)$. Hence, the system is time invariant because the output does not depend on the particular time the input is applied.

The fundamental result in LTI system theory is that any LTI system can be characterized entirely by a single function called the system's impulse response. The output of the system is simply the convolution of the input to the system with the system's impulse response. This method of analysis is often called the *time domain* point-of-view. The same result is true of discrete-time linear shift-invariant systems in which signals are discrete-time samples, and convolution is defined on sequences.

Equivalently, any LTI system can be characterized in the *frequency domain* by the system's transfer function, which is the Laplace transform of the system's impulse response (or Z transform in the case of discrete-time systems). As a result of the properties of these transforms, the output of the system in the frequency domain is the product of the transfer function and the transform of the input. In other words, convolution in the time domain is equivalent to multiplication in the frequency domain.



Relationship between the **time domain** and the **frequency domain**

For all LTI systems, the eigenfunctions, and the basis functions of the transforms, are complex exponentials. This is, if the input to a system is the complex waveform Ae^{st} for some complex amplitude A and complex frequency s, the output will be some complex constant times the input, say Be^{st} for some new complex amplitude B. The ratio B/A is the transfer function at frequency s.

Because sinusoids are a sum of complex exponentials with complex-conjugate frequencies, if the input to the system is a sinusoid, then the output of the system will also be a sinusoid, perhaps with a different amplitude and a different phase, but always with the same frequency upon reaching steady-state. LTI systems cannot produce frequency components that are not in the input.

LTI system theory is good at describing many important systems. Most LTI systems are considered "easy" to analyze, at least compared to the time-varying and/or nonlinear case. Any system that can be modeled as a linear homogeneous differential equation with constant coefficients is an LTI system. Examples of such systems are electrical circuits made up of resistors, inductors, and capacitors (RLC circuits). Ideal spring–mass–damper systems are also LTI systems, and are mathematically equivalent to RLC circuits.

Most LTI system concepts are similar between the continuous-time and discrete-time (linear shift-invariant) cases. In image processing, the time variable is replaced with two space variables, and the notion of time invariance is replaced by two-dimensional shift invariance. When analyzing filter banks and MIMO systems, it is often useful to consider vectors of signals.

A linear system that is not time-invariant can be solved using other approaches such as the Green function method. The same method must be used when the initial conditions of the problem are not null.

Continuous-time systems

Impulse response and convolution

The behavior of a linear, continuous-time, time-invariant system with input signal x(t) and output signal y(t) is described by the convolution integral,[2] **:**

$$\begin{aligned} y(t) = x(t) * h(t) &\overset{\text{def}}{=} \int_{-\infty}^{\infty} x(t-\tau)\cdot h(\tau)\,\mathrm{d}\tau \\ &= \int_{-\infty}^{\infty} x(\tau)\cdot h(t-\tau)\,\mathrm{d}\tau, \qquad \text{(using commutativity)} \end{aligned}$$

where $h(t)$ is the system's response to an impulse**:** $x(\tau)=\delta(\tau)$. $y(t)$ is therefore proportional to a weighted average of the input function $x(\tau)$. The weighting function is $h(-\tau)$, simply shifted by amount t. As t changes, the weighting function emphasizes different parts of the input function. When $h(\tau)$ is zero for all negative τ, $y(t)$ depends only on values of x prior to time t, and the system is said to be causal.

To understand why the convolution produces the output of an LTI system, let the notation $\{x(u-\tau);\ u\}$ represent the function $x(u-\tau)$ with variable u and constant τ. And let the shorter notation $\{x\}$ represent $\{x(u);\ u\}$. Then a continuous-time system transforms an input function, $\{x\}$, into an output function, $\{y\}$. And in general, every value of the output can depend on every value of the input. This concept is represented by**:**

$$y(t) \overset{\text{def}}{=} O_t\{x\},$$

where O_t is the transformation operator for time t. In a typical system, $y(t)$ depends most heavily on the values of x that occurred near time t. Unless the transform itself changes with t, the output function is just constant, and the system is uninteresting.

For a linear system, O must satisfy **Eq.1:**

$$O_t\left\{\int_{-\infty}^{\infty} c_\tau\ x_\tau(u)\,\mathrm{d}\tau;\ u\right\} = \int_{-\infty}^{\infty} c_\tau\ \underbrace{y_\tau(t)}_{O_t\{x_\tau\}}\ \mathrm{d}\tau. \qquad \textbf{(\textit{Eq.2})}$$

And the time-invariance requirement is**:**

$$\begin{aligned} O_t\{x(u-\tau);\ u\} &= y(t-\tau) \\ &\overset{\text{def}}{=} O_{t-\tau}\{x\}. \end{aligned} \qquad \textbf{(\textit{Eq.3})}$$

In this notation, we can write the **impulse response** as $h(t) \overset{\text{def}}{=} O_t\{\delta(u);\ u\}$.

Similarly**:**

$$\begin{aligned} h(t-\tau) &\overset{\text{def}}{=} O_{t-\tau}\{\delta(u);\ u\} \\ &= O_t\{\delta(u-\tau);\ u\}. \qquad \text{(using } \textbf{Eq.3}\text{)} \end{aligned}$$

Substituting this result into the convolution integral**:**

$$\begin{aligned} x(t) * h(t) &= \int_{-\infty}^{\infty} x(\tau)\cdot h(t-\tau)\,\mathrm{d}\tau \\ &= \int_{-\infty}^{\infty} x(\tau)\cdot O_t\{\delta(u-\tau);\ u\}\,\mathrm{d}\tau, \end{aligned}$$

which has the form of the right side of **Eq.2** for the case $c_\tau = x(\tau)$ and $x_\tau(u) = \delta(u-\tau)$.

Eq.2 then allows this continuation**:**

$$\begin{aligned} x(t) * h(t) &= O_t\left\{\int_{-\infty}^{\infty} x(\tau)\cdot\delta(u-\tau)\,\mathrm{d}\tau;\ u\right\} \\ &= O_t\left\{x(u);\ u\right\} \\ &\stackrel{\text{def}}{=} y(t). \end{aligned}$$

In summary, the input function, $\{x\}$, can be represented by a continuum of time-shifted impulse functions, combined "linearly", as shown at ***Eq.1***. The system's linearity property allows the system's response to be represented by the corresponding continuum of impulse responses, combined in the same way. And the time-invariance property allows that combination to be represented by the convolution integral.

The mathematical operations above have a simple graphical simulation.[3]

Exponentials as eigenfunctions

An eigenfunction is a function for which the output of the operator is a scaled version of the same function. That is,

$$\mathcal{H}f = \lambda f,$$

where f is the eigenfunction and λ is the eigenvalue, a constant.

The exponential functions Ae^{st}, where $A, s \in \mathbb{C}$, are eigenfunctions of a linear, time-invariant operator. A simple proof illustrates this concept. Suppose the input is $x(t) = Ae^{st}$. The output of the system with impulse response $h(t)$ is then

$$\int_{-\infty}^{\infty} h(t-\tau)Ae^{s\tau}\,\mathrm{d}\tau$$

which, by the commutative property of convolution, is equivalent to

$$\begin{aligned} \overbrace{\int_{-\infty}^{\infty} h(\tau)\,Ae^{s(t-\tau)}\,\mathrm{d}\tau}^{\mathcal{H}f} &= \int_{-\infty}^{\infty} h(\tau)\,Ae^{st}e^{-s\tau}\,\mathrm{d}\tau \quad = Ae^{st}\int_{-\infty}^{\infty} h(\tau)\,e^{-s\tau}\,\mathrm{d}\tau \\ &= \underbrace{\overbrace{Ae^{st}}^{f}}_{\text{Input}}\underbrace{\overbrace{H(s)}^{\lambda}}_{\text{Scalar}}, \end{aligned}$$

where the scalar

$$H(s) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} h(t)e^{-st}\,\mathrm{d}t$$

is dependent only on the parameter s.

So the system's response is a scaled version of the input. In particular, for any $A, s \in \mathbb{C}$, the system output is the product of the input Ae^{st} and the constant $H(s)$. Hence, Ae^{st} is an eigenfunction of an LTI system, and the corresponding eigenvalue is $H(s)$.

Direct proof

It is also possible to directly derive complex exponentials as eigenfunctions of LTI systems.

Let's set $v(t) = e^{i\omega t}$ some complex exponential and $v_a(t) = e^{i\omega(t+a)}$ a time-shifted version of it.

$H[v_a](t) = e^{i\omega a}H[v](t)$ by linearity with respect to the constant $e^{i\omega a}$.

$H[v_a](t) = H[v](t+a)$ by time invariance of H .

So $H[v](t+a) = e^{i\omega a}H[v](t)$. Setting $t = 0$ and renamming we get :

$$H[v](\tau) = e^{i\omega\tau}H[v](0)$$

i.e. that a complex exponential $e^{i\omega\tau}$ as input will give a complex exponential of same frequency as output.

Fourier and Laplace transforms

The eigenfunction property of exponentials is very useful for both analysis and insight into LTI systems. The Laplace transform

$$H(s) \stackrel{\text{def}}{=} \mathcal{L}\{h(t)\} \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} h(t)e^{-st}\,\mathrm{d}t$$

is exactly the way to get the eigenvalues from the impulse response. Of particular interest are pure sinusoids (i.e., exponential functions of the form $e^{j\omega t}$ where $\omega \in \mathbb{R}$ and $j \stackrel{\text{def}}{=} \sqrt{-1}$). These are generally called complex exponentials even though the argument is purely imaginary. The Fourier transform $H(j\omega) = \mathcal{F}\{h(t)\}$ gives the eigenvalues for pure complex sinusoids. Both of $H(s)$ and $H(j\omega)$ are called the *system function*, *system response*, or *transfer function*.

The Laplace transform is usually used in the context of one-sided signals, i.e. signals that are zero for all values of *t* less than some value. Usually, this "start time" is set to zero, for convenience and without loss of generality, with the transform integral being taken from zero to infinity (the transform shown above with lower limit of integration of negative infinity is formally known as the bilateral Laplace transform).

The Fourier transform is used for analyzing systems that process signals that are infinite in extent, such as modulated sinusoids, even though it cannot be directly applied to input and output signals that are not square integrable. The Laplace transform actually works directly for these signals if they are zero before a start time, even if they are not square integrable, for stable systems. The Fourier transform is often applied to spectra of infinite signals via the Wiener–Khinchin theorem even when Fourier transforms of the signals do not exist.

Due to the convolution property of both of these transforms, the convolution that gives the output of the system can be transformed to a multiplication in the transform domain, given signals for which the transforms exist

$$y(t) = (h*x)(t) \stackrel{\text{def}}{=} \int_{-\infty}^{\infty} h(t-\tau)x(\tau)\,\mathrm{d}\tau \stackrel{\text{def}}{=} \mathcal{L}^{-1}\{H(s)X(s)\}.$$

Not only is it often easier to do the transforms, multiplication, and inverse transform than the original convolution, but one can also gain insight into the behavior of the system from the system response. One can look at the modulus of the system function |*H*(*s*)| to see whether the input $\exp(st)$ is *passed* (let through) the system or *rejected* or *attenuated* by the system (not let through).

Examples

- A simple example of an LTI operator is the derivative.
  - $\frac{\mathrm{d}}{\mathrm{d}t}\left(c_1 x_1(t) + c_2 x_2(t)\right) = c_1 x_1'(t) + c_2 x_2'(t)$ (i.e., it is linear)
  - $\frac{\mathrm{d}}{\mathrm{d}t} x(t-\tau) = x'(t-\tau)$ (i.e., it is time invariant)

  When the Laplace transform of the derivative is taken, it transforms to a simple multiplication by the Laplace variable *s*.

$$\mathcal{L}\left\{\frac{\mathrm{d}}{\mathrm{d}t} x(t)\right\} = sX(s)$$

  That the derivative has such a simple Laplace transform partly explains the utility of the transform.

- Another simple LTI operator is an averaging operator

$$\mathcal{A}\left\{x(t)\right\} \overset{\text{def}}{=} \int_{t-a}^{t+a} x(\lambda)\,\mathrm{d}\lambda.$$

  By the linearity of integration,

$$\begin{aligned}\mathcal{A}\left\{c_1 x_1(t) + c_2 x_2(t)\right\} &= \int_{t-a}^{t+a} \left(c_1 x_1(\lambda) + c_2 x_2(\lambda)\right)\,\mathrm{d}\lambda \\ &= c_1 \int_{t-a}^{t+a} x_1(\lambda)\,\mathrm{d}\lambda + c_2 \int_{t-a}^{t+a} x_2(\lambda)\,\mathrm{d}\lambda \\ &= c_1 \mathcal{A}\left\{x_1(t)\right\} + c_2 \mathcal{A}\left\{x_2(t)\right\},\end{aligned}$$

  it is linear. Additionally, because

$$\begin{aligned}\mathcal{A}\left\{x(t-\tau)\right\} &= \int_{t-a}^{t+a} x(\lambda-\tau)\,\mathrm{d}\lambda \\ &= \int_{(t-\tau)-a}^{(t-\tau)+a} x(\xi)\,\mathrm{d}\xi \\ &= \mathcal{A}\{x\}(t-\tau),\end{aligned}$$

  it is time invariant. In fact, $\mathcal{A}$ can be written as a convolution with the boxcar function $\Pi(t)$. That is,

$$\mathcal{A}\left\{x(t)\right\} = \int_{-\infty}^{\infty} \Pi\left(\frac{\lambda - t}{2a}\right) x(\lambda)\,\mathrm{d}\lambda,$$

  where the boxcar function

$$\Pi(t) \overset{\text{def}}{=} \begin{cases} 1 & \text{if } |t| < \frac{1}{2}, \\ 0 & \text{if } |t| > \frac{1}{2}. \end{cases}$$

Important system properties

Some of the most important properties of a system are causality and stability. Causality is a necessity if the independent variable is time, but not all systems have time as an independent variable. For example, a system that processes still images does not need to be causal. Non-causal systems can be built and can be useful in many circumstances. Even non-real systems can be built and are very useful in many contexts.

Causality

Main article: Causal system

A system is causal if the output depends only on present and past, but not future inputs. A necessary and sufficient condition for causality is

$$h(t) = 0 \quad \forall t < 0,$$

where $h(t)$ is the impulse response. It is not possible in general to determine causality from the Laplace transform, because the inverse transform is not unique. When a region of convergence is specified, then causality can be determined.

Stability

Main article: BIBO stability

A system is **bounded-input, bounded-output stable** (BIBO stable) if, for every bounded input, the output is finite. Mathematically, if every input satisfying

$$\|x(t)\|_{\infty} < \infty$$

leads to an output satisfying

$$\|y(t)\|_{\infty} < \infty$$

(that is, a finite maximum absolute value of $x(t)$ implies a finite maximum absolute value of $y(t)$), then the system is stable. A necessary and sufficient condition is that $h(t)$, the impulse response, is in L^1 (has a finite L^1 norm):

$$\|h(t)\|_1 = \int_{-\infty}^{\infty} |h(t)| \,\mathrm{d}t < \infty.$$

In the frequency domain, the region of convergence must contain the imaginary axis $s = j\omega$.

As an example, the ideal low-pass filter with impulse response equal to a sinc function is not BIBO stable, because the sinc function does not have a finite L^1 norm. Thus, for some bounded input, the output of the ideal low-pass filter is unbounded. In particular, if the input is zero for $t < 0$ and equal to a sinusoid at the cut-off frequency for $t > 0$, then the output will be unbounded for all times other than the zero crossings.

Discrete-time systems

Almost everything in continuous-time systems has a counterpart in discrete-time systems.

Discrete-time systems from continuous-time systems

In many contexts, a discrete time (DT) system is really part of a larger continuous time (CT) system. For example, a digital recording system takes an analog sound, digitizes it, possibly processes the digital signals, and plays back an analog sound for people to listen to.

Formally, the DT signals studied are almost always uniformly sampled versions of CT signals. If $x(t)$ is a CT signal, then an analog to digital converter will transform it to the DT signal:

$$x[n] \overset{\text{def}}{=} x(nT) \qquad \forall n \in \mathbb{Z},$$

where *T* is the sampling period. It is very important to limit the range of frequencies in the input signal for faithful representation in the DT signal, since then the sampling theorem guarantees that no information about the CT signal is lost. A DT signal can only contain a frequency range of $1/(2T)$; other frequencies are aliased to the same range.

Impulse response and convolution

Let $\{x[m-k];\ m\}$ represent the sequence $\{x[m-k];\ \text{for all integer values of m}\}$.
And let the shorter notation $\{x\}$ represent $\{x[m];\ m\}.$
A discrete system transforms an input sequence, $\{x\}$ into an output sequence, $\{y\}.$ In general, every element of the output can depend on every element of the input. Representing the transformation operator by O, we can write:

$$y[n] \overset{\text{def}}{=} O_n\{x\}.$$

Note that unless the transform itself changes with **n**, the output sequence is just constant, and the system is uninteresting. (Thus the subscript, **n**.) In a typical system, **y[n]** depends most heavily on the elements of **x** whose indices are near **n**.

For the special case of the Kronecker delta function, $x[m] = \delta[m],$ the output sequence is the **impulse response:**

$$h[n] \overset{\text{def}}{=} O_n\{\delta[m];\ m\}.$$

For a linear system, O must satisfy:

$$O_n\left\{\sum_{k=-\infty}^{\infty} c_k \cdot x_k[m];\ m\right\} = \sum_{k=-\infty}^{\infty} c_k \cdot O_n\{x_k\}.$$

(Eq.4)

And the time-invariance requirement is:

$$\begin{aligned} O_n\{x[m-k];\ m\} &= y[n-k] \\ &\overset{\text{def}}{=} O_{n-k}\{x\}. \end{aligned}$$

(Eq.5)

In such a system, the impulse response, $\{h\},$ characterizes the system completely. I.e., for any input sequence, the output sequence can be calculated in terms of the input and the impulse response. To see how that is done, consider the identity:

$$x[m] \equiv \sum_{k=-\infty}^{\infty} x[k] \cdot \delta[m-k],$$

which expresses $\{x\}$ in terms of a sum of weighted delta functions.
Therefore:

$$\begin{aligned} y[n] = O_n\{x\} &= O_n\left\{\sum_{k=-\infty}^{\infty} x[k] \cdot \delta[m-k];\ m\right\} \\ &= \sum_{k=-\infty}^{\infty} x[k] \cdot O_n\{\delta[m-k];\ m\}, \end{aligned}$$

where we have invoked **Eq.4** for the case $c_k = x[k]$ and $x_k[m] = \delta[m-k].$
And because of **Eq.5**, we may write:

$$\begin{aligned} O_n\{\delta[m-k];\ m\} &= O_{n-k}\{\delta[m];\ m\} \\ &\overset{\text{def}}{=} h[n-k]. \end{aligned}$$

Therefore:

$$y[n] = \sum_{k=-\infty}^{\infty} x[k] \cdot h[n-k]$$

$$= \sum_{k=-\infty}^{\infty} x[n-k] \cdot h[k], \qquad \text{(commutativity)}$$

which is the familiar discrete convolution formula. The operator O_n can therefore be interpreted as proportional to a weighted average of the function **x[k]**. The weighting function is **h[-k]**, simply shifted by amount **n**. As **n** changes, the weighting function emphasizes different parts of the input function. Equivalently, the system's response to an impulse at **n=0** is a "time" reversed copy of the unshifted weighting function. When **h[k]** is zero for all negative **k**, the system is said to be causal.

Exponentials as eigenfunctions

An eigenfunction is a function for which the output of the operator is the same function, just scaled by some amount. In symbols,

$$\mathcal{H}f = \lambda f,$$

where *f* is the eigenfunction and λ is the eigenvalue, a constant.

The exponential functions $z^n = e^{sTn}$, where $n \in \mathbb{Z}$, are eigenfunctions of a linear, time-invariant operator. $T \in \mathbb{R}$ is the sampling interval, and $z = e^{sT}, \ z, s \in \mathbb{C}$. A simple proof illustrates this concept.

Suppose the input is $x[n] = z^n$. The output of the system with impulse response $h[n]$ is then

$$\sum_{m=-\infty}^{\infty} h[n-m]\, z^m$$

which is equivalent to the following by the commutative property of convolution

$$\sum_{m=-\infty}^{\infty} h[m]\, z^{(n-m)} = z^n \sum_{m=-\infty}^{\infty} h[m]\, z^{-m} = z^n H(z)$$

where

$$H(z) \stackrel{\text{def}}{=} \sum_{m=-\infty}^{\infty} h[m] z^{-m}$$

is dependent only on the parameter *z*.

So z^n is an eigenfunction of an LTI system because the system response is the same as the input times the constant $H(z)$.

Z and discrete-time Fourier transforms

The eigenfunction property of exponentials is very useful for both analysis and insight into LTI systems. The Z transform

$$H(z) = \mathcal{Z}\{h[n]\} = \sum_{n=-\infty}^{\infty} h[n] z^{-n}$$

is exactly the way to get the eigenvalues from the impulse response. Of particular interest are pure sinusoids, i.e. exponentials of the form $e^{j\omega n}$, where $\omega \in \mathbb{R}$. These can also be written as z^n with $z = e^{j\omega}$. These are generally called complex exponentials even though the argument is purely imaginary. The Discrete-time Fourier transform (DTFT) $H(e^{j\omega}) = \mathcal{F}\{h[n]\}$ gives the eigenvalues of pure sinusoids. Both of $H(z)$ and $H(e^{j\omega})$ are called the *system function*, *system response*, or *transfer function*'.

The Z transform is usually used in the context of one-sided signals, i.e. signals that are zero for all values of t less than some value. Usually, this "start time" is set to zero, for convenience and without loss of generality. The Fourier transform is used for analyzing signals that are infinite in extent.

Due to the convolution property of both of these transforms, the convolution that gives the output of the system can be transformed to a multiplication in the transform domain. That is,

$$y[n]=(h*x)[n]=\sum_{m=-\infty}^{\infty}h[n-m]x[m]=\mathcal{Z}^{-1}\{H(z)X(z)\}.$$

Just as with the Laplace transform transfer function in continuous-time system analysis, the Z transform makes it easier to analyze systems and gain insight into their behavior. One can look at the modulus of the system function |*H(z)*| to see whether the input z^n is *passed* (let through) by the system, or *rejected* or *attenuated* by the system (not let through).

Examples

- A simple example of an LTI operator is the delay operator $D\{x[n]\} \stackrel{\text{def}}{=} x[n-1]$.
  - $D\left(c_1x_1[n]+c_2x_2[n]\right)=c_1x_1[n-1]+c_2x_2[n-1]=c_1Dx_1[n]+c_2Dx_2[n]$ (i.e., it is linear)
  - $D\{x[n-m]\}=x[n-m-1]=x[(n-1)-m]=D\{x\}[n-m]$ (i.e., it is time invariant)

  The Z transform of the delay operator is a simple multiplication by z^{-1}. That is,

  $$\mathcal{Z}\left\{Dx[n]\right\}=z^{-1}X(z).$$
- Another simple LTI operator is the averaging operator

  $$\mathcal{A}\left\{x[n]\right\} \stackrel{\text{def}}{=} \sum_{k=n-a}^{n+a}x[k].$$

  Because of the linearity of sums,

  $$\begin{aligned}\mathcal{A}\left\{c_1x_1[n]+c_2x_2[n]\right\}&=\sum_{k=n-a}^{n+a}\left(c_1x_1[k]+c_2x_2[k]\right)\\&=c_1\sum_{k=n-a}^{n+a}x_1[k]+c_2\sum_{k=n-a}^{n+a}x_2[k]\\&=c_1\mathcal{A}\left\{x_1[n]\right\}+c_2\mathcal{A}\left\{x_2[n]\right\},\end{aligned}$$

  and so it is linear. Because,

  $$\begin{aligned}\mathcal{A}\left\{x[n-m]\right\}&=\sum_{k=n-a}^{n+a}x[k-m]\\&=\sum_{k'=(n-m)-a}^{(n-m)+a}x[k']\\&=\mathcal{A}\left\{x\right\}[n-m],\end{aligned}$$

  it is also time invariant.

Important system properties

The input-output characteristics of discrete-time LTI system are completely described by its impulse response $h[n]$. Some of the most important properties of a system are causality and stability. Unlike CT systems, non-causal DT systems can be realized. It is trivial to make an acausal FIR system causal by adding delays. It is even possible to make acausal IIR systems.[4] Non-stable systems can be built and can be useful in many circumstances. Even non-real systems can be built and are very useful in many contexts.

Causality

Main article: Causal system

A discrete-time LTI system is causal if the current value of the output depends on only the current value and past values of the input.,[5] A necessary and sufficient condition for causality is

$$h[n] = 0 \ \forall n < 0,$$

where $h[n]$ is the impulse response. It is not possible in general to determine causality from the Z transform, because the inverse transform is not unique. When a region of convergence is specified, then causality can be determined.

Stability

Main article: BIBO stability

A system is **bounded input, bounded output stable** (BIBO stable) if, for every bounded input, the output is finite. Mathematically, if

$$\|x[n]\|_{\infty} < \infty$$

implies that

$$\|y[n]\|_{\infty} < \infty$$

(that is, if bounded input implies bounded output, in the sense that the maximum absolute values of $x[n]$ and $y[n]$ are finite), then the system is stable. A necessary and sufficient condition is that $h[n]$, the impulse response, satisfies

$$\|h[n]\|_1 \overset{\text{def}}{=} \sum_{n=-\infty}^{\infty} |h[n]| < \infty.$$

In the frequency domain, the region of convergence must contain the unit circle (i.e., the locus satisfying $|z| = 1$ for complex *z*).

Notes

[1] Hespanha 2009, p. 78.

[2] Crutchfield[web]

[3] Crutchfield

[4] Vaidyanathan,1995

[5] Phillips 2007, p. 508.

References

- Phillips, C.l., Parr, J.M., & Riskin, E.A (2007). *Signals, systems and Transforms*. Prentice Hall. ISBN 0-13-041207-4.
- Hespanha,J.P. (2009). *Linear System Theory*. Princeton university press. ISBN 0-691-14021-9.
- Crutchfield, Steve (October 12, 2010), "The Joy of Convolution" (http://www.jhu.edu/signals/convolve/index.html), *Johns Hopkins University*, retrieved November 21, 2010
- Vaidyanathan, P. P.; Chen, T. (May 1995). "Role of anticausal inverses in multirate filter banks — Part I: system theoretic fundamentals". *IEEE Trans. Signal Proc.* **43** (6): 1090. Bibcode: 1995ITSP...43.1090V (http://adsabs.harvard.edu/abs/1995ITSP...43.1090V). doi: 10.1109/78.382395 (http://dx.doi.org/10.1109/78.382395).

Further reading

- Porat, Boaz (1997). *A Course in Digital Signal Processing*. New York: John Wiley. ISBN 978-0-471-14961-3.
- Vaidyanathan, P. P.; Chen, T. (May 1995). "Role of anticausal inverses in multirate filter banks — Part I: system theoretic fundamentals". *IEEE Trans. Signal Proc.* **43** (5): 1090. Bibcode: 1995ITSP...43.1090V (http://adsabs.harvard.edu/abs/1995ITSP...43.1090V). doi: 10.1109/78.382395 (http://dx.doi.org/10.1109/78.382395).

External links

- ECE 209: Review of Circuits as LTI Systems (http://www.tedpavlic.com/teaching/osu/ece209/support/circuits_sys_review.pdf) – Short primer on the mathematical analysis of (electrical) LTI systems.
- ECE 209: Sources of Phase Shift (http://www.tedpavlic.com/teaching/osu/ece209/lab3_opamp_FO/lab3_opamp_FO_phase_shift.pdf) – Gives an intuitive explanation of the source of phase shift in two common electrical LTI systems.
- JHU 520.214 Signals and Systems course notes (http://www.ece.jhu.edu/~cooper/courses/214/signalsandsystemsnotes.pdf). An encapsulated course on LTI system theory. Adequate for self teaching.

Transfer function

In engineering, a **transfer function** (also known as the **system function**[1] or **network function** and, when plotted as a graph, **transfer curve**) is a mathematical representation, in terms of spatial or temporal frequency, of the relation between the input and output of a linear time-invariant system with zero initial conditions and zero-point equilibrium.[2] With optical imaging devices, for example, it is the Fourier transform of the point spread function (hence a function of spatial frequency) i.e. the intensity distribution caused by a point object in the field of view.

Explanation

Transfer functions are commonly used in the analysis of systems such as single-input single-output filters, typically within the fields of signal processing, communication theory, and control theory. The term is often used exclusively to refer to linear, time-invariant systems (LTI), as covered in this article. Most real systems have non-linear input/output characteristics, but many systems, when operated within nominal parameters (not "over-driven") have behavior that is close enough to linear that LTI system theory is an acceptable representation of the input/output behavior.

The descriptions below are given in terms of a complex variable, s = σ + j*ω, which bears a brief explanation. In many applications, it is sufficient to define σ=0 (and s = j*ω), which reduces the Laplace transforms with complex arguments to Fourier transforms with real argument ω. The applications where this is common are ones where there is interest only in the steady-state response of an LTI system, not the fleeting turn-on and turn-off behaviors or stability issues. That is usually the case for signal processing and communication theory.

Thus, for continuous-time input signal $x(t)$ and output $y(t)$, the transfer function $H(s)$ is the linear mapping of the Laplace transform of the input, $X(s) = \mathcal{L}\left\{x(t)\right\}$, to the Laplace transform of the output $Y(s) = \mathcal{L}\left\{y(t)\right\}$:

$$Y(s) = H(s)\;X(s)$$

or

$$H(s) = \frac{Y(s)}{X(s)} = \frac{ \mathcal{L}\left\{y(t)\right\} }{ \mathcal{L}\left\{x(t)\right\} }.$$

In discrete-time systems, the relation between an input signal $x(t)$and output $y(t)$is dealt with using the z-transform, and then the transfer function is similarly written as $H(z) = \frac{Y(z)}{X(z)}$and this is often referred to as the pulse-transfer function.Wikipedia:Citation needed

Direct derivation from differential equations

Consider a linear differential equation with constant coefficients

$$L[u] = \frac{d^n u}{dt^n} + a_1 \frac{d^{n-1}u}{dt^{n-1}} + \dots + a_{n-1}\frac{du}{dt} + a_n u = r(t)$$

where *u* and *r* are suitably smooth functions of *t*, and *L* is the operator defined on the relevant function space, that transforms *u* into *r*. That kind of equation can be used to constrain the output function *u* in terms of the *forcing* function *r*. The transfer function, written as an operator $F[r] = u$, is the right inverse of *L*, since $L[F[r]] = r$.

Solutions of the *homogeneous*, constant-coefficient differential equation $L[u] = 0$can be found by trying $u = e^{\lambda t}$. That substitution yields the *characteristic polynomial*

$$p_L(\lambda) = \lambda^n + a_1\lambda^{n-1} + \dots + a_{n-1}\lambda + a_n$$

The inhomogeneous case can be easily solved if the input function *r* is also of the form $r(t) = e^{st}$. In that case, by substituting $u = H(s)e^{st}$ one finds that $L[H(s)e^{st}] = e^{st}$ if and only if

$$H(s) = \frac{1}{p_L(s)}, \qquad p_L(s) \neq 0.$$

Taking that as the definition of the *transfer function* requires careful disambiguation between complex vs. real values, which is traditionally influenced by the interpretation of *abs(H(s))* as the gain and *-atan(H(s))* as the phase lag. Other definitions of the transfer function are used: for example $1/p_L(ik)$.

Signal processing

Let $x(t)$be the input to a general linear time-invariant system, and $y(t)$be the output, and the bilateral Laplace transform of $x(t)$and $y(t)$be

$$X(s) = \mathcal{L}\left\{x(t)\right\} \stackrel{\mathrm{def}}{=} \int_{-\infty}^{\infty} x(t)e^{-st}\, dt$$
$$Y(s) = \mathcal{L}\left\{y(t)\right\} \stackrel{\mathrm{def}}{=} \int_{-\infty}^{\infty} y(t)e^{-st}\, dt .$$

Then the output is related to the input by the transfer function $H(s)$ as

$$Y(s) = H(s)X(s)$$

and the transfer function itself is therefore

$$H(s) = \frac{Y(s)}{X(s)}.$$

In particular, if a complex harmonic signal with a sinusoidal component with amplitude $|X|$, angular frequency ω and phase $\arg(X)$

$$x(t) = Xe^{j\omega t} = |X|e^{j(\omega t + \arg(X))}$$
$$\text{where } X = |X|e^{j\arg(X)}$$

is input to a linear time-invariant system, then the corresponding component in the output is:

$$y(t) = Ye^{j\omega t} = |Y|e^{j(\omega t + \arg(Y))}$$
$$Y = |Y|e^{j\arg(Y)}.$$

Note that, in a linear time-invariant system, the input frequency ω has not changed, only the amplitude and the phase angle of the sinusoid has been changed by the system. The frequency response $H(j\omega)$ describes this change for every frequency ω in terms of *gain*:

$$G(\omega) = \frac{|Y|}{|X|} = |H(j\omega)|$$

and *phase shift*:

$$\phi(\omega) = \arg(Y) - \arg(X) = \arg(H(j\omega)).$$

The phase delay (i.e., the frequency-dependent amount of delay introduced to the sinusoid by the transfer function) is:

$$\tau_{\phi}(\omega) = -\frac{\phi(\omega)}{\omega}.$$

The group delay (i.e., the frequency-dependent amount of delay introduced to the envelope of the sinusoid by the transfer function) is found by computing the derivative of the phase shift with respect to angular frequency ω ,

$$\tau_{g}(\omega) = -\frac{d\phi(\omega)}{d\omega}.$$

The transfer function can also be shown using the Fourier transform which is only a special case of the bilateral Laplace transform for the case where $s = j\omega$.

Common transfer function families

While any LTI system can be described by some transfer function or another, there are certain "families" of special transfer functions that are commonly used. Typical infinite impulse response filters are designed to implement one of these special transfer functions.

Some common transfer function families and their particular characteristics are:

- Butterworth filter – maximally flat in passband and stopband for the given order
- Chebyshev filter (Type I) – maximally flat in stopband, sharper cutoff than Butterworth of same order
- Chebyshev filter (Type II) – maximally flat in passband, sharper cutoff than Butterworth of same order
- Bessel filter – best pulse response for a given order because it has no group delay ripple
- Elliptic filter – sharpest cutoff (narrowest transition between pass band and stop band) for the given order
- Optimum "L" filter
- Gaussian filter – minimum group delay; gives no overshoot to a step function.
- Hourglass filter
- Raised-cosine filter

Control engineering

In control engineering and control theory the transfer function is derived using the Laplace transform.

The transfer function was the primary tool used in classical control engineering. However, it has proven to be unwieldy for the analysis of multiple-input multiple-output (MIMO) systems, and has been largely supplanted by state space representations for such systems. In spite of this, a transfer matrix can be always obtained for any linear system, in order to analyze its dynamics and other properties: each element of a transfer matrix is a transfer function relating a particular input variable to an output variable.

Optics

In optics, modulation transfer function indicates the capability of optical contrast transmission.

For example, when observing a series of black-white-light fringes drawn with a specific spatial frequency, the image quality may decay. White fringes fade while black ones turn brighter.

The modulation transfer function in a specific spatial frequency is defined by:

$$\mathrm{MTF}(f) = \frac{M(\mathrm{image})}{M(\mathrm{source})}$$

Where modulation (M) is computed from the following image or light brightness:

$$M = \frac{(L_{\mathrm{max}} - L_{\mathrm{min}})}{(L_{\mathrm{max}} + L_{\mathrm{min}})}$$

References

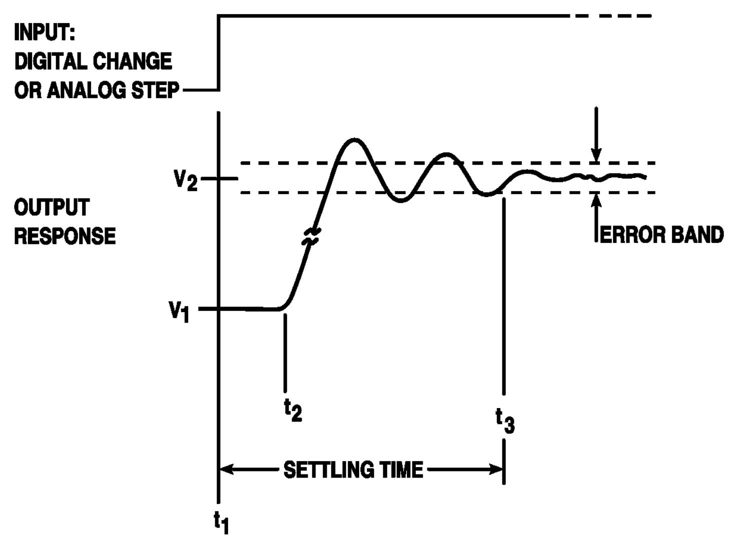
[1] Bernd Girod, Rudolf Rabenstein, Alexander Stenger, *Signals and systems*, 2nd ed., Wiley, 2001, ISBN 0-471-98800-6 p. 50

[2] The Oxford Dictionary of English, 3rd ed., "Transfer function"

External links

- Transfer function (http://planetmath.org/?op=getobj&amp;from=objects&amp;id=5394) at PlanetMath.org.
- ECE 209: Review of Circuits as LTI Systems (http://www.tedpavlic.com/teaching/osu/ece209/support/circuits_sys_review.pdf) — Short primer on the mathematical analysis of (electrical) LTI systems.
- ECE 209: Sources of Phase Shift (http://www.tedpavlic.com/teaching/osu/ece209/lab3_opamp_FO/lab3_opamp_FO_phase_shift.pdf) — Gives an intuitive explanation of the source of phase shift in two simple LTI systems. Also verifies simple transfer functions by using trigonometric identities.
- Transfer function model in Mathematica (http://reference.wolfram.com/mathematica/ref/TransferFunctionModel.html)

Step response



A typical step response for a second order system, illustrating overshoot, followed by ringing, all subsiding within a settling time.

The **step response** of a system in a given initial state consists of the time evolution of its outputs when its control inputs are Heaviside step functions. In electronic engineering and control theory, step response is the time behaviour of the outputs of a general system when its inputs change from zero to one in a very short time. The concept can be extended to the abstract mathematical notion of a dynamical system using an evolution parameter.

From a practical standpoint, knowing how the system responds to a sudden input is important because large and possibly fast deviations from the long term steady state may have extreme effects on the component itself and on other portions of the overall system dependent on this component. In addition, the overall system cannot act until the component's output settles down to some vicinity of its final state, delaying the overall system response. Formally, knowing the step response of a dynamical system gives information on the stability of such a system, and on its ability to reach one stationary state when starting from another.

Time domain *versus* frequency domain

Instead of frequency response, system performance may be specified in terms of parameters describing time-dependence of response. The step response can be described by the following quantities related to its **time behavior**,

- overshoot
- rise time
- settling time
- ringing

In the case of linear dynamic systems, much can be inferred about the system from these characteristics. Below the step response of a simple two-pole amplifier is presented, and some of these terms are illustrated.

Step response of feedback amplifiers

This section describes the step response of a simple negative feedback amplifier shown in Figure 1. The feedback amplifier consists of a main **open-loop** amplifier of gain A_{OL} and a feedback loop governed by a **feedback factor** β. This feedback amplifier is analyzed to determine how its step response depends upon the time constants governing the response of the main amplifier, and upon the amount of feedback used.

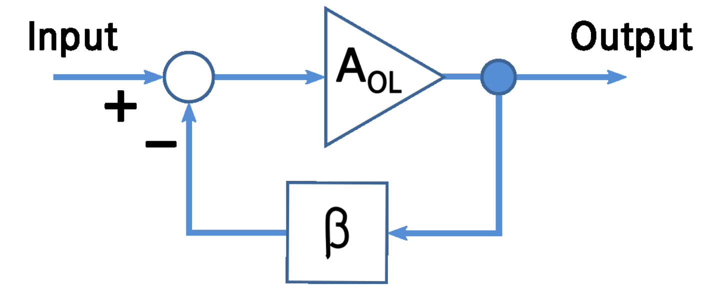


Figure 1: Ideal negative feedback model; open loop gain is A_{OL} and feedback factor is β.

Analysis

A negative feedback amplifier has gain given by (see negative feedback amplifier):

$$A_{FB} = \frac{A_{OL}}{1+\beta A_{OL}},$$

where A_{OL} = **open-loop** gain, A_{FB} = **closed-loop** gain (the gain with negative feedback present) and β = **feedback factor**. The step response of such an amplifier is easily handled in the case that the open-loop gain has two poles (two time constants, τ_1, τ_2), that is, the open-loop gain is given by:

$$A_{OL} = \frac{A_0}{(1+j\omega\tau_1)(1+j\omega\tau_2)},$$

with zero-frequency gain A_0 and angular frequency ω = 2π*f*, which leads to the closed-loop gain:

$$A_{FB} = \frac{A_0}{1+\beta A_0} \bullet \frac{1}{1+j\omega\frac{\tau_1+\tau_2}{1+\beta A_0}+(j\omega)^2\frac{\tau_1\tau_2}{1+\beta A_0}}.$$

The time dependence of the amplifier is easy to discover by switching variables to $s = j\omega$, whereupon the gain becomes:

$$A_{FB} = \frac{A_0}{\tau_1\tau_2} \bullet \frac{1}{s^2+s\left(\frac{1}{\tau_1}+\frac{1}{\tau_2}\right)+\frac{1+\beta A_0}{\tau_1\tau_2}}$$

The poles of this expression (that is, the zeros of the denominator) occur at:

$$2s = -\left(\frac{1}{\tau_1}+\frac{1}{\tau_2}\right) \pm\sqrt{\left(\frac{1}{\tau_1}-\frac{1}{\tau_2}\right)^2-\frac{4\beta A_0}{\tau_1\tau_2}},$$

which shows for large enough values of βA_0 the square root becomes the square root of a negative number, that is the square root becomes imaginary, and the pole positions are complex conjugate numbers, either s_+ or s_-; see Figure 2:

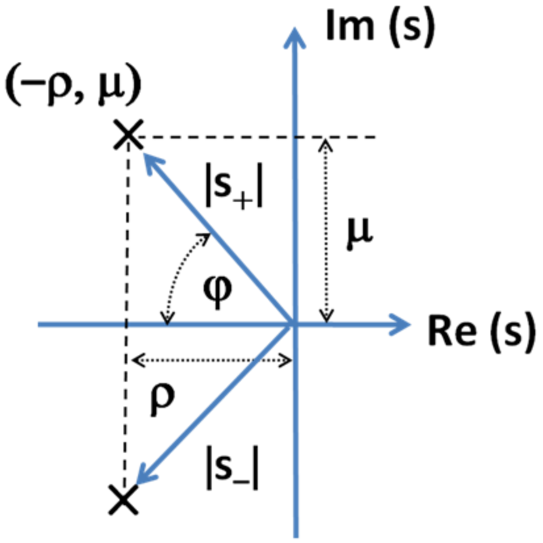


Figure 2: Conjugate pole locations for a two-pole feedback amplifier; *Re*(s) = real axis and *Im*(s) = imaginary axis.

$$s_{\pm} = -\rho \pm j\mu,$$

with

$$\rho = \frac{1}{2}\left(\frac{1}{\tau_1} + \frac{1}{\tau_2}\right),$$

and

$$\mu = \frac{1}{2}\sqrt{\frac{4\beta A_0}{\tau_1 \tau_2} - \left(\frac{1}{\tau_1} - \frac{1}{\tau_2}\right)^2}.$$

Using polar coordinates with the magnitude of the radius to the roots given by |s| (Figure 2):

$$|s| = |s_{\pm}| = \sqrt{\rho^2 + \mu^2},$$

and the angular coordinate φ is given by:

$$\cos\phi = \frac{\rho}{|s|}, \sin\phi = \frac{\mu}{|s|}.$$

Tables of Laplace transforms show that the time response of such a system is composed of combinations of the two functions:

$$e^{-\rho t}\sin(\mu t) \quad \text{and } e^{-\rho t}\cos(\mu t),$$

which is to say, the solutions are damped oscillations in time. In particular, the unit step response of the system is:

$$S(t) = \left(\frac{A_{OL}}{1+\beta A_{OL}}\right)\left(1 - e^{-\rho t}\,\frac{\sin\left(\mu t + \phi\right)}{\sin(\phi)}\right) ,$$

which simplifies to

$$S(t) = 1 - e^{-\rho t}\,\frac{\sin\left(\mu t + \phi\right)}{\sin(\phi)}$$

when A_{OL} tends to infinity and the feedback factor is one.

Notice that the damping of the response is set by ρ, that is, by the time constants of the open-loop amplifier. In contrast, the frequency of oscillation is set by μ, that is, by the feedback parameter through βA_0. Because ρ is a sum of reciprocals of time constants, it is interesting to notice that ρ is dominated by the *shorter* of the two.

Results

Figure 3 shows the time response to a unit step input for three values of the parameter μ. It can be seen that the frequency of oscillation increases with μ, but the oscillations are contained between the two asymptotes set by the exponentials [1 – exp (–ρt)] and [1 + exp(–ρt)]. These asymptotes are determined by ρ and therefore by the time constants of the open-loop amplifier, independent of feedback.

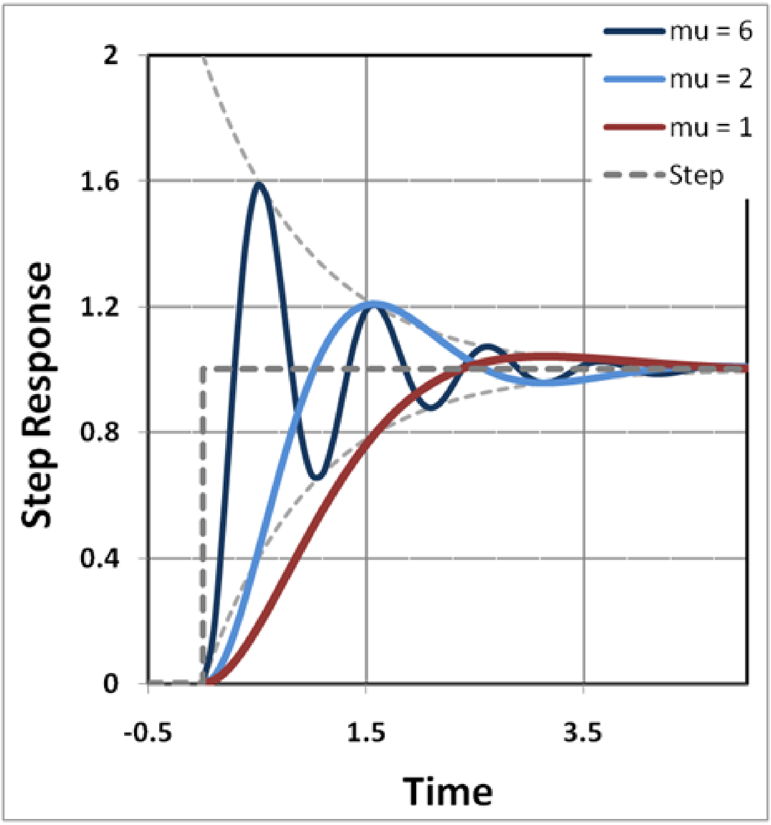


Figure 3: Step-response of a linear two-pole feedback amplifier; time is in units of 1/ρ, that is, in terms of the time constants of A_{OL}; curves are plotted for three values of *mu* = *μ*, which is controlled by β.

The phenomena of oscillation about final value is called **ringing**. The **overshoot** is the maximum swing above final value, and clearly increases with μ. Likewise, the **undershoot** is the minimum swing below final value, again increasing with μ. The **settling time** is the time for departures from final value to sink below some specified level, say 10% of final value.

The dependence of settling time upon μ is not obvious, and the approximation of a two-pole system probably is not accurate enough to make any real-world conclusions about feedback dependence of settling time. However, the asymptotes [1 – exp (–ρt)] and [1 + exp (–ρt)] clearly impact settling time, and they are controlled by the time constants of the open-loop amplifier, particularly the shorter of the two time constants. That suggests that a specification on settling time must be met by appropriate design of the open-loop amplifier.

The two major conclusions from this analysis are:

1. Feedback controls the amplitude of oscillation about final value for a given open-loop amplifier and given values of open-loop time constants, τ_1 and τ_2.
2. The open-loop amplifier decides settling time. It sets the time scale of Figure 3, and the faster the open-loop amplifier, the faster this time scale.

As an aside, it may be noted that real-world departures from this linear two-pole model occur due to two major complications: first, real amplifiers have more than two poles, as well as zeros; and second, real amplifiers are nonlinear, so their step response changes with signal amplitude.

Control of overshoot

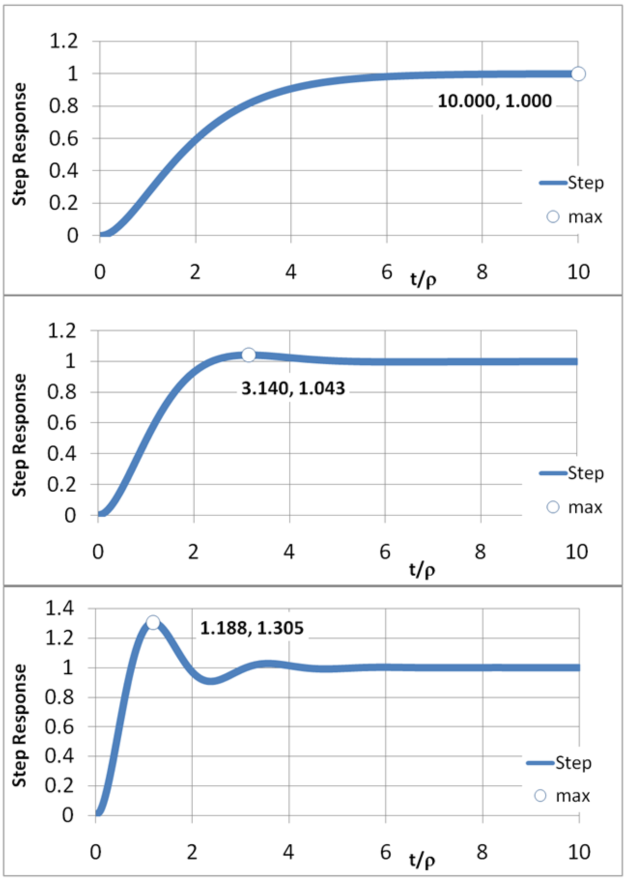


Figure 4: Step response for three values of α. Top: α = 4; Center: α = 2; Bottom: α = 0.5. As α is reduced the pole separation reduces, and the overshoot increases.

How overshoot may be controlled by appropriate parameter choices is discussed next.

Using the equations above, the amount of overshoot can be found by differentiating the step response and finding its maximum value. The result for maximum step response $S_{\max}$ is:

$$S_{\max} = 1 + \exp\left(-\pi \frac{\rho}{\mu}\right).$$

The final value of the step response is 1, so the exponential is the actual overshoot itself. It is clear the overshoot is zero if μ = 0, which is the condition:

$$\frac{4\beta A_0}{\tau_1 \tau_2} = \left(\frac{1}{\tau_1} - \frac{1}{\tau_2}\right)^2.$$

This quadratic is solved for the ratio of time constants by setting $x = (\tau_1 / \tau_2)^{1/2}$ with the result

$$x = \sqrt{\beta A_0} + \sqrt{\beta A_0 + 1}.$$

Because $\beta\ A_0 >> 1$, the 1 in the square root can be dropped, and the result is

$$\frac{\tau_1}{\tau_2} = 4\beta A_0.$$

In words, the first time constant must be much larger than the second. To be more adventurous than a design allowing for no overshoot we can introduce a factor α in the above relation:

$$\frac{\tau_1}{\tau_2} = \alpha \beta A_0,$$

and let α be set by the amount of overshoot that is acceptable.

Figure 4 illustrates the procedure. Comparing the top panel (α = 4) with the lower panel (α = 0.5) shows lower values for α increase the rate of response, but increase overshoot. The case α = 2 (center panel) is the *maximally flat* design that shows no peaking in the Bode gain vs. frequency plot. That design has the rule of thumb built-in safety margin to deal with non-ideal realities like multiple poles (or zeros), nonlinearity (signal amplitude dependence) and manufacturing variations, any of which can lead to too much overshoot. The adjustment of the pole separation (that is, setting α) is the subject of frequency compensation, and one such method is pole splitting.

Control of settling time

The amplitude of ringing in the step response in Figure 3 is governed by the damping factor exp (−ρ t). That is, if we specify some acceptable step response deviation from final value, say Δ, that is:

$$S(t) \leq 1 + \Delta,$$

this condition is satisfied regardless of the value of $\beta\ A_{OL}$ provided the time is longer than the settling time, say t_S, given by:[1]

$$\Delta = e^{-\rho t_S} \text{ or } t_S = \frac{\ln\left(\frac{1}{\Delta}\right)}{\rho} = \tau_2 \frac{2\ln\left(\frac{1}{\Delta}\right)}{1 + \frac{\tau_2}{\tau_1}} \approx 2\tau_2 \ln\left(\frac{1}{\Delta}\right),$$

where the approximation $\tau_1 >> \tau_2$ is applicable because of the overshoot control condition, which makes $\tau_1 = \alpha\beta A_{OL}$ τ_2. Often the settling time condition is referred to by saying the settling period is inversely proportional to the unity gain bandwidth, because $1/(2\pi\,\tau_2)$ is close to this bandwidth for an amplifier with typical dominant pole compensation. However, this result is more precise than this rule of thumb. As an example of this formula, if $\Delta = 1/e^4 = 1.8$ %, the settling time condition is $t_S = 8\,\tau_2$.

In general, control of overshoot sets the time constant ratio, and settling time t_S sets τ_2.[2]

Phase margin

Next, the choice of pole ratio τ_1/τ_2 is related to the phase margin of the feedback amplifier.[3] The procedure outlined in the Bode plot article is followed. Figure 5 is the Bode gain plot for the two-pole amplifier in the range of frequencies up to the second pole position. The assumption behind Figure 5 is that the frequency $f_{0\text{ dB}}$ lies between the lowest pole at $f_1 = 1/(2\pi\tau_1)$ and the second pole at $f_2 = 1/(2\pi\tau_2)$. As indicated in Figure 5, this condition is satisfied for values of $\alpha \geq 1$.

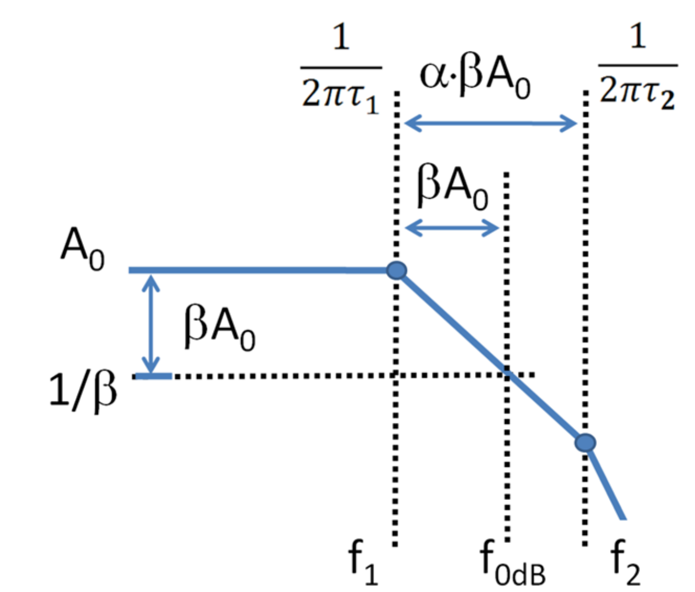


Figure 5: Bode gain plot to find phase margin; scales are logarithmic, so labeled separations are multiplicative factors. For example, $f_{0\text{ dB}} = \beta A_0 \times f_1$.

Using Figure 5 the frequency (denoted by $f_{0\text{ dB}}$) is found where the loop gain βA_0 satisfies the unity gain or 0 dB condition, as defined by:

$$|\beta A_{\mathrm{OL}}(f_{0\text{ db}})| = 1.$$

The slope of the downward leg of the gain plot is (20 dB/decade); for every factor of ten increase in frequency, the gain drops by the same factor:

$$f_{0\text{ dB}} = \beta A_0 f_1.$$

The phase margin is the departure of the phase at $f_{0\text{ dB}}$ from −180°. Thus, the margin is:

$$\phi_m = 180° - \arctan(f_{0\text{ dB}}/f_1) - \arctan(f_{0\text{ dB}}/f_2).$$

Because $f_{0\text{ dB}}/f_1 = \beta A_0 >> 1$, the term in f_1 is 90°. That makes the phase margin:

$$\begin{aligned}\phi_m &= 90° - \arctan(f_{0\text{ dB}}/f_2)\\ &= 90° - \arctan\left(\frac{\beta A_0 f_1}{\alpha\beta A_0 f_1}\right)\\ &= 90° - \arctan\left(\frac{1}{\alpha}\right) = \arctan\left(\alpha\right).\end{aligned}$$

In particular, for case $\alpha = 1$, $\varphi_m = 45°$, and for $\alpha = 2$, $\varphi_m = 63.4°$. Sansen recommends $\alpha = 3$, $\varphi_m = 71.6°$ as a "good safety position to start with".

If α is increased by shortening τ_2, the settling time t_S also is shortened. If α is increased by lengthening τ_1, the settling time t_S is little altered. More commonly, both τ_1 *and* τ_2 change, for example if the technique of pole splitting is used.

As an aside, for an amplifier with more than two poles, the diagram of Figure 5 still may be made to fit the Bode plots by making f_2 a fitting parameter, referred to as an "equivalent second pole" position.

Formal mathematical description

This section provides a formal mathematical definition of step response in terms of the abstract mathematical concept of a dynamical system $\mathfrak{S}$: all notations and assumptions required for the following description are listed here.

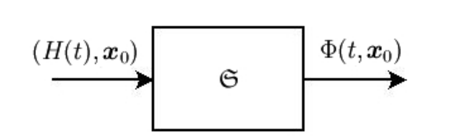


Figure 4: Black box representation of a dynamical system, its input and its step response.

- $t \in T$ is the evolution parameter of the system, called "time" for the sake of simplicity,
- $\boldsymbol{x}|_t \in M$ is the state of the system at time t , called "output" for the sake of simplicity,
- $\Phi : T \times M \longrightarrow M$ is the dynamical system evolution function,
- $\Phi(0, \boldsymbol{x}) = \boldsymbol{x}_0 \in M$ is the dynamical system initial state,
- $H(t)$ is the Heaviside step function

Nonlinear dynamical system

For a general dynamical system, the step response is defined as follows:

$$\boldsymbol{x}|_t = \Phi_{\{H(t)\}(t, \boldsymbol{x}_0)}.$$

It is the evolution function when the control inputs (or source term, or forcing inputs) are Heaviside functions: the notation emphasizes this concept showing *H*(*t*) as a subscript.

Linear dynamical system

For a linear time-invariant black box, let $\mathfrak{S} \equiv S$ for notational convenience: the step response can be obtained by convolution of the Heaviside step function control and the impulse response *h*(*t*) of the system itself

$$a(t) = h * H(t) = H * h(t) = \int_{-\infty}^{+\infty} h(\tau) H(t - \tau)\, d\tau = \int_{-\infty}^{t} h(\tau)\, d\tau.$$

References and notes

[1] This estimate is a bit conservative (long) because the factor 1 /sin(φ) in the overshoot contribution to *S* (*t*) has been replaced by 1 /sin(φ) ≈ 1.

[2] According to Johns and Martin, *op. cit.*, settling time is significant in switched-capacitor circuits, for example, where an op amp settling time must be less than half a clock period for sufficiently rapid charge transfer.

[3] The gain margin of the amplifier cannot be found using a two-pole model, because gain margin requires determination of the frequency f_{180} where the gain flips sign, and this never happens in a two-pole system. If we know f_{180} for the amplifier at hand, the gain margin can be found approximately, but f_{180} then depends on the third and higher pole positions, as does the gain margin, unlike the estimate of phase margin, which is a two-pole estimate.

Further reading

- Robert I. Demrow *Settling time of operational amplifiers* (http://www.analog.com/static/imported-files/application_notes/466359863287538299597392756AN359.pdf)
- Cezmi Kayabasi *Settling time measurement techniques achieving high precision at high speeds* (http://www.wpi.edu/Pubs/ETD/Available/etd-050505-140358/unrestricted/ckayabasi.pdf)
- Vladimir Igorevic Arnol'd "Ordinary differential equations", various editions from MIT Press and from Springer Verlag, chapter 1 "Fundamental concepts"

External links

- Kuo power point slides; Chapter 7 especially (http://bcs.wiley.com/he-bcs/Books?action=resource&bcsId=2357&itemId=0471134767&resourceId=5596)

BIBO stability

In signal processing, specifically control theory, **BIBO stability** is a form of stability for linear signals and systems that take inputs. BIBO stands for *Bounded-Input Bounded-Output*. If a system is BIBO stable, then the output will be bounded for every input to the system that is bounded.

A signal is bounded if there is a finite value $B > 0$ such that the signal magnitude never exceeds B, that is

$|y[n]| \leq B \quad \forall n \in \mathbb{Z}$ for discrete-time signals, or

$|y(t)| \leq B \quad \forall t \in \mathbb{R}$ for continuous-time signals.

Time-domain condition for linear time invariant systems

Continuous-time necessary and sufficient condition

For a continuous time linear time invariant (LTI) system, the condition for BIBO stability is that the impulse response be absolutely integrable, i.e., its L^1 norm exists.

$$\int_{-\infty}^{\infty} |h(t)| \; \mathrm{d}t = \|h\|_1 < \infty$$

Discrete-time sufficient condition

For a discrete time LTI system, the condition for BIBO stability is that the impulse response be absolutely summable, i.e., its ℓ^1 norm exists.

$$\sum_{n=-\infty}^{\infty} |h[n]| = \|h\|_1 < \infty$$

Proof of sufficiency

Given a discrete time LTI system with impulse response $h[n]$ the relationship between the input $x[n]$ and the output $y[n]$ is

$$y[n] = h[n] * x[n]$$

where $*$ denotes convolution. Then it follows by the definition of convolution

$$y[n] = \sum_{k=-\infty}^{\infty} h[k]x[n-k]$$

Let $\|x\|_{\infty}$ be the maximum value of $|x[n]|$, i.e., the L_{∞}-norm.

$$|y[n]| = \left| \sum_{k=-\infty}^{\infty} h[n-k]x[k] \right|$$

$$\leq \sum_{k=-\infty}^{\infty} |h[n-k]| \, |x[k]| \text{ (by the triangle inequality)}$$

$$\leq \sum_{k=-\infty}^{\infty} |h[n-k]| \, \|x\|_{\infty}$$

$$= \|x\|_\infty \sum_{k=-\infty}^{\infty} |h[n-k]|$$

$$= \|x\|_\infty \sum_{k=-\infty}^{\infty} |h[k]|$$

If $h[n]$ is absolutely summable, then $\sum_{k=-\infty}^{\infty} |h[k]| = \|h\|_1 < \infty$ and

$$\|x\|_\infty \sum_{k=-\infty}^{\infty} |h[k]| = \|x\|_\infty \|h\|_1$$

So if $h[n]$ is absolutely summable and $|x[n]|$ is bounded, then $|y[n]|$ is bounded as well because $\|x\|_\infty \|h\|_1 < \infty$.

The proof for continuous-time follows the same arguments.

Frequency-domain condition for linear time invariant systems

Continuous-time signals

For a rational and continuous-time system, the condition for stability is that the region of convergence (ROC) of the Laplace transform includes the imaginary axis. When the system is causal, the ROC is the open region to the right of a vertical line whose abscissa is the real part of the "largest pole", or the pole that has the greatest real part of any pole in the system. The real part of the largest pole defining the ROC is called the abscissa of convergence. Therefore, all poles of the system must be in the strict left half of the s-plane for BIBO stability.

This stability condition can be derived from the above time-domain condition as follows :

$$\int_{-\infty}^{\infty} |h(t)| \, \mathrm{d}t$$

$$= \int_{-\infty}^{\infty} |h(t)| \left|e^{-j\omega t}\right| dt$$

$$= \int_{-\infty}^{\infty} \left|h(t)(1 \cdot e)^{-j\omega t}\right| dt$$

$$= \int_{-\infty}^{\infty} \left|h(t)(e^{\sigma+j\omega})^{-t}\right| dt$$

$$= \int_{-\infty}^{\infty} \left|h(t)e^{-st}\right| dt$$

where $s = \sigma + j\omega$ and $\mathrm{Re}(s) = \sigma = 0$.

The region of convergence must therefore include the imaginary axis.

Discrete-time signals

For a rational and discrete time system, the condition for stability is that the region of convergence (ROC) of the z-transform includes the unit circle. When the system is causal, the ROC is the open region outside a circle whose radius is the magnitude of the pole with largest magnitude. Therefore, all poles of the system must be inside the unit circle in the z-plane for BIBO(bounded input-bounded output) stability.

This stability condition can be derived in a similar fashion to the continuous-time derivation:

$$\sum_{n=-\infty}^{\infty} |h[n]| = \sum_{n=-\infty}^{\infty} |h[n]| \left|e^{-j\omega n}\right|$$

$$= \sum_{n=-\infty}^{\infty} \left|h[n](1 \cdot e)^{-j\omega n}\right|$$

$$= \sum_{n=-\infty}^{\infty} \left| h[n](re^{j\omega})^{-n} \right|$$
$$= \sum_{n=-\infty}^{\infty} \left| h[n]z^{-n} \right|$$

where $z = re^{j\omega}$ and $r = |z| = 1$.

The region of convergence must therefore include the unit circle.

Further reading

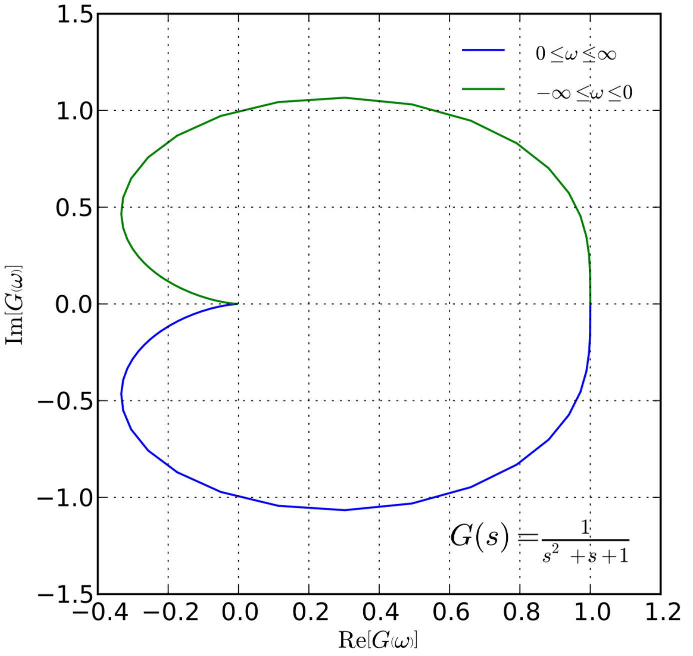
- Gordon E. Carlson *Signal and Linear Systems Analysis with Matlab* second edition, Wiley, 1998, ISBN 0-471-12465-6
- John G. Proakis and Dimitris G. Manolakis *Digital Signal Processing Principals, Algorithms and Applications* third edition, Prentice Hall, 1996, ISBN 0-13-373762-4
- D. Ronald Fannin, William H. Tranter, and Rodger E. Ziemer *Signals & Systems Continuous and Discrete* fourth edition, Prentice Hall, 1998, ISBN 0-13-496456-X
- Proof of the necessary conditions for BIBO stability. [1]
- Christophe Basso *Designing Control Loops for Linear and Switching Power Supplies: A Tutorial Guide* first edition, Artech House, 2012, 978-1608075577

References

[1] http://cnx.org/content/m12319/latest/

Nyquist stability criterion

In control theory and stability theory, the **Nyquist stability criterion**, discovered by Swedish-American electrical engineer Harry Nyquist at Bell Telephone Laboratories in 1932,[1] is a graphical technique for determining the stability of a dynamical system. Because it only looks at the Nyquist plot of the open loop systems, it can be applied without explicitly computing the poles and zeros of either the closed-loop or open-loop system (although the number of each type of right-half-plane singularities must be known). As a result, it can be applied to systems defined by non-rational functions, such as systems with delays. In contrast to Bode plots, it can handle transfer functions with right half-plane singularities. In addition, there is a natural generalization to more complex systems with multiple inputs and multiple outputs, such as control systems for airplanes.



The Nyquist plot for $G(s) = \frac{1}{s^2 + s + 1}$.

The Nyquist criterion is widely used in electronics and control system engineering, as well as other fields, for designing and analyzing systems with feedback. While Nyquist is one of the most general stability tests, it is still restricted to linear, time-invariant (LTI) systems. Non-linear systems must use more complex stability criteria, such as Lyapunov or the circle criterion. While Nyquist is a graphical technique, it only provides a limited amount of intuition for why a system is stable or unstable, or how to modify an unstable system to be stable. Techniques like Bode plots, while less general, are sometimes a more useful design tool.

Background

We consider a system whose open loop transfer function (OLTF) is $G(s)$; when placed in a closed loop with negative feedback $H(s)$, the closed loop transfer function (CLTF) then becomes $G/(1+GH)$. Stability can be determined by examining the roots of the polynomial $1+GH$, e.g. using the Routh array, but this method is somewhat tedious. Conclusions can also be reached by examining the OLTF, using its Bode plots or, as here, polar plot of the OLTF using the Nyquist criterion, as follows.

Any Laplace domain transfer function $\mathcal{T}(s)$can be expressed as the ratio of two polynomials:

$$\mathcal{T}(s) = \frac{N(s)}{D(s)}.$$

The roots of $N(s)$are called the *zeros* of $\mathcal{T}(s)$, and the roots of $D(s)$are the *poles* of $\mathcal{T}(s)$. The poles of $\mathcal{T}(s)$are also said to be the roots of the "characteristic equation" $D(s) = 0$.

The stability of $\mathcal{T}(s)$is determined by the values of its poles: for stability, the real part of every pole must be negative. If $\mathcal{T}(s)$is formed by closing a negative unity feedback loop around the open-loop transfer function $G(s) = A(s)/B(s)$, then the roots of the characteristic equation are also the zeros of $1 + G(s)$, or simply the roots of $A(s) + B(s) = 0$.

Cauchy's argument principle

From complex analysis, specifically the argument principle, we know that a contour $\Gamma_s$drawn in the complex s plane, encompassing but not passing through any number of zeros and poles of a function $F(s)$, can be mapped to another plane (the $F(s)$plane) by the function $F(s)$. The Nyquist plot of $F(s)$, which is the contour $\Gamma_{F(s)} = F(\Gamma_s)$will encircle the point $s = -1/k$ of the $F(s)$plane N times, where $N = Z - P$. Here are Z and Prespectively the number of zeros of $1 + kF(s)$and poles of $F(s)$inside the contour Γ_s. Note that we count encirclements in the $F(s)$plane in the same sense as the contour $\Gamma_s$and that encirclements in the opposite direction are *negative* encirclements. That is, we consider clockwise encirclements to be positive and counterclockwise encirclements to be negative.

Instead of Cauchy's argument principle, the original paper by Harry Nyquist in 1932 uses a less elegant approach. The approach explained here is similar to the approach used by Leroy MacColl (Fundamental theory of servomechanisms 1945) or by Hendrik Bode (Network analysis and feedback amplifier design 1945), both of whom also worked for Bell Laboratories. This approach appears in most modern textbooks on control theory.

The Nyquist criterion

We first construct **The Nyquist Contour**, a contour that encompasses the right-half of the complex plane:

- a path traveling up the $j\omega$ axis, from $0 - j\infty$ to $0 + j\infty$.
- a semicircular arc, with radius $r \rightarrow \infty$, that starts at $0 + j\infty$ and travels clock-wise to $0 - j\infty$.

The Nyquist Contour mapped through the function $1 + G(s)$ yields a plot of $1 + G(s)$ in the complex plane. By the Argument Principle, the number of clock-wise encirclements of the origin must be the number of zeros of $1 + G(s)$ in the right-half complex plane minus the poles of $1 + G(s)$ in the right-half complex plane. If instead, the contour is mapped through the open-loop transfer function $G(s)$, the result is the Nyquist Plot of $G(s)$. By counting the resulting contour's encirclements of -1, we find the difference between the number of poles and zeros in the right-half complex plane of $1 + G(s)$. Recalling that the zeros of $1 + G(s)$ are the poles of the closed-loop system, and noting that the poles of $1 + G(s)$ are same as the poles of $G(s)$, we now state **The Nyquist Criterion**:

Given a Nyquist contour Γ_s, let P be the number of poles of $G(s)$ encircled by Γ_s, and Z be the number of zeros of $1 + G(s)$ encircled by Γ_s. Alternatively, and more importantly, Z is the number of poles of the closed loop system in the right half plane. The resultant contour in the $G(s)$-plane, $\Gamma_{G(s)}$ shall encircle (clock-wise) the point $(-1 + j0)$ N times such that $N = Z - P$.

If the system is originally open-loop unstable, feedback is necessary to stabilize the system. Right-half-plane (RHP) poles represent that instability. For closed-loop stability of a system, the number of closed-loop roots in the right half of the s-plane must be zero. Hence, the number of counter-clockwise encirclements about $-1 + j0$ must be equal to the number of open-loop poles in the RHP. Any clockwise encirclements of the critical point by the open-loop frequency response (when judged from low frequency to high frequency) would indicate that the feedback control system would be destabilizing if the loop were closed. (Using RHP zeros to "cancel out" RHP poles does not remove the instability, but rather ensures that the system will remain unstable even in the presence of feedback, since the closed-loop roots travel between open-loop poles and zeros in the presence of feedback. In fact, the RHP zero can make the unstable pole unobservable and therefore not stabilizable through feedback.)

The Nyquist criterion for systems with poles on the imaginary axis

The above consideration was conducted with an assumption that the open-loop transfer function $G(s)$ does not have any pole on the imaginary axis (i.e. poles of the form $0 + j\omega$). This results from the requirement of the argument principle that the contour cannot pass through any pole of the mapping function. The most common case are systems with integrators (poles at zero).

To be able to analyze systems with poles on the imaginary axis, the Nyquist Contour can be modified to avoid passing through the point $0 + j\omega$. One way to do it is to construct a semicircular arc with radius $r \rightarrow 0$ around $0 + j\omega$, that starts at $0 + j(\omega - r)$ and travels anticlockwise to $0 + j(\omega + r)$. Such a modification implies that the phasor $G(s)$ travels along an arc of infinite radius by $-l\pi$, where l is the multiplicity of the pole on the imaginary axis.

Mathematical Derivation

Our goal is to, through this process, check for the stability of the transfer function of our unity feedback system with gain k, which is given by

$$T(s) = \frac{kG(s)}{1+kG(s)}$$

That is, we would like to check whether the characteristic equation of the above transfer function, given by

$$D(s) = 1 + kG(s) = 0$$

has zeros outside the open left-half-plane (commonly initialized as the OLHP).

We suppose that we have a clockwise (i.e. negatively oriented) contour Γ_s enclosing the right hand plane, with indentations as needed to avoid passing through zeros or poles of the function $G(s)$. Cauchy's argument principle states that

$$-\frac{1}{2\pi i}\oint_{\Gamma_s}\frac{D'(s)}{D(s)}\,ds = N = Z - P$$

Where Z denotes the number of zeros of $D(s)$ enclosed by the contour and P denotes the number of poles of $D(s)$ by the same contour. Rearranging, we have $Z = N + P$, which is to say

$$Z = -\frac{1}{2\pi i}\oint_{\Gamma_s}\frac{D'(s)}{D(s)}\,ds + P$$

We then note that $D(s) = 1 + kG(s)$ has exactly the same poles as $G(s)$. Thus, we may find P by counting the poles of $G(s)$ that appear within the contour, that is, within the open right half plane (ORHP).
We will now rearrange the above integral via substitution. That is, setting $u(s) = D(s)$, we have

$$N = -\frac{1}{2\pi i}\oint_{\Gamma_s}\frac{D'(s)}{D(s)}\,ds = -\frac{1}{2\pi i}\oint_{u(\Gamma_s)}\frac{1}{u}\,du$$

We then make a further substitution, setting $v(u) = \frac{u-1}{k}$. This gives us

$$N = -\frac{1}{2\pi i}\oint_{u(\Gamma_s)}\frac{1}{u}\,du = -\frac{1}{2\pi i}\oint_{v(u(\Gamma_s))}\frac{1}{v+1/k}\,dv$$

We now note that $v(u(\Gamma_s)) = \frac{D(\Gamma_s)-1}{k} = G(\Gamma_s)$ gives us the image of our contour under $G(s)$, which is to say our Nyquist Plot. We may further reduce the integral

$$N = -\frac{1}{2\pi i}\oint_{G(\Gamma_s))}\frac{1}{v+1/k}\,dv$$

by applying Cauchy's integral formula. In fact, we find that the above integral corresponds precisely to the number of times the Nyquist Plot encircles the point $-1/k$ clockwise. Thus, we may finally state that

$$Z = N{+}P = \text{(number of times the Nyquist plot encircles -1/k clockwise)}{+}\text{(number of poles of G(s) in ORHP)}$$

We thus find that $T(s)$ as defined above corresponds to a stable unity-feedback system when Z, as evaluated above, is equal to 0.

Summary

- If the open-loop transfer function $G(s)$ has a zero pole of multiplicity l , then the Nyquist plot has a discontinuity at $\omega = 0$. During further analysis it should be assumed that the phasor travels l times clock-wise along a semicircle of infinite radius. After applying this rule, the zero poles should be neglected, i.e. if there are no other unstable poles, then the open-loop transfer function $G(s)$ should be considered stable.
- If the open-loop transfer function $G(s)$ is stable, then the closed-loop system is unstable for *any* encirclement of the point -1.
- If the open-loop transfer function $G(s)$ is *unstable*, then there must be one *counter* clock-wise encirclement of -1 for each pole of $G(s)$ in the right-half of the complex plane.
- The number of surplus encirclements (greater than N+P) is exactly the number of unstable poles of the closed-loop system
- However, if the graph happens to pass through the point $-1 + j0$, then deciding upon even the marginal stability of the system becomes difficult and the only conclusion that can be drawn from the graph is that there exist zeros on the $j\omega$ axis.

References

- Faulkner, E.A. (1969): *Introduction to the Theory of Linear Systems*; Chapman & Hall; ISBN 0-412-09400-2
- Pippard, A.B. (1985): *Response & Stability*; Cambridge University Press; ISBN 0-521-31994-3
- Gessing, R. (2004): *Control fundamentals*; Silesian University of Technology; ISBN 83-7335-176-0
- Franklin, G. (2002): *Feedback Control of Dynamic Systems*; Prentice Hall, ISBN 0-13-032393-4

Notes

[1] on Alcatel-Lucent website (http://www.alcatel-lucent.com)

Routh–Hurwitz stability criterion

In control system theory, the **Routh–Hurwitz stability criterion** is a mathematical test that is a necessary and sufficient condition for the stability of a linear time invariant (LTI) control system. The Routh test is an efficient recursive algorithm that English mathematician Edward John Routh proposed in 1876 to determine whether all the roots of the characteristic polynomial of a linear system have negative real parts. German mathematician Adolf Hurwitz independently proposed in 1895 to arrange the coefficients of the polynomial into a square matrix, called the Hurwitz matrix, and showed that the polynomial is stable if and only if the sequence of determinants of its principal submatrices are all positive. The two procedures are equivalent, with the Routh test providing a more efficient way to compute the Hurwitz determinants than computing them directly. A polynomial satisfying the Routh-Hurwitz criterion is called a Hurwitz polynomial.

The importance of the criterion is that the roots ***p*** of the characteristic equation of a linear system with negative real parts represent solutions e^{pt} of the system that are stable (bounded). Thus the criterion provides a way to determine if the equations of motion of a linear system have only stable solutions, without solving the system directly. For discrete systems, the corresponding stability test can be handled by the Schur-Cohn criterion, the Jury test and the Bistritz test. With the advent of computers, the criterion has become less widely used, as an alternative is to solve the polynomial numerically, obtaining approximations to the roots directly.

The Routh test can be derived through the use of the Euclidean algorithm and Sturm's theorem in evaluating Cauchy indices. Hurwitz derived his conditions differently.

Using Euclid's algorithm

The criterion is related to Routh–Hurwitz theorem. Indeed, from the statement of that theorem, we have $p-q=w(+\infty)-w(-\infty)$ where:

- p is the number of roots of the polynomial $f(z)$ with negative Real Part;
- q is the number of roots of the polynomial $f(z)$ with positive Real Part (let us remind ourselves that f is supposed to have no roots lying on the imaginary line);
- $w(x)$ is the number of variations of the generalized Sturm chain obtained from $P_0(y)$ and $P_1(y)$ (by successive Euclidean divisions) where $f(iy)=P_0(y)+iP_1(y)$ for a real y.

By the fundamental theorem of algebra, each polynomial of degree n must have n roots in the complex plane (i.e., for an f with no roots on the imaginary line, $p + q = n$). Thus, we have the condition that f is a (Hurwitz) stable polynomial if and only if $p - q = n$ (the proof is given below). Using the Routh–Hurwitz theorem, we can replace the condition on p and q by a condition on the generalized Sturm chain, which will give in turn a condition on the coefficients of f.

Using matrices

Let $f(z)$ be a complex polynomial. The process is as follows:

1. Compute the polynomials $P_0(y)$ and $P_1(y)$ such that $f(iy)=P_0(y)+iP_1(y)$ where y is a real number.
2. Compute the Sylvester matrix associated to $P_0(y)$ and $P_1(y)$.
3. Rearrange each row in such a way that an odd row and the following one have the same number of leading zeros.
4. Compute each principal minor of that matrix.
5. If at least one of the minors is negative (or zero), then the polynomial f is not stable.

Example

- Let $f(z) = az^2 + bz + c$ (for the sake of simplicity we take real coefficients) where $c \neq 0$(to avoid a root in zero so that we can use the Routh–Hurwitz theorem). First, we have to calculate the real polynomials $P_0(y)$ and $P_1(y)$:
$$f(iy) = -ay^2 + iby + c = P_0(y) + iP_1(y) = -ay^2 + c + i(by).$$
Next, we divide those polynomials to obtain the generalized Sturm chain:
  - $P_0(y) = ((-a/b)y)P_1(y) + c,$yields $P_2(y) = -c,$
  - $P_1(y) = ((-b/c)y)P_2(y),$yields $P_3(y) = 0$and the Euclidean division stops.

Notice that we had to suppose *b* different from zero in the first division. The generalized Sturm chain is in this case $(P_0(y), P_1(y), P_2(y)) = (c - ay^2, by, -c)$. Putting $y = +\infty$, the sign of $c - ay^2$is the opposite sign of *a* and the sign of *by* is the sign of *b*. When we put $y = -\infty$, the sign of the first element of the chain is again the opposite sign of *a* and the sign of *by* is the opposite sign of *b*. Finally, -*c* has always the opposite sign of *c*.

Suppose now that *f* is Hurwitz-stable. This means that $w(+\infty) - w(-\infty) = 2$(the degree of *f*). By the properties of the function *w*, this is the same as $w(+\infty) = 2$and $w(-\infty) = 0$. Thus, *a*, *b* and *c* must have the same sign. We have thus found the necessary condition of stability for polynomials of degree 2.

Routh–Hurwitz criterion for second, third, and fourth-order polynomials

In the following, we assume the coefficient of the highest order (e.g. $a_2$in a second order polynomial) to be positive. If necessary, this can always be achieved by multiplication of the polynomial with -1.

- For a second-order polynomial, $P(s) = a_2s^2 + a_1s + a_0 = 0$, all the roots are in the left half plane (and the system with characteristic equation $P(s)$is stable) if all the coefficients satisfy $a_n > 0$.
- For a third-order polynomial $P(s) = a_3s^3 + a_2s^2 + a_1s + a_0 = 0$, all the coefficients must satisfy $a_n > 0$, and $a_2a_1 > a_3a_0$
- For a fourth-order polynomial $P(s) = a_4s^4 + a_3s^3 + a_2s^2 + a_1s + a_0 = 0$, all the coefficients must satisfy $a_n > 0$, and $a_3a_2 > a_4a_1$and $a_3a_2a_1 > a_4a_1^2 + a_3^2a_0$

Higher-order example

A tabular method can be used to determine the stability when the roots of a higher order characteristic polynomial are difficult to obtain. For an *n*th-degree polynomial

- $D(s) = a_ns^n + a_{n-1}s^{n-1} + \cdots + a_1s + a_0$

the table has *n* + 1 rows and the following structure:

a_n	a_{n-2}	a_{n-4}	$\cdots$
a_{n-1}	a_{n-3}	a_{n-5}	$\cdots$
b_1	b_2	b_3	$\cdots$
c_1	c_2	c_3	$\cdots$
$\vdots$	$\vdots$	$\vdots$	$\ddots$

where the elements b_i and c_i can be computed as follows:

- $b_i = \dfrac{a_{n-1} \times a_{n-2i} - a_n \times a_{n-2i-1}}{a_{n-1}}.$
- $c_i = \dfrac{b_1 \times a_{n-2i-1} - a_{n-1} \times b_{i+1}}{b_1}.$

When completed, the number of sign changes in the first column will be the number of non-negative poles.

Consider a system with a characteristic polynomial

- $D(s) = s^5 + 4s^4 + 2s^3 + 5s^2 + 3s + 6.$

We have the following table:

1	2	3	0
4	5	6	0
0.75	1.5	0	0
−3	6	0	
3	0		
6	0		

In the first column, there are two sign changes (0.75 → −3, and −3 → 3), thus there are two non-negative roots where the system is unstable.

Sometimes the presence of poles on the imaginary axis creates a situation of marginal stability. In that case the coefficients of the "Routh Array" in a whole row become zero and thus further solution of the polynomial for finding changes in sign is not possible. Then another approach comes into play. The row of polynomial which is just above the row containing the zeroes is called "Auxiliary Polynomial".

- $s^6 + 2s^5 + 8s^4 + 12s^3 + 20s^2 + 16s + 16 = 0.$

We have the following table:

1	8	20	16
2	12	16	0
2	12	16	0
0	0	0	0

In such a case the Auxiliary polynomial is $A(s) = 2s^4 + 12s^2 + 16.$ which is again equal to zero. The next step is to differentiate the above equation which yields the following polynomial. $B(s) = 8s^3 + 24s^1.$. The coefficients of the row containing zero now become "8" and "24". The process of Routh array is proceeded using these values which yield two points on the imaginary axis. These two points on the imaginary axis are the prime cause of marginal stability.

References

- Hurwitz, A. (1895). "Über die Bedingungen unter welchen eine Gleichung nur Wurzeln mit negativen reellen Teilen besitzt". *Math. Ann. 46, 273-284 (English translation* On the conditions under which an equation has only roots with negative real parts" by H. G. Bergmann in Selected Papers on Mathematical Trends in Control Theory R. Bellman and R. Kalaba Eds. New York: Dover, 1964 pp. 70-82.).
- Routh, E. J. (1877). *A Treatise on the Stability of a Given State of Motion: Particularly Steady Motion*. Macmillan and co.
- Gantmacher, F. R. (1959). "Applications of the Theory of Matrices". *Interscience, New York* **641** (9): 1–8.
- Pippard, A. B.; Dicke, R. H. (1986). "Response and Stability, An Introduction to the Physical Theory" (http://link.aip.org/link/?AJPIAS/54/1052/1). *American Journal of Physics* **54** (11): 1052. Bibcode: 1986AmJPh..54.1052P (http://adsabs.harvard.edu/abs/1986AmJPh..54.1052P). doi: 10.1119/1.14826 (http://dx.doi.org/10.1119/1.14826). Retrieved 2008-05-07.
- Richard C. Dorf, Robert H. Bishop (2001). *Modern Control Systems, 9th Edition*. Prentice Hall. ISBN 0-13-030660-6.

- Rahman, Q. I.; Schmeisser, G. (2002). *Analytic theory of polynomials*. London Mathematical Society Monographs. New Series **26**. Oxford: Oxford University Press. ISBN 0-19-853493-0. Zbl 1072.30006 (http://www.zentralblatt-math.org/zmath/en/search/?format=complete&q=an:1072.30006).
- Weisstein, Eric W. "Routh-Hurwitz Theorem." (http://mathworld.wolfram.com/Routh-HurwitzTheorem.html). *MathWorld--A Wolfram Web Resource*.

External links

- A MATLAB script implementing the Routh-Hurwitz test (http://www.mathworks.com/matlabcentral/fileexchange/25956-routh-hurwitz-stability-test)

Root locus

In control theory and stability theory, **root locus analysis** is a graphical method for examining how the roots of a system change with variation of a certain system parameter, commonly a gain within a feedback system. This is a technique used as a stability criterion in the field of control systems developed by Walter R. Evans which can determine stability of the system. The root locus plots the poles of the closed loop transfer function as a function of a gain parameter.

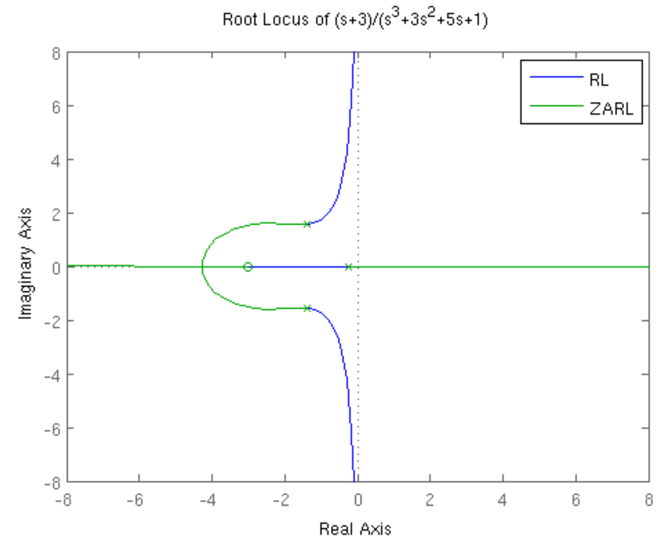
Uses

In addition to determining the stability of the system, the root locus can be used to design the damping ratio and natural frequency of a feedback system. Lines of constant damping ratio can be drawn radially from the origin and lines of constant natural frequency can be drawn as arcs whose center points coincide with the origin. By selecting a point along the root locus that coincides with a desired damping ratio and natural frequency a gain, K, can be calculated and implemented in the controller. More elaborate techniques of controller design using the root locus are available in most control textbooks: for instance, lag, lead, PI, PD and PID controllers can be designed approximately with this technique.

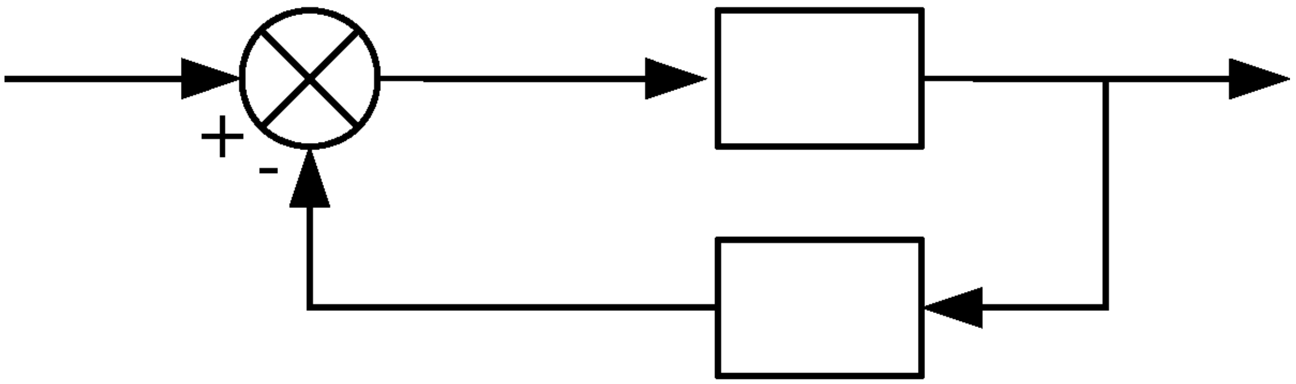
The definition of the damping ratio and natural frequency presumes that the overall feedback system is well approximated by a second order system; i.e. the system has a dominant pair of poles. This is often not the case, so it is good practice to simulate the final design to check if the project goals are satisfied.

Example

Suppose there is a feedback system whose input is the signal $X(s)$ and output is $Y(s)$. The feedback system forward path gain is $G(s)$; the feedback path gain is $H(s)$.



RL = root locus; ZARL = zero angle root locus



For this system, the overall transfer function is given by[1]

$$T(s) = \frac{Y(s)}{X(s)} = \frac{G(s)}{1 + G(s)H(s)}$$

Thus the closed-loop poles (roots of the characteristic equation) of the transfer function are the solutions to the equation $1 + G(s)H(s) = 0$. The principal feature of this equation is that roots may be found wherever $G(s)H(s) = -1$.

In systems without pure delay, the product $G(s)H(s) = -1$ is a rational polynomial function and may be expressed as[2]

$$G(s)H(s) = \frac{K(s+z_1)(s+z_2)\dots(s+z_m)}{(s+p_1)(s+p_2)\dots(s+p_{m+n})}$$

where the $-z_i$ are the m zeros, the $-p_i$ are the $m+n$ poles, and K is a scalar gain. Typically, a root locus diagram will indicate the transfer function's pole locations for varying values of K. A root locus plot will be all those points in the s-plane where $G(s)H(s) = -1$ for any value of K.

The factoring of K and the use of simple monomials means the evaluation of the rational polynomial can be done with vector techniques that add or subtract angles and multiply or divide magnitudes. The vector formulation arises from the fact that each monomial term in the factored $G(s)H(s)$, $(s-a)$ for example, represents the vector from a to s. The polynomial can be evaluated by considering the magnitudes and angles of each of these vectors. According to vector mathematics, the angle of the result is the sum of all the angles in the numerator add minus the sum of all the angles in the denominator. Similarly, the magnitude of the result is the product of all the magnitudes in the numerator divided by the product of all the magnitudes in the denominator. It turns out that the calculation of the magnitude is not needed because *K* varies; one of its values may result in a root. So to test whether a point in the *s*-plane is on the root locus, only the angles to all the open loop poles and zeros need be considered. A graphical method that uses a special protractor called a "Spirule" was once used to determine angles and draw the root loci.

From the function *T*(*s*), it can be seen that the value of *K* does not affect the location of the zeros.Wikipedia:Citation needed The root locus only gives the location of closed loop poles as the gain *K* is varied. The zeros of a system do not move.

Using a few basic rules, the root locus method can plot the overall shape of the path (locus) traversed by the roots as the value of *K* varies. The plot of the root locus then gives an idea of the stability and dynamics of this feedback system for different values of *K*.

Sketching root locus

- Mark open-loop poles and zeros
- Mark real axis portion to the left of an odd number of poles and zeros
- Find asymptotes

Let *P* be the number of poles and *Z* be the number of zeros:

$$P - Z = \text{number of asymptotes}$$

The asymptotes intersect the real axis at α (which is called the centroid) and depart at angle ϕ given by:

$$\phi_l = \frac{180^\circ + (l-1)360^\circ}{P-Z}, l = 1, 2, ..., P-Z$$

$$\alpha = \frac{\sum_P - \sum_Z}{P-Z}$$

where $\sum_P$ is the sum of all the locations of the poles, and $\sum_Z$ is the sum of all the locations of the explicit zeros.

- Phase condition on test point to find angle of departure
- Compute breakaway/break-in points

The breakaway points are located at the roots of the following equation:

$$\frac{dG(s)H(s)}{ds} = 0 \text{ or } \frac{d\overline{GH}(z)}{dz} = 0$$

Once you solve for *z*, the real roots give you the breakaway/reentry points. Complex roots correspond to a lack of breakaway/reentry.

The break-away (break-in) points are obtained by solving a polynomial equation

z-plane versus *s*-plane

The root locus can also be computed in the *z*-plane, the discrete counterpart of the *s*-plane. An equation ($z = e^{sT}$) maps continuous *s*-plane poles (not zeros) into the *z*-domain, where *T* is the sampling period. The stable, left half *s*-plane maps into the interior of the unit circle of the *z*-plane, with the *s*-plane origin equating to $|z| = 1$ (because $e^0 = 1$). A diagonal line of constant damping in the *s*-plane maps around a spiral from (1,0) in the *z* plane as it curves in toward the origin. Note also that the Nyquist aliasing criteria is expressed graphically in the *z*-plane by the *x*-axis, where ($wnT = \pi$). The line of constant damping just described spirals in indefinitely but in sampled data systems, frequency content is aliased down to lower frequencies by integral multiples of the Nyquist frequency. That is, the sampled response appears as a lower frequency and better damped as well since the root in the *z*-plane maps equally well to the first loop of a different, better damped spiral curve of constant damping. Many other interesting and relevant mapping properties can be described, not least that z-plane controllers, having the property that they may be directly implemented from the z-plane transfer function (zero/pole ratio of polynomials), can be imagined graphically on a z-plane plot of the open loop transfer function, and immediately analyzed utilizing root locus.

Since root locus is a graphical angle technique, root locus rules work the same in the *z* and *s* planes.

The idea of a root locus can be applied to many systems where a single parameter *K* is varied. For example, it is useful to sweep any system parameter for which the exact value is uncertain, in order to determine its behavior.

References

[1] Kuo 1967, p. 331.

[2] Kuo 1967, p. 332.

- Kuo, Benjamin C. (1967), "Root Locus Technique", *Automatic Control Systems* (second ed.), Englewood Cliffs, NJ: Prentice-Hall, pp. 329–388, ASIN B000KPT04C (http://www.amazon.com/dp/B000KPT04C), LCCN 67016388 (http://lccn.loc.gov/67016388), OCLC 3805225 (http://www.worldcat.org/oclc/3805225)

Further reading

- Ash, R. H.; Ash, G. H. (October 1968), "Numerical Computation of Root Loci Using the Newton-Raphson Technique", *IEEE Trans. Automatic Control* **13** (5), doi: 10.1109/TAC.1968.1098980 (http://dx.doi.org/10.1109/TAC.1968.1098980)
- Williamson, S. E. (May 1968), "Design Data to assist the Plotting of Root Loci (Part I)", *Control Magazine* **12** (119): 404–407
- Williamson, S. E. (June 1968), "Design Data to assist the Plotting of Root Loci (Part II)", *Control Magazine* **12** (120): 556–559
- Williamson, S. E. (July 1968), "Design Data to assist the Plotting of Root Loci (Part III)", *Control Magazine* **12** (121): 645–647
- Williamson, S. E. (May 15, 1969), "Computer Program to Obtain the Time Response of Sampled Data Systems", *IEE Electronics Letters* **5** (10): 209–210, doi: 10.1049/el:19690159 (http://dx.doi.org/10.1049/el:19690159)
- Williamson, S. E. (July 1969), "Accurate Root Locus Plotting Including the Effects of Pure Time Delay" (http://ieeexplore.ieee.org/Xplore/login.jsp?url=http://ieeexplore.ieee.org/iel5/5247218/5248781/05249527.pdf?tp=&arnumber=5249527&punumber=5247218&authDecision=-203), *Proc. IEE* **116** (7): 1269–1271, doi: 10.1049/piee.1969.0235 (http://dx.doi.org/10.1049/piee.1969.0235)

External links

- Wikibooks: Control Systems/Root Locus (http://en.wikibooks.org/wiki/Control_Systems/Root_Locus)
- Carnegie Mellon / University of Michigan Tutorial (http://www.engin.umich.edu/group/ctm/rlocus/rlocus.html)
- Excellent examples. Start with example 5 and proceed backwards through 4 to 1. Also visit the main page (http://www.swarthmore.edu/NatSci/echeeve1/Ref/LPSA/Root_Locus/RLocusExamples.html#ex5)
- The root-locus method: Drawing by hand techniques (http://www.atp.ruhr-uni-bochum.de/rt1/syscontrol/node46.html)
- "RootLocs": A free multi-featured root-locus plotter for Mac and Windows platforms (http://www.coppice.myzen.co.uk)
- "Root Locus": A free root-locus plotter/analyzer for Windows (http://web.archive.org/web/20091027092528/http://geocities.com/aseldawy/root_locus.html)
- Root Locus at ControlTheoryPro.com (http://wikis.controltheorypro.com/index.php?title=Root_Locus)

- Root Locus Analysis of Control Systems (http://www.roymech.co.uk/Related/Control/root_locus.html)
- MATLAB function for computing root locus of a SISO open-loop model (http://www.mathworks.com/help/toolbox/control/ref/rlocus.html)
- Wechsler, E. R. (January–March 1983), *Root Locus Algorithms for Programmable Pocket Calculators* (http://ipnpr.jpl.nasa.gov/progress_report/42-73/73F.PDF), NASA, pp. 60–64, TDA Progress Report 42-73
- Mathematica function for plotting the root locus (http://reference.wolfram.com/mathematica/ref/RootLocusPlot.html)

Bode plot

A **Bode plot** /ˈboʊdi/ is a graph of the transfer function of a linear, time-invariant system versus frequency, plotted with a log-frequency axis, to show the system's frequency response. It is usually a combination of a **Bode magnitude plot,** expressing the magnitude of the frequency response gain, and a **Bode phase plot,** expressing the frequency response phase shift.

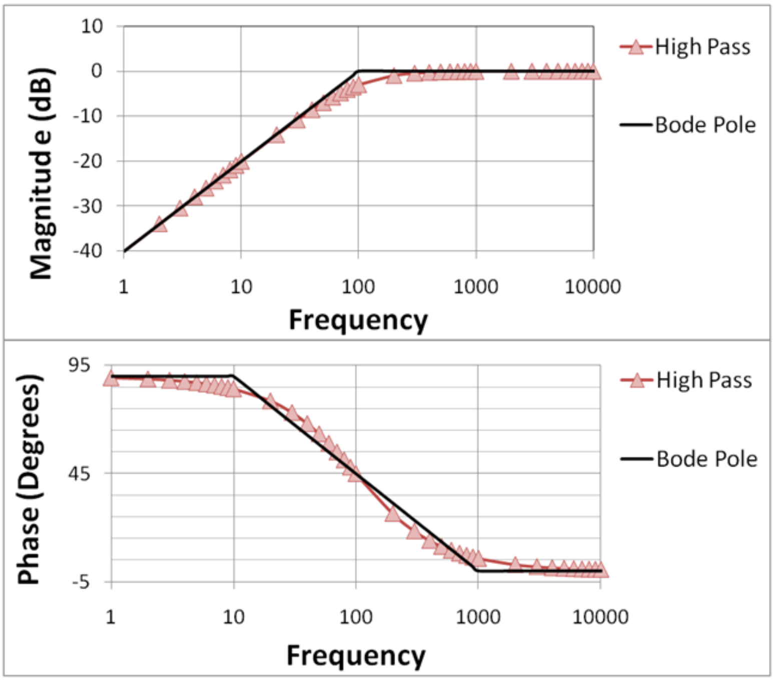


Figure 1(a): The Bode plot for a first-order (one-pole) highpass filter; the straight-line approximations are labeled "Bode pole"; phase varies from 90° at low frequencies (due to the contribution of the numerator, which is 90° at all frequencies) to 0° at high frequencies (where the phase contribution of the denominator is −90° and cancels the contribution of the numerator).

Overview

Among his several important contributions to circuit theory and control theory, engineer Hendrik Wade Bode (1905–1982), while working at Bell Labs in the United States in the 1930s, devised a simple but accurate method for graphing gain and phase-shift plots. These bear his name, *Bode gain plot* and *Bode phase plot*. "Bode" is pronounced /ˈboʊdi/ ***BOH**-dee* (Dutch: [ˈbodə]).[1]

The magnitude axis of the Bode plot is usually expressed as decibels of power, that is by the 20 log rule: 20 times the common (base 10) logarithm of the amplitude gain. With the magnitude gain being logarithmic, Bode plots make multiplication of magnitudes a simple matter of adding distances on the graph (in decibels), since

$$\log(a \cdot b) = \log(a) + \log(b).$$

A **Bode phase plot** is a graph of phase versus frequency, also plotted on a log-frequency axis, usually used in conjunction with the magnitude plot, to evaluate how much a signal will be phase-shifted. For example a signal described by: $A\sin(\omega t)$ may be attenuated but also phase-shifted. If the system attenuates it by a factor x and phase shifts it by $-\Phi$ the signal out of the system will be $(A/x)\sin(\omega t - \Phi)$. The phase shift Φ is generally a function of frequency.

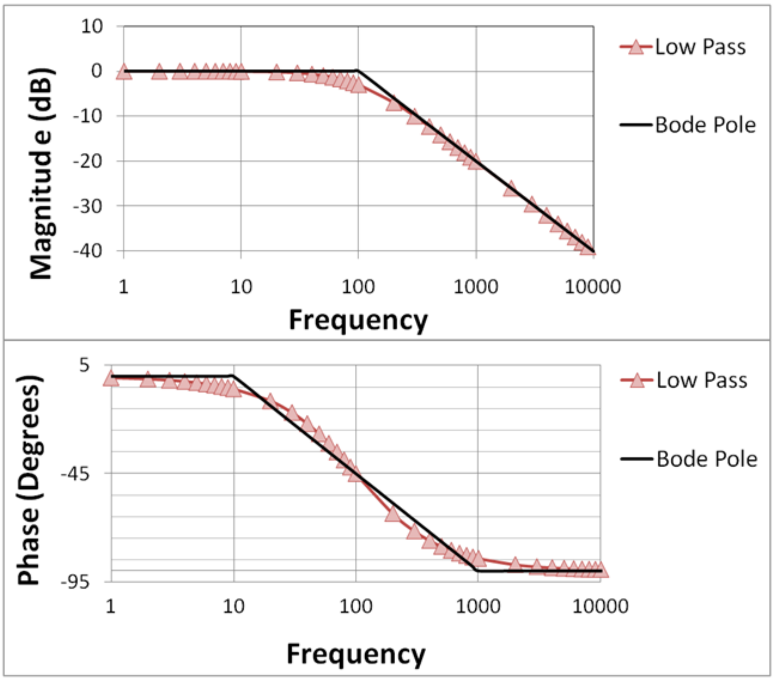


Figure 1(b): The Bode plot for a first-order (one-pole) lowpass filter; the straight-line approximations are labeled "Bode pole"; phase is 90° lower than for Figure 1(a) because the phase contribution of the numerator is 0° at all frequencies.

Phase can also be added directly from the graphical values, a fact that is mathematically clear when phase is seen as the imaginary part of the complex logarithm of a complex gain.

In Figure 1(a), the Bode plots are shown for the one-pole highpass filter function:

$$\mathrm{T}_{\mathrm{High}}(f) = \frac{jf/f_1}{1+jf/f_1},$$

where f is the frequency in Hz, and f_1 is the pole position in Hz, f_1 = 100 Hz in the figure. Using the rules for complex numbers, the magnitude of this function is

$$|\ \mathrm{T}_{\mathrm{High}}(f)\ | = \frac{f/f_1}{\sqrt{1+(f/f_1)^2}},$$

while the phase is:

$$\varphi_{T_{\mathrm{High}}} = 90^\circ - \tan^{-1}(f/f_1).$$

Care must be taken that the inverse tangent is set up to return *degrees*, not radians. On the Bode magnitude plot, decibels are used, and the plotted magnitude is:

$$20\log_{10}|\ \mathrm{T}_{\mathrm{High}}(f)\ | = 20\log_{10}(f/f_1) \\ -20\log_{10}\left(\sqrt{1+(f/f_1)^2}\right).$$

In Figure 1(b), the Bode plots are shown for the one-pole lowpass filter function:

$$\mathrm{T}_{\mathrm{Low}}(f) = \frac{1}{1+jf/f_1}.$$

Also shown in Figure 1(a) and 1(b) are the straight-line approximations to the Bode plots that are used in hand analysis, and described later.

The magnitude and phase Bode plots can seldom be changed independently of each other — changing the amplitude response of the system will most likely change the phase characteristics and vice versa. For minimum-phase systems the phase and amplitude characteristics can be obtained from each other with the use of the Hilbert transform.

If the transfer function is a rational function with real poles and zeros, then the Bode plot can be approximated with straight lines. These asymptotic approximations are called **straight line Bode plots** or **uncorrected Bode plots** and are useful because they can be drawn by hand following a few simple rules. Simple plots can even be predicted without drawing them.

The approximation can be taken further by *correcting* the value at each cutoff frequency. The plot is then called a **corrected Bode plot**.

Rules for handmade Bode plot

The premise of a Bode plot is that one can consider the log of a function in the form:

$$f(x) = A\prod(x - c_n)^{a_n}$$

as a sum of the logs of its poles and zeros:

$$\log(f(x)) = \log(A) + \sum a_n \log(x - c_n).$$

This idea is used explicitly in the method for drawing phase diagrams. The method for drawing amplitude plots implicitly uses this idea, but since the log of the amplitude of each pole or zero always starts at zero and only has one asymptote change (the straight lines), the method can be simplified.

Straight-line amplitude plot

Amplitude decibels is usually done using $\mathbf{dB} = 20\log_{10}(X)$ to define decibels. Given a transfer function in the form

$$H(s) = A\prod\frac{(s - x_n)^{a_n}}{(s - y_n)^{b_n}}$$

where x_n and y_n are constants, $s = j\omega$, $a_n, b_n > 0$, and *H* is the transfer function:

- at every value of s where $\omega = x_n$ (a zero), **increase** the slope of the line by $20 \cdot a_n\ dB$ per decade.
- at every value of s where $\omega = y_n$ (a pole), **decrease** the slope of the line by $20 \cdot b_n\ dB$ per decade.
- The initial value of the graph depends on the boundaries. The initial point is found by putting the initial angular frequency *ω* into the function and finding |*H*(*j*ω)|.
- The initial slope of the function at the initial value depends on the number and order of zeros and poles that are at values below the initial value, and are found using the first two rules.

To handle irreducible 2nd order polynomials, $ax^2 + bx + c$ can, in many cases, be approximated as $(\sqrt{a}x + \sqrt{c})^2$.

Note that zeros and poles happen when ω is *equal to* a certain x_n or y_n. This is because the function in question is the magnitude of H(jω), and since it is a complex function, $|H(j\omega)| = \sqrt{H \cdot H^*}$. Thus at any place where there is a zero or pole involving the term $(s + x_n)$, the magnitude of that term is $\sqrt{(x_n + j\omega)\cdot(x_n - j\omega)} = \sqrt{x_n^2 + \omega^2}$.

Corrected amplitude plot

To correct a straight-line amplitude plot:

- at every zero, put a point $3 \cdot a_n \ \mathrm{dB}$ **above** the line,
- at every pole, put a point $3 \cdot b_n \ \mathrm{dB}$ **below** the line,
- draw a smooth curve through those points using the straight lines as asymptotes (lines which the curve approaches).

Note that this correction method does not incorporate how to handle complex values of x_n or y_n. In the case of an irreducible polynomial, the best way to correct the plot is to actually calculate the magnitude of the transfer function at the pole or zero corresponding to the irreducible polynomial, and put that dot over or under the line at that pole or zero.

Straight-line phase plot

Given a transfer function in the same form as above:

$$H(s) = A \prod \frac{(s - x_n)^{a_n}}{(s - y_n)^{b_n}}$$

the idea is to draw separate plots for each pole and zero, then add them up. The actual phase curve is given by $-\arctan(\mathrm{Im}[H(s)]/\mathrm{Re}[H(s)])$.

To draw the phase plot, for **each** pole and zero:

- if A is positive, start line (with zero slope) at 0 degrees
- if A is negative, start line (with zero slope) at 180 degrees
- if the sum of the number of unstable zeros and poles is odd, add 180 degrees to that basis.
- at every $\omega = |x_n|$ (for stable zeros $- Re(z) < 0$), **increase** the slope by $45 \cdot a_n$ degrees per decade, beginning one decade before $\omega = |x_n|$ (E.g.: $\frac{|x_n|}{10}$)
- at every $\omega = |y_n|$ (for stable poles $- Re(p) < 0$), **decrease** the slope by $45 \cdot b_n$ degrees per decade, beginning one decade before $\omega = |y_n|$ (E.g.: $\frac{|y_n|}{10}$)
- "unstable" (right half plane) poles and zeros ($Re(s) > 0$) have opposite behavior
- flatten the slope again when the phase has changed by $90 \cdot a_n$ degrees (for a zero) or $90 \cdot b_n$ degrees (for a pole),
- After plotting one line for each pole or zero, add the lines together to obtain the final phase plot; that is, the final phase plot is the superposition of each earlier phase plot.

Example

A passive (unity pass band gain) lowpass RC filter, for instance has the following transfer function expressed in the frequency domain:

$$H(jf) = \frac{1}{1 + j2\pi fRC}.$$

From the transfer function it can be determined that the cutoff frequency point f_c (in hertz) is at the frequency

$$f_\mathrm{c} = \frac{1}{2\pi RC}$$

or (equivalently) at

$\omega_\mathrm{c} = \frac{1}{RC}$ where $\omega_\mathrm{c} = 2\pi f_\mathrm{c}$ is the angular cutoff frequency in radians per second.

The transfer function in terms of the angular frequencies becomes:

$$H(j\omega) = \frac{1}{1 + j\frac{\omega}{\omega_{\mathrm{c}}}}.$$

The above equation is the normalized form of the transfer function. The Bode plot is shown in Figure 1(b) above, and construction of the straight-line approximation is discussed next.

Magnitude plot

The magnitude (in decibels) of the transfer function above, (normalized and converted to angular frequency form), given by the decibel gain expression A_{vdB}:

$$A_{\mathrm{vdB}} = 20\log|H(j\omega)| = 20\log\frac{1}{\left|1 + j\frac{\omega}{\omega_{\mathrm{c}}}\right|}$$
$$= -20\log\left|1 + j\frac{\omega}{\omega_{\mathrm{c}}}\right| = -10\log\left(1 + \frac{\omega^2}{\omega_{\mathrm{c}}^2}\right)$$

when plotted versus input frequency ω on a logarithmic scale, can be approximated by two lines and it forms the asymptotic (approximate) magnitude Bode plot of the transfer function:

- for angular frequencies below ω_{c}it is a horizontal line at 0 dB since at low frequencies the $\frac{\omega}{\omega_{\mathrm{c}}}$term is small and can be neglected, making the decibel gain equation above equal to zero,
- for angular frequencies above ω_{c}it is a line with a slope of −20 dB per decade since at high frequencies the $\frac{\omega}{\omega_{\mathrm{c}}}$ term dominates and the decibel gain expression above simplifies to $-20\log\frac{\omega}{\omega_{\mathrm{c}}}$which is a straight line with a slope of −20 dB per decade.

These two lines meet at the corner frequency. From the plot, it can be seen that for frequencies well below the corner frequency, the circuit has an attenuation of 0 dB, corresponding to a unity pass band gain, i.e. the amplitude of the filter output equals the amplitude of the input. Frequencies above the corner frequency are attenuated – the higher the frequency, the higher the attenuation.

Phase plot

The phase Bode plot is obtained by plotting the phase angle of the transfer function given by

$$\varphi = -\tan^{-1}\frac{\omega}{\omega_{\mathrm{c}}}$$

versus ω , where ω and $\omega_{\mathrm{c}}$are the input and cutoff angular frequencies respectively. For input frequencies much lower than corner, the ratio $\frac{\omega}{\omega_{\mathrm{c}}}$is small and therefore the phase angle is close to zero. As the ratio increases the absolute value of the phase increases and becomes –45 degrees when $\omega = \omega_{\mathrm{c}}$. As the ratio increases for input frequencies much greater than the corner frequency, the phase angle asymptotically approaches –90 degrees. The frequency scale for the phase plot is logarithmic.

Normalized plot

The horizontal frequency axis, in both the magnitude and phase plots, can be replaced by the normalized (nondimensional) frequency ratio $\frac{\omega}{\omega_{\mathbf{c}}}$. In such a case the plot is said to be normalized and units of the frequencies are no longer used since all input frequencies are now expressed as multiples of the cutoff frequency $\omega_{\mathbf{c}}$.

An example with pole and zero

Figures 2-5 further illustrate construction of Bode plots. This example with both a pole and a zero shows how to use superposition. To begin, the components are presented separately.

Figure 2 shows the Bode magnitude plot for a zero and a low-pass pole, and compares the two with the Bode straight line plots. The straight-line plots are horizontal up to the pole (zero) location and then drop (rise) at 20 dB/decade. The second Figure 3 does the same for the phase. The phase plots are horizontal up to a frequency factor of ten below the pole (zero) location and then drop (rise) at 45°/decade until the frequency is ten times higher than the pole (zero) location. The plots then are again horizontal at higher frequencies at a final, total phase change of 90°.

Figure 4 and Figure 5 show how superposition (simple addition) of a pole and zero plot is done. The Bode straight line plots again are compared with the exact plots. The zero has been moved to higher frequency than the pole to make a more interesting example. Notice in Figure 4 that the 20 dB/decade drop of the pole is arrested by the 20 dB/decade rise of the zero resulting in a horizontal magnitude plot for frequencies above the zero location. Notice in Figure 5 in the phase plot that the straight-line approximation is pretty approximate in the region where both pole and zero affect the phase. Notice also in Figure 5 that the range of frequencies where the phase changes in the straight line plot is limited to frequencies a factor of ten above and below the pole (zero) location. Where the phase of the pole and the zero both are present, the straight-line phase plot is horizontal because the 45°/decade drop of the pole is arrested by the overlapping 45°/decade rise of the zero in the limited range of frequencies where both are active contributors to the phase.

Example with pole and zero

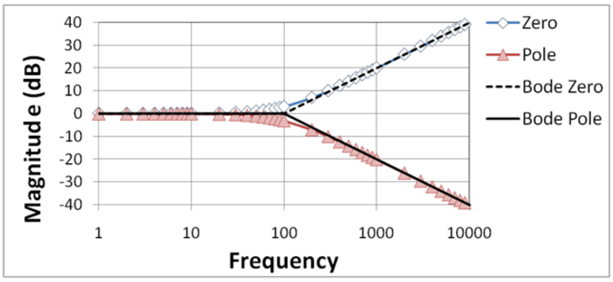


Figure 2: Bode magnitude plot for zero and low-pass pole; curves labeled "Bode" are the straight-line Bode plots

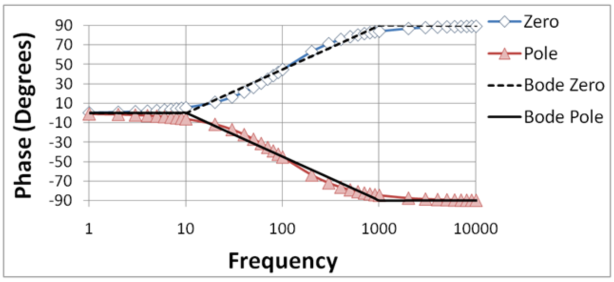


Figure 3: Bode phase plot for zero and low-pass pole; curves labeled "Bode" are the straight-line Bode plots

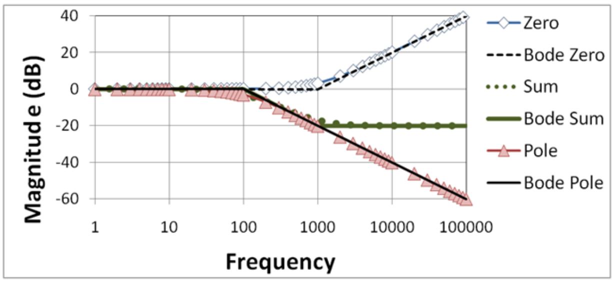


Figure 4: Bode magnitude plot for pole-zero combination; the location of the zero is ten times higher than in Figures 2&3; curves labeled "Bode" are the straight-line Bode plots

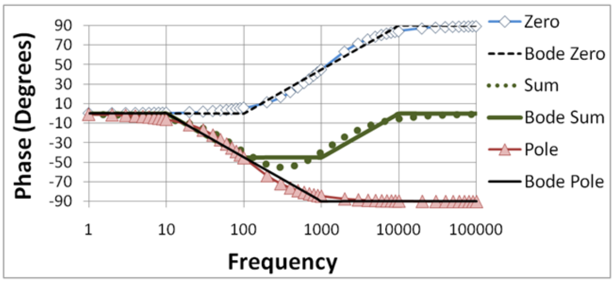


Figure 5: Bode phase plot for pole-zero combination; the location of the zero is ten times higher than in Figures 2&3; curves labeled "Bode" are the straight-line Bode plots

Gain margin and phase margin

See also: Phase margin

Bode plots are used to assess the stability of negative feedback amplifiers by finding the gain and phase margins of an amplifier. The notion of gain and phase margin is based upon the gain expression for a negative feedback amplifier given by

$$A_{FB} = \frac{A_{OL}}{1+\beta A_{OL}} \ ,$$

where A_{FB} is the gain of the amplifier with feedback (the **closed-loop gain**), β is the **feedback factor** and A_{OL} is the gain without feedback (the **open-loop gain**). The gain A_{OL} is a complex function of frequency, with both magnitude and phase.[2] Examination of this relation shows the possibility of infinite gain (interpreted as instability) if the product $\beta A_{OL} = -1$. (That is, the magnitude of βA_{OL} is unity and its phase is −180°, the so-called **Barkhausen stability criterion**). Bode plots are used to determine just how close an amplifier comes to satisfying this condition.

Key to this determination are two frequencies. The first, labeled here as f_{180}, is the frequency where the open-loop gain flips sign. The second, labeled here f_{0dB}, is the frequency where the magnitude of the product $| \beta A_{OL} | = 1$ (in dB, magnitude 1 is 0 dB). That is, frequency f_{180} is determined by the condition:

$$\beta A_{OL}\left(f_{180}\right) = -|\beta A_{OL}\left(f_{180}\right)| = -|\beta A_{OL}|_{180},$$

where vertical bars denote the magnitude of a complex number (for example, $| a + j\, b | = [a^2 + b^2]^{1/2}$), and frequency f_{0dB} is determined by the condition:

$$|\beta A_{OL}\left(f_{0dB}\right)| = 1.$$

One measure of proximity to instability is the **gain margin**. The Bode phase plot locates the frequency where the phase of βA_{OL} reaches −180°, denoted here as frequency f_{180}. Using this frequency, the Bode magnitude plot finds the magnitude of βA_{OL}. If $|\beta A_{OL}|_{180} = 1$, the amplifier is unstable, as mentioned. If $|\beta A_{OL}|_{180} < 1$, instability does not occur, and the separation in dB of the magnitude of $|\beta A_{OL}|_{180}$ from $|\beta A_{OL}| = 1$ is called the *gain margin*. Because a magnitude of one is 0 dB, the gain margin is simply one of the equivalent forms: $20 \log_{10}(|\beta A_{OL}|_{180}) = 20 \log_{10}(|A_{OL}|_{180}) - 20 \log_{10}(1 / \beta)$.

Another equivalent measure of proximity to instability is the **phase margin**. The Bode magnitude plot locates the frequency where the magnitude of $|\beta A_{OL}|$ reaches unity, denoted here as frequency f_{0dB}. Using this frequency, the Bode phase plot finds the phase of βA_{OL}. If the phase of $\beta A_{OL}(f_{0dB}) > -180°$, the instability condition cannot be met at any frequency (because its magnitude is going to be < 1 when $f = f_{180}$), and the distance of the phase at f_{0dB} in degrees above −180° is called the *phase margin*.

If a simple *yes* or *no* on the stability issue is all that is needed, the amplifier is stable if $f_{0dB} < f_{180}$. This criterion is sufficient to predict stability only for amplifiers satisfying some restrictions on their pole and zero positions (minimum phase systems). Although these restrictions usually are met, if they are not another method must be used, such as the Nyquist plot. Optimal gain and phase margins may be computed using Nevanlinna–Pick interpolation theory.

Examples using Bode plots

Figures 6 and 7 illustrate the gain behavior and terminology. For a three-pole amplifier, Figure 6 compares the Bode plot for the gain without feedback (the *open-loop* gain) A_{OL} with the gain with feedback A_{FB} (the *closed-loop* gain). See negative feedback amplifier for more detail.

In this example, A_{OL} = 100 dB at low frequencies, and 1 / β = 58 dB. At low frequencies, A_{FB} ≈ 58 dB as well.

Because the open-loop gain A_{OL} is plotted and not the product β A_{OL}, the condition A_{OL} = 1 / β decides f_{0dB}. The feedback gain at low frequencies and for large A_{OL} is A_{FB} ≈ 1 / β (look at the formula for the feedback gain at the beginning of this section for the case of large gain A_{OL}), so an equivalent way to find f_{0dB} is to look where the feedback gain intersects the open-loop gain. (Frequency f_{0dB} is needed later to find the phase margin.)

Near this crossover of the two gains at f_{0dB}, the Barkhausen criteria are almost satisfied in this example, and the feedback amplifier exhibits a massive peak in gain (it would be infinity if β A_{OL} = −1). Beyond the unity gain frequency f_{0dB}, the open-loop gain is sufficiently small that A_{FB} ≈ A_{OL} (examine the formula at the beginning of this section for the case of small A_{OL}).

Figure 7 shows the corresponding phase comparison: the phase of the feedback amplifier is nearly zero out to the frequency f_{180} where the open-loop gain has a phase of −180°. In this vicinity, the phase of the feedback amplifier plunges abruptly downward to become almost the same as the phase of the open-loop amplifier. (Recall, A_{FB} ≈ A_{OL} for small A_{OL}.)

Comparing the labeled points in Figure 6 and Figure 7, it is seen that the unity gain frequency f_{0dB} and the phase-flip frequency f_{180} are very nearly equal in this amplifier, f_{180} ≈ f_{0dB} ≈ 3.332 kHz, which means the gain margin and phase margin are nearly zero. The amplifier is borderline stable.

Figures 8 and 9 illustrate the gain margin and phase margin for a different amount of feedback β. The feedback factor is chosen smaller than in Figure 6 or 7, moving the condition | β A_{OL} | = 1 to lower frequency. In this example, 1 / β = 77 dB, and at low frequencies A_{FB} ≈ 77 dB as well.

Figure 8 shows the gain plot. From Figure 8, the intersection of 1 / β and A_{OL} occurs at f_{0dB} = 1 kHz. Notice that the peak in the gain A_{FB} near f_{0dB} is almost gone.[3]

Figure 9 is the phase plot. Using the value of f_{0dB} = 1 kHz found above from the magnitude plot of Figure 8, the open-loop phase at f_{0dB} is −135°, which is a phase margin of 45° above −180°.

Using Figure 9, for a phase of −180° the value of f_{180} = 3.332 kHz (the same result as found earlier, of course[4]). The open-loop gain from Figure 8 at f_{180} is 58 dB, and 1 / β = 77 dB, so the gain margin is 19 dB.

Stability is not the sole criterion for amplifier response, and in many applications a more stringent demand than stability is good step response. As a rule of thumb, good step response requires a phase margin of at least 45°, and often a margin of over 70° is advocated, particularly where component variation due to manufacturing tolerances is an issue. See also the discussion of phase margin in the step response article.

Examples

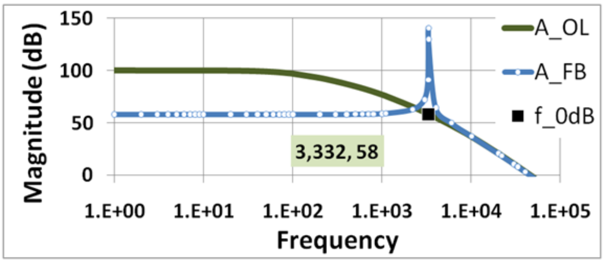


Figure 6: Gain of feedback amplifier A_{FB} in dB and corresponding open-loop amplifier A_{OL}. Parameter 1/β = 58 dB, and at low frequencies A_{FB} ≈ 58 dB as well. The gain margin in this amplifier is nearly zero because $|\beta A_{OL}| = 1$ occurs at almost $f = f_{180°}$.

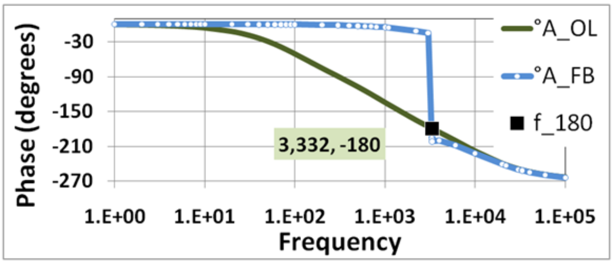


Figure 7: Phase of feedback amplifier $°A_{FB}$ in degrees and corresponding open-loop amplifier $°A_{OL}$. The phase margin in this amplifier is nearly zero because the phase-flip occurs at almost the unity gain frequency $f = f_{0dB}$ where $|\beta A_{OL}| = 1$.

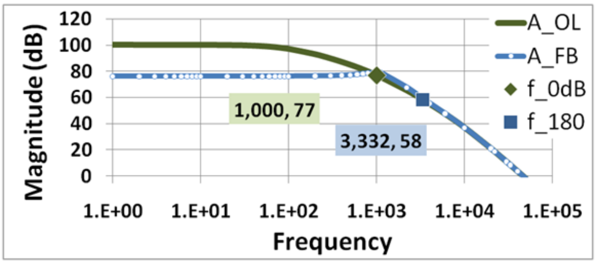


Figure 8: Gain of feedback amplifier A_{FB} in dB and corresponding open-loop amplifier A_{OL}. In this example, 1 / β = 77 dB. The gain margin in this amplifier is 19 dB.

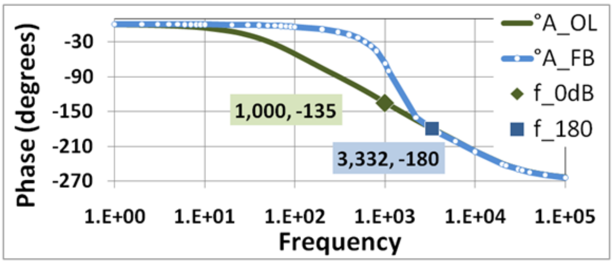


Figure 9: Phase of feedback amplifier A_{FB} in degrees and corresponding open-loop amplifier A_{OL}. The phase margin in this amplifier is 45°.

Bode plotter

The Bode plotter is an electronic instrument resembling an oscilloscope, which produces a Bode diagram, or a graph, of a circuit's voltage gain or phase shift plotted against frequency in a feedback control system or a filter. An example of this is shown in Figure 10. It is extremely useful for analyzing and testing filters and the stability of feedback control systems, through the measurement of corner (cutoff) frequencies and gain and phase margins.

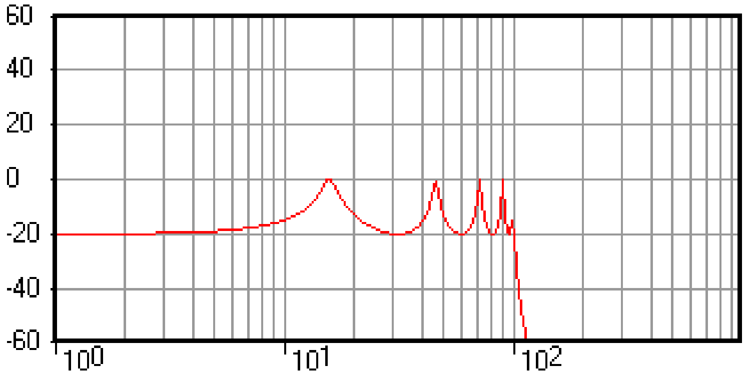


Figure 10: Amplitude diagram of a 10th order Chebyshev filter plotted using a Bode Plotter application. The chebyshev transfer function is defined by poles and zeros which are added by clicking on a graphical complex diagram.

This is identical to the function performed by a vector network analyzer, but the network analyzer is typically used at much higher frequencies.

For education/research purposes, plotting Bode diagrams for given transfer functions facilitates better understanding and getting faster results (see external links).

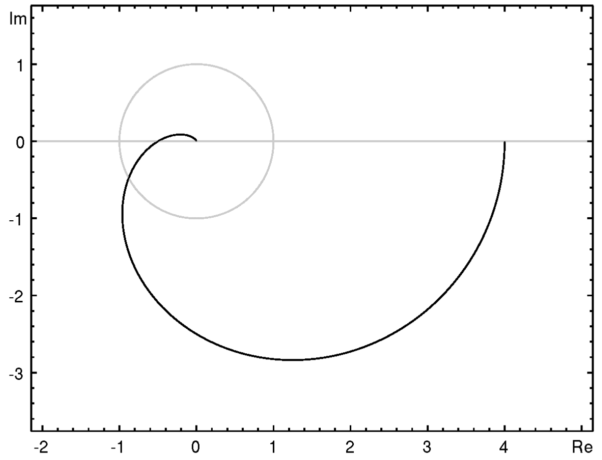
Related plots

Main article: Nyquist plot

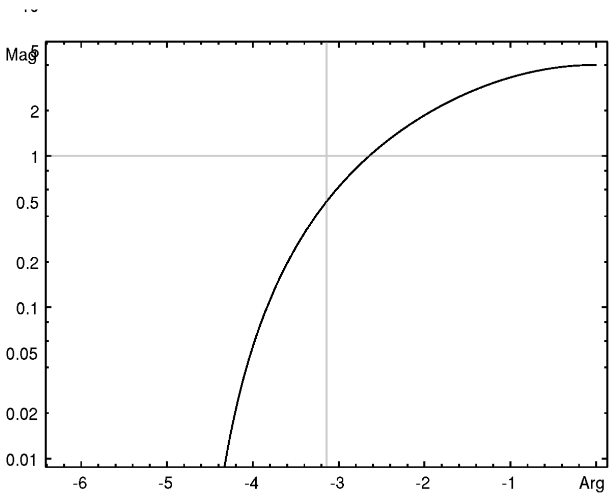
Main article: Nichols plot

Two related plots that display the same data in different coordinate systems are the Nyquist plot and the Nichols plot. These are parametric plots, with frequency as the input and magnitude and phase of the frequency response as the output. The Nyquist plot displays these in polar coordinates, with magnitude mapping to radius and phase to argument (angle). The Nichols plot displays these in rectangular coordinates, on the log scale.

Related Plots



A Nyquist plot.



A Nichols plot of the same response.

Notes

[1] Van Valkenburg, M. E. University of Illinois at Urbana-Champaign, "In memoriam: Hendrik W. Bode (1905-1982)", IEEE Transactions on Automatic Control, Vol. AC-29, No 3., March 1984, pp. 193-194. Quote: "Something should be said about his name. To his colleagues at Bell Laboratories and the generations of engineers that have followed, the pronunciation is boh-dee. The Bode family preferred that the original Dutch be used as boh-dah."

[2] Ordinarily, as frequency increases the magnitude of the gain drops and the phase becomes more negative, although these are only trends and may be reversed in particular frequency ranges. Unusual gain behavior can render the concepts of gain and phase margin inapplicable. Then other methods such as the Nyquist plot have to be used to assess stability.

[3] The critical amount of feedback where the peak in the gain *just* disappears altogether is the *maximally flat* or Butterworth design.

[4] The frequency where the open-loop gain flips sign f_{180} does not change with a change in feedback factor; it is a property of the open-loop gain. The value of the gain at f_{180} also does not change with a change in β. Therefore, we could use the previous values from Figures 6 and 7. However, for clarity the procedure is described using only Figures 8 and 9.

References

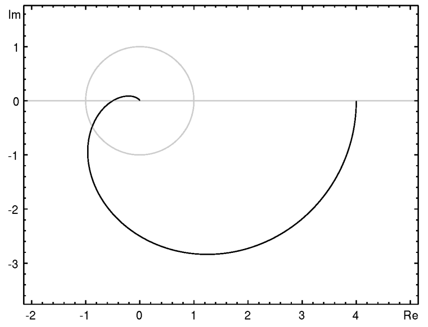
External links

Wikimedia Commons has media related to ***Bode plots***.

- Explanation of Bode plots with movies and examples (http://www.facstaff.bucknell.edu/mastascu/eControlHTML/Freq/Freq5.html)
- How to draw piecewise asymptotic Bode plots (http://lpsa.swarthmore.edu/Bode/BodeHow.html)
- Summarized drawing rules (http://lims.mech.northwestern.edu/~lynch/courses/ME391/2003/bodesketching.pdf) (PDF)
- Bode plot applet (http://www.uwm.edu/People/msw/BodePlot/) - Accepts transfer function coefficients as input, and calculates magnitude and phase response
- Circuit analysis in electrochemistry (http://www.abc.chemistry.bsu.by/vi/fit.htm)
- Tim Green: *Operational amplifier stability* (http://www.en-genius.net/includes/files/acqt_013105.pdf) Includes some Bode plot introduction
- Gnuplot code for generating Bode plot: DIN-A4 printing template (pdf)
- MATLAB function for creating a Bode plot of a system (http://www.mathworks.com/help/control/ref/bode.html)
- MATLAB Tech Talk videos explaining Bode plots and showing how to use them for control design (http://www.mathworks.com/videos/tech-talks/controls/)
- Insert the poles and zeros and this website will draw the asymptotic and accurate Bode plots (http://www.onmyphd.com/?p=bode.plot)
- Mathematica function for creating the Bode plot (http://reference.wolfram.com/mathematica/ref/BodePlot.html)

Nyquist plot

A **Nyquist plot** is a parametric plot of a frequency response used in automatic control and signal processing. The most common use of Nyquist plots is for assessing the stability of a system with feedback. In Cartesian coordinates, the real part of the transfer function is plotted on the X axis. The imaginary part is plotted on the Y axis. The frequency is swept as a parameter, resulting in a plot per frequency. Alternatively, in polar coordinates, the gain of the transfer function is plotted as the radial coordinate, while the phase of the transfer function is plotted as the angular coordinate. The Nyquist plot is named after Harry Nyquist, a former engineer at Bell Laboratories.



A Nyquist plot.

Uses

Assessment of the stability of a closed-loop negative feedback system is done by applying the Nyquist stability criterion to the Nyquist plot of the open-loop system (i.e. the same system without its feedback loop). This method is easily applicable even for systems with delays and other non-rational transfer functions, which may appear difficult to analyze by means of other methods. Stability is determined by looking at the number of encirclements of the point at (-1,0). Range of gains over which the system will be stable can be determined by looking at crossing of the real axis.

The Nyquist plot can provide some information about the shape of the transfer function. For instance, the plot provides information on the difference between the number of poles and zeros of the transfer function[1] by the angle at which the curve approaches the origin.

When drawn by hand, a cartoon version of the Nyquist plot is sometimes used, which shows the shape of the curve, but where coordinates are distorted to show more detail in regions of interest. When plotted computationally, one needs to be careful to cover all frequencies of interest. This typically means that the parameter is swept logarithmically, in order to cover a wide range of values.

References

[1] Nyquist Plots (http://www.facstaff.bucknell.edu/mastascu/econtrolhtml/Freq/Freq6.html)

External links

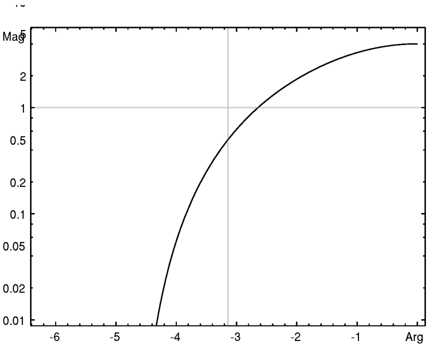
Wikimedia Commons has media related to ***Nyquist plots***.

- Applets with modifiable parameters (http://controlcan.homestead.com/files/idxpages.htm)
- EIS Spectrum Analyser - a freeware program for analysis and simulation of impedance spectra (http://www.abc.chemistry.bsu.by/vi/analyser/)
- MATLAB function (http://www.mathworks.com/help/control/ref/nyquist.html) for creating a Nyquist plot of a frequency response of a dynamic system model.
- PID Nyquist plot shaping (http://www.pidlab.com/en/pid-control-lab-3-1) - free interactive virtual tool, control loop simulator

- Mathematica function for creating the Nyquist plot (http://reference.wolfram.com/mathematica/ref/NyquistPlot.html)

Nichols plot

The **Nichols plot** is a plot used in signal processing and control design, named after American engineer Nathaniel B. Nichols.[1][2][3]



A Nichols plot.

Use in Control Design

Given a transfer function,

$$G(s) = \frac{Y(s)}{X(s)}$$

with the closed-loop transfer function defined as,

$$M(s) = \frac{G(s)}{(1 + G(s))}$$

the Nichols plots displays $20\log_{10}(|G(s)|)$ versus $\arg(G(s))$. Loci of constant $20\log_{10}(|M(s)|)$ and $\arg(M(s))$ are overlaid to allow the designer to obtain the closed loop transfer function directly from the open loop transfer function. Thus, the frequency ω is the parameter along the curve. This plot may be compared to the Bode plot in which the two inter-related graphs - $20\log_{10}(|G(s)|)$ versus $\log_{10}(\omega)$ and $\arg(G(s))$ versus $\log_{10}(\omega)$) - are plotted.

In feedback control design, the plot is useful for assessing the stability and robustness of a linear system. This application of the Nichols plot is central to the Quantitative feedback theory (QFT) of Horowitz and Sidi, which is a well known method for robust control system design.

In most cases, $\arg(G(s))$ refers to the phase of the system's response. Although similar to a Nyquist plot, a Nichols plot is plotted in a Cartesian coordinate system while a Nyquist plot is plotted in a polar coordinate system.

References

[1] Isaac M. Howowitz, *Synthesis of Feedback Systems*, Academic Press, 1963, Lib Congress 63-12033 p. 194-198

[2] Boris J. Lurie and Paul J. Enright, *Classical Feedback Control*, Marcel Dekker, 2000, ISBN 0-8247-0370-7 p. 10

[3] Allen Stubberud, Ivan Williams, and Joseph DeStefano, *Shaums Outline Feedback and Control Systems*, McGraw-Hill, 1995, ISBN 0-07-017052-5 ch. 17

External links

- Mathematica function for creating the Nichols plot (http://reference.wolfram.com/mathematica/ref/NicholsPlot.html)

Phase margin

In electronic amplifiers, the **phase margin** (PM) is the difference between the phase, measured in degrees, and 180°, for an amplifier's output signal (relative to its input), as a function of frequency. Typically the open-loop phase lag (relative to input) varies with frequency, progressively increasing to exceed 180°, at which frequency the output signal becomes inverted, or antiphase in relation to the input. The PM as defined will be positive but decreasing at frequencies less than the frequency at which inversion sets in (at which PM = 0), and PM is negative (PM < 0) at higher frequencies. In the presence of negative feedback, a zero or negative PM at a frequency where the loop gain exceeds unity (1) guarantees instability. Thus positive PM is a "safety margin" that ensures proper (non-oscillatory) operation of the circuit. This applies to amplifier circuits as well as more generally, to active filters, under various load conditions (e.g. reactive loads). In its simplest form, involving ideal negative feedback *voltage* amplifiers with non-reactive feedback, the phase margin is measured at the frequency where the open-loop voltage gain of the amplifier equals the desired closed-loop DC voltage gain.

More generally, PM is defined as that of the amplifier and its feedback network combined (the "loop", normally opened at the amplifier input), measured at a frequency where the loop gain is unity, and prior to the closing of the loop, through tying the output of the open loop to the input source, in such a way as to subtract from it.

In the above loop-gain definition, it is assumed that the amplifier input presents zero load. To make this work for non-zero-load input, the output of the feedback network needs to be loaded with an equivalent load for the purpose of determining the frequency response of the loop gain.

It is also assumed that the graph of gain vs. frequency crosses unity gain with a negative slope and does so only once. This consideration matters only with reactive and active feedback networks, as may be the case with active filters.

Phase margin and its important companion concept, gain margin, are measures of stability in closed-loop, dynamic-control systems. Phase margin indicates relative stability, the tendency to oscillate during its damped response to an input change such as a step function. Gain margin indicates absolute stability and the degree to which the system will oscillate, without limit, given any disturbance.

The output signals of all amplifiers exhibit a time delay when compared to their input signals. This delay causes a phase difference between the amplifier's input and output signals. If there are enough stages in the amplifier, at some frequency, the output signal will lag behind the input signal by one cycle period at that frequency. In this situation, the amplifier's output signal will be in phase with its input signal though lagging behind it by 360°, i.e., the output will have a phase angle of −360°. This lag is of great consequence in amplifiers that use feedback. The reason: the amplifier will oscillate if the fed-back output signal is in phase with the input signal at the frequency at which its open-loop voltage gain equals its closed-loop voltage gain and the open-loop voltage gain is one or greater. The oscillation will occur because the fed-back output signal will then reinforce the input signal at that frequency.[1] In conventional operational amplifiers, the critical output phase angle is −180° because the output is fed back to the input through an inverting input which adds an additional −180°.

In practice, feedback amplifiers must be designed with phase margins substantially in excess of 0°, even though amplifiers with phase margins of, say, 1° are theoretically stable. The reason is that many practical factors can reduce the phase margin below the theoretical minimum. A prime example is when the amplifier's output is connected to a capacitive load. Therefore, operational amplifiers are usually compensated to achieve a minimum phase margin of 45° or so. This means that at the frequency at which the open and closed loop gains meet, the phase angle is −135°. The calculation is: {{{1}}} See Warwick or Stout for a detailed analysis of the techniques and results of compensation to insure adequate phase margins. See also the article "Pole splitting". Often amplifiers are designed to achieve a typical phase margin of 60 degrees. If the typical phase margin is around 60 degrees then the minimum phase margin will typically be greater than 45 degrees. A phase margin of 60 degrees is also a magic number because it allows for the fastest settling time when attempting to follow a voltage step input (a Butterworth design). An amplifier with lower phase margin will ring[2] for longer and an amplifier with more phase margin will take a longer time to rise to the voltage step's final level.

A related measure is gain margin. While phase margin comes from the phase where the loop gain equals one, the gain margin is based upon the gain where the phase equals -180 degrees.

Footnotes

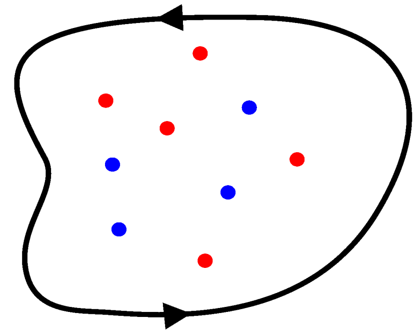
[1] *Ibid*, p. 245.

[2] Ringing is the displaying of a decaying oscillation for a portion of the output signal's cycle; see ringing artifacts.

References

Argument principle

In complex analysis, the **argument principle** (or **Cauchy's argument principle**) relates the difference between the number of zeros and poles of a meromorphic function to a contour integral of the function's logarithmic derivative.



The simple contour C (black), the zeros of f (blue) and the poles of f (red). Here we have $\oint_C \frac{f'(z)}{f(z)}\,dz = 2\pi i(4-5)$.

Specifically, if $f(z)$ is a meromorphic function inside and on some closed contour C, and f has no zeros or poles on C, then

$$\oint_C \frac{f'(z)}{f(z)}\,dz = 2\pi i(N-P)$$

where N and P denote respectively the number of zeros and poles of $f(z)$ inside the contour C, with each zero and pole counted as many times as its multiplicity and order, respectively, indicate. This statement of the theorem assumes that the contour C is simple, that is, without self-intersections, and that it is oriented counter-clockwise.

More generally, suppose that $f(z)$ is a meromorphic function on an open set Ω in the complex plane and that C is a closed curve in Ω which avoids all zeros and poles of f and is contractible to a point inside Ω. For each point $z \in \Omega$, let $n(C,z)$ be the winding number of C around z. Then

$$\oint_C \frac{f'(z)}{f(z)}\,dz = 2\pi i\left(\sum_a n(C,a) - \sum_b n(C,b)\right)$$

where the first summation is over all zeros a of f counted with their multiplicities, and the second summation is over the poles b of f counted with their orders.

Interpretation of the contour integral

The contour integral $\oint_C \frac{f'(z)}{f(z)}\,dz$ can be interpreted in two ways:

- as the total change in the argument of *f*(*z*) as *z* travels around *C*, explaining the name of the theorem; this follows from

$$\frac{d}{dz}\log(f(z)) = \frac{f'(z)}{f(z)}$$

and the relation between arguments and logarithms.

- as 2*πi* times the winding number of the path *f*(*C*) around the origin, using the substitution *w* = *f*(*z*):

$$\oint_C \frac{f'(z)}{f(z)}\,dz = \oint_{f(C)} \frac{1}{w}\,dw$$

Proof of the argument principle

Let z_N be a zero of *f*. We can write $f(z) = (z - z_N)^k g(z)$ where *k* is the multiplicity of the zero, and thus $g(z_N) \neq 0$. We get

$$f'(z) = k(z - z_N)^{k-1} g(z) + (z - z_N)^k g'(z)$$

and

$$\frac{f'(z)}{f(z)} = \frac{k}{z - z_N} + \frac{g'(z)}{g(z)}.$$

Since $g(z_N) \neq 0$, it follows that $g'(z)/g(z)$ has no singularities at z_N, and thus is analytic at z_N, which implies that the residue of $f'(z)/f(z)$ at z_N is *k*.

Let z_P be a pole of *f*. We can write $f(z) = (z - z_\mathrm{P})^{-m} h(z)$ where *m* is the order of the pole, and $h(z_\mathrm{P}) \neq 0$. Then,

$$f'(z) = -m(z - z_P)^{-m-1} h(z) + (z - z_P)^{-m} h'(z).$$

and

$$\frac{f'(z)}{f(z)} = \frac{-m}{z - z_P} + \frac{h'(z)}{h(z)}$$

similarly as above. It follows that $h'(z)/h(z)$ has no singularities at z_P since $h(z_\mathrm{P}) \neq 0$ and thus it is analytic at z_P. We find that the residue of $f'(z)/f(z)$ at z_P is $-m$.

Putting these together, each zero z_N of multiplicity *k* of *f* creates a simple pole for $f'(z)/f(z)$ with the residue being *k*, and each pole z_P of order *m* of *f* creates a simple pole for $f'(z)/f(z)$ with the residue being $-m$. (Here, by a simple pole we mean a pole of order one.) In addition, it can be shown that $f'(z)/f(z)$ has no other poles, and so no other residues.

By the residue theorem we have that the integral about *C* is the product of 2*πi* and the sum of the residues. Together, the sum of the *k* 's for each zero z_N is the number of zeros counting multiplicities of the zeros, and likewise for the poles, and so we have our result.

Applications and consequences

The argument principle can be used to efficiently locate zeros or poles of meromorphic functions on a computer. Even with rounding errors, the expression $\frac{1}{2\pi i}\oint_C \frac{f'(z)}{f(z)}\,dz$ will yield results close to an integer; by determining these integers for different contours *C* one can obtain information about the location of the zeros and poles. Numerical tests of the Riemann hypothesis use this technique to get an upper bound for the number of zeros of Riemann's $\xi(s)$ function inside a rectangle intersecting the critical line.

The proof of Rouché's theorem uses the argument principle.

Modern books on feedback control theory quite frequently use the argument principle to serve as the theoretical basis of the Nyquist stability criterion.

A consequence of the more general formulation of the argument principle is that, under the same hypothesis, if *g* is an analytic function in Ω, then

$$\frac{1}{2\pi i}\oint_C g(z)\frac{f'(z)}{f(z)}\,dz = \sum_a n(C,a)g(a) - \sum_b n(C,b)g(b).$$

For example, if *f* is a polynomial having zeros z_1, ..., z_p inside a simple contour *C*, and *g*(*z*) = z^k, then

$$\frac{1}{2\pi i}\oint_C z^k\frac{f'(z)}{f(z)}\,dz = z_1^k + z_2^k + \ldots + z_p^k,$$

is power sum symmetric polynomial of the roots of *f*.

Another consequence is if we compute the complex integral:

$$\oint_C f(z)\frac{g'(z)}{g(z)}\,dz$$

for an appropriate choice of *g* and *f* we have the Abel–Plana formula:

$$\sum_{n=0}^{\infty} f(n) - \int_0^\infty f(x)\,dx = f(0)/2 + i\int_0^\infty \frac{f(it)-f(-it)}{e^{2\pi t}-1}\,dt$$

which expresses the relationship between a discrete sum and its integral.

Generalized argument principle

There is an immediate generalization of the argument principle. The integral

$$\oint_C \frac{f'(z)}{f(z)}g(z)\,dz$$

is equal to g evaluated at the zeroes, minues g evaluated at the poles.

History

According to the book by Frank Smithies (*Cauchy and the Creation of Complex Function Theory*, Cambridge University Press, 1997, p. 177), Augustin-Louis Cauchy presented a theorem similar to the above on 27 November 1831, during his self-imposed exile in Turin (then capital of the Kingdom of Piedmont-Sardinia) away from France. However, according to this book, only zeroes were mentioned, not poles. This theorem by Cauchy was only published many years later in 1974 in a hand-written form and so is quite difficult to read. Cauchy published a paper with a discussion on both zeroes and poles in 1855, two years before his death.

References

- Rudin, Walter (1986). *Real and Complex Analysis (International Series in Pure and Applied Mathematics)*. McGraw-Hill. ISBN 978-0-07-054234-1.
- Ahlfors, Lars (1979). *Complex analysis: an introduction to the theory of analytic functions of one complex variable*. McGraw-Hill. ISBN 978-0-07-000657-7.
- Churchill, Ruel Vance; Brown, James Ward (1989). *Complex Variables and Applications*. McGraw-Hill. ISBN 978-0-07-010905-6.
- Backlund, R.-J. (1914) Sur les zéros de la fonction zeta(s) de Riemann, C. R. Acad. Sci. Paris 158, 1979-1982.

Article Sources and Contributors

LTI system theory *Source*: https://en.wikipedia.org/w/index.php?oldid=600300203 *Contributors*: AK456, Batsmacks, BenFrantzDale, Bob K, Bodrell, CBM, Cbs228, Cihan, Cxw, DerHexer, Dicklyon, Gbruin, Giftlite, GoingBatty, Grebaldar, GreyCat, Infzy, Interferometrist, J04n, JABoye47x, Jamelan, Jpkotta, Juansempere, Kb9wte, Mako098765, Marlynboke, Mdd, Melcombe, Metacomet, Michael Hardy, Mmeijeri, Mwilde, Natalya, Nejko, Nish-NJITWILL, Niteowlneils, No-body, Nuwewsco, Oli Filth, Onkelringelhuth, Rbj, Rsrikanth05, Schizodelight, SchreiberBike, Smtchahal, Sonick, Splash, Sterrys, TedPavlic, ThePianoMan, Toffile, Zvika, ^musaz, 80 anonymous edits

Transfer function *Source*: https://en.wikipedia.org/w/index.php?oldid=618348959 *Contributors*: A multidimensional liar, Ale2006, Ap, Applebytom, Attilios, AxelBoldt, BenFrantzDale, Bob K, Brad7777, Bryan Derksen, Colin Marquardt, Conversion script, Crazycasta, Cyrius, D1ma5ad, Dbroadwell, Dysprosia, Elwikipedista, Fresheneesz, Frze, Giftlite, Grubber, H2g2bob, HappySophie, HeroTsai, Hypergeek14, Ihatetoregister, Imjustmatthew, Jendem, Jiuguang Wang, Jorge Stolfi, Josh1238, Jwdietrich2, Kku, L Kensington, Light current, Linas, Maitchy, Mastermeister, Mckee, Mebden, Melcombe, Metacomet, Michael Hardy, Mschlindwein, Mwilde, Mwtoews, NathanHagen, Nejko, Oarih, Oleg Alexandrov, Oli Filth, Omegatron, Pol098, Raggot, Rbj, Rhetth, Ro8269, Roadrunner, Rwestafer, Sagaciousuk, SamuelRiv, Smack, Smyth, Spinningspark, Splash, Sterrys, TedPavlic, The Anome, Thermochap, Tijfo098, Tlotoxl, Tobias Hoevekamp, Tuos, UKoch, Vonkje, Wallers, Widr, Wowo53, Xlation, Zfeinst, 119 anonymous edits

Step response *Source*: https://en.wikipedia.org/w/index.php?oldid=607170239 *Contributors*: Billymac00, Brews ohare, Chendy, Cuaxdon, Daniele.tampieri, DexDor, Dicklyon, EagleFan, EdSaWiki, Furby100, Hike395, Jiuguang Wang, Kmellem, Michael Hardy, Nbarth, Paul August, Rhesusmonkeyboy, Ro8269, Rs2, STEVE SURE, Sbyrnes321, Serenthia, Silverfish, TedPavlic, 18 anonymous edits

BIBO stability *Source*: https://en.wikipedia.org/w/index.php?oldid=607308644 *Contributors*: A8UDI, Brews ohare, Bruguiea, CambridgeBayWeather, Cburnett, Damian Yerrick, Dedalus (usurped), Feynmanliang, Giftlite, H, Hakeem.gadi, Hypergeek14, Jiuguang Wang, Jpkotta, Jwy, Lambertch, Michael Hardy, Mike Rosoft, Mkbergman, Mwilde, Novangelis, Oli Filth, R'n'B, Rdrozd, Ro8269, Romberg, Sapphic, TedPavlic, 37 anonymous edits

Nyquist stability criterion *Source*: https://en.wikipedia.org/w/index.php?oldid=617888688 *Contributors*: Aminrahimian, Attilios, Bengski68, Beta16, Brews ohare, Buffetrand, Charles Matthews, Chetvorno, Cuaxdon, Dewritech, DexDor, Dicklyon, Duncharris, Ewlyahoocom, Giftlite, Haditya, Hvn0413, Ikiwaner, Jenblower, Jiuguang Wang, Jm1234567890, Keepitfree, Krishnavedala, LachlanA, Linuxlad, Lupo, Lv131, Mailer diablo, Matutano, Merovingian, Minhlong87, Musically ut, Nillerdk, Novangelis, Pharaoh of the Wizards, Phobulos, PoqVaUSA, Preben Holm, Saung Tadashi, Smalljim, Sun Creator, Toffile, Trusilver, Wknight8111, XMxWx, Zueignung, Пика Пика, 67 anonymous edits

Routh–Hurwitz stability criterion *Source*: https://en.wikipedia.org/w/index.php?oldid=618221578 *Contributors*: Becritical, Bestable, Billymac00, Bletchley, Brews ohare, Catslash, Charles Matthews, Chetvorno, Deltahedron, Dicklyon, Dionyziz, Djenicek, Duncharris, Ekotkie, Fij, FourBlades, Franklin Yu, Fweeky, Giftlite, Insertesla, Jitse Niesen, Jiuguang Wang, Julien Tuerlinckx, Kevinsane, Kmellem, Linuxlad, Mark viking, Mdd, Mdupont, Mhkoepf, Michael Hardy, Nanasid, Novangelis, Olaf, Oleg Alexandrov, Pelotas, Ruddyscent, Saung Tadashi, Spradlig, Tbsmith, Tecgl, Titoxd, User A1, Vivek7de, Wordcotton, Zaxxonal, 53 anonymous edits

Root locus *Source*: https://en.wikipedia.org/w/index.php?oldid=618609616 *Contributors*: Abce2, Adoniscik, AhmedHan, Alphachimp, Angela, Applebytom, Aturevsk, AvicAWB, Bgwhite, Biezl, Brews ohare, Burns512, Celieber, Charles Matthews, Cuaxdon, Daniel fady, David Ayton, Dja25, Donner60, Eng-Mkh, Fredrik, Glrx, Gtnash, Hvn0413, J.delanoy, Kgartner, Kir360, Lpkeys, Marcosagliano, Marcus Cyron, Markrkrebs, Masgatotkaca, Mdd, Melhawar, Michael Hardy, Musically ut, OlavN, Oleg Alexandrov, Panda34, Petrb, Quinwound, Raiker, Rasputin243, RickK, Rs2, Salamurai, Silverfish, Spradlig, Sss41, Topbanana, Trevor MacInnis, Tuvas, Updatehelper, Vanished user 9i39j3, Velella, WhiteOak2006, Wigan25, Xeno, Yamamoto Ichiro, Zfeinst, Zzxterry, 130 anonymous edits

Bode plot *Source*: https://en.wikipedia.org/w/index.php?oldid=614228865 *Contributors*: A2Kafir, Andy.ackland, Anna Lincoln, Anon J, Applebytom, Arbitrarily0, Arcturus4669, Aturevsk, Avicennasis, Avoided, BenFrantzDale, Betacommand, Bobblehead, Brews ohare, Buffetrand, Chaosgate, Chris the speller, Clam0p, Clemens vi, Dbtfz, Deflective, Derdeib, DexDor, Dicklyon, Dr.K., Drew335, Ec5618, ExportRadical, Falcorian, Fresheneesz, Fritsebits, Gaius Cornelius, Ganzor, Giftlite, Gmoose1, Graibeard, Gravix, Heron, Hooperbloob, Huntthetroll, Insaneinside, Jamelan, Jeff P Harris, Jiuguang Wang, Jive Dadson, Joy, Jshen6, KPH2293, Kenyon, Kmellem, Kwamikagami, Lambertch, Lugia2453, Martarius, Mdd, Mehran, Mellery, Michael Devore, Michael Hardy, Miguelgoldstein, Mivey6, Mr. PIM, MusikAnimal, Nabla, Nbarth, Niketmjoshi, Nitin.mehta, Nodlit, Omegatron, Patrick, Peter.scottie, Phosphoricx, Poulpy, Private Pilot, Redheylin, Rehnn83, Requestion, Shadowjams, SimonP, Sommacal alfonso, Spradlig, Stratzvyda, Stw, Tim Starling, Toffile, Tomer shalev, User A1, Vukg, WakingLili, Wikispaghetti, Wtshymanski, Zoomzoom1, 134 anonymous edits

Nyquist plot *Source*: https://en.wikipedia.org/w/index.php?oldid=603542001 *Contributors*: A. B., Applebytom, Attilios, Aturevsk, Bazz, Brews ohare, Cuaxdon, CyrilB, DexDor, Diego Moya, Engelec, Epicgenius, Flehmen, Giftlite, Hemanshu, Heron, Hooperbloob, Jeff P Harris, Jiuguang Wang, John of Reading, JorisvS, Linuxlad, MarkusHagenlocher, Mdd, Mets501, Michael Hardy, Mpassman, Nbarth, Omegatron, Patrick, Pjvpjv, Pladask, PoqVaUSA, Reconsider the static, Ro8269, Rod57, Salasks, Solarra, Spiff, The Anome, Titan, Toffile, Trojancowboy, Vonkje, XMxWx, Xanchester, Yorrak, 46 anonymous edits

Nichols plot *Source*: https://en.wikipedia.org/w/index.php?oldid=602119253 *Contributors*: Applebytom, B4hand, Bazz, Blair Bonnett, Captain Wanga!, Chthonicdaemon, Encyclops, Engelec, Farhoudk, Giftlite, HenryXVII, Hooperbloob, ILUsion, Magicarrow2, Mdd, Michael Hardy, Niceguyedc, Peter17, The Anome, Tuvosi, 4 anonymous edits

Phase margin *Source*: https://en.wikipedia.org/w/index.php?oldid=609621725 *Contributors*: Ahamkah, Anoneditor, Brews ohare, Buffetrand, Depassp, DexDor, Dsignoff, Emote, Fredrik, Giftlite, Gmoose1, Hooperbloob, Jiuguang Wang, Makyen, Nbarth, Rod57, Spradlig, TedPavlic, Theoldanarchist, Thinking of England, Toffile, Toolnut, Wikid77, 15 anonymous edits

Argument principle *Source*: https://en.wikipedia.org/w/index.php?oldid=618166459 *Contributors*: Adriaan Joubert, AxelBoldt, BigJohnHenry, Btyner, Buckeye1973, Cic, Darklilac, DionysosProteus, Doctormatt, Dysprosia, Gaius Cornelius, Gelo71, Geometry guy, Ghirlandajo, Giftlite, Gundamlh, Ideal gas equation, Jamesmcmahon0, Janm67, Jshadias, K9re11, Karl-H, Manifestement, Mazi, Michael Hardy, Mirko vukovic, Oleg Alexandrov, Passw0rd, Pleasantville, R.e.b., RDBury, Rpchase, Sammy1339, Saung Tadashi, Shambolic Entity, Silly rabbit, ThibautLienart, Xnn, 48 anonymous edits

Image Sources, Licenses and Contributors

File:LTI.png *Source*: https://en.wikipedia.org/w/index.php?title=File:LTI.png *License*: Public Domain *Contributors*: Flappiefh, Joelholdsworth, Juanlu001, Ma-Lik, Maksim

File:High accuracy settling time measurements figure 1.png *Source*: https://en.wikipedia.org/w/index.php?title=File:High_accuracy_settling_time_measurements_figure_1.png *License*: Public Domain *Contributors*: Original Howard K. Schoenwetter. Traced by Sagsaw

Image:Block Diagram for Feedback.svg *Source*: https://en.wikipedia.org/w/index.php?title=File:Block_Diagram_for_Feedback.svg *License*: Creative Commons Attribution-Sharealike 3.0,2.5,2.0,1.0 *Contributors*: Block_diagram_for_feedback.PNG: Brews ohare derivative work: Serenthia

Image:Conjugate poles in s-plane.PNG *Source*: https://en.wikipedia.org/w/index.php?title=File:Conjugate_poles_in_s-plane.PNG *License*: Creative Commons Attribution-Share Alike *Contributors*: Brews ohare

Image:Step response for two-pole feedback amplifier.PNG *Source*: https://en.wikipedia.org/w/index.php?title=File:Step_response_for_two-pole_feedback_amplifier.PNG *License*: Creative Commons Attribution-Share Alike *Contributors*: Brews ohare

Image:Overshoot control.PNG *Source*: https://en.wikipedia.org/w/index.php?title=File:Overshoot_control.PNG *License*: Creative Commons Attribution-Share Alike *Contributors*: Brews ohare

Image:Phase for Step Response.PNG *Source*: https://en.wikipedia.org/w/index.php?title=File:Phase_for_Step_Response.PNG *License*: Creative Commons Attribution-Share Alike *Contributors*: Brews ohare

Image:Step response.jpg *Source*: https://en.wikipedia.org/w/index.php?title=File:Step_response.jpg *License*: Public Domain *Contributors*: Daniele.tampieri

File:Nyquist example.svg *Source*: https://en.wikipedia.org/w/index.php?title=File:Nyquist_example.svg *License*: Public Domain *Contributors*: Utkarsh Upadhyay,--Krishnavedala (talk) 20:21, 3 June 2011 (UTC) (python version)

Image:RL&ZARL-(1 3)-(1 3 5 1).png *Source*: https://en.wikipedia.org/w/index.php?title=File:RL&ZARL-(1_3)-(1_3_5_1).png *License*: Public Domain *Contributors*: Mintz l

File:General closed loop feedback system.svg *Source*: https://en.wikipedia.org/w/index.php?title=File:General_closed_loop_feedback_system.svg *License*: Public Domain *Contributors*: Mintz l

Image:Bode High-Pass.PNG *Source*: https://en.wikipedia.org/w/index.php?title=File:Bode_High-Pass.PNG *License*: Creative Commons Attribution-Share Alike *Contributors*: Brews ohare

Image:Bode Low-Pass.PNG *Source*: https://en.wikipedia.org/w/index.php?title=File:Bode_Low-Pass.PNG *License*: Creative Commons Attribution-Share Alike *Contributors*: Brews ohare

Image:Bode Low Pass Magnitude Plot.PNG *Source*: https://en.wikipedia.org/w/index.php?title=File:Bode_Low_Pass_Magnitude_Plot.PNG *License*: Creative Commons Attribution-Share Alike *Contributors*: Brews ohare

Image:Bode Low Pass Phase Plot.PNG *Source*: https://en.wikipedia.org/w/index.php?title=File:Bode_Low_Pass_Phase_Plot.PNG *License*: Creative Commons Attribution-Share Alike *Contributors*: Brews ohare

Image:Bode Pole-Zero Magnitude Plot.PNG *Source*: https://en.wikipedia.org/w/index.php?title=File:Bode_Pole-Zero_Magnitude_Plot.PNG *License*: Creative Commons Attribution-Share Alike *Contributors*: Brews ohare

Image:Bode Pole-Zero Phase Plot.PNG *Source*: https://en.wikipedia.org/w/index.php?title=File:Bode_Pole-Zero_Phase_Plot.PNG *License*: Creative Commons Attribution-Share Alike *Contributors*: Brews ohare

Image:Magnitude of feedback amplifier.PNG *Source*: https://en.wikipedia.org/w/index.php?title=File:Magnitude_of_feedback_amplifier.PNG *License*: Creative Commons Attribution-Share Alike *Contributors*: Brews ohare

Image:Phase of feedback amplifier.PNG *Source*: https://en.wikipedia.org/w/index.php?title=File:Phase_of_feedback_amplifier.PNG *License*: Creative Commons Attribution-Share Alike *Contributors*: Brews ohare

Image:Gain Margin.PNG *Source*: https://en.wikipedia.org/w/index.php?title=File:Gain_Margin.PNG *License*: Creative Commons Attribution-Share Alike *Contributors*: Brews ohare

Image:Phase Margin.PNG *Source*: https://en.wikipedia.org/w/index.php?title=File:Phase_Margin.PNG *License*: Creative Commons Attribution-Share Alike *Contributors*: Brews ohare

Image:Bodeplot.png *Source*: https://en.wikipedia.org/w/index.php?title=File:Bodeplot.png *License*: Public Domain *Contributors*: derivative work: Stw (talk) Bodeplot.JPG: Vukg

Image:Nyquist.svg *Source*: https://en.wikipedia.org/w/index.php?title=File:Nyquist.svg *License*: Public Domain *Contributors*: Engelec

Image:Nichols.svg *Source*: https://en.wikipedia.org/w/index.php?title=File:Nichols.svg *License*: Public domain *Contributors*: Engelec (talk)

Image:Commons-logo.svg *Source*: https://en.wikipedia.org/w/index.php?title=File:Commons-logo.svg *License*: logo *Contributors*: Anomie

File:Nichols.svg *Source*: https://en.wikipedia.org/w/index.php?title=File:Nichols.svg *License*: Public domain *Contributors*: Engelec (talk)

Image:argument principle1.svg *Source*: https://en.wikipedia.org/w/index.php?title=File:Argument_principle1.svg *License*: Public Domain *Contributors*: Archibald

License

Creative Commons Attribution-Share Alike 3.0
//creativecommons.org/licenses/by-sa/3.0/