

The Laplace Transform

The two main techniques in signal processing, convolution and Fourier analysis, teach that a linear system can be completely understood from its impulse or frequency response. This is a very generalized approach, since the impulse and frequency responses can be of nearly any shape or form. In fact, it is *too* general for many applications in science and engineering. Many of the parameters in our universe interact through *differential equations*. For example, the voltage across an inductor is proportional to the derivative of the current through the device. Likewise, the force applied to a mass is proportional to the derivative of its velocity. Physics is filled with these kinds of relations. The frequency and impulse responses of these systems cannot be arbitrary, but must be consistent with the solution of these differential equations. This means that their impulse responses can only consist of *exponentials* and *sinusoids*. The Laplace transform is a technique for analyzing these special systems when the signals are *continuous*. The z-transform is a similar technique used in the *discrete* case.

The Nature of the s-Domain

The Laplace transform is a well established mathematical technique for solving differential equations. It is named in honor of the great French mathematician, Pierre Simon De Laplace (1749-1827). Like all transforms, the Laplace transform changes one signal into another according to some fixed set of rules or equations. As illustrated in Fig. 32-1, the Laplace transform changes a signal in the time domain into a signal in the **s-domain**, also called the **s-plane**. The time domain signal is continuous, extends to both positive and negative infinity, and may be either periodic or aperiodic. The Laplace transform allows the time domain to be *complex*; however, this is seldom needed in signal processing. In this discussion, and nearly all practical applications, the time domain signal is completely real.

As shown in Fig. 32-1, the s-domain is a complex plane, i.e., there are real numbers along the horizontal axis and imaginary numbers along the vertical axis. The distance along the real axis is expressed by the variable, σ , a lower

case Greek sigma. Likewise, the imaginary axis uses the variable, ω , the natural frequency. This coordinate system allows the location of any point to be specified by providing values for σ and ω . Using complex notation, each location is represented by the complex variable, s , where: $s = \sigma + j\omega$. Just as with the Fourier transform, signals in the s-domain are represented by capital letters. For example, a time domain signal, $x(t)$, is transformed into an s-domain signal, $X(s)$, or alternatively, $X(\sigma, \omega)$. The s-plane is continuous, and extends to infinity in all four directions.

In addition to having a *location* defined by a complex number, each point in the s-domain has a *value* that is a complex number. In other words, each location in the s-plane has a real part and an imaginary part. As with all complex numbers, the real & imaginary parts can alternatively be expressed as the magnitude & phase.

Just as the Fourier transform analyzes signals in terms of sinusoids, the Laplace transform analyzes signals in terms of sinusoids *and* exponentials. From a mathematical standpoint, this makes the Fourier transform a *subset* of the more elaborate Laplace transform. Figure 32-1 shows a graphical description of how the s-domain is related to the time domain. To find the values along a vertical line in the s-plane (the values at a particular σ), the time domain signal is first multiplied by the exponential curve: $e^{-\sigma t}$. The left half of the s-plane multiplies the time domain with exponentials that *increase* with time ($\sigma < 0$), while in the right half the exponentials *decrease* with time ($\sigma > 0$). Next, take the complex Fourier transform of the exponentially weighted signal. The resulting spectrum is placed along a vertical line in the s-plane, with the top half of the s-plane containing the positive frequencies and the bottom half containing the negative frequencies. Take special note that the values on the y-axis of the s-plane ($\sigma = 0$) are exactly equal to the Fourier transform of the time domain signal.

As discussed in the last chapter, the complex Fourier Transform is given by:

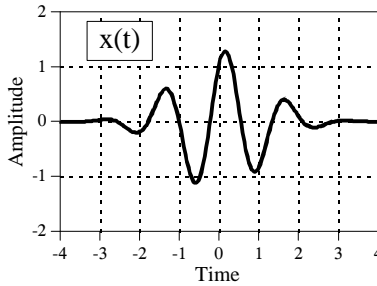
$$X(\omega) = \int_{-\infty}^{\infty} x(t) e^{-j\omega t} dt$$

This can be expanded into the Laplace transform by first multiplying the time domain signal by the exponential term:

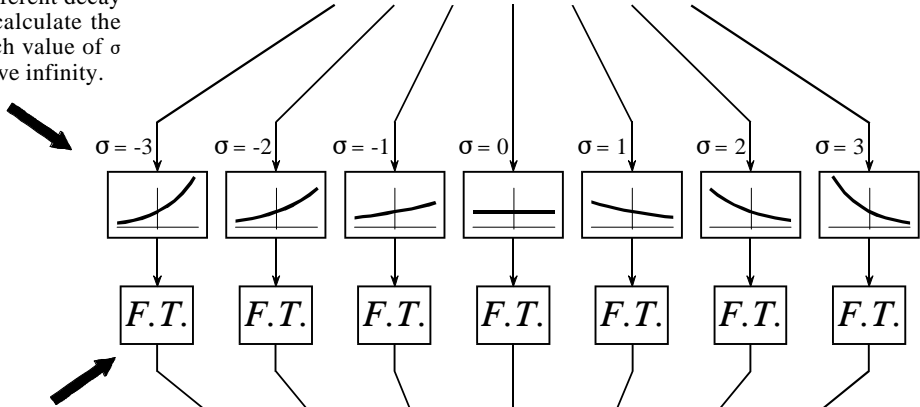
$$X(\sigma, \omega) = \int_{-\infty}^{\infty} [x(t) e^{-\sigma t}] e^{-j\omega t} dt$$

While this is not the simplest form of the Laplace transform, it is probably the best description of the strategy and operation of the technique. To

STEP 1
Start with the time domain signal called $x(t)$



STEP 2
Multiply the time domain signal by an infinite number of exponential curves, each with a different decay constant, σ . That is, calculate the signal: $x(t) e^{-\sigma t}$ for each value of σ from negative to positive infinity.



STEP 3
Take the complex Fourier Transform of each exponentially weighted time domain signal. That is, calculate:

$$\int_{-\infty}^{\infty} [x(t) e^{-\sigma t}] e^{-j\omega t} dt$$

for each value of σ from negative to positive infinity.

STEP 4
Arrange each spectrum along a vertical line in the s -plane. The positive frequencies are in the upper half of the s -plane while the negative frequencies are in the lower half.

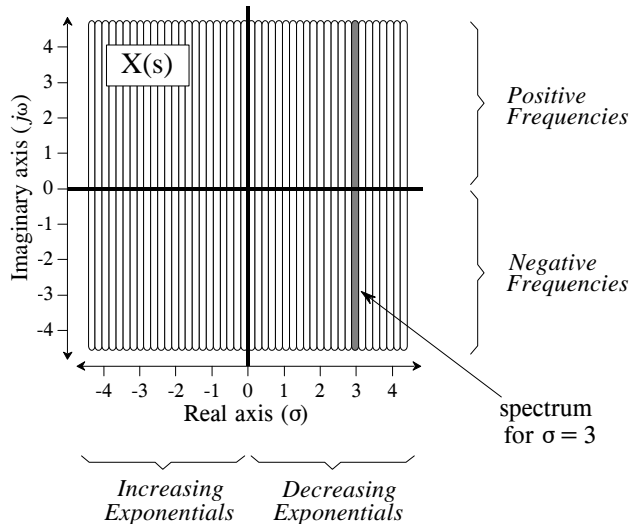


FIGURE 32-1

The Laplace transform. The Laplace transform converts a signal in the *time domain*, $x(t)$, into a signal in the *s-domain*, $X(s)$ or $X(\sigma, \omega)$. The values along each vertical line in the *s-domain* can be found by multiplying the time domain signal by an exponential curve with a decay constant σ , and taking the complex Fourier transform. When the time domain is entirely real, the upper half of the *s-plane* is a mirror image of the lower half.

place the equation in a shorter form, the two exponential terms can be combined:

$$X(\sigma, \omega) = \int_{-\infty}^{\infty} x(t) e^{-(\sigma + j\omega)t} dt$$

Finally, the *location* in the complex plane can be represented by the complex variable, s , where $s = \sigma + j\omega$. This allows the equation to be reduced to an even more compact expression:

EQUATION 32-1

The Laplace transform. This equation defines how a time domain signal, $x(t)$, is related to an s-domain signal, $X(s)$. The s-domain variables, s , and $X(\)$, are complex. While the time domain *may* be complex, it is usually real.

$$X(s) = \int_{-\infty}^{\infty} x(t) e^{-st} dt$$

This is the final form of the Laplace transform, one of the most important equations in signal processing and electronics. Pay special attention to the term: e^{-st} , called a *complex exponential*. As shown by the above derivation, complex exponentials are a compact way of representing both sinusoids and exponentials in a single expression.

Although we have explained the Laplace transform as a two stage process (multiplication by an exponential curve followed by the Fourier transform), keep in mind that this is only a teaching aid, a way of breaking Eq. 32-1 into simpler components. The Laplace transform is a single equation relating $x(t)$ and $X(s)$, not a step-by-step procedure. Equation 32-1 describes how to calculate each *point* in the s-plane (identified by its values for σ and ω) based on the values of σ , ω , and the time domain signal, $x(t)$. Using the Fourier transform to *simultaneously* calculate all the points along a vertical line is merely a convenience, not a requirement. However, it is very important to remember that the values in the s-plane along the y-axis ($\sigma = 0$) are *exactly* equal to the Fourier transform. As explained later in this chapter, this is a key part of why the Laplace transform is useful.

To explore the nature of Eq. 32-1 further, let's look at several individual points in the s-domain and examine how the values at these locations are related to the time domain signal. To start, recall how individual points in the *frequency domain* are related to the time domain signal. Each point in the frequency domain, identified by a specific value of ω , corresponds to two sinusoids, $\cos(\omega t)$ and $\sin(\omega t)$. The real part is found by multiplying the time domain signal by the cosine wave, and then integrating from $-\infty$ to ∞ . The imaginary part is found in the same way, except the sine wave is used. If we are dealing with the *complex* Fourier transform, the values at the corresponding negative frequency, $-\omega$, will be the complex conjugate (same real part, negative imaginary part) of the values at ω . The Laplace transform is just an extension of these same concepts.

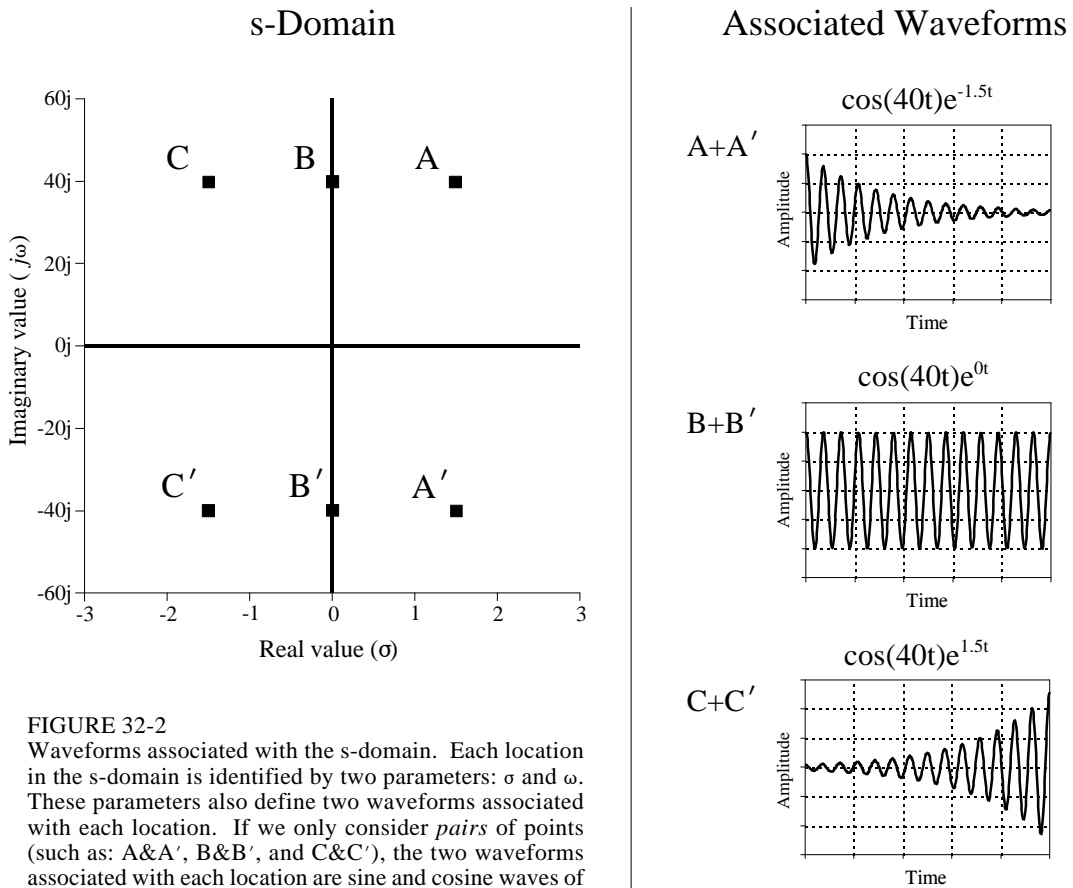


FIGURE 32-2

Waveforms associated with the s-domain. Each location in the s-domain is identified by two parameters: σ and ω . These parameters also define two waveforms associated with each location. If we only consider *pairs* of points (such as: A&A', B&B', and C&C'), the two waveforms associated with each location are sine and cosine waves of frequency ω , with an exponentially changing amplitude controlled by σ .

Figure 32-2 shows three *pairs* of points in the s-plane: A&A', B&B', and C&C'. Just as in the complex frequency spectrum, the points at A, B, & C (the positive frequencies) are the complex conjugates of the points at A', B', & C' (the negative frequencies). The top half of the s-plane is a mirror image of the lower half, and both halves are needed to correspond with a real time domain signal. In other words, treating these points in pairs bypasses the complex math, allowing us to operate in the time domain with only real numbers.

Since each of these pairs has specific values for σ and $\pm\omega$, there are two waveforms associated with each pair: $\cos(\omega t)e^{-\sigma t}$ and $\sin(\omega t)e^{-\sigma t}$. For instance, points A&A' are at a location of $\sigma = 1.5$ and $\omega = \pm 40$, and therefore correspond to the waveforms: $\cos(40t)e^{-1.5t}$ and $\sin(40t)e^{-1.5t}$. As shown in Fig. 32-2, these are sinusoids that exponentially *decreases* in amplitude as time progresses. In this same way, the sine and cosine waves associated with B&B' have a *constant* amplitude, resulting from the value of σ being zero. Likewise, the sine and cosine waves that are associated with locations C&C' exponentially *increases* in amplitude, since σ is negative.

The value at each location in the s-plane consists of a *real part* and an *imaginary part*. The real part is found by multiplying the time domain signal by the exponentially weighted cosine wave and then integrated from $-\infty$ to ∞ . The imaginary part is found in the same way, except the exponentially weighted sine wave is used instead. It looks like this in equation form, using the real part of A&A' as an example:

$$ReX(\sigma=1.5, \omega=\pm 40) = \int_{-\infty}^{\infty} x(t) \cos(40t) e^{-1.5t} dt$$

Figure 32-3 shows an example of a time domain waveform, its frequency spectrum, and its s-domain representation. The example time domain signal is a rectangular pulse of width *two* and height *one*. As shown, the complex Fourier transform of this signal is a sinc function in the real part, and an entirely zero signal in the imaginary part. The s-domain is an undulating two-dimensional signal, displayed here as topographical surfaces of the real and imaginary parts. The mathematics works like this:

$$X(s) = \int_{-\infty}^{\infty} x(t) e^{-st} dt = \int_{-1}^1 1 e^{-st} dt$$

In words, we start with the definition of the Laplace transform (Eq. 32-1), plug in the unity value for $x(t)$, and change the limits to match the length of the nonzero portion of the time domain signal. Evaluating this integral provides the s-domain signal, expressed in terms of the complex location, s , and the complex value, $X(s)$:

$$X(s) = \frac{e^s - e^{-s}}{s}$$

While this is the most compact form of the answer, the use of complex variables makes it difficult to understand, and impossible to generate a visual display, such as Fig. 32-3. The solution is to replace the complex variable, s , with $\sigma + j\omega$, and then separate the real and imaginary parts:

$$ReX(\sigma, \omega) = \frac{\sigma \cos(\omega) [e^\sigma - e^{-\sigma}] + \omega \sin(\omega) [e^\sigma + e^{-\sigma}]}{\sigma^2 + \omega^2}$$

$$ImX(\sigma, \omega) = \frac{\sigma \sin(\omega) [e^\sigma + e^{-\sigma}] - \omega \cos(\omega) [e^\sigma - e^{-\sigma}]}{\sigma^2 + \omega^2}$$

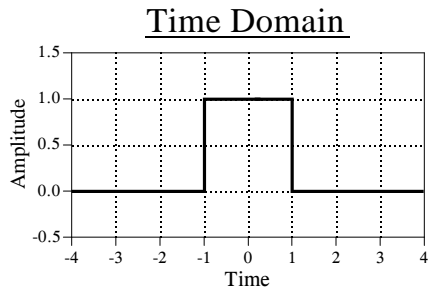
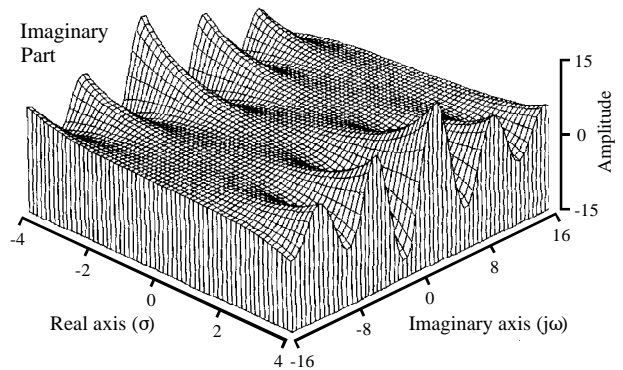
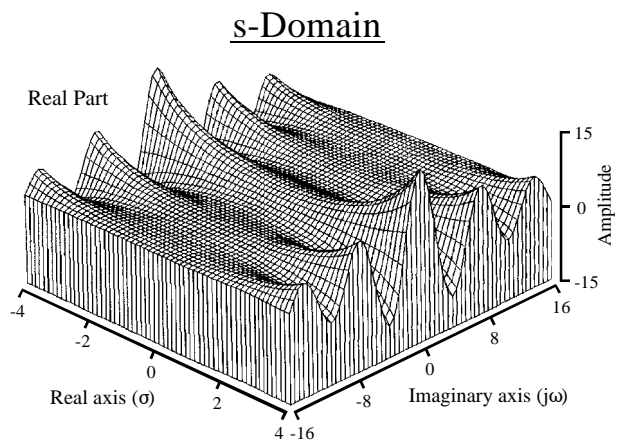
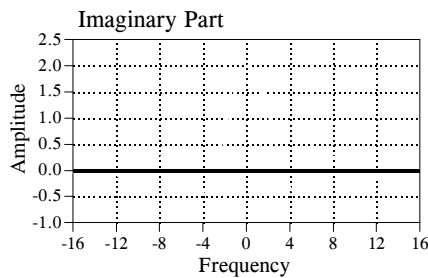
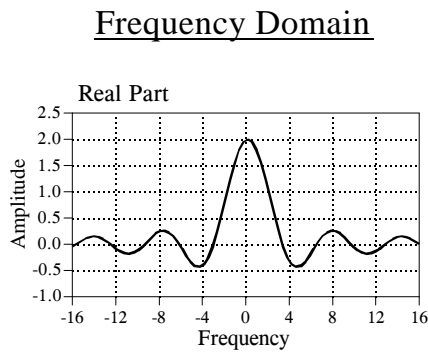
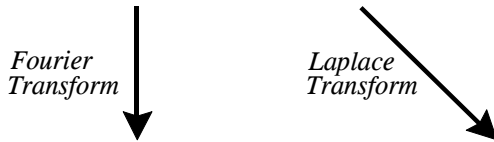


FIGURE 32-3
Time, frequency and s-domains. A time domain signal (the rectangular pulse) is transformed into the frequency domain using the Fourier transform, and into the s-domain using the Laplace transform.



The topographical surfaces in Fig. 32-3 are graphs of these equations. These equations are quite long and the mathematics to derive them is very tedious. This brings up a practical issue: with algebra of this complexity, how do we know that we haven't made an error in the calculations? One check is to verify

that these equations reduce to the *Fourier transform* along the y-axis. This is done by setting σ to zero in the equations, and simplifying:

$$\left. \operatorname{Re} X(\sigma, \omega) \right|_{\sigma=0} = \frac{2 \sin(\omega)}{\omega} \qquad \left. \operatorname{Im} X(\sigma, \omega) \right|_{\sigma=0} = 0$$

As illustrated in Fig. 32-3, these are the correct frequency domain signals, the same as found by directly taking the Fourier transform of the time domain waveform.

Strategy of the Laplace Transform

An analogy will help in explaining how the Laplace transform is used in signal processing. Imagine you are traveling by train at night between two cities. Your map indicates that the path is very straight, but the night is so dark you cannot see any of the surrounding countryside. With nothing better to do, you notice an altimeter on the wall of the passenger car and decide to keep track of the elevation changes along the route.

Being bored after a few hours, you strike up a conversation with the conductor: "Interesting terrain," you say. "It seems we are generally increasing in elevation, but there are a few interesting irregularities that I have observed." Ignoring the conductor's obvious disinterest, you continue: "Near the start of our journey, we passed through some sort of abrupt rise, followed by an equally abrupt descent. Later we encountered a shallow depression." Thinking you might be dangerous or demented, the conductor decides to respond: "Yes, I guess that is true. Our destination is located at the base of a large mountain range, accounting for the general increase in elevation. However, along the way we pass on the outskirts of a large mountain and through the center of a valley."

Now, think about how you understand the relationship between elevation and distance along the train route, compared to that of the conductor. Since you have directly measured the elevation along the way, you can rightly claim that you know *everything* about the relationship. In comparison, the conductor knows this same complete information, but in a simpler and more intuitive form: the location of the hills and valleys that *cause* the dips and humps along the path. While your description of the signal might consist of thousands of individual measurements, the conductor's description of the signal will contain only a few parameters.

To show how this is analogous to signal processing, imagine we are trying to understand the characteristics of some electric circuit. To aid in our investigation, we carefully measure the impulse response and/or the frequency response. As discussed in previous chapters, the impulse and frequency responses contain *complete* information about this linear system.

However, this does not mean that you know the information in the *simplest* way. In particular, you understand the frequency response as a set of values that change with frequency. Just as in our train analogy, the frequency response can be more easily understood in terms of the terrain *surrounding* the frequency response. That is, by the characteristics of the s-plane.

With the train analogy in mind, look back at Fig. 32-3, and ask: how does the shape of this s-domain aid in understanding the frequency response? The answer is, it doesn't! The s-plane in this example makes a nice graph, but it provides no insight into why the frequency domain behaves as it does. This is because the Laplace transform is designed to analyze a specific class of time domain signals: *impulse responses that consist of sinusoids and exponentials*. If the Laplace transform is taken of some other waveform (such as the rectangular pulse in Fig. 32-3), the resulting s-domain is meaningless.

As mentioned in the introduction, systems that belong to this class are extremely common in science and engineering. This is because sinusoids and exponentials are solutions to *differential equations*, the mathematics that controls much of our physical world. For example, all of the following systems are governed by differential equations: electric circuits, wave propagation, linear and rotational motion, electric and magnetic fields, heat flow, etc.

Imagine we are trying to understand some linear system that is controlled by differential equations, such as an electric circuit. Solving the differential equations provides a mathematical way to find the impulse response. Alternatively, we could measure the impulse response using suitable pulse generators, oscilloscopes, data recorders, etc. Before we inspect the newly found impulse response, we ask ourselves what we *expect* to find. There are several characteristics of the waveform that we know without even looking. First, the impulse response must be *causal*. In other words, the impulse response must have a value of zero until the input becomes nonzero at $t = 0$. This is the cause and effect that our universe is based upon.

The second thing we know about the impulse response is that it will be composed of *sinusoids and exponentials*, because these are the solutions to the differential equations that govern the system. Try as we might, we will never find this type of system having an impulse response that is, for example, a square pulse or triangular waveform. Third, the impulse response will be *infinite* in length. That is, it has nonzero values that extend from $t = 0$ to $t = +\infty$. This is because sine and cosine waves have a constant amplitude, and exponentials decay toward zero without ever actually reaching it. If the system we are investigating is **stable**, the amplitude of the impulse response will become smaller as time increases, reaching a value of zero at $t = +\infty$. There is also the possibility that the system is **unstable**, for example, an amplifier that spontaneously oscillates due to an excessive amount of feedback. In this case, the impulse response will *increase* in amplitude as time increases, becoming infinitely large. Even the smallest disturbance to this system will produce an unbounded output.

The general mathematics of the Laplace transform is very similar to that of the Fourier transform. In both cases, predetermined waveforms are multiplied by the time domain signal, and the result integrated. At first glance, it would appear that the strategy of the Laplace transform is the same as the Fourier transform: correlate the time domain signal with a set of basis functions to decompose the waveform. Not true! Even though the mathematics is much the same, the rationale behind the two techniques is very different. The Laplace transform *probes* the time domain waveform to identify its key features: the *frequencies* of the sinusoids, and the *decay constants* of the exponentials. An example will show how this works.

The center column in Fig. 32-5 shows the impulse response of the *RLC* notch filter discussed in Chapter 30. It contains an impulse at $t = 0$, followed by an exponentially decaying sinusoid. As illustrated in (a) through (e), we will *probe* this impulse response with various exponentially decaying sinusoids. Each of these probing waveforms is characterized by two parameters: ω , that determines the sinusoidal frequency, and σ , that determines the decay rate. In other words, each probing waveform corresponds to a different location in the *s*-plane, as shown by the *s*-plane diagram in Fig. 32-4. The impulse response is probed by *multiplying* it with these waveforms, and then integrating the result from $t = -\infty$ to $+\infty$. This action is shown in the right column. Our goal is to find combinations of σ and ω that exactly *cancel* the impulse response being investigated. This cancellation can occur in two forms: the area under the curve can be either *zero*, or just *barely infinite*. All other results are uninteresting and can be ignored. Locations in the *s*-plane that produce a zero cancellation are called **zeros** of the system. Likewise, locations that produce the "just barely infinite" type of cancellation are called **poles**. Poles and zeros are analogous to the mountains and valleys in our train story, representing the terrain "around" the frequency response.

To start, consider what happens when the probing waveform decreases in amplitude as time advances, as shown in (a). This will occur whenever $\sigma > 0$ (the right half of the *s*-plane). Since both the impulse response and the probe becomes smaller with increasing time, the product of the two will also have this same characteristic. When the product of the two waveforms is integrated from negative to positive infinity, the result will be some number that is not especially interesting. In particular, a decreasing probe

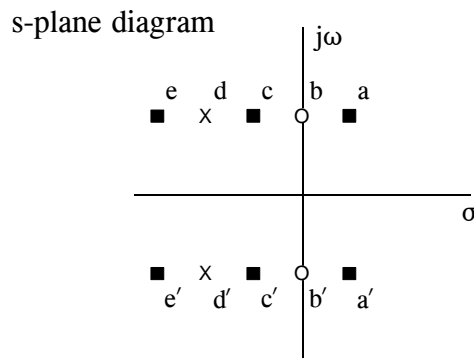


FIGURE 32-4 Pole-zero example. The notch filter has two poles (represented by \times) and two zeros (represented by \circ). This *s*-plane diagram shows the five locations we will "probe" in this example to analyze this system. (Figure 30-5 is a continuation of this example).

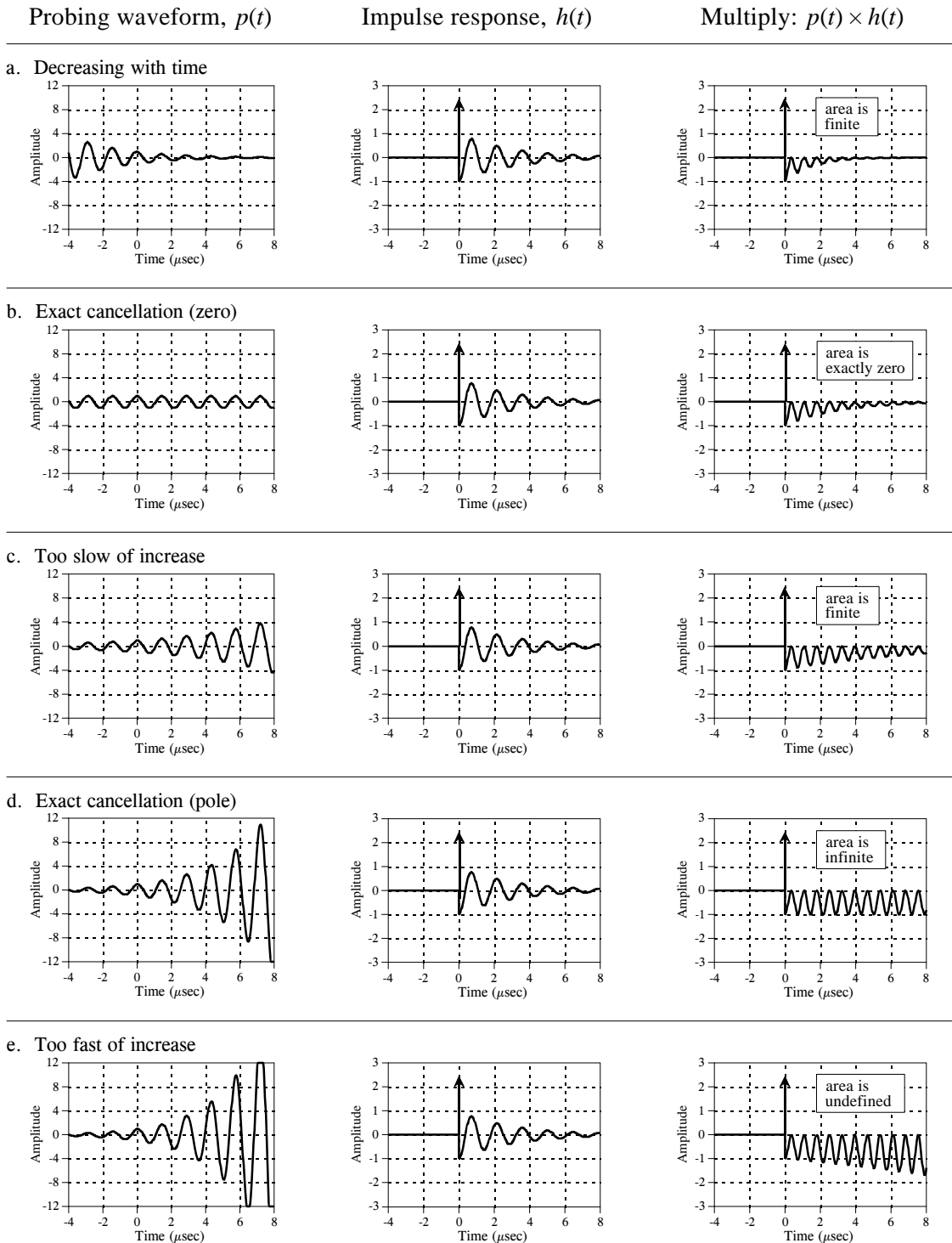


FIGURE 32-5

Probing the impulse response. The Laplace transform can be viewed as probing the system's impulse response with various exponentially decaying sinusoids. Probing waveforms that produce a cancellation are called *poles* and *zeros*. This illustration shows five probing waveforms (left column) being applied to the impulse response of a notch filter (center column). The locations in the s -plane that correspond to these five waveforms are shown in Fig. 32-4.

cannot cancel a decreasing impulse response. This means that a stable system will not have any poles with $\sigma > 0$. In other words, all of the poles in a stable system are confined to the left half of the s-plane. In fact, poles in the right half of the s-plane show that the system is unstable (i.e., an impulse response that *increases* with time).

Figure (b) shows one of the special cases we have been looking for. When this waveform is multiplied by the impulse response, the resulting integral has a value of zero. This occurs because the area above the x-axis (from the delta function) is exactly equal to the area below (from the rectified sinusoid). The values for σ and ω that produce this type of cancellation are called a *zero* of the system. As shown in the s-plane diagram of Fig. 32-4, zeros are indicated by small circles (\circ).

Figure (c) shows the next probe we can try. Here we are using a sinusoid that exponentially *increases* with time, but at a rate *slower* than the impulse response is *decreasing* with time. This results in the product of the two waveforms also decreasing as time advances. As in (a), this makes the integral of the product some uninteresting real number. The important point being that no type of exact cancellation occurs.

Jumping out of order, look at (e), a probing waveform that increases at a *faster* rate than the impulse response decays. When multiplied, the resulting signal increases in amplitude as time advances. This means that the area under the curve becomes larger with increasing time, and the total area from $t = -\infty$ to $+\infty$ is not defined. In mathematical jargon, the *integral does not converge*. In other words, not all areas of the s-plane have a defined value. The portion of the s-plane where the integral is defined is called the **region-of-convergence**. In some mathematical techniques it is important to know what portions of the s-plane are within the region-of-convergence. However, this information is not needed for the applications in this book. Only the exact cancellations are of interest for this discussion.

In (d), the probing waveform increases at exactly the same rate that the impulse response decreases. This makes the product of the two waveforms have a constant amplitude. In other words, this is the dividing line between (c) and (e), resulting in a total area that is just *barely undefined* (if the mathematicians will forgive this loose description). In more exact terms, this point is on the borderline of the region of convergence. As mentioned, values for σ and ω that produce this type of exact cancellation are called *poles* of the system. Poles are indicated in the s-plane by crosses (\times).

Analysis of Electric Circuits

We have introduced the Laplace transform in graphical terms, describing what the waveforms look like and how they are manipulated. This is the most intuitive way of understanding the approach, but is very different from how it is actually used. The Laplace transform is inherently a mathematical technique; it is used by writing and manipulating *equations*. The problem is, it is easy to

become lost in the abstract nature of the complex algebra and lose all connection to the real world. Your task is to merge the two views together. The Laplace transform is the primary method for analyzing electric circuits. Keep in mind that *any* system governed by differential equations can be handled the same way; electric circuits are just an example we are using.

The brute force approach is to solve the differential equations controlling the system, providing the system's impulse response. The impulse response can then be converted into the s -domain via Eq. 32-1. Fortunately, there is a better way: transform each of the individual components into the s -domain, and *then* account for how they interact. This is very similar to the *phasor transform* presented in Chapter 30, where resistors, inductors and capacitors are represented by R , $j\omega L$, and $1/j\omega C$, respectively. In the Laplace transform, resistors, inductors and capacitors become the complex variables: R , sL , and $1/sC$. Notice that the phasor transform is a *subset* of the Laplace transform. That is, when σ is set to zero in $s = \sigma + j\omega$, R becomes R , sL becomes $j\omega L$, and $1/sC$ becomes $1/j\omega C$.

Just as in Chapter 30, we will treat each of the three components as an individual system, with the current waveform being the input signal, and the voltage waveform being the output signal. When we say that resistors, inductors and capacitors become R , sL , and $1/sC$ in the s -domain, this refers to the output divided by the input. In other words, the Laplace transform of the *voltage waveform* divided by the Laplace transform of the *current waveform* is equal to these expressions.

As an example of this, imagine we force the current through an inductor to be a unity amplitude cosine wave with a frequency given by ω_0 . The resulting *voltage* waveform across the inductor can be found by solving the differential equation that governs its operation:

$$v(t) = L \frac{d}{dt} i(t) = L \frac{d}{dt} \cos(\omega_0 t) = -\omega_0 L \sin(\omega_0 t)$$

If we start the current waveform at $t = 0$, the voltage waveform will also start at this same time (i.e., $i(t) = 0$ and $v(t) = 0$ for $t < 0$). These voltage and current waveforms are converted into the s -domain by Eq. 32-1:

$$I(s) = \int_0^{\infty} \cos(\omega_0 t) e^{-st} dt = \frac{\omega_0}{\omega_0^2 + s^2}$$

$$V(s) = \int_0^{\infty} -\omega_0 L \sin(\omega_0 t) e^{-st} dt = -\frac{\omega_0 L s}{\omega_0^2 + s^2}$$

To complete this example, we will divide the s-domain voltage by the s-domain current, just as if we were using Ohm's law ($R = V/I$):

$$\frac{V(s)}{I(s)} = \frac{\frac{\omega_0 L s}{\omega_0^2 + s^2}}{\frac{\omega_0}{\omega_0^2 + s^2}} = sL$$

We find that the s-domain representation of the voltage across the inductor, divided by the s-domain representation of the current through the inductor, is equal to sL . This is *always* the case, regardless of the current waveform we start with. In a similar way, the ratio of s-domain voltage to s-domain current is always equal to R for resistors, and $1/sC$ for capacitors.

Figure 32-6 shows an example circuit we will analyze with the Laplace transform, the RLC notch filter discussed in Chapter 30. Since this analysis is the same for all electric circuits, we will outline it in steps.

Step 1. Transform each of the components into the s-domain. In other words, replace the value of each resistor with R , each inductor with sL , and each capacitor with $1/sC$. This is shown in Fig. 32-6.

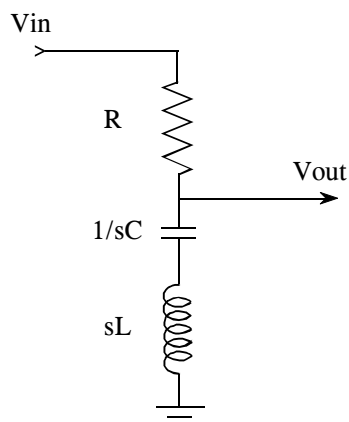
Step 2: Find $H(s)$, the output divided by the input. As described in Chapter 30, this is done by treating each of the components as if they obey Ohm's law, with the "resistances" given by: R , sL , and $1/sC$. This allows us to use the standard equations for resistors in series, resistors in parallel, voltage dividers, etc. Treating the RLC circuit in this example as a voltage divider (just as in Chapter 30), $H(s)$ is found:

$$H(s) = \frac{V_{out}(s)}{V_{in}(s)} = \frac{sL + 1/sC}{R + sL + 1/sC} = \frac{sL + 1/sC}{R + sL + 1/sC} \left[\frac{s}{s} \right] = \frac{Ls^2 + 1/C}{Ls^2 + Rs + 1/C}$$

As you recall from *Fourier* analysis, the frequency spectrum of the output signal divided by the frequency spectrum of the input signal is equal to the system's *frequency response*, given the symbol, $H(\omega)$. The above equation is an extension of this into the s-domain. The signal, $H(s)$, is called the system's **transfer function**, and is equal to the s-domain representation of the output signal divided by the s-domain representation of the input signal. Further, $H(s)$ is equal to the *Laplace transform* of the impulse response, just the same as $H(\omega)$ is equal to the *Fourier transform* of the impulse response.

FIGURE 32-6

Notch filter analysis in the s-domain. The first step in this procedure is to replace the resistor, inductor & capacitor values with their s-domain equivalents.



So far, this is identical to the techniques of the last chapter, except for using s instead of $j\omega$. The difference between the two methods is what happens from this point on. This is as far as we can go with $j\omega$. We might graph the frequency response, or examining it in some other way; however, this is a mathematical dead end. In comparison, the interesting aspects of the Laplace transform have just begun. Finding $H(s)$ is the key to Laplace analysis; however, it must be expressed in a particular form to be useful. This requires the algebraic manipulation of the next two steps.

Step 3: Arrange $H(s)$ to be one polynomial over another. This makes the transfer function written as:

EQUATION 32-2

Transfer function in polynomial form.

$$H(s) = \frac{as^2 + bs + c}{as^2 + bs + c}$$

It is always possible to express the transfer function in this form *if* the system is controlled by differential equations. For example, the rectangular pulse shown in Fig. 32-3 is not the solution to a differential equation and its Laplace transform cannot be written in this way. In comparison, any electric circuit composed of resistors, capacitors, and inductors can be written in this form. For the *RLC* notch filter used in this example, the algebra shown in *step 2* has already placed the transfer function in the correct form, that is:

$$H(s) = \frac{as^2 + bs + c}{as^2 + bs + c} = \frac{Ls^2 + 1/C}{Ls^2 + Rs + 1/C}$$

where: $a = L$, $b = 0$, $c = 1/C$; and $a = L$, $b = R$, $c = 1/C$

Step 4: Factor the numerator and denominator polynomials. That is, break the numerator and denominator polynomials into components that each contain

a single s . When the components are multiplied together, they must equal the original numerator and denominator. In other words, the equation is placed into the form:

EQUATION 32-3

The factored s-domain. This form allows the s-domain to be expressed as poles and zeros.

$$H(s) = \frac{(s - z_1)(s - z_2)(s - z_3)\dots}{(s - p_1)(s - p_2)(s - p_3)\dots}$$

The roots of the numerator, $z_1, z_2, z_3 \dots$, are the **zeros** of the equation, while the roots of the denominator, $p_1, p_2, p_3 \dots$, are the **poles**. These are the same zeros and poles we encountered earlier in this chapter, and we will discuss how they are used in the next section.

Factoring an s-domain expression is straightforward if the numerator and denominator are *second-order polynomials*, or less. In other words, we can easily handle the terms: s and s^2 , but not: s^3, s^4, s^5, \dots . This is because the roots of a second-order polynomial, $ax^2 + bx + c$, can be found by using the quadratic equation: $x = -b \pm \sqrt{b^2 - 4ac} / 2a$. With this method, the transfer function of the example notch filter is factored into:

$$H(s) = \frac{(s - z_1)(s - z_2)}{(s - p_1)(s - p_2)}$$

where:

$$\begin{aligned} z_1 &= j/\sqrt{LC} & p_1 &= \frac{-R + \sqrt{R^2 - 4L/C}}{2L} \\ z_2 &= -j/\sqrt{LC} & p_2 &= \frac{-R - \sqrt{R^2 - 4L/C}}{2L} \end{aligned}$$

As in this example, a second-order system has a maximum of two zeros and two poles. The number of poles in a system is equal to the number of independent energy storing components. For instance, inductors and capacitors store energy, while resistors do not. The number of zeros will be equal to, or less than, the number of poles.

Polynomials greater than second order cannot generally be factored using algebra, requiring more complicated numerical methods. As an alternative, circuits can be constructed as a *cascade* of *second-order* stages. A good example is the family of analog filters presented in Chapter 3. For instance, an eight pole filter is designed by cascading four stages of two poles each. The important point is that this multistage approach is used to overcome limitations in the *mathematics*, not limitations in the *electronics*.

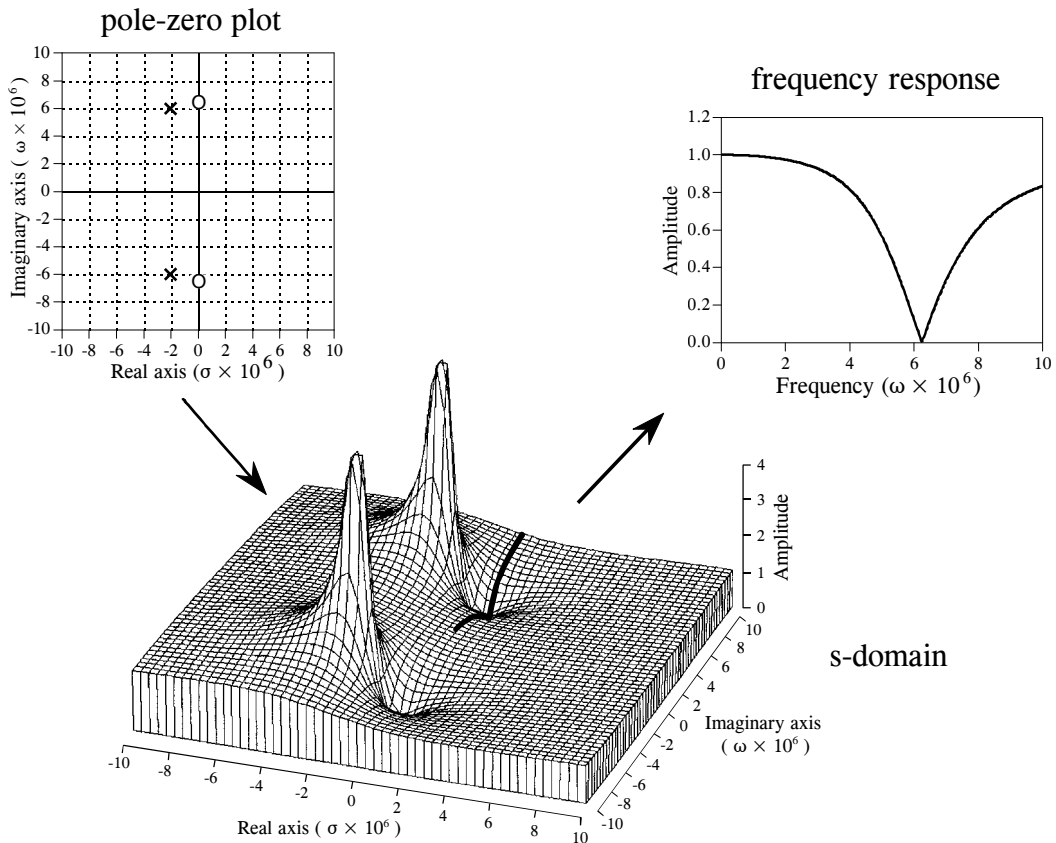


FIGURE 32-7

Poles and zeros in the s -domain. These illustrations show the relationship between the pole-zero plot, the s -domain, and the frequency response. The notch filter component values used in these graphs are: $R=220\ \Omega$, $C=470\ \mu\text{F}$, and $L=54\ \mu\text{H}$. These values place the center of the notch at $\omega = 6.277$ million, i.e., a frequency of approximately 1 MHz.

The Importance of Poles and Zeros

To make this less abstract, we will use actual component values for the notch filter we just analyzed: $R = 220\ \Omega$, $L = 54\ \mu\text{H}$, $C = 470\ \mu\text{F}$. Plugging these values into the above equations, places the poles and zeros at:

$$\begin{aligned} z_1 &= 0 + j 6.277 \times 10^6 & p_1 &= -2.037 \times 10^6 + j 5.937 \times 10^6 \\ z_2 &= 0 - j 6.277 \times 10^6 & p_2 &= -2.037 \times 10^6 - j 5.937 \times 10^6 \end{aligned}$$

These pole and zero locations are shown in Fig. 32-7. Each zero is represented by a circle, while each pole is represented by a cross. This is called a **pole-zero diagram**, and is the most common way that s -domain data are displayed. Figure 32-7 also shows a topographical display of the s -plane. For simplicity, only the magnitude is shown, but don't forget that

there is a corresponding phase. Just as mountains and valleys determine the shape of the surface of the earth, the poles and zeros determine the shape of the s -plane. Unlike mountains and valleys, every pole and zero is exactly the same shape and size as every other pole and zero. The only unique characteristic a pole or zero has is its *location*. Poles and zeros are important because they provide a concise representation of the value at *any point in the s -plane*. That is, we can completely describe the characteristics of the system using only a *few parameters*. In the case of the *RLC* notch filter, we only need to specify four complex parameters to represent the system: z_1, z_2, p_1, p_2 (each consisting of a real and an imaginary part).

To better understand poles and zeros, imagine an ant crawling around the s -plane. At any particular location the ant happens to be (i.e., some value of s), there is a corresponding value of the transfer function, $H(s)$. This value is a complex number that can be expressed as the magnitude & phase, or as the real & imaginary parts. Now, let the ant carry us to one of the zeros in the s -plane. The value we measure for the real and imaginary parts will be *zero* at this location. This can be understood by examining the mathematical equation for $H(s)$ in Eq. 32-3. If the location, s , is equal to any of the zeros, one of the terms in the numerator will be zero. This makes the entire expression equal to zero, regardless of the other values.

Next, our ant journey takes us to one of the poles, where we again measure the value of the real and imaginary parts of $H(s)$. The measured value becomes larger and larger as we come close to the exact location of the pole (hence the name). This can also be understood from Eq. 32-3. If the location, s , is equal to any of the p 's, the denominator will be equal to zero, and the division by zero makes the entire expression infinity large.

Having explored the unique locations, our ant journey now moves randomly throughout the s -plane. The value of $H(s)$ at each location depends entirely on the positioning of the poles and the zeros, because there are *no* other types of features allowed in this strange terrain. If we are near a pole, the value will be large; if we are near a zero, the value will be small.

Equation 32-3 also describes how *multiple* poles and zeros interact to form the s -domain signal. Remember, subtracting two complex numbers provides the *distance* between them in the complex plane. For example, $(s - z_n)$ is the distance between the arbitrary location, s , and the zero located at z_n . Therefore, Eq. 32-3 specifies that the value at each location, s , is equal to the distance to all of the zeros *multiplied*, divided by the distance to all of the poles *multiplied*.

This brings us to the heart of this chapter: how the location of the poles & zeros provides a deeper understanding of the system's *frequency response*. The frequency response is equal to the values of $H(s)$ along the imaginary axis, signified by the dark line in the topographical plot of Fig. 32-7. Imagine our ant starting at the origin and crawling along this path. Near the origin, the distance to the zeros is approximately equal to the distance to the poles. This makes the numerator and denominator in Eq. 32-3

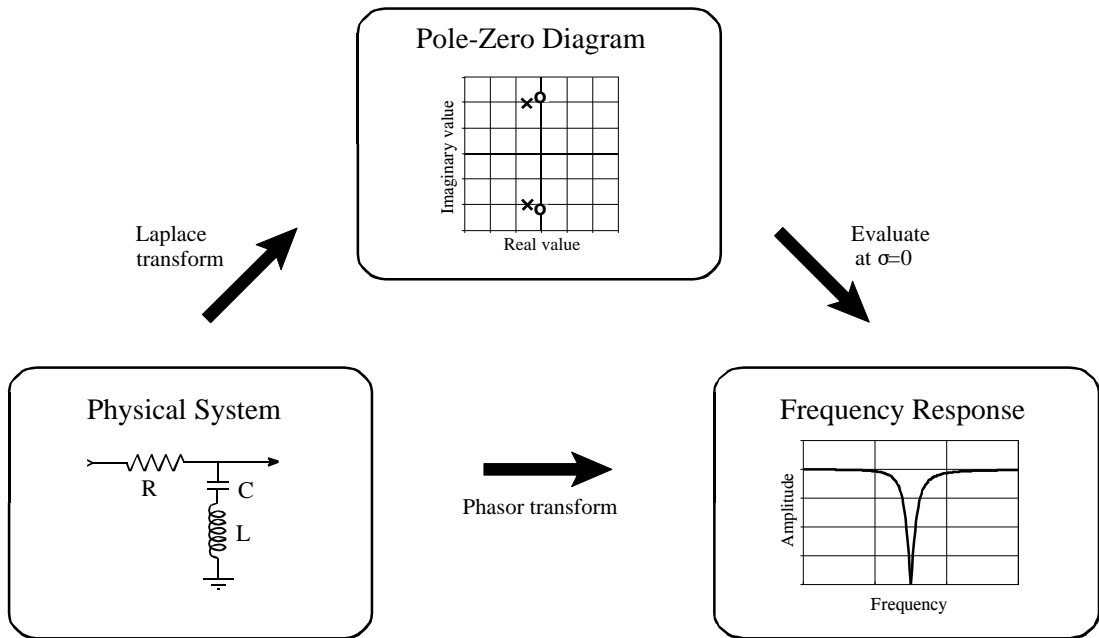


FIGURE 32-8

Strategy for using the Laplace transform. The phasor transform presented in Chapter 30 (the method using R , $j\omega L$, & $-j/\omega C$) allows the frequency response to be directly calculated from the parameters of the physical system. In comparison, the Laplace transform calculates an s -domain representation from the physical system, usually displayed in the form of a pole-zero diagram. In turn, the frequency response can be obtained from the s -domain by evaluating the transfer function along the imaginary axis. While both methods provide the same end result, the intermediate step of the s -domain provides insight into why the frequency response behaves as it does.

cancel, providing a unity frequency response at low frequencies. The situation doesn't change significantly until the ant moves near the pole and zero location. When approaching the zero, the value of $H(s)$ drops suddenly, becoming *zero* when the ant is upon the *zero*. As the ant moves past the pole and zero pair, the value of $H(s)$ again returns to unity. Using this type of visualization, it can be seen that the width of the notch depends on the distance between the pole and zero.

Figure 32-8 summarizes how the Laplace transform is used. We start with a physical system, such as an electric circuit. If we desire, the phasor transform can directly provide the frequency response of the system, as described in Chapter 30. An alternative is to take the Laplace transform using the four step method previously outlined. This results in a mathematical expression for the transfer function, $H(s)$, which can be represented in a pole-zero diagram. The frequency response can then be found by evaluating the transfer function along the imaginary axis, that is, by replacing each s with $j\omega$. While both methods provide the same result, the intermediate pole-zero diagram provides an understanding of *why* the system behaves as it does, and how it can be changed.

Filter Design in the s-Domain

The most powerful application of the Laplace transform is the design of systems *directly* in the s-domain. This involves two steps: First, the s-domain is designed by specifying the number and location of the poles and zeros. This is a pure mathematical problem, with the goal of obtaining the best frequency response. In the second step, an electronic circuit is derived that provides this s-domain representation. This is something of an art, since there are many circuit configurations that have a given pole-zero diagram.

As previously mentioned, *step 4* of the Laplace transform method is very difficult if the system contains more than two poles or two zeros. A common solution is to implement multiple poles and zeros in *successive stages*. For example, a 6 pole filter is implemented as three successive stages, with each stage containing up to two poles and two zeros. Since each of these stages can be represented in the s-domain by a quadratic numerator divided by a quadratic denominator, this approach is called designing with **biquads**.

Figure 32-9 shows a common biquad circuit, the one used in the filter design method of Chapter 3. This is called the **Sallen-Key** circuit, after R.P. Sallen and E.L. Key, authors of a paper that described this technique in the mid 1950s. While there are several variations, the most common circuit uses two resistors of equal value, two capacitors of equal value, and an amplifier with an amplification of between 1 and 3. Although not available to Sallen and Key, the amplifiers can now be made with low-cost op amps with appropriate feedback resistors. Going through the four step circuit analysis procedure, the location of this circuit's two poles can be related to the component values:

EQUATION 32-4
Sallen-Key pole locations. These equations relate the pole position, ω and σ , to the amplifier gain, A , the resistor, R , and capacitor, C .

$$\sigma = \frac{A-3}{2RC}$$

$$\omega = \frac{\pm\sqrt{-A^2 + 6A - 5}}{2RC}$$

These equations show that the two poles always lie somewhere on a circle of radius: $1/RC$. The exact position along the circle depends on the gain of the amplifier. As shown in (a), an amplification of 1 places both of the poles on the real axis. The frequency response of this configuration is a low-pass filter with a relatively smooth transition between the passband and stopband. The -3dB (0.707) cutoff frequency of this circuit, denoted by ω_0 , is where the circle intersects the imaginary axis, i.e., $\omega_0 = 1/RC$.

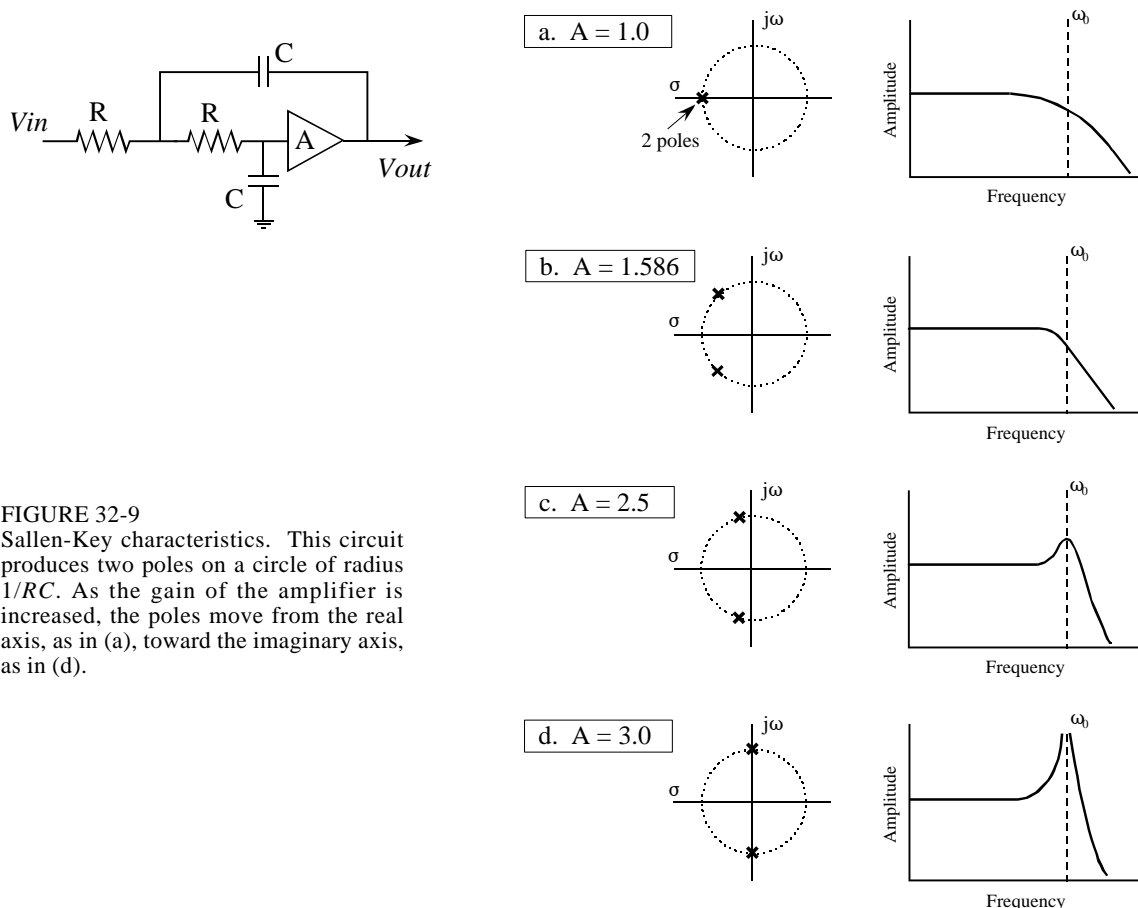


FIGURE 32-9 Sallen-Key characteristics. This circuit produces two poles on a circle of radius $1/RC$. As the gain of the amplifier is increased, the poles move from the real axis, as in (a), toward the imaginary axis, as in (d).

As the amplification is increased, the poles move along the circle, with a corresponding change in the frequency response. As shown in (b), an amplification of 1.586 places the poles at 45 degree angles, resulting in the frequency response having a sharper transition. Increasing the amplification further moves the poles even closer to the imaginary axis, resulting in the frequency response showing a peaked curve. This condition is illustrated in (c), where the amplification is set at 2.5. The amplitude of the peak continues to grow as the amplification is increased, until a gain of 3 is reached. As shown in (d), this is a special case that places the poles directly on the imaginary axis. The corresponding frequency response now has an infinity large value at the peak. In practical terms, this means the circuit has turned into an oscillator. Increasing the gain further pushes the poles deeper into the right half of the s-plane. As mentioned before, this correspond to the system being unstable (spontaneous oscillation).

Using the Sallen-Key circuit as a building block, a wide variety of filter types can be constructed. For example, a low-pass **Butterworth filter** is designed by placing a selected number of poles evenly around the left-half of the circle, as shown in Fig. 32-10. Each two poles in this configuration requires one

Sallen-Key stage. As described in Chapter 3, the Butterworth filter is maximally flat, that is, it has the sharpest transition between the passband and stopband *without peaking* in the frequency response. The more poles used, the faster the transition. Since all the poles in the Butterworth filter lie on the same circle, all the cascaded stages use the same values for R and C . The only thing different between the stages is the amplification. Why does this circular pattern of poles provide the optimally flat response? Don't look for an obvious or intuitive answer to this question; it just falls out of the mathematics.

Figure 32-11 shows how the pole positions of the Butterworth filter can be modified to produce the **Chebyshev filter**. As discussed in Chapter 3, the Chebyshev filter achieves a sharper transition than the Butterworth at the expense of ripple being allowed into the passband. In the s -domain, this corresponds to the circle of poles being flattened into an *ellipse*. The more flattened the ellipse, the more ripple in the passband, and the sharper the transition. When formed from a cascade of Sallen-Key stages, this requires different values of resistors and capacitors in each stage.

Figure 32-11 also shows the next level of sophistication in filter design strategy: the **elliptic filter**. The elliptic filter achieves the sharpest possible transition by allowing ripple in both the passband and the stopband. In the s -domain, this corresponds to placing zeros directly on the real axis, with the first one near the cutoff frequency. Elliptic filters come in several varieties and are significantly more difficult to design than Butterworth and Chebyshev configurations. This is because the poles and zeros of the elliptic filter do not lie in a simple geometric pattern, but in a mathematical arrangement involving elliptic functions and integrals (hence the name).

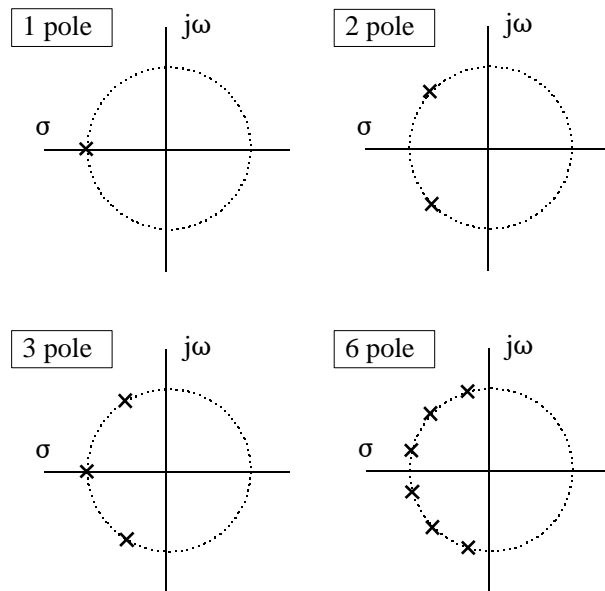
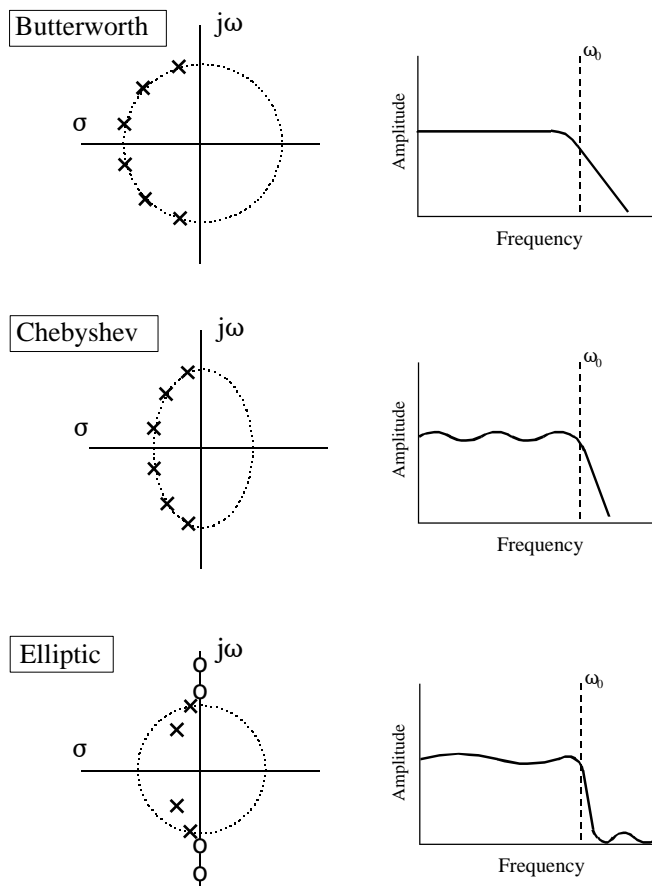


FIGURE 32-10

The Butterworth s -plane. The low-pass Butterworth filter is created by placing poles equally around the left-half of a circle. The more poles used in the filter, the faster the roll-off.

FIGURE 32-11

Classic pole-zero patterns. These are the three classic pole-zero patterns in filter design. Butterworth filters have poles equally spaced around a circle, resulting in a maximally flat response. Chebyshev filters have poles placed on an ellipse, providing a sharper transition, but at the cost of ripple in the passband. Elliptic filters add zeros to the stopband. This results in a faster transition, but with ripple in the passband *and* stopband.



Since each biquad produces two poles, **even order filters** (2 pole, 4 pole, 6 pole, etc.) can be constructed by cascading biquad stages. However, **odd order filters** (1 pole, 3 pole, 5 pole, etc.) require something that the biquad cannot provide: a single pole on the imaginary axis. This turns out to be nothing more than a simple RC circuit added to the cascade. For example, a 9 pole filter can be constructed from 5 stages: 4 Sallen-Key biquads, plus one stage consisting of a single capacitor and resistor.

These classic pole-zero patterns are for low-pass filters; however, they can be modified for other frequency responses. This is done by designing a low-pass filter, and then performing a mathematical transformation in the s-domain. We start by calculating the low-pass filter pole locations, and then writing the transfer function, $H(s)$, in the form of Eq. 32-3. The transfer function of the corresponding high-pass filter is found by replacing each "s" with " $1/s$ ", and then rearranging the expression to again be in the pole-zero form of Eq. 32-3. This defines new pole and zero locations that implement the high-pass filter. More complicated s-domain transforms can create band-pass and band-reject filters from an initial low-pass design. This type of mathematical manipulation in the s-domain is the central theme of filter design, and entire books are

devoted to the subject. Analog filter design is 90% *mathematics*, and only 10% *electronics*.

Fortunately, the design of high-pass filters using Sallen-Key stages doesn't require this mathematical manipulation. The " $1/s$ " for " s " replacement in the s-domain corresponds to swapping the resistors and capacitors in the circuit. In the s-plane, this swap places the poles at a new position, and adds two zeros directly at the origin. This results in the frequency response having a value of zero at DC (zero frequency), just as you would expect for a high-pass filter. This brings the Sallen-Key circuit to its full potential: the implementation of two poles *and* two zeros.