Preface

Complex Numbers

Definition of Complex Numbers

A complex number z can be written in *rectangular form* (also called *Cartesian form*)

$$z = x + jy, \tag{P-1}$$

where x, y are real numbers and $j = \sqrt{-1}$ is the *imaginary unit.*

On its own, x is called the *real part* of z, and y is called the *imaginary part* of z. This can be written as

$$\mathrm{Re}(z) = x \tag{P-2}$$
$$\mathrm{Im}(z) = y. \tag{P-3}$$

Alternatively, z can be written in *polar form*

$$z = |z|e^{j\theta} = |z|\underline{/\theta}, \tag{P-4}$$

where

$$|z| = \sqrt{x^2 + y^2} \tag{P-5}$$

is the magnitude of z, and

$$\theta = \arg(z) = \arg(x + jy) = \operatorname{atan2}(y, x) = \begin{cases} \arctan\left(\dfrac{y}{x}\right), & x > 0 \\ \arctan\left(\dfrac{y}{x}\right) \pm \pi, & x < 0 \\ \operatorname{sgn}(y) \cdot \dfrac{\pi}{2}, & x = 0 \text{ and } y \neq 0 \\ \text{undefined}, & x = 0 \text{ and } y = 0 \end{cases} \tag{P-6}$$

is its phase angle, calculated from the four-quadrant inverse tangent function atan2.

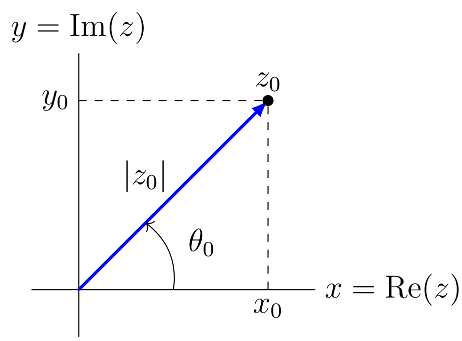
From *Euler's formula*,

$$e^{j\theta} = \cos\theta + j\sin\theta. \tag{P-7}$$

Combining Eqs. (1), (4), and (7), it also then follows that

$$x = |z|\cos(\theta), \tag{P-8}$$
$$y = |z|\sin(\theta). \tag{P-9}$$

Graphically, a complex number z_0 can be drawn on the *complex plane* as a vector from the origin to point (x_0, y_0), with $x = \text{Re}(z)$ as the horizontal axis and $y = \text{Im}(z)$ as the vertical axis.



Operations with Complex Numbers

The *complex conjugate* of z is given by

$$z^* = (x + jy)^* = x - jy. \tag{P-10}$$

This is achieved by substituting $j \leftarrow (-j)$ and can be written as

$$z^* = \big[z\big|_{j\leftarrow(-j)} = \big[x + jy\big|_{j\leftarrow(-j)} = x - jy. \tag{P-11}$$

It then follows that the norm of z is

$$|z| = \sqrt{zz^*} = \sqrt{x^2 + y^2}. \tag{P-12}$$

Some operations between two complex numbers

$$z_1 = x_1 + jy_1 = |z_1|e^{j\theta_1}$$
$$z_2 = x_2 + jy_2 = |z_2|e^{j\theta_2}$$

are defined in the following list:

- Addition/subtraction: $z_1 \pm z_2 = (x_1 \pm x_2) + j(y_1 \pm y_2)$
- Multiplication: $z_1 z_2 = |z_1||z_2|e^{j(\theta_1+\theta_2)}$
- Division: $\dfrac{z_1}{z_2} = \dfrac{|z_1|}{|z_2|}e^{j(\theta_1-\theta_2)}$

Recall from *Euler's formula* that $e^{j\theta} = \cos\theta + j\sin\theta$. The sinusoids can be rewritten such that

$$\cos\theta = \frac{e^{j\theta} + e^{-j\theta}}{2} \tag{P-13}$$

$$\sin\theta = \frac{e^{j\theta} - e^{-j\theta}}{j2}. \tag{P-14}$$

Generalizing Euler's formula, exponentiation of a complex number gives

$$e^z = e^{x+jy} = e^x e^{jy} = e^x(\cos\theta + j\sin\theta). \tag{P-15}$$

De Moivre's formula defines the n^{th} power of a complex number z, for n is a positive integer:

$$z^n = [|z|(\cos\theta + j\sin\theta)]^n = |z|^n[\cos(n\theta) + j\sin(n\theta)]. \tag{P-16}$$

In a similar vein, the n^{th} root of a complex number z (for n is a positive integer) is

$$\begin{aligned} z^{1/n} &= [|z|(\cos\theta + j\sin\theta)]^{1/n} \\ &= |z|^{1/n}\left[\cos\left(\frac{\theta + 2k\pi}{n}\right) + j\sin\left(\frac{\theta + 2k\pi}{n}\right)\right], \text{ for } k = 0, 1, 2, ..., n-1. \end{aligned} \tag{P-17}$$

Solutions to the equation $z^n = 1$ are called n^{th} *roots of unity* and are defined as

$$z = \cos\left(\frac{2k\pi}{n}\right) + j\sin\left(\frac{2k\pi}{n}\right) = e^{j2k\pi/n}, \text{ for } k = 0, 1, 2, ..., n-1. \tag{P-18}$$

The following are common equivalences for powers of imaginary unit j:

$$j = \sqrt{-1} = e^{j\pi/2} \tag{P-19}$$

$$j^2 = -1 = e^{-j\pi} \tag{P-20}$$

$$j^3 = -j = e^{-j\pi/2} \tag{P-21}$$

$$j^4 = 1 \tag{P-22}$$

$$\sqrt{j} = \pm e^{j\pi/4} = \pm\frac{(1+j)}{\sqrt{2}} \tag{P-23}$$

$$\sqrt{-j} = \pm e^{-j\pi/4} = \pm\frac{(1-j)}{\sqrt{2}} \tag{P-24}$$

Regions in the Complex Plane

Let z be a complex variable and $z_0 = x_0 + jy_0$ be a complex number. Then

$$|z - z_0| = r \tag{P-25}$$

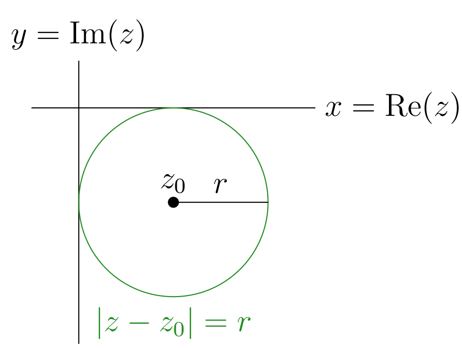
graphically represents a circle of radius 2 centered at (x_0, y_0). In the inequality form,

$$|z - z_0| < r \tag{P-26}$$

represents the interior of the circle, whereas

$$|z - z_0| > r \tag{P-27}$$

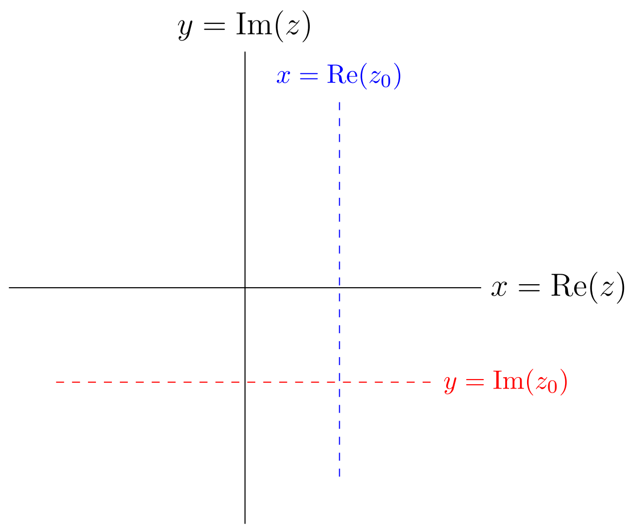
represents the entire region of the complex plane outside of the circle.



This can be proven by plugging in $z = x + jy$ and $z_0 = x_0 + jy_0$ to get the rectangular equation of a circle

$$\sqrt{(x - x_0)^2 + (y - y_0)^2} = r \Longrightarrow (x - x_0)^2 + (y - y_0)^2 = r^2. \tag{P-28}$$

$\mathrm{Re}(z_0)$ represents a vertical line at $x = x_0$, and $\mathrm{Im}(z_0)$ represents a horizontal line at $y = y_0$. The inequality forms are evident in the rectangular form.



In general, any region described by a complex equation or inequality can be identified by solving in rectangular form and using candidate values of z to test the regions of validity.

Complex Functions of Time

Just as with complex numbers and complex variables, complex functions of time $z(t)$ also follow many of the same properties. In rectangular form,

$$z(t) = x(t) + jy(t) \tag{P-29}$$
$$\mathrm{Re}[z(t)] = x(t) \tag{P-30}$$
$$\mathrm{Im}[z(t)] = y(t). \tag{P-31}$$

In polar form,

$$z(t) = |z(t)| \exp[j\theta(t)] \tag{P-32}$$
$$|z(t)| = \sqrt{x^2(t) + y^2(t)} \tag{P-33}$$
$$\theta(t) = \arg[x(t) + jy(t)]. \tag{P-34}$$

Here, while the Steinmetz phasor notation $\angle\theta(t)$ can be used as an equivalent notation for $exp[j\theta(t)]$, it can also be used as a phase operator, where

$$\angle z(t) = \theta(t). \tag{P-35}$$

The phase operator can also be extended to complex numbers and variables, with $\angle z = \theta$.

Just as before, $x(t), y(t), |z(t)|$, and $\angle z(t)$ are all real-valued functions of time.

The complex conjugate function $z^*(t)$ has the following properties:

$$z^*(t) = [z(t)|_{j\leftarrow(-j)} \tag{P-36}$$
$$= [x(t) + jy(t)|_{j\leftarrow(-j)} = x(t) - jy(t) \tag{P-37}$$
$$|z(t)|^2 = z(t)z^*(t) = z^*(t)z(t). \tag{P-38}$$

While Euler's and De Moivre's formulas also apply to complex functions of time, we are not particularly interested in their applications. Often, the polar form of complex functions of time are the only complex equations of interest.

Operations as Functions

Here, we introduce some functions that do not necessarily represent signals, but rather they serve as mathematical operations for analysis.

The *indicator function*, also called the *characteristic function*, is a Boolean function given by

$$\mathbb{1}_A(a) = \begin{cases} 1, & a \in A \\ 0, & otherwise \end{cases}. \tag{P-39}$$

The *signum (or sign) function* captures the sign of the value a:

$$\mathrm{sgn}(a) = \begin{cases} -1, & a < 1 \\ 0, & a = 0 \\ 1, & a > 1 \end{cases}. \tag{P-40}$$

The *floor function* lowers the value of a to the nearest integer below a:

$$\mathrm{floor}(a) = \begin{cases} a, & a \in \mathbb{Z} \\ \mathrm{RoundDown}(a), & otherwise \end{cases} \tag{P-41}$$

The *ceiling function* raises the value of a to the nearest integer above a:

$$\mathrm{ceil}(a) = \begin{cases} a, & a \in \mathbb{Z} \\ \mathrm{RoundUp}(a), & otherwise \end{cases} \tag{P-42}$$

Note that

$$\mathrm{floor}(a) \leq a \leq \mathrm{ceil}(a). \tag{P-43}$$

The *remainder function* returns the remainder of division a/b and is borrowed from the integer modulo operation:

$$\mathrm{mod}(a, b) = (a \mod b) = a - b \cdot \mathrm{floor}(a/b). \tag{P-44}$$

Chapter 1

Signals

Signals are essentially functions that convey information about some (usually physical) phenomenon.

1.1 Classification of Signals Based on Mapping

Since signals are functions, signals can be classified by their domain and codomain.

1.1.1 Multichannel vs Multidimensional

Signals that can be expressed as vector-valued functions are called *multichannel signals*. These signals typically are generated from multiple sources or multiple sensors. An *N-channel* signal may be written as

$$\mathbf{S}(t) = \begin{bmatrix} S_1(t) \\ S_2(t) \\ \vdots \\ S_N(t) \end{bmatrix}, \text{ for } t \in \mathbb{R}. \tag{1.1}$$

An example of a multichannel signal is the electrocardiogram (ECG or EKG), which utilizes either 3 leads or 12 leads to measure the electrical activity of the heart. As a result, 3-channel or 12-channel signals are recorded.

Phased arrays, commonly used for radar and 5G MIMO communications, also transmit multichannel signals. Each element in the antenna array is applied different phase shifts such that the collective beam is steered to a specific direction. Here, each element represents a different channel.

Signals that can be expressed as multivariate functions are called *multidimensional signals*. An *M-dimensional* signal may be written as

$$S(x_1, x_2, ..., x_M), \text{ for } \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{bmatrix} \in \mathbb{R}^M. \tag{1.2}$$

An example of a multidimensional signal is a black-and-white image. The intensity or brightness of each point in an image is a function of two variables, often notated as $I(x, y)$. This can be extended to 3-dimensional signals in the form of black-and-white videos, notated as $I(x, y, t)$. After all, a single frame of a video is merely just an image.

Some signals can be both multichannel and multidimensional. That is,

$$\mathbf{S}(x_1, x_2, ...x_M) = \begin{bmatrix} S_1(x_1, x_2, ...x_M) \\ S_2(x_1, x_2, ...x_M) \\ \vdots \\ S_N(x_1, x_2, ...x_M) \end{bmatrix}, \text{ for } \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{bmatrix} \in \mathbb{R}^M. \tag{1.3}$$

With the introduction of color in digital media, the RGB (red, green, blue) model becomes prevalent in creating techological devices that can display color such as monitors and mobile phones. As a result, color images are both 3-channel and 2-dimensional, with intensity denoted as

$$\mathbf{I}(x, y) = \begin{bmatrix} I_R(x, y) \\ I_G(x, y) \\ I_B(x, y) \end{bmatrix}, \tag{1.4}$$

and color videos are 3-channel and 3-dimensional, with intensity denoted as

$$\mathbf{I}(x, y, t) = \begin{bmatrix} I_R(x, y, t) \\ I_G(x, y, t) \\ I_B(x, y, t) \end{bmatrix}. \tag{1.5}$$

For the rest of this text, we will only discuss single-channel, one-dimensional signals (specifically time-varying signals) – these will simply be referred to as signals. There are countless examples of these signals, some of which include audio (mono, as stereo is 2-channel), speech, circuit voltages, and transmitted messages.

1.1.2 Continuous-Time vs Discrete-Time

Continuous-time signals (or CT signals for short) are defined for every value of time $t \in \mathbb{R}$ – whether or not many of these values end up being zero depends on the physical nature of the signal itself. Another way to rephrase is that continuous-time signals operate in the continuous time domain. These signals are often notated with parentheses, $x(t)$.

Discrete-time signals (or DT signals for short) are defined only at specific values of time, usually integers $n \in \mathbb{Z}$. That is, discrete-time signals operate in the discrete time domain. Sampling a continuous-time signal at periodic intervals produces a discrete-time signal. Discrete-time signals are often notated with brackets, $x[n]$, and are represented as sequences.

Note that the terms continuous-time and discrete-time describe signals that vary with time. For multidimensional signals that vary with location such as images, the terms *continuous-space* and *discrete-space* are more appropriate.

1.1.3 Continuous-Valued vs Discrete-Valued

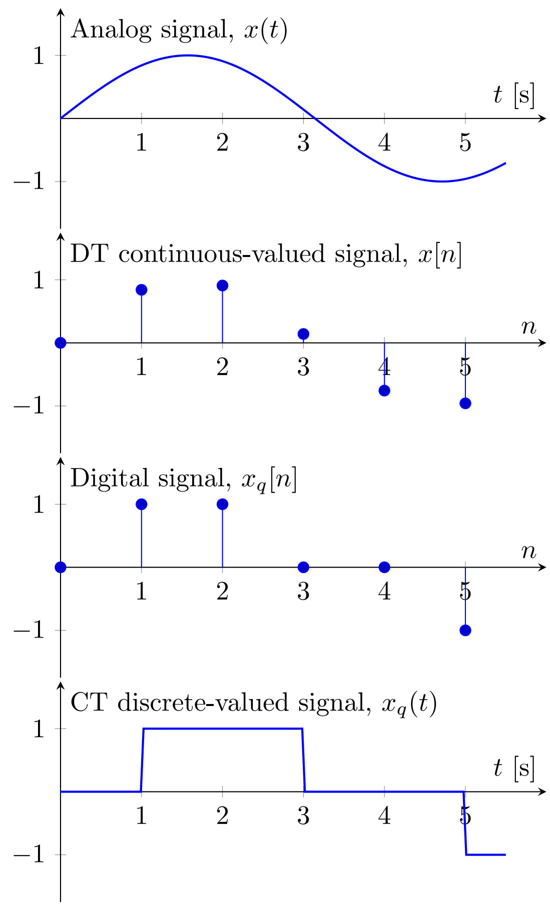
If the amplitude of a signal can take on any value within a finite or infinite range, the signal is *continuous-valued.* Otherwise, if a signal can only take on a finite set of values within the range as its amplitude, the signal is said to be *discrete-valued.*

For instance, given a range of $[0, 1]$, if the signal can capture values of 0.234567, $\pi/4$, $0.8999...$, and every number between those values, then it is continuous-valued. A discrete-valued signal may only be able to take on $\{0, 0.5\}$, or perhaps $\{0, 0.25, 0.5, 0.75\}$, or $\{0, 1/8, ..., 7/8\}$ – usually there is a logic to what intermediate values these countable sets are limited to. An N-bit quantizer would influence the precision of the values within 2^{-N}.

Furthermore, a signal that is both continuous-valued and continuous-time is called an *analog signal*. Similarly, a signal that is both discrete-valued and discrete-time is called a *digital signal*.

	Continuous-Valued	Discrete-Valued
Continuous-Time (CT)	Analog	—
Discrete-Time (DT)	—	Digital

Figure 1.1: Signals based on time and range.



In Figure 1.1, an analog signal $x(t)$ is first plotted. When *sampled*, a discrete-time continuous-valued signal $x[n]$ is formed. Because of the precision that modern processors offer, sometimes these signals are also confusingly referred to as digital signals; however, while $x[n]$ can have infinite precision, a digital signal $x_q[n]$ has limited precision. Digital signal $x_q[n]$ is plotted, where in this version, the signal has limited *quantization levels* $\{-1, 0, 1\}$. This *quantized* signal $x_q[n]$ can then be "*held*" such that a continuous-time discrete-valued signal $x_q(t)$ is formed.

For the rest of this text, only analog signals are of interest.

1.1.4 Deterministic vs Random

Lastly, if a signal can be reproduced without randomness dictating the future of the signal, then it is said to be a *deterministic signal*. Otherwise, it is classified as a *random signal*.

While one could argue a noisy sine wave is irreproducible, it is merely a sum of two signals: a deterministic sine wave and random noise. Each counterpart could be studied on its own. For the rest of this text, only deterministic signals are of interest.

Collectively, we will only be looking into signals with the following characteristics for the remainder of this text:

- single-channel, one-dimensional
- continuous-time, continuous-valued (analog)
- deterministic

1.2 Signal Transformations

One of the simplest ways a signal can be modified is by an *affine transformation on the independent variable*. That is, $x(t) \mapsto x(at+b)$.

1.2.1 Time Shifting

A CT signal that is time-shifted by T seconds can be expressed as

$$y(t) = x(t-T). \tag{1.6}$$

If $T > 0$, then $y(t)$ is said to be delayed by T seconds relative to $x(t)$. If $T < 0$, then $y(t)$ is advanced by T seconds.

1.2.2 Time Scaling

A CT signal that is time-scaled by some factor a can be expressed as

$$y(t) = x(at). \tag{1.7}$$

If $|a| > 1$, then $y(t)$ is a temporally compressed version of $x(t)$. If $|a| < 1$, then $y(t)$ is a temporally expanded version of $x(t)$.

1.2.3 Time Reversal

A time reversal is a reflection of some signal over the vertical axis and can be expressed as

$$y(t) = x(-t). \tag{1.8}$$

1.2.4 Combined Transformation

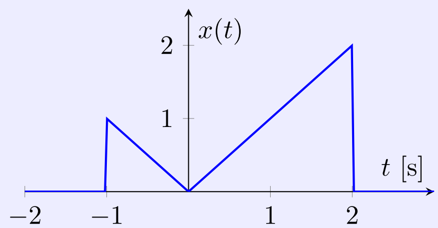
Let

$$y(t) = x(at-b) = x\left(a\left(t-\frac{b}{a}\right)\right). \tag{1.9}$$

There are two approaches to transforming the signal $x(t)$ to $y(t)$.

Approach 1	Approach 2
1. Time scale by $\lvert a \rvert$. Then reflect if $a < 0$. 2. Shift by b/a units.	1. Shift by b units. 2. Time scale by $\lvert a \rvert$. Then reflect if $a < 0$.

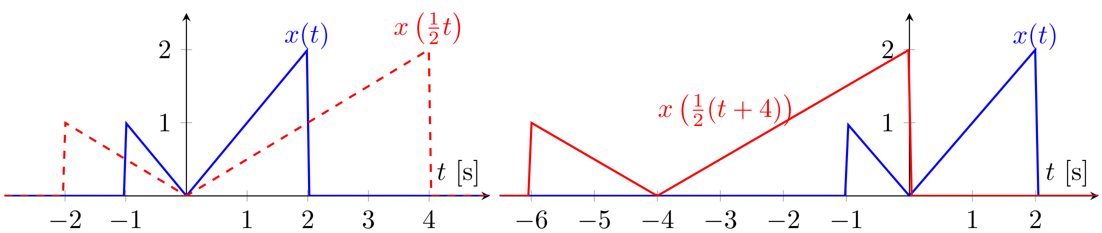
Example 1.2.1. Let $x(t)$ be characterized by the following plot.



Plot $y(t) = x\left(\frac{1}{2}(t+4)\right)$.

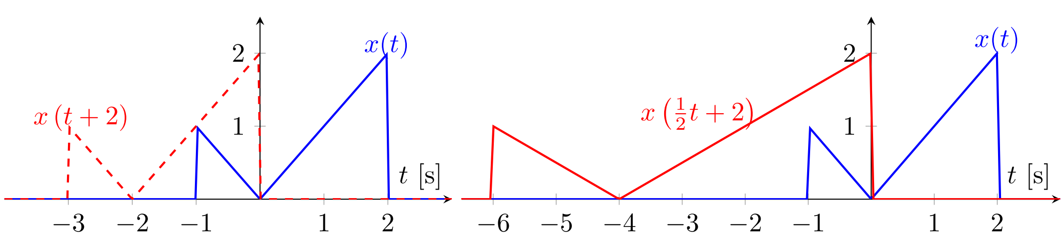
SOLUTION

Using Approach 1:



Here, the plot is first expanded to twice the duration. Then the expanded plot is translated 4 units to the left.

Using Approach 2, for $y(t) = x\left(\frac{1}{2}(t+4)\right) = x\left(\frac{1}{2}t+2\right)$:



Here, the plot is first translated 2 units to the left. Then the shifted plot is expanded to twice the duration, relative to time $t = 0$.

■

1.3 Waveforms

1.3.1 Exponential and Sinusoidal Signals

Complex exponential functions are periodic functions that can be described as

$$x(t) = Ae^{j\omega_0 t} = A\exp(j\omega_0 t), \tag{1.10}$$

with some complex scalar $A = |A|e^{j\theta}$ and *fundamental angular frequency* ω_0. As such, these functions can be rewritten as

$$x(t) = |A|e^{j(\omega_0 t+\theta)}. \tag{1.11}$$

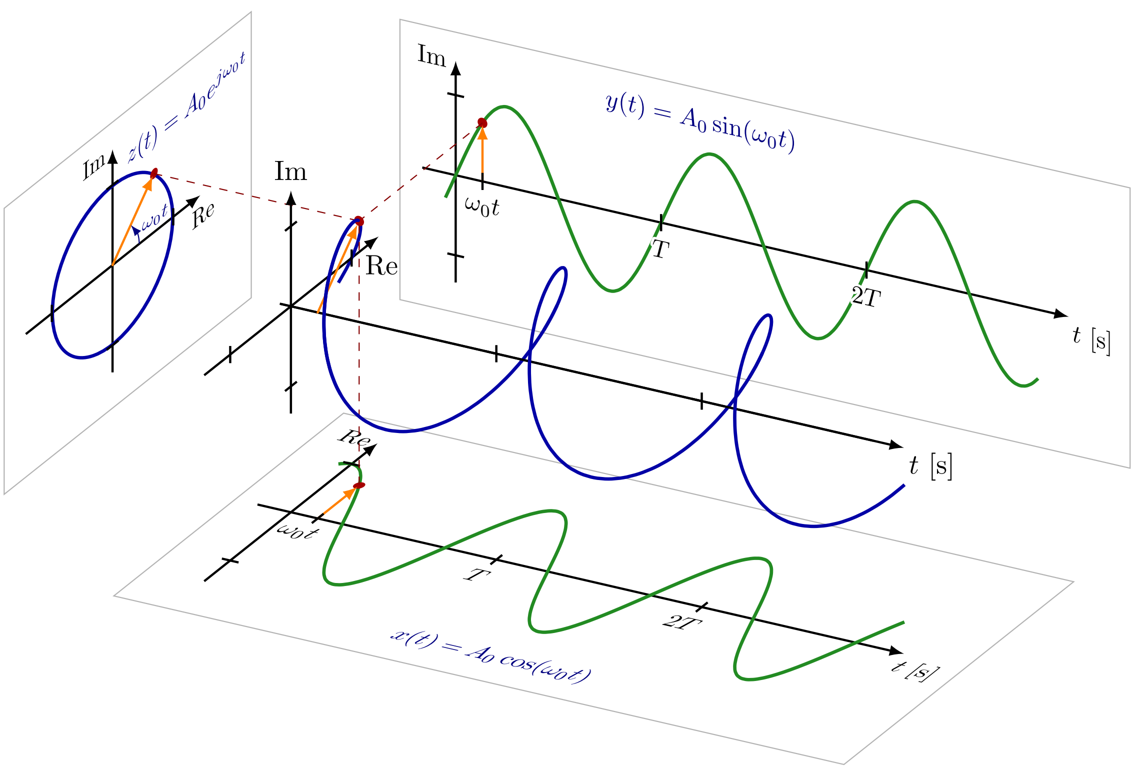
The real and imaginary parts of the exponential function can be isolated to obtain *sinusoids*:

$$\mathrm{Re}(Ae^{j\omega_0 t}) = |A|\cos(\omega_0 t + \theta) \tag{1.12}$$

$$\mathrm{Im}(Ae^{j\omega_0 t}) = |A|\sin(\omega_0 t + \theta) \tag{1.13}$$

As complex exponential functions are periodic, it follows that both the real and imaginary parts are periodic as well. Therefore, both cosine and sine functions are periodic with fundamental angular frequency ω_0.

Figure 1.2: Complex exponential function and its components.



1.3.2 Singularity Signals

Singularity functions are a class of discontinuous functions that are continuous everywhere except at a singularity point. That is, a function which either is not continuous everywhere or has a some-ordered derivative that is not continuous everywhere is considered a singularity function. When not translated, the singularity point is at $t = 0$, where the value is undefined.

The *unit impulse function*, also called the *Dirac delta function*, is defined as

$$\delta(t-T) = \begin{cases} +\infty, & t = T \\ 0, & t \neq T \end{cases}, \tag{1.14}$$

$$\text{with} \int_{-\infty}^{+\infty} \delta(t-T)\,dt = 1. \tag{1.15}$$

Graphically, $\delta(t - T)$ is plotted with a vertical arrow of length 1 pointing up from the t-axis at time $t = T$, with length 1 representing the area under the curve of an impulse function is 1. When multiplied by a scalar k, the plot of $k \cdot \delta(t-T)$ has an arrow length $|k|$ and may point down if $k < 0$. Interestingly, when time-scaled by a factor a,

$$\delta(at) = \frac{1}{|a|} \cdot \delta(t). \tag{1.16}$$

The impulse function also has a unique *sampling property*, also called *sifting property*, and is defined as

$$\int_{-\infty}^{+\infty} x(t)\delta(t-T)\,dt = x(T), \tag{1.17}$$

$$x(t)\delta(t-T) = x(T)\delta(t-T). \tag{1.18}$$

The antiderivative of the unit impulse function is the *unit step function*, also called the *Heaviside step function*, and is defined as

$$u(t-T) = \begin{cases} 0, & t < T \\ 1, & t > T \end{cases}. \tag{1.19}$$

Essentially, the unit step function can be treated as an "on switch", where at time $t = T$, the signal being multiplied gets turned on. In this case, since $u(t - T) = 1 \cdot u(t - T)$, the constant function is inactive up until time $t = T$, at which the constant function is finally turned on. This concept becomes important in creating other signals.

Similarly, the time-reversed step function $u(T - t)$ can be treated as an "off switch", where at time $t = T$, the signal being multiplied gets turned off.

The antiderivative of the unit step function is the *unit ramp function*, which is defined as

$$r(t-T) = \mathrm{ramp}(t-T) = (t-T)u(t-T) \tag{1.20}$$

$$= \begin{cases} 0, & t < T \\ t-T, & t > T \end{cases}. \tag{1.21}$$

Notice that the ramp function can be written as a product of the linear function and the unit step function. Using the "on switch" concept, the linear function is suppressed for all times before T; however, once time $t = T$ has arrived, the linear function is activated.

The antiderivative of the unit ramp function is the *unit parabolic function*, which is defined as

$$\operatorname{quad}(t-T) = \frac{1}{2}(t-T)^2 u(t-T) \tag{1.22}$$

$$= \begin{cases} 0, & t < T \\ \frac{1}{2}(t-T)^2, & t > T \end{cases}. \tag{1.23}$$

1.3.3 Modifying Exponential, Sinusoidal, and Singularity Signals

The *unit exponential decay* with time constant τ is defined as

$$\exp[-(t-T)/\tau]\ u(t-T),\ \text{for}\ \tau > 0. \tag{1.24}$$

The *rectangular pulse* is defined as

$$\operatorname{rect}(t) = \Pi(t) = u\left(t + \frac{1}{2}\right) - u\left(t - \frac{1}{2}\right) \tag{1.25}$$

$$= \begin{cases} 1, & |t| < \frac{1}{2} \\ 0, & otherwise \end{cases}. \tag{1.26}$$

More commonly, when written as $\operatorname{rect}\left(\frac{t-T}{\tau}\right)$, the rectangular pulse has a width τ that is centered at $t = T$.

The *unnormalized triangular pulse*, or simply the *triangular pulse*, is defined as

$$\operatorname{tri}(t) = \Lambda(t) = \begin{cases} 1 - |t|, & |t| < 1 \\ 0, & otherwise \end{cases}. \tag{1.27}$$

When written as $\operatorname{tri}\left(\frac{t-T}{\tau}\right)$, the unnormalized triangular pulse has a width 2τ that is centered at $t = T$.

The *normalized triangular pulse* is defined as

$$\overline{\operatorname{tri}}(t) = \overline{\Lambda}(t) = \begin{cases} 1 - 2|t|, & |t| < \frac{1}{2} \\ 0, & otherwise \end{cases}. \tag{1.28}$$

When written as $\overline{\operatorname{tri}}\left(\frac{t-T}{\tau}\right)$, the normalized triangular pulse has a width τ that is centered at $t = T$.

The *signum function*, also called the *sign function*, is defined as

$$\operatorname{sgn}(t) = u(t) - u(-t) \tag{1.29}$$

$$= 2u(t) - 1 \tag{1.30}$$

$$= \begin{cases} -1, & t < 0 \\ 0, & t = 0 \\ 1, & t > 0 \end{cases}. \tag{1.31}$$

The *unnormalized sinc function*, also called the *sampling function*, is defined as

$$\operatorname{sinc}_u(t) = \operatorname{Sa}(t) = \frac{\sin(t)}{t},\ \text{for}\ \operatorname{sinc}_u(0) = 1. \tag{1.32}$$

The *normalized sinc function*, or simply the *sinc function*, is defined as

$$\mathrm{sinc}(t) = \frac{\sin(\pi t)}{\pi t}, \text{ for } \mathrm{sinc}(0) = 1. \tag{1.33}$$

The waveforms of all signals described in this section can be found in Table 1.1.

Example 1.3.1. Given $x(t) = u(t-3)u(5-t)$, express $x(-2t-1)$ as a piecewise function.

SOLUTION

For $b > a$, the expression $u(t-a)u(b-t)$ can rewritten as $u(t-a) - u(t-b)$. Therefore,

$$\begin{aligned} x(t) = u(t-3)u(5-t) &= u(t-3) - u(t-5) \\ &= \begin{cases} 1, & 3 < t < 5 \\ 0, & \textit{otherwise} \end{cases} \end{aligned}$$

Substituting $t \leftarrow (-2t-1)$,

$$\begin{aligned} x(-2t-1) &= \begin{cases} 1, & 3 < -2t-1 < 5 \\ 0, & \textit{otherwise} \end{cases} \\ &= \begin{cases} 1, & -3 < t < -2 \\ 0, & \textit{otherwise} \end{cases} \\ &= u(t+3) - u(t+2). \end{aligned}$$

■

1.4 Linear Combination of Signals

Just as mathematical expressions can be summed, signals (which essentially are functions) can be summed as well.

While any and all real signals are valid candidates for summation, this section will only focus on the unit impulse, unit step, unit ramp, rectangular pulse, and unit exponential decay as addends of interest. In particular, only signals defined for $t \geq 0$ will be synthesized.

Ignoring points of discontinuity, these signals graphically represent either slope changes or vertical offsets, which happen at time of *excitation* – for rectangular pulses, they happen strictly between time of excitation and time of *termination*. These graphical properties can be seen in Table 1.2.

Table 1.1: Common waveforms.

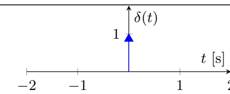
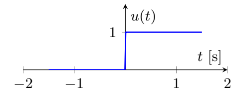
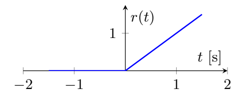
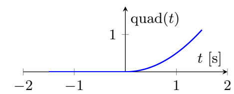
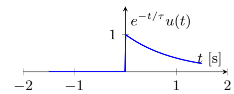
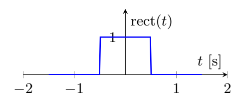
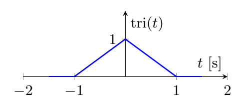
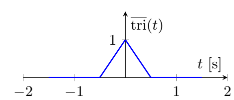
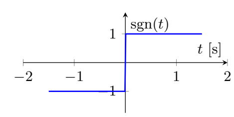
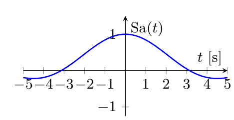
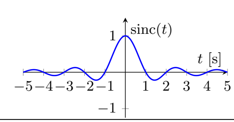
Function	Expression	General Shape
Unit impulse	$\delta(t-T)=\begin{cases}+\infty, & t=T\\ 0, & t\neq T\end{cases}$	
Unit step	$u(t-T)=\begin{cases}0, & t<T\\ 1, & t>T\end{cases}$	
Unit ramp	$r(t-T)=\mathrm{ramp}(t-T)=(t-T)u(t-T)$	
Unit parabolic	$\mathrm{quad}(t-T)=\frac{1}{2}(t-T)^2u(t-T)$	
Unit exponential decay	$\exp[-(t-T)/\tau]\ u(t-T)$	
Rectangular pulse	$\mathrm{rect}\left(\frac{t-T}{\tau}\right)=\Pi\left(\frac{t-T}{\tau}\right)=u(t-T_1)-u(t-T_2)$ $T_1=T-\frac{\tau}{2},\quad T_2=T+\frac{\tau}{2}$	
(Unnormalized) triangular pulse	$\mathrm{tri}\left(\frac{t-T}{\tau}\right)=\Lambda\left(\frac{t-T}{\tau}\right)=r(t-T_1)-2r(t-T)-r(t-T_2)$ $T_1=T-\tau,\quad T_2=T+\tau$	
Normalized triangular pulse	$\overline{\mathrm{tri}}\left(\frac{t-T}{\tau}\right)=\overline{\Lambda}\left(\frac{t-T}{\tau}\right)=2r(t-T_1)-4r(t-T)-2r(t-T_2)$ $T_1=T-\frac{\tau}{2},\quad T_2=T+\frac{\tau}{2}$	
Signum function	$\mathrm{sgn}(t)=u(t)-u(-t)$	
Sampling function	$\mathrm{sinc}_u(t)=\mathrm{Sa}(t)=\frac{\sin(t)}{t},$ for $\mathrm{sinc}_u(0)=1$	
Sinc function	$\mathrm{sinc}(t)=\frac{\sin(\pi t)}{\pi t},$ for $\mathrm{sinc}(0)=1$	

Table 1.2: Graphical properties of select waveforms.

	Slope Change, $\Delta m(T) = m(T^+) - m(T^-)$	Offset, $\Delta y(T) = y(T^+) - y(T^-)$
$k \cdot \delta(t-T)$	Vertical arrow at $t = T$ of length k	None
$k \cdot u(t-T)$	None	$\Delta y(T) = k$
$k \cdot \mathrm{rect}\left(\frac{t-T}{\tau}\right)$	None	$\Delta y(T - \tau/2) = k$, $\Delta y(T + \tau/2) = -k$
$k \cdot r(t-T)$	$\Delta m(T) = k$	None
$k \cdot \exp\left(-\frac{t-T}{\tau}\right)$	Exponentially decay rate $1/\tau$ for $t > T$	$\Delta y(T) = k$

Waveform synthesis tips:

- Sort the addends by their respective excitation times
- Mark all excitation and termination times on the time axis
- Starting at zero slope and zero offset, make changes at each marked time, going from left to right along the time axis

Example 1.4.1. Given $x(t) = 2\,\mathrm{rect}\left(\frac{t-2}{2}\right) + r(t-3)u(4-t) + e^{-(t-4)}u(t-4) + \delta(t-5)$, plot $x(-2t+2)$.

SOLUTION

We need to plot $x(t)$ before its transformed version. Since the addends are already ordered, we determine the excitation and termination times, labeled t_0 and t_f.

$$\begin{aligned} 2\,\mathrm{rect}\left(\frac{t-2}{2}\right) &\Longrightarrow t_0 = 1, t_f = 3 \\ r(t-3)u(4-t) &\Longrightarrow t_0 = 3, t_f = 4 \\ e^{-(t-4)}u(t-4) &\Longrightarrow t_0 = 4 \\ \delta(t-5) &\Longrightarrow t_0 = 5 \end{aligned}$$

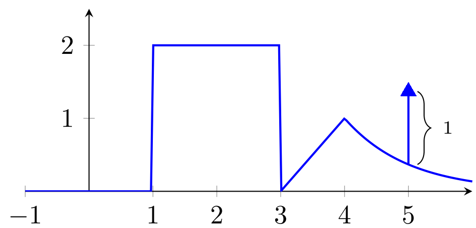
Next, we make a note of all slope changes Δm and offsets Δy due to each addend at each time.

$$\begin{aligned} 2\,\mathrm{rect}\left(\frac{t-2}{2}\right) &\Longrightarrow \Delta y = 2 \text{ at } t_0 = 1, \Delta y = -2 \text{ at } t_f = 3 \\ r(t-3)u(4-t) &\Longrightarrow \Delta m = 1 \text{ at } t_0 = 3, \Delta m = -1 \text{ and } \Delta y = -1 \text{ at } t_f = 4 \\ e^{-(t-4)}u(t-4) &\Longrightarrow \text{exponential decay and } \Delta y = 1 \text{ at } t_0 = 4 \\ \delta(t-5) &\Longrightarrow \text{arrow length 1 at } t_0 = 5 \end{aligned}$$

Collecting the changes with respect to time:

$$\begin{aligned}
t = 1 &\Longrightarrow \Delta y = 2 \\
t = 3 &\Longrightarrow \Delta m = 1, \Delta y = -2 \\
t = 4 &\Longrightarrow \text{exponential decay and } \Delta y = -1 + 1 = 0 \\
t = 5 &\Longrightarrow \text{arrow length 1 at } t_0 = 5
\end{aligned}$$

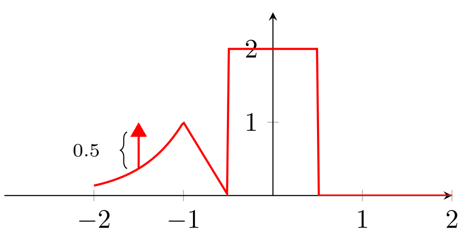
Using the collection of changes, plot $x(t)$ from left-to-right.



When applying signal transformations, note that for the transformed impulse function:

$$\begin{aligned}
\delta((-2t + 2) - 5) = \delta(-2t - 3) &= \delta(-2(t + 1.5)) \\
&= \frac{1}{|-2|} \cdot \delta(t + 1.5) = \frac{1}{2} \cdot \delta(t + 1.5)
\end{aligned}$$

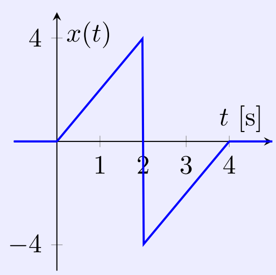
The new transformed signal is plotted below:



Waveform deconstruction tips:

- Mark all times when slope changes or offsets happen, and also note when there are impulse or exponential behaviors
- If there is a slope change $\Delta m = m_2 - m_1$ at $t = T$, then use addend $\Delta m \cdot r(t - T)$.
- If there is an instantaneous jump $\Delta y = y_2 - y_1$ at $t = T$, then use addend $\Delta y \cdot u(t - T)$. If the jump is followed by exponential decay, then use $\Delta y \cdot \exp(-(t - T)/\tau)$ instead.
- If there is a vertical arrowhead of length k at $t = T$, then use $k \cdot \delta(t - T)$.

Example 1.4.2. Given the plot of $x(t)$ below, find an expression for $x(t)$.



SOLUTION

First, make note of all times that introduce some behavioral change.

$$\begin{aligned} t = 0 &\implies \Delta m = 2 - 0 = 2 \\ t = 2 &\implies \Delta y = -4 - 4 = -8 \\ t = 4 &\implies \Delta m = 0 - 2 = -2 \end{aligned}$$

From above, we are likely dealing with a ramp function at $t = 0$, a step function at $t = 2$, and another ramp function at $t = 4$. That is,

$$x(t) = 2r(t) - 8u(t - 2) - 2r(t - 4).$$

■

1.5 Classification of Signals Based on Properties

Signals can also be classified by the content of the signals themselves.

1.5.1 Causality

In the real world, only causal signals can ever occur, though theoretical types can be defined.

- Causal signals: $x(t) = 0$ for $t < 0$
- Noncausal signals: $x(t) \neq 0$ for $t < 0$
- Anticausal signals: $x(t) = 0$ for $t > 0$

1.5.2 Symmetry

A signal could have even symmetry, odd symmetry, or no symmetry.

- Even signals: $x(t) = x(-t)$
- Odd signals: $x(t) = -x(-t)$, or alternatively $x(-t) = -x(t)$

One way to visualize symmetry is to imagine "folding" the graph of the signal. If the graph aligns with itself when folded along the vertical axis, then it has *even symmetry* (sometimes called *mirror symmetry*). If not, then if the graph aligns with itself when folded along the vertical and then the horizontal (time) axis, then it has *odd symmetry* (sometimes called *rotation symmetry*). If neither the 1-fold nor 2-fold tests pass, then it has no symmetry.

When multiplying symmetric functions, take note of the following properties:

- (even) $\times$ (even) $=$ even
- (even) $\times$ (odd) $=$ odd
- (odd) $\times$ (odd) $=$ even

This is analagous to using $+1$ for even and -1 for odd.

Lastly, while a signal might not have any symmetry, any signal can be expressed as a sum of two component signals: one with even symmetry and the other with odd symmetry. Sometimes it is easier to analyze each individual component rather than analyze the whole signal on its own. For

$$x(t) = x_e(t) + x_o(t), \tag{1.34}$$

the even and odd components are given by

$$x_e(t) = \frac{1}{2}\left[x(t) + x(-t)\right] \tag{1.35}$$

$$x_o(t) = \frac{1}{2}\left[x(t) - x(-t)\right] \tag{1.36}$$

Example 1.5.1. Given $x(t) = e^{2t}$, find its even and odd components.

SOLUTION

The even component is given by

$$\begin{aligned} x_e(t) &= \frac{1}{2}\left[x(t) + x(-t)\right] \\ &= \frac{1}{2}\left[e^{2t} + e^{-2t}\right] = \cosh(2t), \end{aligned}$$

and the odd component is given by

$$\begin{aligned} x_e(t) &= \frac{1}{2}\left[x(t) - x(-t)\right] \\ &= \frac{1}{2}\left[e^{2t} - e^{-2t}\right] = \sinh(2t). \end{aligned}$$

■

1.5.3 Periodicity

A signal $x(t)$ is said to be *periodic* if $x(t) = x(t+nT_0)$ for all integers n and time t with fundamental period T_0. Note that $\omega_0 = 2\pi/T_0$ is the fundamental angular frequency, and $f_0 = 1/T_0 = \omega_0/2\pi$ is the fundamental (linear) frequency.

> The sum of N periodic signals are itself periodic if:
>
> - $\omega_0 = \mathbf{GCD}(\omega_1, \omega_2, ..., \omega_N)$ exists, where ω_k is the fundamental angular frequency of the k^{th} signal
> - $\forall n \in \left\{\frac{\omega_1}{\omega_0}, \frac{\omega_2}{\omega_0}, ..., \frac{\omega_N}{\omega_0}\right\}$ are integers, with each n called the n^{th} harmonic (or mode) of fundamental angular frequency ω_0 (or of fundamental frequency f_0)

Example 1.5.2. Determine if $x(t) = \sin(\frac{5\pi}{6}t) + \cos(\frac{3\pi}{4}t) - \exp\left(j\frac{\pi}{3}t\right)$ is periodic or not.

SOLUTION

Notice that each addend in $x(t)$ is periodic itself. Let $\omega_1 = 5\pi/6, \omega_2 = 3\pi/4, \omega_3 = \pi/3$. Then

$$\begin{aligned}
\omega_0 &= \mathbf{GCD}(\omega_1, \omega_2, \omega_3) \\
&= \mathbf{GCD}\left(\frac{5\pi}{6}, \frac{3\pi}{4}, \frac{\pi}{3}\right) \\
&= \mathbf{GCD}\left(\frac{10\pi}{12}, \frac{9\pi}{12}, \frac{4\pi}{12}\right) = \frac{\pi}{12}.
\end{aligned}$$

The corresponding harmonics can be calculated

$$\begin{aligned}
\omega_1/\omega_0 &= 10 \\
\omega_2/\omega_0 &= 9 \\
\omega_3/\omega_0 &= 4
\end{aligned}$$

and are shown to be all integers.

Therefore, $x(t) = \sin(10\omega_0 t) + \cos(9\omega_0 t) + e^{j4\omega_0 t}$ is periodic.

■

1.5.4 Signal Power and Energy

The total *energy* of a signal $x(t)$ is given by

$$E = \int_{-\infty}^{+\infty} |x(t)|^2 \, dt, \tag{1.37}$$

whereas the *time-average power* of a signal is given by

$$P_{av} = \lim_{T\to\infty} \frac{1}{T} \int_{-T/2}^{+T/2} |x(t)|^2 \, dt. \tag{1.38}$$

Furthermore, if a signal $x(t)$ is periodic with fundamental period T_0, then

$$P_{av} = \frac{1}{T_0} \int_{-T_0/2}^{+T_0/2} |x(t)|^2 \, dt. \tag{1.39}$$

In fact, if $x(t) = A \sin(\omega_0 t + \theta)$ for A is real, then

$$P_{av} = \frac{A^2}{2}. \tag{1.40}$$

It can then be said that all periodic signals are power signals.

A signal can be classified based on the values of P_{av} and E.

- Power signals: P_{av} is finite and $E \to \infty$
- Energy signals: $P_{av} = 0$ and E is finite
- Non-physical signals: $P_{av} \to \infty$ and $E \to \infty$

Example 1.5.3. Find the total energy of the signal

$$x(t) = \begin{cases} 0, & t \geq 0 \\ 3t, & 0 \leq t \leq 2 \\ 6e^{-(t-2)}, & t \geq 2 \end{cases}$$

SOLUTION

$$\begin{aligned} E = \int_{-\infty}^{+\infty} |x(t)|^2 \, dt = \int_0^2 |3t|^2 \, dt + \int_2^{+\infty} |6e^{-(t-2)}|^2 \, dt \\ &= \left[\frac{9t^3}{3} \right|_0^2 + 36e^4 \left[\frac{-e^{-2t}}{2} \right|_2^{+\infty} \\ &= (24 - 0) + 36e^4 \left(0 + \frac{e^{-4}}{2} \right) \\ &= 24 + 18 = 42. \end{aligned}$$

■

Example 1.5.4. Fully describe what type of signal $x(t) = (3 - j4)e^{j\pi t/3}$ is.

SOLUTION

Since $x(t)$ is a complex exponential function, it is periodic with

$$\omega_0 = \frac{\pi}{3} = \frac{2\pi}{T_0} \Longrightarrow T_0 = 6.$$

Therefore, $x(t)$ is also a power signal with average power

$$\begin{aligned} P_{av} = \frac{1}{T_0}\int_{-T_0/2}^{+T_0/2} |x(t)|^2\, dt &= \frac{1}{6}\int_{-3}^{3} x(t)x^*(t)\, dt \\ &= \frac{1}{6}\int_{-3}^{3} [(3-j4)e^{j\pi t/3}][(3+j4)e^{-j\pi t/3}]\, dt \\ &= \frac{1}{6}\int_{-3}^{3} 25\, dt = 25. \end{aligned}$$

Checking for symmetry, we find there is none since

$$\begin{aligned} x(-t) &= (3-j4)e^{-j\pi t/3} \neq x(t), \\ -x(-t) &= -(3-j4)e^{-j\pi t/3} \neq x(t). \end{aligned}$$

Therefore, $x(t)$ is:

- an analog signal,
- deterministic,
- noncausal,
- asymmetric,
- periodic,
- and a power signal.

■

The signal processing equations for signal average power and energy are borrowed from physics itself. The physics equation for instantaneous power is given by

$$p(t) = i^2(t)R = \frac{v^2(t)}{R} \tag{1.41}$$

It then follows that the physics equation for energy is

$$E = \int_{-\infty}^{+\infty} p(t)\, dt \tag{1.42}$$

and the physics equation for average power for periodic signals is

$$P_{av} = \frac{1}{T}\int_{-T/2}^{+T/2} p(t)\, dt \tag{1.43}$$

In this case, we take $|x(t)|^2$ to be analogous to $p(t)$. With signal processing, the equations for signal average power and energy are actually normalized. If the signal of interest $x(t)$ ends up being a voltage or current signal, the units must be accounted for when unnormalizing the values.

For cases where normalized signal power and energy need to be converted to unnormalized power and energy describing physical systems such as circuits, we redefine

$$E_{norm} = \int_{-\infty}^{+\infty} |x(t)|^2 \, dt \tag{1.44}$$

$$P_{av,norm} = \lim_{T \to \infty} \frac{1}{T} \int_{-T/2}^{+T/2} |x(t)|^2 \, dt \tag{1.45}$$

If $x(t) = v(t)$ is the voltage across a resistor, then it follows that

$$E = \frac{E_{norm}}{R} \tag{1.46}$$

$$P_{av} = \frac{P_{av,norm}}{R} \tag{1.47}$$

Otherwise if $x(t) = i(t)$ is the current through a resistor, then it follows that

$$E = E_{norm} R \tag{1.48}$$

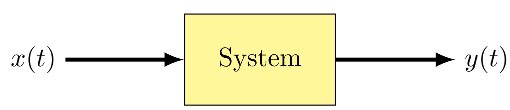
$$P_{av} = P_{av,norm} R \tag{1.49}$$

Chapter 2

Systems

From the previous chapter, we introduce the idea of modifying signals by an affine transformation on the independent variable. In fact, this was the very first introduction to the concept of *systems*. A system essentially transforms an *input signal* into an *output signal*. Furthermore, a system can be modeled as an operator that maps an input function to an output function.

Given an input signal $x(t)$ and an output signal $y(t)$, a system can be depicted using a block diagram:



There are many practical examples of systems in the real world, including circuits, signal processing systems, communications systems, and feedback control systems, those of which are depicted in Figures 2.1 – 2.4.

Figure 2.1: A simple RC circuit.

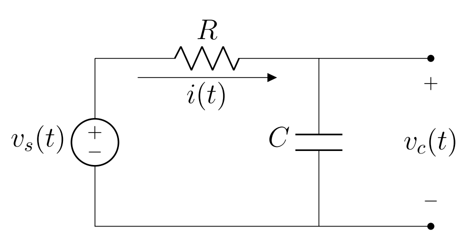


Figure 2.2: Signal processing system.

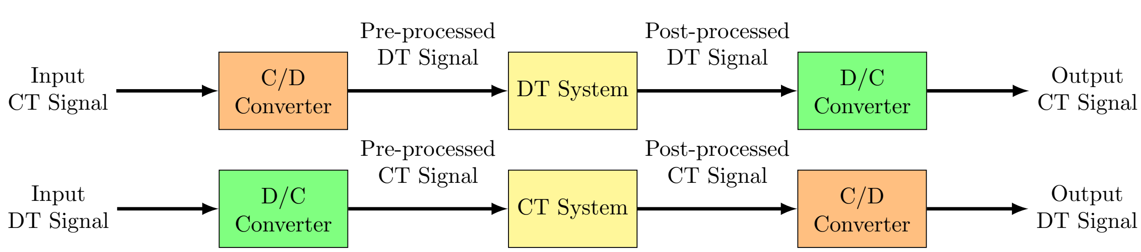


Figure 2.3: Communications system.

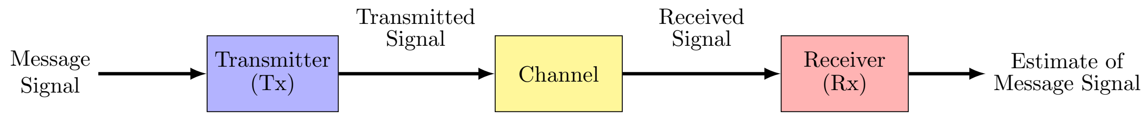
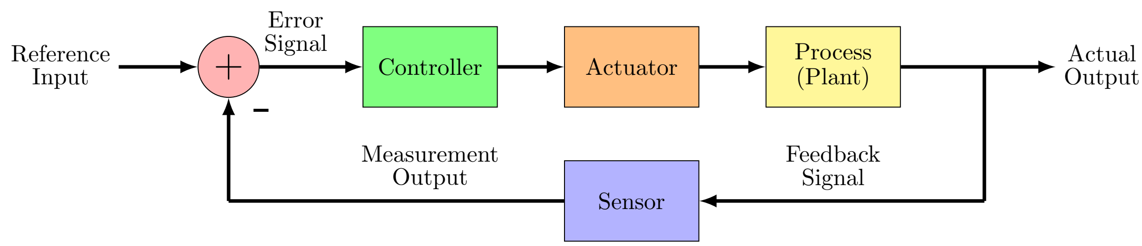


Figure 2.4: Feedback control system.



This chapter will go over different types of systems and why we are particularly interested in a class of systems called *linear time-invariant (LTI) systems*.

2.1 Classification of Systems Based on Mapping

Systems can be classified by the type of signals that are being inputted and outputted.

2.1.1 Multichannel vs Multidimensional

A system that deals with multichannel input and output systems is called a *multichannel system*.

There are four types of multichannel systems, described by the number of channels in the input and output signals and aptly named so:

- Single-input single-output (SISO) systems
- Single-input multiple-output (SIMO) systems
- Multiple-input single-output (MISO) systems
- Multiple-input multiple-output (MIMO) systems

A system that deals with multidimensional input and output signals is called a *multidimensional system*.

The rest of this text will cover one-dimensional SISO systems.

2.1.2 Continuous-Time vs Discrete-Time

A system that handles CT input and output signals is called a *continuous-time system*, whereas a system that deals with DT input and output signals is called a *discrete-time system*. Furthermore, a system that handles analog input and output signals is called a *analog system*, and a system that deals with digital input and output signals is called a *digital system*.

Interestingly, there are cases where a system maps a CT input signal to a DT output signal, or a DT input signal to a CT output signal. The most common example is the analog-to-digital and digital-to-analog converters. Such systems are called *hybrid systems*.

The rest of this text will cover only analog systems.

2.1.3 Deterministic vs Stochastic

A system that maps a deterministic input signal to a *predictable* deterministic output signal is called a *deterministic system*. A system that maps a random input signal to an *unpredictable* random signal is called a *stochastic system*; stochastic systems may be modeled by some probability distribution but cannot be precisely predicted.

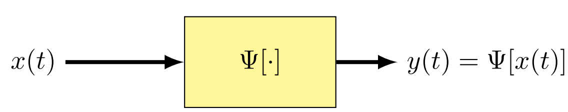
A system that deals with deterministic input signals and unpredictable output signals is called a *chaotic system*.

The rest of this text will cover only deterministic systems.

2.2 Classification of Systems Based on Properties

Systems can also be classified by *how* it maps inputs to outputs. Here, the content of the output signal with respect to the input signal is analyzed.

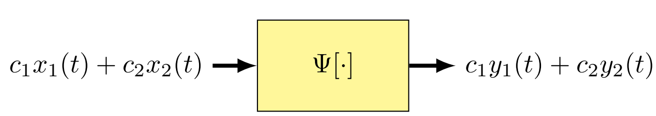
For this chapter, let $\Psi[\cdot]$ represent some operator that the system resembles, with the new block diagram shown.



From the previous chapter, an affine transformation on the independent variable can be represented as a system with $y(t) = \Psi[x(t)] = x(at - b)$.

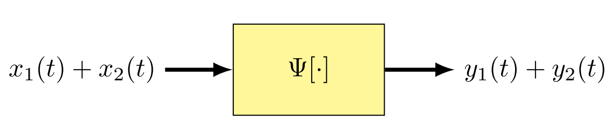
2.2.1 Linear vs Nonlinear

A system is said to be *linear* if it follows the *superposition principle*, as depicted in the block diagram below with constants c_1, c_2, input signal addends $x_1(t), x_2(t)$, and output signal addends $y_1(t), y_2(t)$:



In practice, if the superposition principle is met, then a sum of N input addends will generate a sum of N output addends.

The superposition principle can be broken down into two properties: additivity and scalability. A system has *additive property* if for $y_1(t) = \Psi[x_1(t)]$ and $y_2(t) = \Psi[x_2(t)]$, the system-generated output from the summed inputs is the sum of the outputs $\Psi[x_1(t) + x_2(t)] = y_1(t) + y_2(t)$. That is:



A system has *scaling property* if for some constant c, the system-generated output for a scaled input is an appropriately scaled output $\Psi[c \cdot x(t)] = c \cdot y(t)$. That is:

$$c \cdot x(t) \longrightarrow \boxed{\Psi[\cdot]} \longrightarrow c \cdot y(t)$$

If a system has both scaling and additive properties, then the superposition principle is met, and the system is linear. If the superposition principle is not met, then the system is said to be *nonlinear*.

Example 2.2.1. Determine if the system characterized by the following equation is linear:

$$\frac{d^2y(t)}{dt} + 2\frac{dy(t)}{dt} + 3y(t) = 4\frac{dx(t)}{dt} + 5x(t)$$

SOLUTION

First, determine if the system is additive. Consider the following sum:

$$\begin{array}{rl} & \left[\dfrac{d^2y_1(t)}{dt} + 2\dfrac{dy_1(t)}{dt} + 3y_1(t) = 4\dfrac{dx_1(t)}{dt} + 5x_1(t)\right] \\ + & \left[\dfrac{d^2y_2(t)}{dt} + 2\dfrac{dy_2(t)}{dt} + 3y_2(t) = 4\dfrac{dx_2(t)}{dt} + 5x_2(t)\right] \\ \hline \end{array}$$

$$\frac{d^2[y_1(t)+y_2(t)]}{dt} + 2\frac{d[y_1(t)+y_2(t)]}{dt} + 3[y_1(t)+y_2(t)] = 4\frac{d[x_1(t)+x_2(t)]}{dt} + 5[x_1(t)+x_2(t)]$$

We are able to combine input components and output components due to the additive nature of the derivative. Here we can see that $\Psi[x_1(t)+x_2(t)] = \Psi[x_1(t)] + \Psi[x_2(t)] = y_1(t) + y_2(t)$. Therefore, the system is additive.

Lastly, determine if the system is scalable. Consider the following product:

$$\begin{aligned} &\frac{d^2[c \cdot y(t)]}{dt} + 2\frac{d[c \cdot y(t)]}{dt} + 3[c \cdot y(t)] = 4\frac{d[c \cdot x(t)]}{dt} + 5[c \cdot x(t)] \\ &\implies c \cdot \left[\frac{d^2y(t)}{dt} + 2\frac{dy(t)}{dt} + 3y(t)\right] = c \cdot \left[4\frac{dx(t)}{dt} + 5x(t)\right] \end{aligned}$$

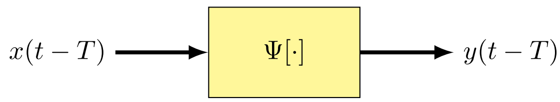
We are able to factor out c because of the scaling property of the derivative. Here we can see that $\Psi[c \cdot x(t)] = c \cdot \Psi[x(t)] = c \cdot y(t)$. Therefore, the system is scalable.

Since the system is both additive and scalable, the superposition principle is met, and the system is linear.

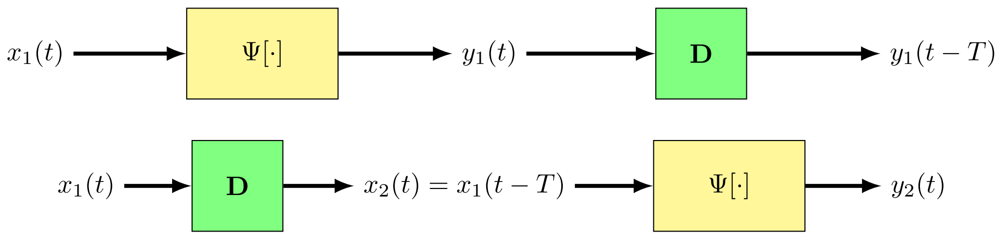
■

2.2.2 Time-Invariant vs Time-Variant

A system is said to be *time-invariant* if a delay in the input signal results in a corresponding delay in the output signal such that $\Psi[x(t-T)] = y(t-T)$. That is:



One way to check is to consider the following two block diagrams, with the "D"-block representing a delay of T seconds:



From the outputs above, if $y_2(t) = y_1(t-T)$, then the system is time-invariant. Otherwise, if the outputs are not equal, then the system is *time-variant*, in which $\Psi[x(t-T), T] = y(t-T)$ results in an additional dependence on T.

Example 2.2.2. Determine if the system characterized by the following equation is time-invariant.

$$y(t) = \frac{t \cdot x(t+2)}{x(t-1)}$$

SOLUTION

If we feed the input into the system first then delay the output, then we get:

$$y_1(t) = \frac{t \cdot x_1(t+2)}{x_1(t-1)}$$
$$y_1(t-T) = \frac{(t-T) \cdot x_1((t-T)+2)}{x_1((t-T)-1)}$$

If we delay the input first then feed the delayed input into the system, then we get:

$$x_2(t) = x_1(t-T)$$
$$y_2(t) = \frac{t \cdot x_2(t+2)}{x_2(t-1)} = \frac{t \cdot x_1((t-T)+2)}{x_1((t-T)-1)}$$

Since $y_2(t) \neq y_1(t-T)$, the system is time-variant. ■

Systems that are both linear and time-invariant are called linear time-invariant (LTI) systems. We will explore LTI systems more later.

2.2.3 Dynamic vs Memoryless

A system is said to be *memoryless* or *static* if the output $y(t)$ at time t depends only on the input $x(t)$ at time t. The only LTI system that is also memoryless has the form $y(t) = ax(t)$ for a is a constant.

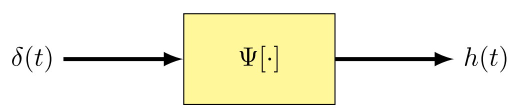
Otherwise, a system whose output additionally depends on past and/or future values of the input is a *dynamic* system; that is, the output at t depends on the input at $t - T$ for any $T \neq 0$. Most systems in the physical world are dynamic.

2.2.4 Causal vs Noncausal

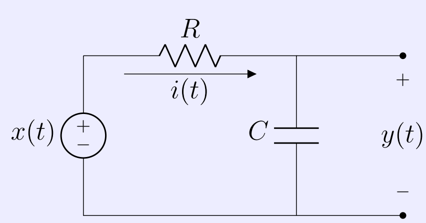
Causal systems, also called *physically realizable systems*, are systems whose output at time t only depend on the past and present values of the input at time $t - T$ for $T > 0$ and time t, respectively. Otherwise, a system that anticipates a future value of the input at time $t + T$ before generating an output at time t is said to be *noncausal*.

Essentially, noncausal systems see the future, which is impossible in the physical world. Generally, from the affine transformation on the independent variable, any system with a time advance or a temporal expansion of the input signal is classified as noncausal.

One way to determine causality is by using an impulse signal $\delta(t)$ as input and observing its output $h(t)$, also called the *impulse response*. If $h(t) = 0$ for all $t < 0$, then the system is causal; otherwise it is noncausal.



Example 2.2.3. Determine if the following circuit is a causal system.



SOLUTION

From Kirchhoff's voltage law (KVL):

$$R \cdot i(t) + y(t) = x(t)$$
$$\implies RC \cdot \frac{dy(t)}{dt} + y(t) = x(t).$$

Let $\tau = RC$. Then

$$\frac{dy(t)}{dt} + \frac{1}{\tau} \cdot y(t) = \frac{1}{\tau} \cdot x(t).$$

Since the integrating factor is $\exp\left(\int 1/\tau \, dt\right) = \exp(t/\tau)$, multiply the whole equation by $\exp(t/\tau)$ such that

$$\begin{aligned} e^{t/\tau} \cdot \frac{dy(t)}{dt} + \frac{e^{t/\tau}}{\tau} \cdot y(t) &= \frac{e^{t/\tau}}{\tau} \cdot x(t) \\ \frac{d}{dt}\left[e^{t/\tau} y(t)\right] &= \frac{e^{t/\tau}}{\tau} \cdot x(t). \end{aligned}$$

Inputting in an impulse function for the input, we get

$$\frac{d}{dt}\left[e^{t/\tau} h(t)\right] = \frac{e^{t/\tau}}{\tau} \cdot \delta(t).$$

Integrating both sides:

$$\begin{aligned} \int_{0^-}^{T} \left(\frac{d}{dt}\left[e^{t/\tau} h(t)\right]\right) dt &= \int_{0^-}^{T} \left(\frac{e^{t/\tau}}{\tau} \cdot \delta(t)\right) dt \\ e^{T/\tau} h(T) &= \int_{0^-}^{T} \left(\frac{e^{0}}{\tau} \cdot \delta(t)\right) dt \\ e^{T/\tau} h(T) &= \frac{1}{\tau} u(T). \end{aligned}$$

Rearranging the equation and substituting $T \leftarrow t$, we get

$$h(t) = \frac{e^{-t/\tau}}{\tau} u(t).$$

Since $h(t)$ is causal, the circuit is a causal system.

■

2.2.5 BIBO Stable vs Unstable

A signal $x(t)$ is said to be *bounded* if there exists some constant C such that

$$|x(t)| \leq C, \text{ for } \forall t. \tag{2.1}$$

A system with a *bounded input, bounded output* (BIBO, for short) is said to be *BIBO stable*; that is, every bounded input signal results in a bounded output signal. Otherwise, if a system produces an unbounded output for a bounded input, it is said to be *unstable*.

For LTI systems, we can particularly look at the impulse response $h(t)$. An LTI system is BIBO stable if and only if $h(t)$ is *absolutely integrable* such that

$$\int_{-\infty}^{+\infty} |h(t)| \, dt = C, \text{ for } C \text{ is finite.} \tag{2.2}$$

Example 2.2.4. Determine if an LTI system with impulse response $h(t) = e^{-|t|}$ is stable.

SOLUTION

$$\begin{aligned}\int_{-\infty}^{+\infty} |h(t)|\, dt = \int_{-\infty}^{+\infty} |e^{-|t|}|\, dt = 2\int_{0}^{+\infty} e^{-t}\, dt \\ = 2\left[-e^{-t}\right|_0^{+\infty} = 2[0-(-1)] = 2 < \infty\end{aligned}$$

Since $h(t)$ is absolutely integrable, the system is BIBO stable.

■

Example 2.2.5. Let $C = A + jB$ and $\gamma = \alpha + j\beta$ be complex values. Show that all LTI systems with exponential decay $h(t) = Ce^{\gamma t}u(t)$ as impulse responses are BIBO stable if and only if $\mathrm{Re}(\gamma) < 0$.

SOLUTION

First, let $\alpha < 0$. Then

$$\begin{aligned}\int_{-\infty}^{+\infty} |h(t)|\, dt &= \int_{-\infty}^{+\infty} |Ce^{\gamma t}u(t)|\, dt = \int_{0}^{+\infty} |C|\cdot|e^{\alpha t}|\cdot|e^{j\beta t}|\, dt \\ &= |C|\int_0^{+\infty} e^{\alpha t}\, dt = |C|\int_0^{+\infty} e^{-|\alpha| t}\, dt = |C|\left[-\frac{e^{-|\alpha|t}}{|\alpha|}\right|_0^{+\infty} = \frac{|C|}{|\alpha|} < \infty.\end{aligned}$$

Now, let $\alpha = 0$. Then

$$\int_{-\infty}^{+\infty} |h(t)|\, dt = \left[|C|\int_0^{+\infty} e^{\alpha t}\, dt\right|_{\alpha=0} = |C|\int_0^{+\infty} 1\, dt = |C|\left[t\right|_0^{+\infty} = \infty.$$

Lastly, let $\alpha > 0$. Then

$$\int_{-\infty}^{+\infty} |h(t)|\, dt = |C|\int_0^{+\infty} e^{\alpha t}\, dt = |C|\left[\frac{e^{|\alpha|t}}{|\alpha|}\right|_0^{+\infty} = \infty.$$

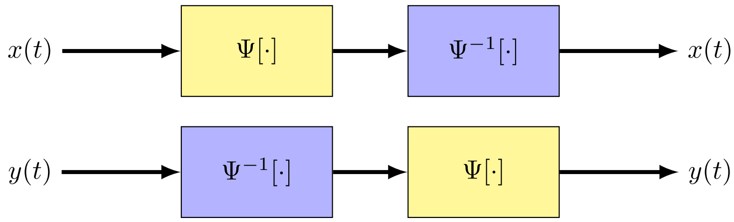
Since $h(t)$ is only absolutely integrable when $\alpha < 0$, the system is only stable when $\mathrm{Re}(\gamma) < 0$.

■

2.2.6 Invertible vs Non-Invertible

A system is said to be *invertible* if it generates unique output signals for every unique input signal. That is, an invertible system has a one-to-one mapping between inputs and outputs. If a system has a many-to-one mapping between inputs and outputs, then it is said to be *non-invertible*.

Alternatively, an invertible system has an inverse system $\Psi^{-1}[\cdot]$ that maps outputs back to the inputs of the forward system $\Psi[\cdot]$ – it may be delayed but the shape of the waveform is maintained.



While a system can be invertible, it does not mean it is implementable. For this very reason, when inverting LTI systems, generally we are also interested in the properties of the inverse system; that is, if the original system is BIBO stable and causal, we would want the inverse system to be BIBO stable and causal as well.

Example 2.2.6. Determine if the system characterized by $y(t) = x(2t-3)$ is invertible.

SOLUTION

Here we can solve for $x(t)$. First, define $y(\tau) = x(2\tau - 3)$. Then

$$t = 2\tau - 3 \Longrightarrow \tau = \frac{1}{2}(t+3).$$

Substituting this relationship in, we get

$$x(t) = y\left(\frac{1}{2}(t+3)\right).$$

Since both forward and inverse systems are affine transformations on the independent variable, the system is invertible.

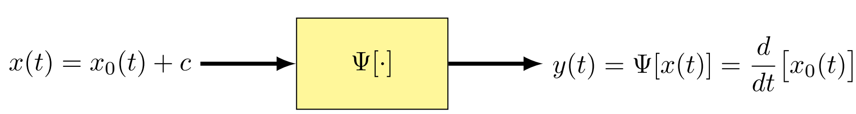
■

Example 2.2.7. Determine if the system characterized by $y(t) = \frac{d}{dt}\big[x(t)\big]$ is invertible.

SOLUTION

Let $x(t) = x_0(t)+c$, where $x_0(t)$ is the non-constant portion of the function and c is the constant in the function.

Regardless of the value of c, the function $x(t)$ will always map to the derivative of $x_0(t)$ since the derivative of a constant is zero.



This is a many-to-one mapping. Therefore the system is non-invertible.

■

2.3 LTI Systems

By now, we know that an *LTI system* is a system that is both linear and time-invariant. But why are we particularly study in LTI systems? We study LTI systems because of their predictable nature. As long as we know the system response to a few select input signals, we can accurately predict the output for all input signals.

Although many systems in the real world are not LTI systems, they might be "locally nearly LTI", where the behavior of the system could possibly be approximated with an LTI system within a small analytic region. Because we are modeling real systems using LTI systems, we would need to choose an LTI model that is also dynamic, causal, and stable.

2.3.1 LTI System Response to Singularity Signals

As prefaced before, the *impulse response* $h(t)$ of a system is the system response to an inputted impulse signal $\delta(t)$, given zero initial conditions. The impulse response of an LTI system is depicted in the following block diagram:

$$\delta(t) \longrightarrow \boxed{\text{LTI}} \longrightarrow h(t)$$

For LTI systems, the impulse response plays an important role. An LTI system can be characterized by its impulse response such that any output signal can be predicted by performing an operation called *convolution* between the input signal and the impulse response. The block diagram of an LTI system can be depicted as:

$$x(t) \longrightarrow \boxed{h(t)} \longrightarrow y(t)$$

Convolution is an operation defined by the *convolution integral*

$$x(t) * h(t) = \int_{-\infty}^{+\infty} x(\tau)h(t-\tau)\,d\tau \tag{2.3}$$

$$= \int_{-\infty}^{+\infty} x(t-\tau)h(\tau)\,d\tau. \tag{2.4}$$

All LTI systems characterized by $h(t)$ abide by convolution. Convolution will be explored further in the next chapter.

The *step response* $y_{step}(t)$ of a system is the system response to an inputted unit step signal $u(t)$, given zero initial conditions. The step response of an LTI system is depicted in the following block diagram:

$$u(t) \longrightarrow \boxed{\text{LTI}} \longrightarrow y_{step}(t)$$

While not as prevalent in signal processing applications, the step response appears in primarily control-based applications, where we are interested in how well a system "tracks" a step input.

Interestingly, just as $u(t)$ is the antiderivative of $\delta(t)$, the step response $y_{step}(t)$ is the antiderivative of the impulse response $h(t)$. That is,

$$h(t) = \frac{dy_{step}(t)}{dt}, \tag{2.5}$$

$$y_{step}(t) = \int_{-\infty}^{t} h(\tau)\, d\tau. \tag{2.6}$$

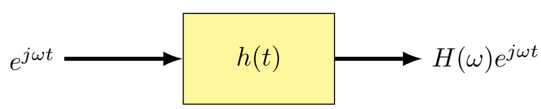
For both the impulse response and step response, all initial conditions must be zero in order to generalize the performance of any system modeled by that particular LTI system. If initial conditions were left as nonzero, we must deal with case-specific systems, where an LTI model with nonzero initial conditions can only be applicable for one system with those specific initial conditions.

2.3.2 LTI System Response to Exponential and Sinusoidal Signals

Using the convolution integral, we can generalize the LTI system response to a complex exponential signal $e^{j\omega t}$.

$$\begin{aligned} x(t) * h(t) &= \int_{-\infty}^{+\infty} x(t-\tau)h(\tau)\, d\tau \\ \implies e^{j\omega t} * h(t) &= \int_{-\infty}^{+\infty} e^{j\omega(t-\tau)}h(\tau)\, d\tau \\ &= e^{j\omega t} \underbrace{\int_{-\infty}^{+\infty} h(\tau)e^{-j\omega\tau}\, d\tau}_{H(\omega)}. \end{aligned}$$

Therefore, the LTI system response to a complex exponential can be depicted as:



Here, $H(\omega)$ is called the *frequency response* of the system. Note that $H(\omega)$ is independent of time t, is a complex function of $j\omega$, and is defined for an everlasting complex exponential. $H(\omega)$ also has conjugate symmetry, in which

$$H^*(\omega) = H(-\omega) \tag{2.7}$$

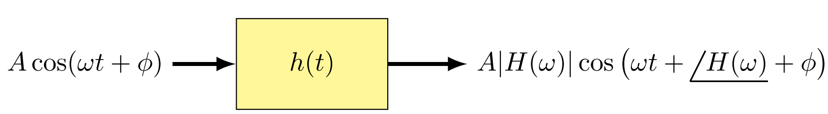
It then also follows that

$$|H(\omega)| = |H(-\omega)| \tag{2.8}$$

$$\angle H(-\omega) = -\angle H(\omega) \tag{2.9}$$

In a later chapter, the relationship between the impulse reponse and the frequency response will be explored. There exists some operator in which the impulse response can be mapped to the frequency response and vice versa.

Since sinusoids form the real and imaginary parts of a complex exponential, the LTI sinusoidal response can also be depicted:



Example 2.3.1. The input and output to an LTI system is given by

$$x(t) = u(t) + 2\cos(2t),$$
$$y(t) = u(t) - e^{-2t}u(t) + \sqrt{2}\cos(2t - 45°).$$

Suppose a different signal is now inputted to the very same LTI system and is defined as

$$x(t) = 5u(t-3) + 3\sqrt{2}\cos(2t - 60°).$$

Determine the new output $y(t)$.

SOLUTION

From the superposition principle, we know that as long as we know the response to each addend of the input signal, we can combine each of the responses to get the overall output.

By observation, we first analyze just the unit step addends.

$$u(t) \longrightarrow \boxed{\text{LTI}} \longrightarrow (1 - e^{-2t})u(t)$$

Since this is an LTI system, any delay in the input will have the same corresponding delay in the output. Additionally, an input signal that is scaled by some factor will be scaled by the same factor in the output. Therefore,

$$5u(t-3) \longrightarrow \boxed{\text{LTI}} \longrightarrow 5(1 - e^{-2(t-3)})u(t-3)$$

Next, we analyze just the sinusoidal addends.

$$2\cos(2t) \longrightarrow \boxed{\text{LTI}} \longrightarrow \sqrt{2}\cos(2t - 45°)$$

From the LTI sinusoidal response, it follows that

$$|H(\omega)| = \frac{\sqrt{2}}{2},$$
$$\angle H(\omega) = -45°.$$

Therefore, the LTI sinusoidal response of the new addend is given by

$$3\sqrt{2}\cos(2t - 60°) \longrightarrow \boxed{\text{LTI}} \longrightarrow 3\sqrt{2}\left(\frac{\sqrt{2}}{2}\right)\cos(2t - 60° - 45°)$$

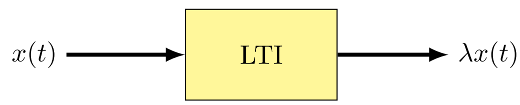
Summing the two LTI system responses, the overall output is given by

$$y(t) = 5[1 - e^{-2(t-3)}]u(t-3) + 3\cos(2t - 105°).$$

■

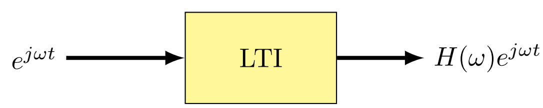
2.3.3 Eigenfunctions of LTI Systems

An *eigenfunction* of an LTI system is some function $x(t)$ such that the system response is $\lambda x(t)$ for some scalar λ (independent of t). That is,



For LTI systems characterized by $h(t)$, we can use the convolution integral to determine if a function is an eigenfunction of the system.

Previously, we saw that for a complex exponential input $e^{j\omega t}$, we get a scaled output $H(\omega)e^{j\omega t}$. Therefore, the complex exponential is an eigenfunction of the system. In fact, this particular eigenfunction yields the frequency response function $H(\omega)$, which is crucial for Fourier transforms in a later chapter.

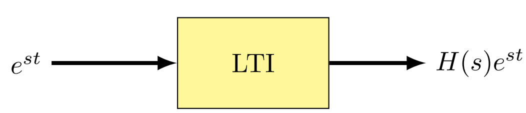


Example 2.3.2. Determine if e^{st} is an eigenfunction of LTI systems, for $s = \sigma + j\omega$.

SOLUTION

$$e^{st} * h(t) = \int_{-\infty}^{+\infty} e^{s(t-\tau)} h(\tau)\, d\tau = e^{st} \underbrace{\int_{-\infty}^{+\infty} h(\tau) e^{-s\tau}\, d\tau}_{H(s)}$$

Since $H(s)$ is a constant for a given s and is scaling e^{st}, the function e^{st} is an eigenfunction of LTI systems. In fact, it yields the transfer function $H(s)$, which is crucial for Laplace transforms in a later chapter.



■

Example 2.3.3. Determine if $3^t u(t)$ is an eigenfunction of LTI systems.

SOLUTION

$$3^t u(t) * h(t) = \int_{-\infty}^{+\infty} 3^{t-\tau} u(t-\tau) h(\tau)\, d\tau = 3^t \int_{-\infty}^{t} 3^{-\tau} h(\tau)\, d\tau$$

Since the integral is dependent on time t, it is not a constant. Therefore $3^t u(t)$ is not an eigenfunction of LTI systems.

■

2.3.4 Linear Constant-Coefficient Differential Equations (LCCDE)

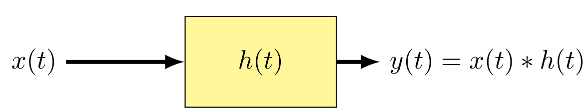
LTI systems can be characterized by a class of differential equations called *linear constant-coefficient differential equations* (LCCDE):

$$\sum_{k=0}^{N} a_{N-k} \frac{d^k y(t)}{dt^k} = \sum_{k=0}^{M} b_{M-k} \frac{d^k x(t)}{dt^k}. \tag{2.10}$$

However, higher-order nonhomogeneous LCCDEs are generally difficult to solve, which is why it is generally not covered until after Laplace and/or Fourier transforms are studied. Because of that, the LTI system response of LCCDEs will not be addressed until a later chapter.

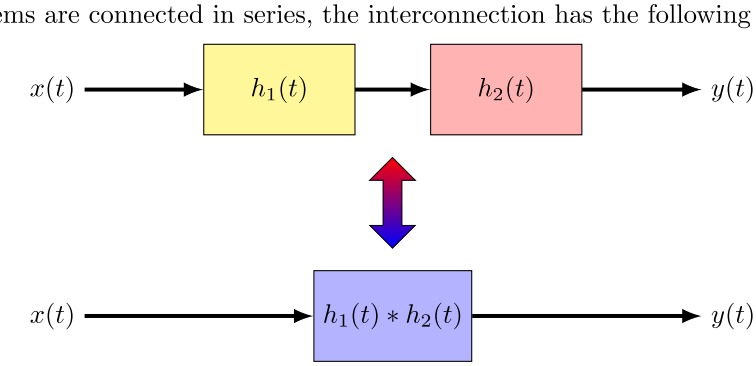
2.4 Interconnection of LTI Systems

We have briefly seen that the convolution of input $x(t)$ and impulse response $h(t)$ of an LTI system gives us the output $y(t)$ of the system.

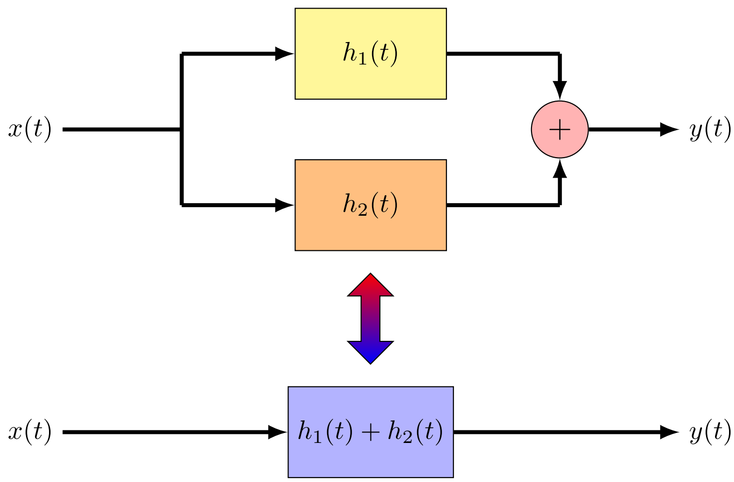


Suppose we begin introducing multiple LTI systems that are somehow connected to each other. When we are only interested in the corresponding output signal to an input signal, we do not particularly care about an individual LTI system – rather, we are more interested in the interconnection of LTI systems as a whole. In these cases, we want to find an equivalent impulse response that characterizes the interconnection and not just a single LTI system.

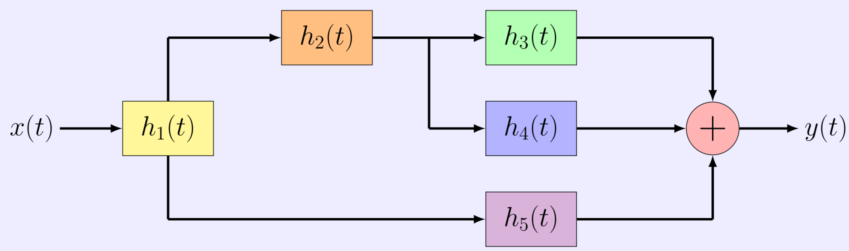
When LTI systems are connected in series, the interconnection has the following equivalence:



When LTI systems are connected in parallel, the interconnection has the following equivalence:



Example 2.4.1. Consider the following interconnection of LTI systems. Find the equivalent impulse response for the interconnection.



SOLUTION

In the top-right corner of the interconnection, two systems are in parallel. This parallel set is itself in series with $h_2(t)$. Therefore, the top branch can be rewritten as

$$h_2(t) * [h_3(t) + h_4(t)].$$

The top branch equivalent and the bottom branch are in parallel and can be rewritten as

$$h_2(t) * [h_3(t) + h_4(t)] + h_5(t).$$

Lastly, the two-branch equivalent and $h_1(t)$ are in series such that the overall equivalent impulse response is

$$h(t) = h_1(t) * \{h_2(t) * [h_3(t) + h_4(t)] + h_5(t)\}.$$

■

Chapter 3

Convolution

In the previous chapter, we were introduced to the concept of *convolution*, an operation ($*$) defined by the *convolution integral*

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

We have seen that the operation is commutative since by change of variables,

$$\begin{aligned} x(t) * h(t) &= \int_{-\infty}^{+\infty} x(\tau)h(t-\tau)\,d\tau \\ &= \int_{-\infty}^{+\infty} x(t-\tau)h(\tau)\,d\tau = h(t) * x(t). \end{aligned}$$

We have also already seen another convolution property in effect with the sampling property of the impulse function.

$$\begin{aligned} &\int_{-\infty}^{+\infty} x(t)\delta(t-T)\,dt = x(T) \\ \Longrightarrow &\int_{-\infty}^{+\infty} x(\tau)\delta(\tau-t)\,d\tau = \int_{-\infty}^{+\infty} x(\tau)\delta(t-\tau)\,d\tau = x(t) \\ &x(t) * \delta(t) = x(t). \end{aligned}$$

We also introduced the concept that any LTI system characterized by impulse response $h(t)$ will always have predictable outputs since the output is simply the convolution of the input signal and the impulse response, provided that all initial conditions of the system are zero. That is,

$$y(t) = x(t) * h(t) = \int_{-\infty}^{+\infty} x(\tau)h(t-\tau)\,d\tau. \tag{3.1}$$

We can even take this one step further. Since the physical world can only have causal signals and causal systems (with $x(t)$ and $h(t)$ both being causal) the output $y(t)$ must be causal as well. This is reflected in a modified version of the convolution integral,

$$y(t) = u(t)\int_{0}^{t} x(\tau)h(t-\tau)\,d\tau = u(t)\int_{0}^{t} h(\tau)x(t-\tau)\,d\tau. \tag{3.2}$$

In this chapter, different techniques to computing the convolution of two functions are explored.

3.1 Analytical Convolution

One way to compute the convolution is to directly solve the convolution integral itself. Refer to Table 3.1 for some convolution properties.

Table 3.1: Convolution properties.

Property	Description
Commutativity	$x(t) * h(t) = h(t) * x(t)$
Associativity	$x(t) * [h(t) * g(t)] = [x(t) * h(t)] * g(t)$
Distributivity	$x(t) * [h_1(t) + \cdots + h_N(t)] = [x(t) * h_1(t)] + \cdots + [x(t) * h_N(t)]$
Time-shift	$x(t - T_1) * h(t - T_2) = y(t - T_1 - T_2)$
Convolution with impulse	$x(t) * \delta(t) = x(t)$
Convolution with step	$x(t) * u(t) = \int_{-\infty}^{t} x(\tau)\, d\tau$ (ideal integrator)
Causal signals and systems	$y(t) = u(t) \int_{0}^{t} h(\tau) x(t - \tau)\, d\tau$
Width	$\text{Width}[y(t)] = \text{Width}[x(t)] + \text{Width}[h(t)]$
Area	$\text{Area}[y(t)] = \text{Area}[x(t)] + \text{Area}[h(t)]$
Differentation	$\left(\frac{d^m x(t)}{dt^m}\right) * \left(\frac{d^n h(t)}{dt^n}\right) = \frac{d^{m+n} y(t)}{dt^{m+n}}$
Integration	$\int_{-\infty}^{t} y(\tau)\, d\tau = x(t) * \left[\int_{-\infty}^{t} h(\tau)\, d\tau\right] = \left[\int_{-\infty}^{t} x(\tau)\, d\tau\right] * h(t)$

When integrating two causal functions, one of the functions will become time-reversed in the convolution integral. We can use this to our advantage in redefining the limits of integration. That is, for $f(\tau)$ representing the non-step factors,

$$\int_{-\infty}^{+\infty} f(\tau) \cdot \underbrace{u(\tau - a)}_{\tau - a > 0} \cdot \underbrace{u(b - \tau)}_{b - \tau > 0}\, d\tau = u(b - a) \int_{a}^{b} f(\tau)\, d\tau. \tag{3.3}$$

The inequalities above are defined for when the step functions are "switched on", and combining the two inequalities gives us an interval of validity $a < \tau < b$, which is what is used to update the convolution integral.

Example 3.1.1. Find $y(t) = x(t) * h(t)$ for

$$\begin{aligned} x(t) &= te^{-2t}u(t), \\ h(t) &= e^{-3t}u(t). \end{aligned}$$

SOLUTION

Using the convolution integral, it follows that

$$\begin{aligned} y(t) = [te^{-2t}u(t)] * [e^{-3t}u(t)] &= \int_{-\infty}^{+\infty} \tau e^{-2\tau} \underbrace{u(\tau)}_{\tau > 0} e^{-3(t-\tau)} \underbrace{u(t-\tau)}_{\tau < t} \, d\tau \\ &= u(t-0) \int_0^t \tau e^{-2\tau} e^{-3(t-\tau)} \, d\tau \\ &= e^{-3t}u(t) \int_0^t \tau e^{\tau} \, d\tau = e^{-3t}[te^t - e^t + 1]u(t). \end{aligned}$$

■

Example 3.1.2. Find $y(t) = x(t) * h(t)$ for

$$\begin{aligned} x(t) &= u(t-2) - u(t-4), \\ h(t) &= u(t-3) - u(t-5). \end{aligned}$$

SOLUTION

Using the convolution integral and the distributive property, it follows that

$$\begin{aligned} y(t) &= [u(t-2) - u(t-4)] * [u(t-3) - u(t-5)] \\ &= u(t-2) * u(t-3) - u(t-2) * u(t-5) - u(t-4) * u(t-3) + u(t-4) * u(t-5) \\ &= \int_{-\infty}^{+\infty} \underbrace{u(\tau-2)}_{\tau > 2} \underbrace{u((t-3)-\tau)}_{\tau < t-3} \, d\tau - \int_{-\infty}^{+\infty} \underbrace{u(\tau-2)}_{\tau > 2} \underbrace{u((t-5)-\tau)}_{\tau < t-5} \, d\tau \\ &\quad - \int_{-\infty}^{+\infty} \underbrace{u(\tau-4)}_{\tau > 4} \underbrace{u((t-3)-\tau)}_{\tau < t-3} \, d\tau + \int_{-\infty}^{+\infty} \underbrace{u(\tau-4)}_{\tau > 4} \underbrace{u((t-5)-\tau)}_{\tau < t-5} \, d\tau. \end{aligned}$$

Simplifying the integrals, we get

$$\begin{aligned} y(t) &= u((t-3)-2)\int_{2}^{t-3} d\tau - u((t-5)-2)\int_{2}^{t-5} d\tau \\ &\quad - u((t-3)-4)\int_{4}^{t-3} d\tau + u((t-5)-4)\int_{4}^{t-5} d\tau \\ &= (t-5)u(t-5) - 2(t-7)u(t-7) + (t-9)u(t-9) \\ &= r(t-5) - 2r(t-7) + r(t-9). \end{aligned}$$

■

While we are mostly interested in intervals of validity that follow causal signals and systems, there are other possible intervals of validity for different types of systems that may or may not be causal. Refer to Table 3.2 for these integral redefinitions.

Table 3.2: Integral redefinitions based on unit step functions.

$$\int_{-\infty}^{+\infty} f(\tau)u(\tau-a)\,d\tau = \int_{a}^{+\infty} f(\tau)\,d\tau \tag{3.4}$$

$$\int_{-\infty}^{+\infty} f(\tau)u(b-\tau)\,d\tau = \int_{-\infty}^{b} f(\tau)\,d\tau \tag{3.5}$$

$$\int_{-\infty}^{+\infty} f(\tau)u(\tau-a)u(b-\tau)\,d\tau = u(b-a)\int_{a}^{b} f(\tau)\,d\tau \tag{3.6}$$

$$\begin{aligned} \int_{-\infty}^{+\infty} f(\tau)u(\tau-a)u(\tau-b)\,d\tau &= \int_{\max(a,b)}^{+\infty} f(\tau)\,d\tau \\ &= u(a-b)\int_{a}^{+\infty} f(\tau)\,d\tau + u(b-a)\int_{b}^{+\infty} f(\tau)\,d\tau \end{aligned} \tag{3.7}$$

$$\begin{aligned} \int_{-\infty}^{+\infty} f(\tau)u(a-\tau)u(b-\tau)\,d\tau &= \int_{-\infty}^{\min(a,b)} f(\tau)\,d\tau \\ &= u(b-a)\int_{-\infty}^{a} f(\tau)\,d\tau + u(a-b)\int_{-\infty}^{b} f(\tau)\,d\tau \end{aligned} \tag{3.8}$$

Example 3.1.3. Find $y(t) = x(t) * h(t)$ for

$$\begin{aligned} x(t) &= e^{-t}u(t-3), \\ h(t) &= u(-t). \end{aligned}$$

SOLUTION

Using the convolution integral, it follows that

$$\begin{aligned}
y(t) = [e^{-t}u(t-3)] * [u(-t)] &= \int_{-\infty}^{+\infty} e^{-\tau} \underbrace{u(\tau-3)}_{\tau>3} \underbrace{u(\tau-t)}_{\tau>t} \, d\tau \\
&= \int_{\max(3,t)}^{+\infty} e^{-\tau} u(\tau-3) u(\tau-t) \, d\tau \\
&= u(3-t) \int_{3}^{+\infty} e^{-\tau} \, d\tau + u(t-3) \int_{t}^{+\infty} e^{-\tau} \, d\tau \\
&= e^{-3}u(3-t) + e^{-t}u(t-3).
\end{aligned}$$

■

Table 3.3: Commonly encountered convolutions.

$$u(t) * u(t) = t \cdot u(t) \tag{3.9}$$

$$e^{at}u(t) * u(t) = \left(\frac{e^{at}-1}{a}\right) u(t) \tag{3.10}$$

$$e^{at}u(t) * e^{bt}u(t) = \left(\frac{e^{at}-e^{bt}}{a-b}\right) u(t) \tag{3.11}$$

$$e^{at}u(t) * e^{at}u(t) = te^{at}u(t) \tag{3.12}$$

$$te^{at}u(t) * e^{bt}u(t) = \left(\frac{e^{bt}-e^{at}+(a-b)te^{at}}{(a-b)^2}\right) u(t) \tag{3.13}$$

$$te^{at}u(t) * e^{at}u(t) = \frac{1}{2} \cdot t^2 e^{at} u(t) \tag{3.14}$$

$$\delta(t-T_1) * \delta(t-T_2) = \delta(t-T_1-T_2) \tag{3.15}$$

3.2 Pointwise Graphical Convolution

Another way to compute the convolution is to interpret the convolution graphically. Here, we will first introduce the pointwise method, best used when only a few samples of the output are of interest. Since

$$y(t) = x(t) * h(t) = \int_{-\infty}^{+\infty} x(\tau)h(t-\tau) \, d\tau,$$

the output at a certain single time t is given by the area under the resultant curve formed by multiplying $x(\tau)$ and $h(t-\tau)$, which is a time-reversed and translated version of $h(\tau)$.

POINTWISE GRAPHICAL CONVOLUTION.

Step 1: Determine the set of time values of interest $t_0, ..., t_N$.

Step 2: Plot $x(\tau)$ and $h(-\tau)$ along the τ-axis.

Step 3: Let $t = t_0$ and shift $h(-\tau)$ to the right by t units to represent $h(t - \tau)$.

Step 4: Using the region of overlap between $x(\tau)$ and $h(t - \tau)$, there are many options to solve:

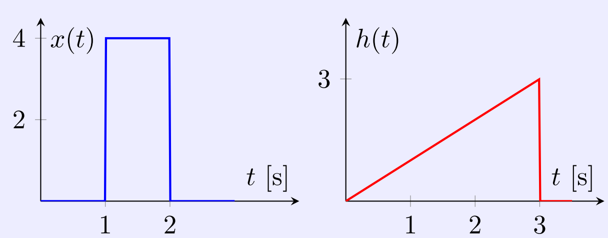
(a) if one curve is constant during the overlap and the other curve has a simple area under the curve, then $y(t)$ is the area under the curve but scaled by the constant.

(b) if both curves are linear during the overlap, then point-by-point multiplication can be performed, and the area under the new curve is taken to be $y(t)$.

(c) in general, integration during the overlap will always work.

Step 5: Repeat Steps 2–4 for the remaining time values of interest.

Example 3.2.1. Find $y(t) = x(t) * h(t)$, given the following plots.



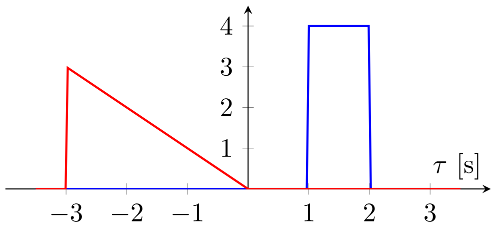
SOLUTION

From the width property of convolution, we know that the width of the output is the sum of the two widths, which is 5 seconds.

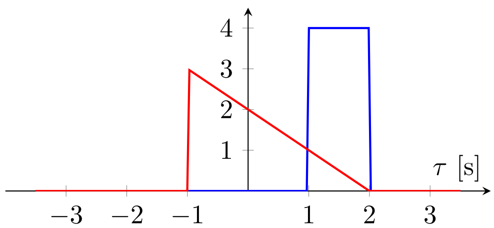
We also know that from the time-shift property of the convolution, since $x(t)$ is excited at $t = 0$ and $h(t)$ is excited at $t = 1$, the output $y(t)$ must be excited at $t = 1$.

From this information alone, we can already deduce that the output is only nonzero when $1 < t < 6$. Therefore, it is wise to choose values during that interval.

Suppose we are interested in only finding $y(2)$. First, we plot $x(\tau)$ and $h(-\tau)$ on the τ-axis.

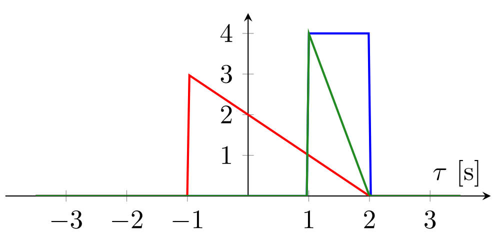


Since we want to observe at time $t = 2$, we shift the red plot to the right by 2 units:



Here, we notice that the region of overlap occurs during $1 < \tau < 2$. Note that the area under the red curve during that time is $\frac{1}{2} \times 1 \times 1 = \frac{1}{2}$, and the constant value of the blue plot during that time is 4. Therefore, $y(2)$ is the scaled area under the curve during the region of overlap. That is, $y(2) = \frac{1}{2} \times 4 = 2$.

Alternatively, we could have performed point-by-point multiplication and find that the area under the green curve in the following plot is still 2.



Of course, integration generally always works as well. During the region of overlap, we can let $x(\tau) = 4$ and $h(2 - \tau) = 2 - \tau$ such that

$$y(2) = \int_1^2 x(\tau)h(2-\tau)\, d\tau = \int_1^2 4(2-\tau)\, d\tau = 2.$$

We can repeat the process as many times as needed for other values of $t \in (1, 6)$ until we get enough samples to sketch $y(t)$.

■

3.3 Region-Based Graphical Convolution

The pointwise method is impractical if we are interested in the output response over a larger, if not entire, time domain. Here, we introduce the region-based graphical convolution, which utilizes a signal's *leading and trailing edges*, if any.

REGION-BASED GRAPHICAL CONVOLUTION.

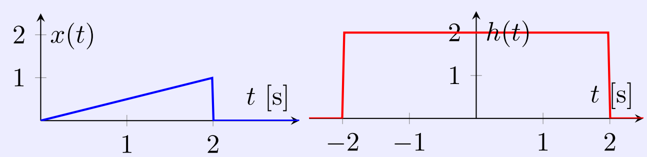
Step 1: Plot $x(\tau)$ and $h(-\tau)$ along the τ-axis, with $h(-\tau)$ being the simpler function to flip.

Step 2: Label the leading and trailing edges of $h(-\tau)$.

Step 3: Identify a region of overlap and use the edges to determine the interval of t-values. Then evaluate the convolution integral during that region of overlap.

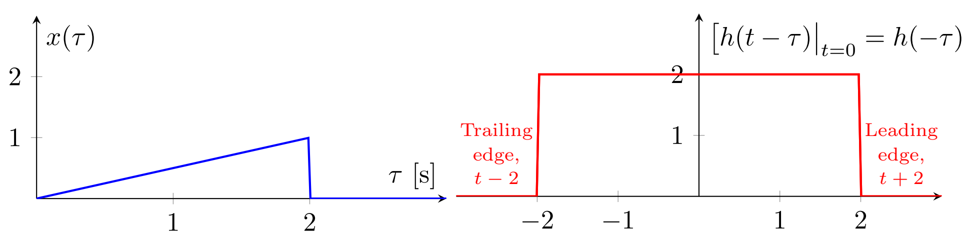
Step 4: Repeat Steps 2–3 for any other regions of overlap when shifting $h(-\tau)$ from left-to-right along the τ-axis. Keep track of the width of $h(t-\tau)$.

Example 3.3.1. Find $y(t) = x(t) * h(t)$, given the following plots.



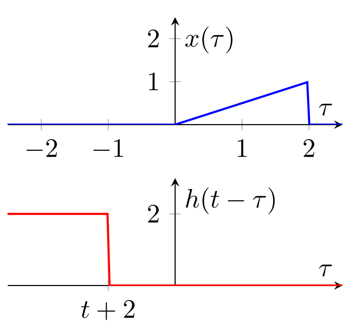
SOLUTION

First, plot $x(\tau)$ and $h(-\tau)$. Then identify the leading and trailing edges of $h(-\tau)$.



From $h(-\tau)$ above, we see that the leading edge is $\tau = t + 2$ and the trailing edge is $\tau = t - 2$. Now we place $h(t-\tau)$ to the furthest left and gradually shift to the right.

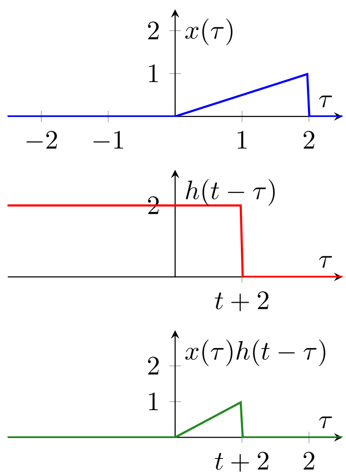
<u>Region 1: No overlap</u>



Leading edge: $t + 2 < 0 \Rightarrow t < -2$

No overlap, so $y(t) = 0$ for $t < -2$.

Region 2: Partial overlap

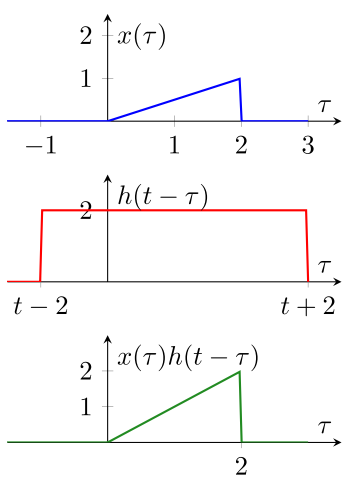


Leading edge: $0 < t+2 < 2 \Rightarrow -2 < t < 0$

$$y(t) = \int_{-\infty}^{+\infty} x(\tau)h(t-\tau)\,d\tau$$

$$\Longrightarrow y(t) = \int_{0}^{t+2} \left[\frac{1}{2}\tau\right][2]\,d\tau = \frac{1}{2}(t+2)^2$$

Region 3: Complete overlap

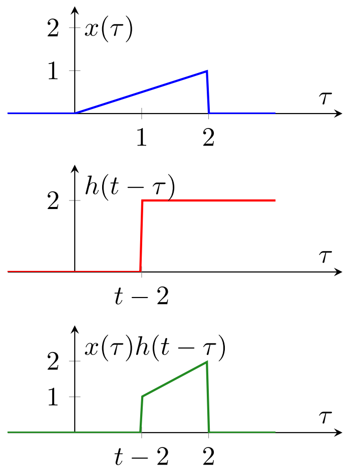


Leading edge: $t+2 > 2 \Rightarrow t > 0$
Trailing edge: $t-2 < 0 \Rightarrow t < 2$

$$\Longrightarrow 0 < t < 2$$

$$\Longrightarrow y(t) = \int_{0}^{2} \left[\frac{1}{2}\tau\right][2]\,d\tau = 2$$

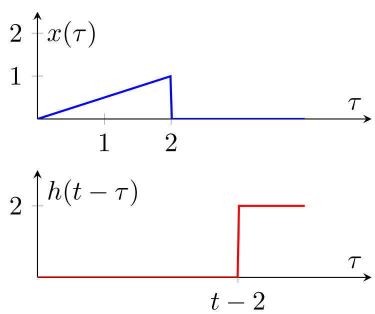
Region 4: Partial overlap



Trailing edge: $0 < t-2 < 2 \Rightarrow 2 < t < 4$

$$\Longrightarrow y(t) = \int_{t-2}^{2} \left[\frac{1}{2}\tau\right][2]\,d\tau = 2 - \frac{1}{2}(t-2)^2$$

Region 5: No overlap



Trailing edge: $t-2 > 2 \Rightarrow t > 4$

No overlap, so $y(t) = 0$ for $t > 4$.

Aggregating the outputs from the five regions, we can write a piecewise function for the overall output:

$$y(t) = \begin{cases} 0, & t < -2 \\ \frac{1}{2}(t+2)^2, & -2 < t < 0 \\ 2, & 0 < t < 2 \\ 2 - \frac{1}{2}(t-2)^2, & 2 < t < 4 \\ 0, & t > 4 \end{cases}.$$

This can also be written with unit step functions:

$$y(t) = \frac{1}{2}(t+2)^2[u(t+2) - u(t)] + 2[u(t) - u(t-2)] + \left[2 - \frac{1}{2}(t-2)^2\right][u(t-2) - u(t-4)].$$

Chapter 4

Laplace Transform

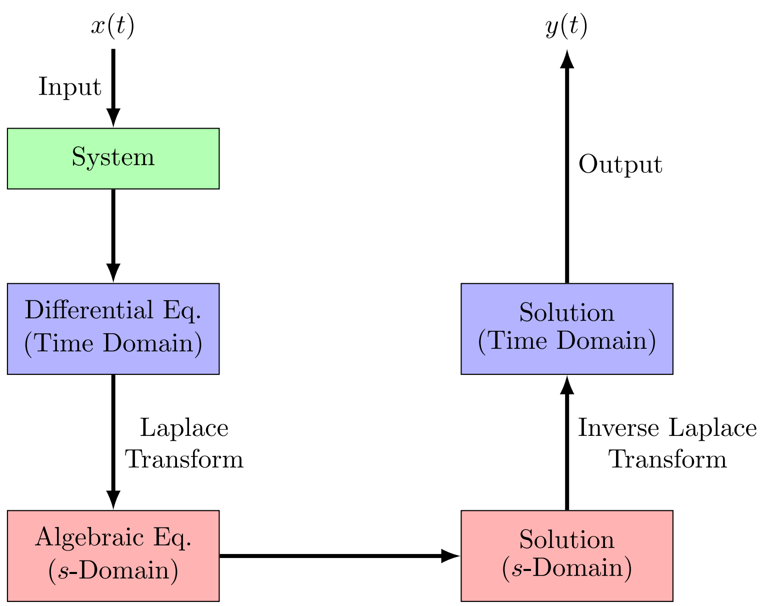
Previously, we have seen that LTI systems can be characterized by their impulse responses $h(t)$, and the outputs of those systems can be computed by taking the convolution of the input signal and the impulse response.

In fact, all LTI systems can be characterized by some impulse response. Going back to the affine transformation on the independent variable, its impulse response is given by

$$h(t) = \delta(at - b) = \frac{1}{|a|}\delta\left(t - \frac{b}{a}\right). \tag{4.1}$$

As seen in the first two chapters, some LTI systems can be represented by LCCDEs. While one could find the system response by solving the differential equation, it is more difficult to solve it that way. Here, we introduce a concept called the Laplace transform which offers a simpler way to solve differential equations by mapping them to a domain that uses algebraic equations.

As signals and systems in the physical world are causal, we are particularly interested in the unilateral (one-sided) Laplace transform.



4.1 Unilateral Laplace Transform

The *unilateral Laplace transform* of a signal $x(t)$ is defined as

$$X(s) = \mathcal{L}[x(t)] = \int_{0^-}^{\infty} x(t)e^{-st}\,dt, \tag{4.2}$$

where $s = \sigma + j\omega$ is the *complex frequency.*

The Laplace transform essentially is an operator that maps a signal defined in the *time domain* to another signal defined in the *s-domain.* $x(t)$ and $X(s)$ constitute a *Laplace transform pair.* This relationship can be written as

$$x(t) \iff X(s). \tag{4.3}$$

Because of this relationship, there exists an *inverse Laplace transform* defined by

$$x(t) = \mathcal{L}^{-1}[X(s)] = \frac{1}{j2\pi}\int_{\sigma-j\infty}^{\sigma+j\infty} X(s)e^{+st}\,ds, \tag{4.4}$$

though direct computation of the inverse integral is often more cumbersome than referencing a table of Laplace transform pairs.

Additionally, for signal $x(t)$ to be transformable, the unilateral Laplace integral must be absolutely integrable such that

$$\begin{aligned}\int_{0^-}^{\infty} |x(t)e^{-st}|\,dt &= \int_{0^-}^{\infty} |x(t)||e^{-(\sigma+j\omega)t}|\,dt \\ &= \int_{0^-}^{\infty} |x(t)||e^{-\sigma t}||e^{-j\omega t}|\,dt < \infty.\end{aligned} \tag{4.5}$$

From above, $|e^{-j\omega t}| < 1$ and σ is real. Therefore, the integrable condition simplifies to

$$\int_{0^-}^{\infty} |x(t)|e^{-\sigma t}\,dt < \infty. \tag{4.6}$$

The *region of convergence* (ROC) is then defined as the values for s for which the unilateral Laplace integral is absolutely integrable for a given signal $x(t)$ for all t. Here, the ROC is defined by a set of values for $\sigma = \mathrm{Re}(s)$.

For the unilateral Laplace transform, the ROC informs if there exists a unilateral Laplace transform for some signal $x(t)$ that can be mapped back to the time domain. Since we are more interested in causal signals and systems, signals that are physically realizable always have a unilateral Laplace transform.

Since all physically realizable signals and systems have a Laplace transform, the ROC is not necessary when modeling those entities. In fact, all causal signals have a unique unilateral Laplace transform.

Table 4.1 lists the properties of the unilateral Laplace transform.

Table 4.1: Properties of the unilateral Laplace transform.

Property	$x(t)$	$X(s) = \mathcal{L}[x(t)]$
Superposition	$K_1 x_1(t) + K_2 x_2(t)$	$K_1 X_1(s) + K_2 X_2(s)$
Time scaling	$x(at),\ a > 0$	$\frac{1}{a} X\left(\frac{s}{a}\right)$
Time shift	$x(t-T)u(t-T),\ T > 0$	$e^{-Ts} X(s)$
Frequency shift	$e^{-at} x(t)$	$X(s+a)$
Time 1st derivative	$x'(t) = \frac{dx(t)}{dt}$	$sX(s) - x(0^-)$
Time 2nd derivative	$x''(t) = \frac{d^2x(t)}{dt^2}$	$s^2 X(s) - sx(0^-) - x'(0^-)$
Time nth derivative	$x^{(n)}(t) = \frac{d^n x(t)}{dt^n}$	$s^n X(s) - \sum_{k=1}^{n} s^{n-k} x^{(k-1)}(0^-)$
Time integral	$\int_{0^-}^{t} x(\tau)\, d\tau$	$\frac{1}{s} X(s)$
Frequency derivative	$t \cdot x(t)$	$-\frac{dX(s)}{ds} = -X'(s)$
Frequency integral	$\frac{x(t)}{t}$	$\int_{s}^{\infty} X(\gamma)\, d\gamma$
Convolution	$x_1(t) * x_2(t)$	$X_1(s) X_2(s)$
Multiplication	$x_1(t) x_2(t)$	$X_1(s) * X_2(s)$
Initial value theorem	$x(0^+)$	$\lim_{s \to \infty} X(s)$
Final value theorem	$x(\infty)$	$\lim_{s \to 0} X(s)$

Example 4.1.1. Find the unilateral Laplace transform of $x(t) = e^{-at}u(t)$, for $a > 0$.

SOLUTION

Using the unilateral Laplace integral,

$$X(s) = \mathcal{L}[x(t)] = \int_{0^-}^{\infty} x(t) e^{-st}\, dt,$$

we get

$$X(s) = \int_{0^-}^{\infty} [e^{-at}u(t)]e^{-st}\,dt = \int_{0^-}^{\infty} e^{-(s+a)t}\,dt = \left[\frac{e^{-(s+a)t}}{-(s+a)}\right|_{0^-}^{\infty} = \frac{1}{s+a}.$$

■

Table 4.2 lists some common unilateral Laplace transform pairs. While one could find the inverse Laplace transform using the inverse Laplace integral, it is much simpler to reference a table of transform pairs.

Example 4.1.2. Suppose a signal $x(t) = u(t-2) - u(t-4)$ is inputted into an LTI system characterized by $h(t) = t \cdot u(t)$. Find the unilateral Laplace transform of the output signal $y(t)$.

SOLUTION

First, let $x_0(t) = u(t)$. From the convolution property, we can find the unilateral Laplace transforms of both $x_0(t)$ and $h(t)$, then take the product $X_0(s)H(s)$.

$$\begin{aligned} X_0(s) &= \int_{0^-}^{\infty} u(t)e^{-st}\,dt = \left[\frac{e^{-st}}{-s}\right|_{0^-}^{\infty} = \frac{1}{s} \\ H(s) &= \int_{0^-}^{\infty} t \cdot u(t)e^{-st}\,dt = \left[\frac{e^{-st}}{s^2}(-st-1)\right|_{0^-}^{\infty} = \frac{1}{s^2} \\ \implies Y_0(s) &= X_0(s)H(s) = \frac{1}{s^3} \end{aligned}$$

Using the table of unilateral transform pairs, we inverse-transform the product back to the time domain to get

$$\begin{aligned} y_0(t) = \mathcal{L}^{-1}[Y_0(s)] &= \mathcal{L}^{-1}\left[\frac{1}{s^3}\right] \\ &= \mathcal{L}^{-1}\left[\frac{1}{2} \cdot \frac{2}{s^3}\right] \\ &= \frac{1}{2} \cdot \mathcal{L}^{-1}\left[\frac{2}{s^3}\right] = \frac{1}{2} \cdot t^2 u(t). \end{aligned}$$

By the superposition principle for LTI systems, it follows that

$$x(t) = x_0(t-2) - x_0(t-4) \longrightarrow \boxed{h(t)} \longrightarrow y(t) = y_0(t-2) - y_0(t-4)$$

Therefore,

$$\begin{aligned} y(t) &= y_0(t-2) - y_0(t-4) \\ &= \frac{1}{2} \cdot (t-2)^2 u(t-2) - \frac{1}{2} \cdot (t-4)^2 u(t-4). \end{aligned}$$

■

Table 4.2: Unilateral Laplace transform pairs.

$x(t)$	$X(s) = \mathcal{L}[x(t)]$
$\delta(t)$	1
$u(t)$	$\frac{1}{s}$
$e^{-at}u(t)$	$\frac{1}{s+a}$
$t \cdot u(t)$	$\frac{1}{s^2}$
$t^2 \cdot u(t)$	$\frac{2}{s^3}$
$te^{-at}u(t)$	$\frac{1}{(s+a)^2}$
$t^2e^{-at}u(t)$	$\frac{2}{(s+a)^3}$
$t^{n-1}e^{-at}u(t)$	$\frac{(n-1)!}{(s+a)^n}$
$\sin(\omega_0 t)u(t)$	$\frac{\omega_0}{s^2+\omega_0^2}$
$\sin(\omega_0 t+\theta)u(t)$	$\frac{s\sin\theta+\omega_0\cos\theta}{s^2+\omega_0^2}$
$\cos(\omega_0 t)u(t)$	$\frac{s}{s^2+\omega_0^2}$
$\cos(\omega_0 t+\theta)u(t)$	$\frac{s\cos\theta-\omega_0\sin\theta}{s^2+\omega_0^2}$
$e^{-at}\sin(\omega_0 t)u(t)$	$\frac{\omega_0}{(s+a)^2+\omega_0^2}$
$e^{-at}\cos(\omega_0 t)u(t)$	$\frac{s+a}{(s+a)^2+\omega_0^2}$
$2e^{-at}\cos(bt-\theta)u(t)$	$\frac{e^{+j\theta}}{s+a+jb}+\frac{e^{-j\theta}}{s+a-jb}$
$e^{-at}\cos(bt-\theta)u(t)$	$\frac{(s+a)\cos\theta+b\sin\theta}{(s+a)^2+b^2}$
$\frac{2t^{n-1}}{(n-1)!}e^{-at}\cos(bt-\theta)u(t)$	$\frac{e^{+j\theta}}{(s+a+jb)^n}+\frac{e^{-j\theta}}{(s+a-jb)^n}$

4.2 Bilateral Laplace Transform

The *bilateral Laplace transform* of a signal $x(t)$ is defined as

$$X(s) = \mathcal{L}[x(t)] = \int_{-\infty}^{+\infty} x(t)e^{-st}\,dt. \tag{4.7}$$

Similar to the unilateral Laplace transform, for a signal $x(t)$ to be transformable, the bilateral Laplace integral must be absolutely integrable such that the simplified integrable condition given by

$$\int_{-\infty}^{+\infty} |x(t)|e^{-\sigma t}\,dt < \infty \tag{4.8}$$

holds true. However, the region of convergence (ROC) plays a bigger role for the bilateral Laplace transform.

While the unilateral Laplace transform refers to the ROC to validate the existence of a unilateral transform pair, the bilateral Laplace transform needs the ROC to correctly map the s-domain solution to the time domain. That is, the ROC informs the mapping for the inverse bilateral transform. Two different signals may have the same Laplace transform and only differ by their ROCs.

For instance, consider the causal signal $x(t) = e^{-at}u(t)$ with the integrable condition

$$\int_{-\infty}^{+\infty} |e^{-at}u(t)|e^{-\sigma t}\,dt = \int_{0^-}^{\infty} e^{-(a+\sigma)t}\,dt < \infty.$$

If $\mathrm{Re}(s) > -a$, then it follows that

$$\int_{0^-}^{\infty} e^{-(a+\sigma)t}\,dt = \left[\frac{e^{-(a+\sigma)t}}{-(a+\sigma)}\right|_{0^-}^{\infty} < \infty.$$

If $\mathrm{Re}(s) \leq -a$, then for $\epsilon = -(a+\sigma) \geq 0$, it follows that

$$\int_{0^-}^{\infty} e^{-(a+\sigma)t}\,dt = \int_{0^-}^{\infty} e^{+\epsilon t}\,dt = \left[\frac{e^{\epsilon t}}{\epsilon}\right|_{0^-}^{\infty} = \infty.$$

However, the anticausal signal $x_A(t) = -e^{-at}u(-t)$ has an ROC of $\mathrm{Re}(s) \leq -a$ such that

$$\int_{-\infty}^{+\infty} |-e^{-at}u(-t)|e^{-(a+\sigma)t}\,dt = \int_{-\infty}^{0^+} e^{-(a+\sigma)t}\,dt = \left[\frac{e^{-(a+\sigma)t}}{-(a+\sigma)}\right|_{-\infty}^{0^+} < \infty.$$

Here, we see two signals – one of which is causal and the other anticausal – that would have the same bilateral Laplace transform $X(s) = \frac{1}{s+a}$ but with different ROCs.

Generally, while the unilateral Laplace transform is convenient for solving differential equations with initial conditions, the bilateral Laplace transform provides insight into system characteristics such as stability, causality, and frequency response.

However, we can avoid using the bilateral Laplace transform (and avoid ROC calculations) when analyzing system characteristics by setting zero initial conditions and/or using other tools such as the Fourier transform, which will be covered in a later chapter.

For the rest of the text, when referring to the Laplace transform, we will be using the unilateral form.

4.3 Partial Fraction Expansion

When converting a solution from s-domain back to time domain, we will often have a s-domain solution in the form of a rational function that may need to be rewritten as a sum of parts, with each part having a Laplace transform pair.

Consider an s-domain function with the form

$$X(s) = \frac{N(s)}{D(s)}, \tag{4.9}$$

written as a ratio of a polynomial numerator $N(s)$ to a polynomial denominator $D(s)$. The roots of the numerator are called the *zeros* of $X(s)$ and are defined by

$$N(s) = 0, \tag{4.10}$$

whereas the roots of the denominator are called the *poles* of $X(s)$ and are defined by

$$D(s) = 0. \tag{4.11}$$

Before even performing partial fraction expansion, first we examine the orders of the numerator and denominator. Let

$$m = \deg[N(s)], \tag{4.12}$$
$$n = \deg[D(s)]. \tag{4.13}$$

There are three types of rational functions, each with a different task at hand.

- $m < n$: $X(s)$ is *strictly proper*. It is ready for partial fraction expansion, if needed.
- $m = n$: $X(s)$ is *proper*. Perform long division and express $X(s)$ in terms of the quotient and remainder, then determine if the remainder term needs partial fraction expansion.
- $m > n$: $X(s)$ is *improper*. Perform long division and express $X(s)$ in terms of the quotient and remainder, then determine if the remainder term needs partial fraction expansion.

A term would need partial fraction expansion if there is not an easily matching Laplace transform pair provided by the table. Usually this occurs when $\deg[D(s)] > 2$.

Example 4.3.1. Determine if

$$X(s) = \frac{2s^2 + 8s + 6}{s^2 + 2s + 1}$$

needs to undergo partial fraction expansion.

SOLUTION

Since both the numerator and denominator are of the same polynomic order, first we need to long divide.

$$
\begin{array}{r}
2 \\
s^2+2s+1\overline{\big)\; 2s^2+8s+6} \\
-\,2s^2-4s-2 \\
\hline
4s+4
\end{array}
$$

From here, we can rewrite $X(s)$ as

$$X(s) = 2 + \frac{4s+4}{s^2+2s+1} = 2 + \frac{4(s+1)}{(s+1)^2} = 2 + \frac{4}{s+1}.$$

Since both addends are similar to the ones in the Laplace transform pairs table, partial fraction expansion is not needed, and

$$x(t) = \mathcal{L}^{-1}\left[2 + \frac{4}{s+1}\right] = 2\delta(t) + 4e^{-t}u(t).$$

■

Example 4.3.2. Determine if

$$X(s) = \frac{s^4+4}{s^3-3s-1}$$

needs to undergo partial fraction expansion.

SOLUTION

Since both the numerator and denominator are of the same polynomic order, first we need to long divide.

$$
\begin{array}{r}
s \\
s^3-3s-1\overline{\big)\; s^4 \qquad\qquad +4} \\
-\,s^4+3s^2+s \qquad \\
\hline
3s^2+s+4
\end{array}
$$

From here, we can rewrite $X(s)$ as

$$X(s) = s + \frac{3s^2+s+4}{s^3-3s-1}.$$

Since the remainder term does not have a matching Laplace transform pair, partial fraction expansion is needed.

■

Given that a remainder term requires partial fraction expansion, all poles p_n of $X(s)$ must be solved first so that $D(s)$ is expressed in factored form. That is,

$$X(s) = \frac{N(s)}{D(s)} = \frac{N(s)}{\prod_{\ell=1}^{n}(s-p_\ell)}. \tag{4.14}$$

There are four different cases for partial fraction expansion. Depending on the poles, one or more cases may apply. Each case requires using a technique called the *residue method*, which evaluates a modified expression at a pole $s = p$ so that factors cancel and expansion coefficients can be calculated.

Case 1: Distinct real poles.

Suppose there are n distinct real poles in $X(s)$. Then $X(s)$ can be expanded as

$$X(s) = \frac{N(s)}{(s-p_1)(s-p_2)\cdots(s-p_n)} = \frac{A_1}{s-p_1} + \frac{A_2}{s-p_2} + \cdots + \frac{A_n}{s-p_n}. \tag{4.15}$$

The expansion coefficients for Case 1 is then given by

$$A_n = (s-p_n)X(s)\big|_{s=p_n}. \tag{4.16}$$

Example 4.3.3. Use the residue method to expand $X(s) = \dfrac{s^2-4s+3}{s(s+1)(s+3)}$.

SOLUTION

First, set up the expanded form

$$X(s) = \frac{A_1}{s} + \frac{A_2}{s+1} + \frac{A_3}{s+3}.$$

The expansion coefficients are given by

$$A_1 = sX(s)\big|_{s=0} = s\frac{s^2-4s+3}{s(s+1)(s+3)}\bigg|_{s=0} = \frac{s^2-4s+3}{(s+1)(s+3)}\bigg|_{s=0} = 1$$

$$A_2 = (s+1)X(s)\big|_{s=-1} = (s+1)\frac{s^2-4s+3}{s(s+1)(s+3)}\bigg|_{s=-1} = \frac{s^2-4s+3}{s(s+3)}\bigg|_{s=0} = -4$$

$$A_3 = (s+3)X(s)\big|_{s=-3} = (s+3)\frac{s^2-4s+3}{s(s+1)(s+3)}\bigg|_{s=-1} = \frac{s^2-4s+3}{s(s+1)}\bigg|_{s=0} = 4$$

The full expansion is then

$$X(s) = \frac{1}{s} - \frac{4}{s+1} + \frac{4}{s+3}.$$

Here, we see that inverse Laplace transform is

$$\begin{aligned} x(t) &= \mathcal{L}^{-1}\left[\frac{1}{s} - \frac{4}{s+1} + \frac{4}{s+3}\right] \\ &= u(t) - 4e^{-t}u(t) + 4e^{-3t}u(t). \end{aligned}$$

■

Case 2: Repeated real poles.

Suppose there is a repeated pole at $s = p$. Then $X(s)$ can be expanded as

$$X(s) = \frac{N(s)}{(s-p)^m} = \frac{B_1}{s-p} + \frac{B_2}{(s-p)^2} + \cdots + \frac{B_m}{(s-p)^m}, \tag{4.17}$$

with expansion coefficients given by

$$B_m = (s-p)^m X(s)\big|_{s=p} \tag{4.18}$$

$$B_{m-1} = \left\{\frac{d}{ds}\left[(s-p)^m X(s)\right]\right\}\bigg|_{s=p} \tag{4.19}$$

$$B_{m-2} = \left\{\frac{1}{2!} \cdot \frac{d^2}{ds^2}\left[(s-p)^m X(s)\right]\right\}\bigg|_{s=p}, \tag{4.20}$$

or more generally,

$$B_n = \left\{\frac{1}{(m-n)!} \cdot \frac{d^{m-n}}{ds^{m-n}}\left[(s-p)^m X(s)\right]\right\}\bigg|_{s=p}, \quad \text{for } n = 1, 2, ..., m. \tag{4.21}$$

Example 4.3.4. Use the residue method to expand $X(s) = \dfrac{s^2+3s+3}{(s+2)(s+3)^3}$.

SOLUTION

Setting up the expanded form,

$$X(s) = \frac{A_1}{s+2} + \frac{B_1}{s+3} + \frac{B_2}{(s+3)^2} + \frac{B_3}{(s+3)^3}.$$

The expansion coefficients are given by

$$A_1 = sX(s)\big|_{s=-2} = s\frac{s^2+3s+3}{(s+2)(s+3)^3}\bigg|_{s=-2} = \frac{s^2+3s+3}{(s+3)^3}\bigg|_{s=-2} = 1$$

$$B_3 = (s+3)^3X(s)\big|_{s=-3} = (s+3)^3\frac{s^2+3s+3}{(s+2)(s+3)^3}\bigg|_{s=-3} = \frac{s^2+3s+3}{s+2}\bigg|_{s=-3} = -3$$

$$B_2 = \left\{\frac{d}{ds}[(s+3)^3X(s)]\right\}\bigg|_{s=-3} = \frac{d}{ds}\left[\frac{s^2+3s+3}{s+2}\right]\bigg|_{s=-3} = \frac{s^2+4s+3}{(s+2)^2}\bigg|_{s=-3} = 0$$

$$B_1 = \left\{\frac{1}{2!}\cdot\frac{d^2}{ds^2}[(s+3)^3X(s)]\right\}\bigg|_{s=-3} = \frac{1}{2}\cdot\frac{d^2}{ds^2}\left[\frac{s^2+3s+3}{s+2}\right]\bigg|_{s=-3} = \frac{1}{2}\cdot\frac{2}{(s+2)^3}\bigg|_{s=-3} = -1$$

The full expanded form is then

$$X(s) = \frac{1}{s+2} - \frac{1}{s+3} - \frac{3}{(s+3)^3},$$

with inverse Laplace transform

$$\begin{aligned} x(t) &= e^{-2t}u(t) - e^{-3t}u(t) - \frac{3}{2}t^2e^{-3t}u(t) \\ &= \left[e^{-2t} - e^{-3t} - \frac{3}{2}t^2e^{-3t}\right]u(t). \end{aligned}$$

■

Case 3: Distinct complex poles.

If $X(s)$ has a complex pole at $s = p$, it also follows that its complex conjugate $s = p^*$ is also a pole. Interestingly, the expansion coefficients associated with both complex poles are also complex conjugates of each other. That is,

$$X(s) = \frac{N(s)}{s^2+as+b} = \frac{C}{s-p} + \frac{C^*}{s-p^*}, \tag{4.22}$$

with expansion coefficient

$$C = (s-p)X(s)\big|_{s=p}. \tag{4.23}$$

Example 4.3.5. Use the residue method to expand $X(s) = \dfrac{s(s-8)}{(s+2)(s^2+16)}$.

SOLUTION

It follows that for $s^2+16=0$, there are poles at $s = \pm j4$, along with $s = -2$. The expanded form then is given by

$$X(s) = \frac{A_1}{s+2} + \frac{C_1}{s+j4} + \frac{C_1^*}{s-j4},$$

with expansion coefficients

$$A_1 = (s+2)X(s)\big|_{s=-2} = (s+2)\frac{s(s-8)}{(s+2)(s^2+16)}\bigg|_{s=-2} = \frac{s(s-8)}{(s^2+16)}\bigg|_{s=-2} = 1$$

$$C_1 = (s+j4)X(s)\big|_{s=-j4} = \frac{s(s-8)}{(s+2)(s-j4)}\bigg|_{s=-j4} = -j = e^{-j\pi/2}$$

$$C_1^* = C_1\big|_{j\leftarrow(-j)} = e^{+j\pi/2}$$

The full expanded form is then

$$X(s) = \frac{1}{s+2} + \frac{e^{-j\pi/2}}{s+j4} + \frac{e^{+j\pi/2}}{s-j4},$$

with inverse Laplace transform

$$x(t) = e^{-2t}u(t) + 2\cos(4t+\pi/2)u(t).$$

Case 4: Repeated complex poles.

Lastly, for repeated complex poles, the expansion is given by

$$\begin{aligned} X(s) &= \frac{N(s)}{(s^2+as+b)^m} = \frac{N(s)}{(s-p)^m(s-p^*)^m} \\ &= \left(\frac{D_1}{s-p} + \frac{D_2}{(s-p)^2} + \cdots + \frac{D_m}{(s-p)^m}\right) + \left(\frac{D_1^*}{s-p^*} + \frac{D_2^*}{(s-p^*)^2} + \cdots + \frac{D_m^*}{(s-p^*)^m}\right), \end{aligned} \tag{4.24}$$

with expansion coefficients

$$D_n = \left\{\frac{1}{(m-n)!}\cdot\frac{d^{m-n}}{ds^{m-n}}[(s-p)^m X(s)]\right\}\bigg|_{s=p}, \quad \text{for } n = 1, 2, ..., m. \tag{4.25}$$

Table 4.3 lists the corresponding Laplace transform pairs for the four partial fraction expansion cases.

Table 4.3: **Unilateral Laplace transform pairs based on poles.**

Pole	$X(s)$	$x(t) = \mathcal{L}^{-1}[X(s)]$
Distinct real	$\dfrac{A}{s+a}$	$Ae^{-at}u(t)$
Repeated real	$\dfrac{A}{(s+a)^n}$	$\dfrac{At^{n-1}}{(n-1)!}e^{-at}u(t)$
Distinct complex	$\dfrac{Ae^{+j\theta}}{s+a+jb} + \dfrac{Ae^{-j\theta}}{s+a-jb}$	$2Ae^{-at}\cos(bt-\theta)u(t)$
Repeated complex	$\dfrac{Ae^{+j\theta}}{(s+a+jb)^n} + \dfrac{Ae^{-j\theta}}{(s+a-jb)^n}$	$\dfrac{2At^{n-1}}{(n-1)!}e^{-at}\cos(bt-\theta)u(t)$

4.4 Transfer Function

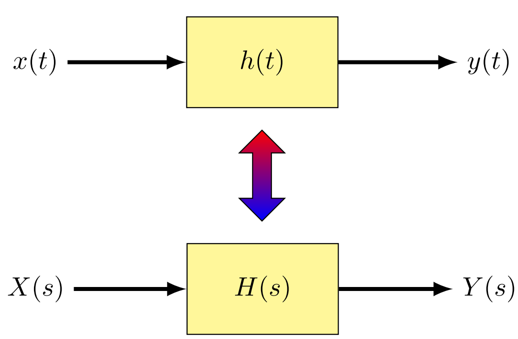
Given that all initial conditions of the input and output signals of an LTI system are zero, the *transfer function* is the Laplace transform of the impulse response and is defined as

$$H(s) = \frac{Y(s)}{X(s)} = \mathcal{L}[h(t)]. \tag{4.26}$$

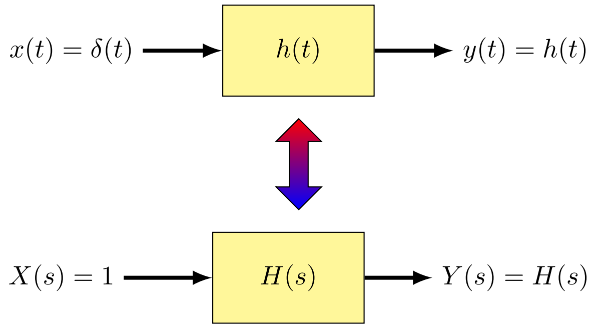
In fact, by the convolution property, it follows that

$$y(t) = x(t) * h(t) \iff Y(s) = X(s)H(s). \tag{4.27}$$

Symbolically,



As the impulse response is the system response to input $x(t) = \delta(t)$, consider the Laplace transform of the impulse signal $\mathcal{L}[\delta(t)] = 1$. Then it also follows that



To find the transfer function of an LTI system, there are two approaches:

1. Find $X(s), Y(s)$. Then calculate $H(s) = Y(s)/X(s)$.
2. Find $X(s), Y(s)$. Then substitute $X(s) = 1$ and $Y(s) = H(s)$.

Typically, the transfer function $H(s)$ is used to analyze the characteristics of the LTI system. Sometimes, when one is interested in representing systems solely for the purpose of finding an output response, the *conditional transfer function* $H_C(s)$ is calculated, where initial conditions are not necessarily zero; this is when the substitution approach works best. However, when including nonzero initial conditions, the system can no longer be called LTI.

Example 4.4.1. Find the transfer function for an LTI system characterized by LCCDE

$$\begin{gathered} \frac{d^2y(t)}{dt^2} + 3\frac{dy(t)}{dt} + 2y(t) = \frac{dx(t)}{dt} + 3x(t), \\ x(0^-) = 0, \\ y(0^-) = y'(0^-) = 1. \end{gathered}$$

Then find the conditional transfer function of the system.

SOLUTION

To find the transfer function $H(s)$, set all initial conditions to zero and take the Laplace transform of the LCCDE.

$$\begin{gathered} \mathcal{L}\left[\frac{d^2y(t)}{dt^2} + 3\frac{dy(t)}{dt} + 2y(t)\right] = \mathcal{L}\left[\frac{dx(t)}{dt} + 3x(t)\right] \\ s^2Y(s) + 3sY(s) + 2Y(s) = sX(s) + 3X(s) \\ Y(s)[s^2 + 3s + 2] = X(s)[s + 3] \end{gathered}$$

Rearranging the equation, we get the transfer function

$$H(s) = \frac{Y(s)}{X(s)} = \frac{s+3}{s^2+3s+2}.$$

Now we consider the initial conditions to solve the conditional transfer function.

$$[s^2Y(s) - s - 1] + 3[sY(s) - 1] + 2Y(s) = [sX(s) - 0] + 3X(s)$$
$$Y(s)[s^2 + 3s + 2] - (s + 4) = X(s)[s + 3].$$

Now let $X(s) = 1$ and $Y(s) = H_C(s)$. The conditional transfer function is given by

$$H_C(s)[s^2 + 3s + 2] - (s + 4) = [1][s + 3]$$
$$\implies H_C(s) = \frac{2s + 7}{s^2 + 3s + 2}$$

For the rest of this text, unless explicitly specified, only transfer functions with zero initial conditions will be considered.

4.4.1 System Stability

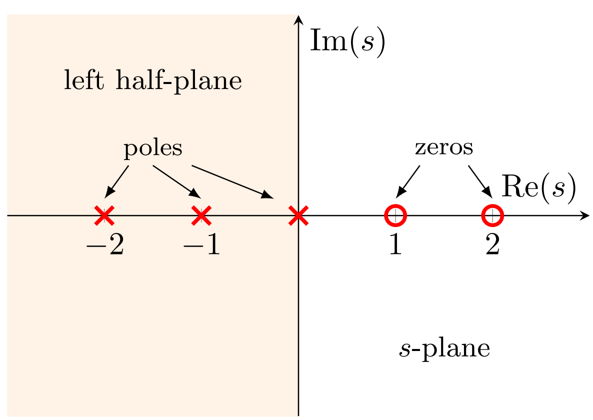
As previously introduced, zeros are the roots of the numerator of a transfer function ($N(s) = 0$), and poles are the roots of the denominator of a transfer function ($D(s) = 0$). We can visualize both zeros and poles on the s-*plane*, which plots $\text{Im}(s)$ versus $\text{Re}(s)$. As seen in Figure 4.1, zeros are plotted with circles, and poles are plotted with crosses.

The *open left half-plane* (OLHP) refers to the region of the s-plane which describes $\text{Re}(s) < 0$.

- If an LTI system with transfer function $H(s)$ is strictly proper or proper, then the system is BIBO stable if and only if its poles reside in the OLHP.
- If $H(s)$ is improper, then the system is unstable.

Note that the imaginary axis is not included in the OLHP; a system with poles on the imaginary axis is not stable.

Figure 4.1: s-plane plot of a system with zeros and poles.



Example 4.4.2. Determine if an LTI system with the following impulse response is BIBO stable:

$$h(t) = (4 + j5)e^{(2+j3)t}u(t) + (4 - j5)e^{(2-j3)t}u(t)$$

SOLUTION

First, find the transfer function by taking the Laplace transform of the impulse response.

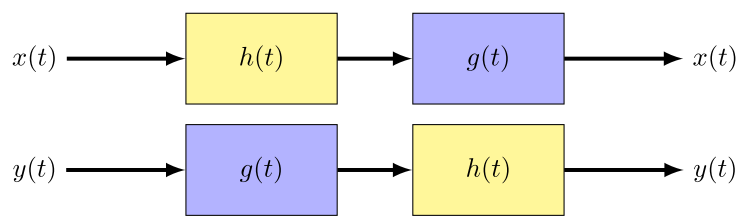
$$\begin{aligned} H(s) = \mathcal{L}[h(t)] &= (4 + j5)\mathcal{L}[e^{(2+j3)t}u(t)] + (4 - j5)\mathcal{L}[e^{(2-j3)t}u(t)] \\ &= \frac{4 + j5}{s - 2 - j3} + \frac{4 - j5}{s - 2 + j3} \\ &= \frac{8s - 46}{(s - 2 - j3)(s - 2 + j3)} \end{aligned}$$

We can see that the poles are given by $p_1 = 2 + j3$ and $p_2 = 2 - j3$, and the zero is given by $z_1 = 46/8$. Since $\text{Re}(p_k) > 0$, the poles are not in the OHLP, and thus the system is unstable.

■

4.4.2 Invertible Systems

As previously introduced, a system is invertible if there exists an inverse system that maps an output signal $y(t)$ back to its input $x(t)$. In the LTI case, an inverse system characterized by $g(t)$ exists such that



While an inverse system $g(t)$ may exist, it does not necessarily mean it is stable. In fact, for

$$G(s) = \frac{1}{H(s)}, \tag{4.28}$$

in order for a BIBO stable and causal LTI system $H(s)$ to have a stable and causal inverse LTI system $G(s)$, the transfer function $H(s)$ must be a proper rational function (i.e., $\deg[N(s)] = \deg[D(s)]$) with both its poles and zeros all residing in the OLHP of the s-plane. Such a system is called a *minimum phase system.*

Example 4.4.3. An LTI system is characterized by its impulse response

$$h(t) = \delta(t) - 4e^{-3t}u(t).$$

Determine if it is a minimum phase system.

SOLUTION

$$\begin{aligned} H(s) = \mathcal{L}[h(t)] &= \mathcal{L}[\delta(t)] - 4\mathcal{L}[e^{-3t}u(t)] \\ &= 1 - \frac{4}{s+3} = \frac{s-1}{s+3}. \end{aligned}$$

From above, there is a pole $p = -3$ and a zero $z = 1$. While the forward system may be stable for $\text{Re}(p) < 0$, the inverse is not since $\text{Re}(z) > 0$. Therefore, the LTI system is not a minimum phase system.

■

4.5 System Response of Second-Order LCCDEs

Consider second-order LCCDEs of the form

$$\frac{d^2y(t)}{dt^2} + a_1\frac{dy(t)}{dt} + a_2y(t) = b_1\frac{dx(t)}{dt} + b_2x(t). \tag{4.29}$$

Setting initial conditions to zero such that the system is LTI, the transfer function is given by

$$\begin{aligned} s^2Y(s) + a_1sY(s) + a_2Y(s) &= b_1sX(s) + b_2X(s) \\ \implies H(s) = \frac{Y(s)}{X(s)} &= \frac{b_1s + b_2}{s^2 + a_1s + a_2s}. \end{aligned} \tag{4.30}$$

Define the following system attributes

- Attenuation coefficient [Np/s]: $\alpha = \dfrac{a_1}{2}$
- Undamped natural frequency [rad/s]: $\omega_0 = \sqrt{a_2}$
- Damping coefficient [unitless]: $\xi = \dfrac{\alpha}{\omega_0} = \dfrac{a_1}{2\sqrt{a_2}}$

such that the solutions to the *characteristic equation*

$$s^2 + a_1s + a_2 = 0 \tag{4.31}$$

can be written as

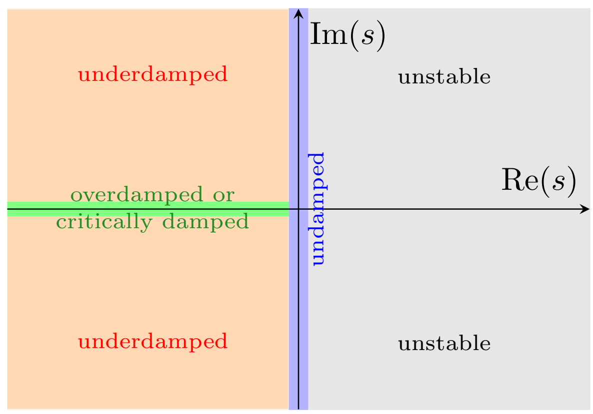
$$s = \frac{-a_1 \pm \sqrt{a_1^2 - 4a_2}}{2} = -\alpha \pm \sqrt{\alpha^2 - \omega_0^2} = \omega_0\left[-\xi \pm \sqrt{\xi^2 - 1}\right]. \tag{4.32}$$

In fact, the solutions to the characteristic equation are the poles p_1, p_2 of the system. Since $\text{Re}(p_k) < 0$ must be true for a system to be BIBO stable, it also holds from above that $a_1 > 0$ and $a_2 > 0$ must be true for the system to be BIBO stable.

Depending on the damping coefficient ξ, there are five possible system responses for an input impulse or step signal:

- $\xi > 1$: overdamped response
- $\xi = 1$: critically damped response
- $0 < \xi < 1$: underdamped response
- $\xi = 0$: undamped response
- $\xi < 0$: unstable response

Figure 4.2: Second-order system response on the s-plane.



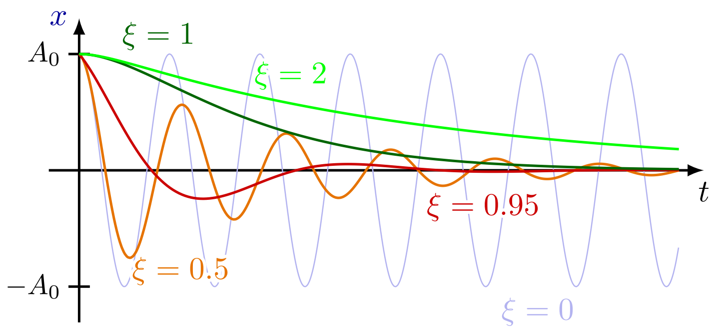
We are particularly interested in the first three damping responses as they describe how a system response reaches *steady-state*, or *equilibrium*. That is, it describes how fast and in what manner a system approaches the final value of the system response.

As seen in Figure 4.3, overdamped responses slowly approach steady-state, whereas underdamped responses quickly but oscillatorily approach steady-state at a *damped natural frequency* of

$$\omega_d = \sqrt{\omega_0^2 - \alpha^2} = \omega\sqrt{1 - \xi^2}. \tag{4.33}$$

Critically damped responses represent the quickest non-oscillatory path to steady-state.

Figure 4.3: Damping responses.



Example 4.5.1. An LTI system is described by the LCCDE

$$\frac{d^2y(t)}{dt^2} + B\frac{dy(t)}{dt} + 25y(t) = \frac{dx(t)}{dt} + 23x(t).$$

Compute the range of values for constant B such that the system impulse response is:

(a) overdamped

(b) critically damped

(c) underdamped

Then let $B = 26$ and compute the impulse and step responses of the system.

SOLUTION

The characteristic equation of the LCCDE is

$$s^2 + Bs + 25 = 0.$$

The damping coefficient is then given by

$$\xi = \frac{a_1}{2\sqrt{a_2}} = \frac{B}{2\sqrt{25}} = \frac{B}{10}.$$

It then follows that

- overdamped: $\xi > 1 \Longrightarrow B > 10$
- critically damped: $\xi = 1 \Longrightarrow B = 10$
- underdamped: $0 < \xi < 1 \Longrightarrow 0 < B < 10$

Letting $B = 26$ and setting initial conditions to zero, the Laplace transform of the LCCDE is

$$Y(s)[s^2 + 26s + 25] = X(s)[s + 23].$$

The transfer function is given by

$$\begin{aligned} H(s) &= \frac{Y(s)}{X(s)} = \frac{s+23}{s^2+26s+25} = \frac{s+23}{(s+1)(s+25)} = \frac{A_1}{s+1} + \frac{A_2}{s+25}, \\ A_1 &= (s+1)H(s)\big|_{s=-1} = 22/24 = 11/12, \\ A_2 &= (s+25)H(s)\big|_{s=-25} = (-2)/(-24) = 1/12. \end{aligned}$$

Therefore the impulse response is

$$h(t) = \mathcal{L}^{-1}[H(s)] = \frac{1}{12}\left[11e^{-t}u(t) + e^{-25t}u(t)\right].$$

The step response can be calculated from

$$y_{step}(t) = \int_{-\infty}^{t} h(\tau)\, d\tau \iff Y_{step}(s) = \frac{1}{s}H(s).$$

$$\begin{aligned} Y_{step}(s) &= \frac{s+23}{s(s+1)(s+25)} = \frac{A_0}{s} + \frac{A_1}{s+1} + \frac{A_2}{s+25}, \\ A_0 &= sY_{step}(s)\big|_{s=0} = 23/25 \\ A_1 &= (s+1)Y_{step}(s)\big|_{s=-1} = 22/(-24) = -11/12, \\ A_2 &= (s+25)Y_{step}(s)\big|_{s=-25} = (-2)/600 = -1/300. \end{aligned}$$

$$y_{step}(t) = \mathcal{L}^{-1}[Y_{step}(s)] = \frac{23}{25}u(t) - \frac{11}{12}e^{-t}u(t) - \frac{1}{300}e^{-25t}u(t).$$

■

4.6 System Response Partitions

Given some system response $y(t)$, there are three different ways to partition the response, with each partition describing a particular characteristic of the system response:

1. Zero-state / zero-input partition
2. Natural / forced partition
3. Transient / steady-state partition

Type 1: Zero-state / zero-input partition.

As the names imply,

- the *zero-state response* (ZSR) is the system response when there are no initial conditions, and
- the *zero-input response* (ZIR) is the system response when there is no input signal (but there are initial conditions).

It follows that

$$y(t) = y_{ZSR}(t) + y_{ZIR}(t). \tag{4.34}$$

Any exponentials $e^{-\lambda_k t}$ in the zero-input response $y_{ZIR}(t)$ are referred to as the *characteristic modes* (also called *natural modes*, or simply modes) of the system. The zero-input response is comprised of only natural modes, whereas the zero-state response has a mix of natural modes and "forced" modes from the input excitation function $x(t)$.

Type 2: Natural / forced partition.

On the discussion of natural modes and forced modes,

- the *natural response* is the part of the system response with only natural modes which describe the natural identity of the system, and
- the *forced response* is the part of the system response with only forced modes which mimic the character of the excitation function $x(t)$.

It follows that

$$y(t) = y_{nat}(t) + y_{forc}(t). \tag{4.35}$$

The natural response is closely related to the zero-input response; since the zero-input response is composed exclusively of modes, we can use it to identify the modes, then collect all terms with those modes to find the natural response. Naturally, the characteristic modes can also be determined by finding the solutions s of the characteristic equation and calculating e^{-st}.

Type 3: Transient / steady-state partition.

The end behavior of the terms of the system response can be analyzed such that

- the *transient response* is the part of the system response that decays to zero as $t \to \infty$, and
- the *steady-state response* is the part of the system response that remains after the transient response goes to zero.

It follows that

$$y(t) = y_{tr}(t) + y_{ss}(t). \tag{4.36}$$

We have seen this behavior before with the damping response of a stable LTI system described by a second-order LCCDE. The curves eventually approach a final value K, with the approach rate and shape dependent on the damping coefficient ξ. The damping response actually can be written as a sum of transient terms and the steady-state response $K \cdot u(t)$.

Again, notice here that we reserve the term "LTI system" for systems that have zero initial conditions, in addition to the superposition principle condition. Partitions can be applied to any LTI or "LTI-like" system response, whether or not the initial conditions are zero.

Example 4.6.1. Suppose a system can be characterized by the LCCDE

$$\begin{aligned} \frac{d^2y(t)}{dt^2} + 5\frac{dy(t)}{dt} + 4y(t) &= \frac{dx(t)}{dt} + 3x(t), \\ y(0) = y'(0) &= 1, \\ x(t) &= u(t). \end{aligned}$$

Identify all six partitioned responses of the system response.

SOLUTION

Taking the Laplace transform of the LCCDE, it follows that

$$\begin{aligned} \mathcal{L}\left[\frac{d^2y(t)}{dt^2} + 3\frac{dy(t)}{dt} + 2y(t)\right] &= \mathcal{L}\left[\frac{dx(t)}{dt} + 3x(t)\right] \\ [s^2Y(s) - s - 1] + 3[sY(s) - 1] + 2Y(s) &= [sX(s) - 0] + 3X(s) \\ Y(s)[s^2 + 3s + 2] - (s+4) &= X(s)[s+3]. \end{aligned}$$

Since $X(s) = \mathcal{L}[u(t)] = \frac{1}{s}$, it follows that

$$\begin{aligned} Y(s)[s^2 + 3s + 2] - (s+4) &= \frac{s+3}{s} \\ Y(s)[(s+1)(s+2)] &= \frac{s^2 + 5s + 3}{s} \\ \implies Y(s) &= \frac{s^2 + 5s + 3}{s(s+1)(s+2)} = \frac{A_1}{s} + \frac{A_2}{s+1} + \frac{A_3}{s+2}. \end{aligned}$$

After partial fraction decomposition, the inverse Laplace transform gives

$$y(t) = 1.5u(t) + e^{-t}u(t) - 1.5e^{-2t}u(t).$$

To find the modes, first solve for $y_{ZIR}(t)$. We can do this by taking the Laplace transform of the LCCDE and setting $X(s) = 0$.

$$\begin{aligned} Y(s)[s^2 + 3s + 2] - (s+4) &= X(s)[s+3] = 0 \\ \implies Y_{ZIR}(s) &= \frac{s+4}{s^2 + 3s + 2} = \frac{s+4}{(s+1)(s+2)} = \frac{3}{s+1} - \frac{2}{s+2}. \end{aligned}$$

Then it follows that

$$\begin{aligned} y_{ZIR}(t) &= 3e^{-t}u(t) - 2e^{-2t}u(t), \\ y_{ZSR}(t) &= y(t) - y_{ZIR}(t) = 1.5u(t) - 2e^{-t}u(t) + 0.5e^{-2t}u(t). \end{aligned}$$

From $y_{ZIR}(t)$, the characteristic modes are given by exponentials $\{e^{-t}, e^{-2t}\}$. Therefore, by grouping natural modes together and forced modes together,

$$\begin{aligned} y_{nat}(t) &= e^{-t}u(t) - 1.5e^{-2t}u(t), \\ y_{forc}(t) &= 1.5u(t). \end{aligned}$$

Lastly, we can group decaying terms together and the remaining terms together such that

$$
\begin{aligned}
y_{tr} &= e^{-t}u(t) - 1.5e^{-2t}u(t), \\
y_{ss} &= 1.5u(t).
\end{aligned}
$$

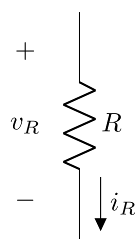
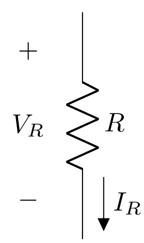
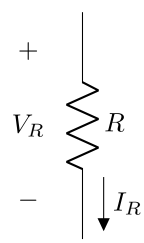
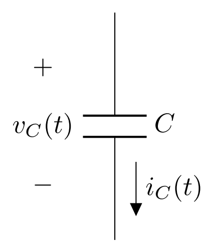
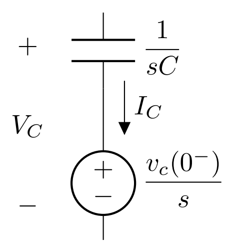
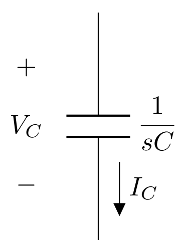
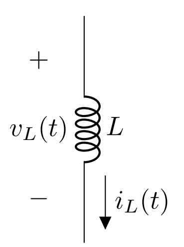
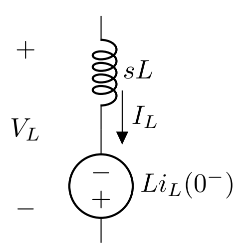
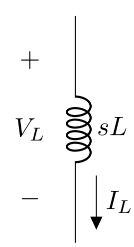
■

4.7 s–Domain Circuit Analysis

Just as differential equations can be solved in the s-domain, the systems represented by differential equations can be solved in the s-domain. For instance, an electric circuit can be transformed to an s-domain equivalent and solved using algebraic equations, with the s-domain electrical components referred to as impedances. The equivalent s-domain models can be seen in Table 4.4.

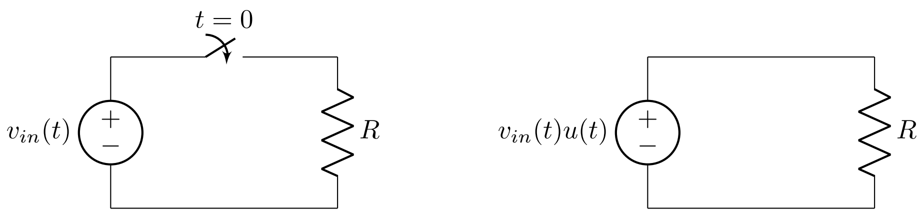
While only Thévenin equivalent models with voltage sources are shown in Table 4.4, Norton equivalent models with current sources can be drawn instead, though the relation between the Thévenin equivalents and the steady-state equivalents are easier to visualize.

Table 4.4: Circuit models for electrical components in the s-domain.

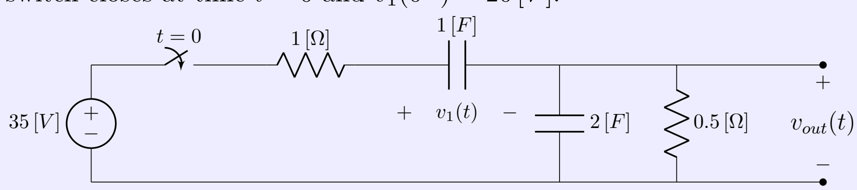
<table>
<tr><th>Electrical Component</th><th>Time Domain</th><th>s-Domain with Initial Conditions</th><th>s-Domain at Steady State</th></tr>
<tr><td>Resistor</td><td>
</td><td>
</td><td>
</td></tr>
<tr><td>Capacitor</td><td>
</td><td>
</td><td>
</td></tr>
<tr><td>Inductor</td><td>
</td><td>
</td><td>
</td></tr>
</table>

Recall that the unit step function $u(t)$ can be treated as an "on switch", where the signal being multiplied gets turned on at time $t = 0$. Similarly here, circuit switches can be represented using signals multiplied by $u(t)$, as seen in Figure 4.4. Because of how switches operate, here we see that the initial condition is zero.

Figure 4.4: Equivalent signal representation of switches.

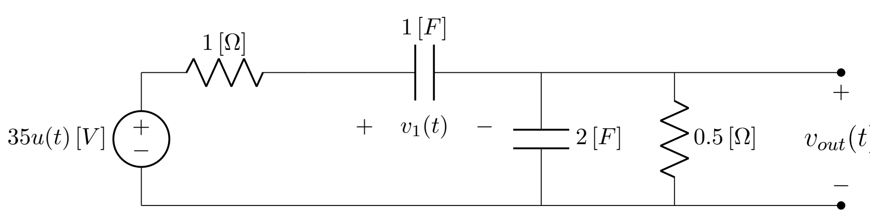


Example 4.7.1. Find the output response $v_{out}(t)$ of the following circuit, given that the switch closes at time $t = 0$ and $v_1(0^-) = 20\,[V]$.

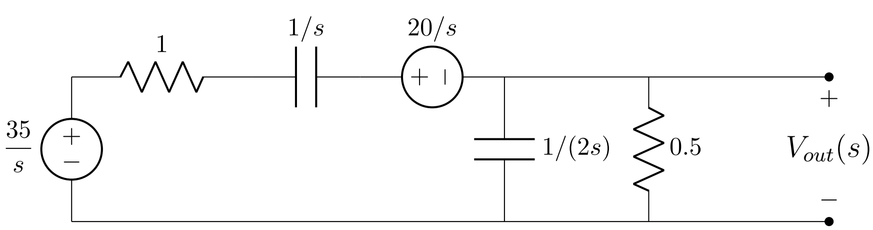


SOLUTION

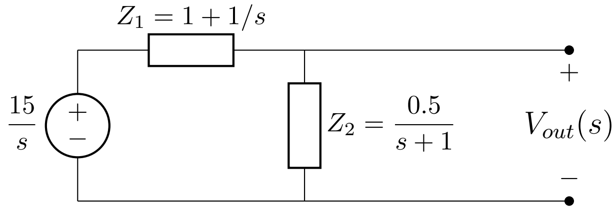
Using the equivalent model for switches, first we redraw the circuit as



Now we can transform the circuit into an s-domain equivalent model:



We can algebraically combine the two voltage sources together, the two series impedances together, and the two parallel impedances together:



From here, the s-domain output response can be calculated using the voltage divider:

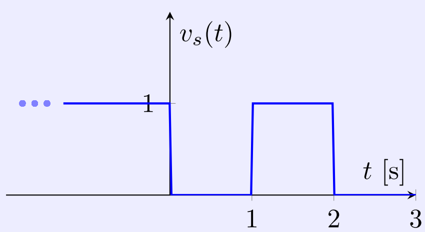
$$V_{out}(s) = \frac{15}{s} \cdot \left[\frac{Z_2}{Z_1 + Z_2}\right] = \frac{7.5}{s^2 + 2.5s + 1} = \frac{5}{s + 0.5} - \frac{5}{s + 2}.$$

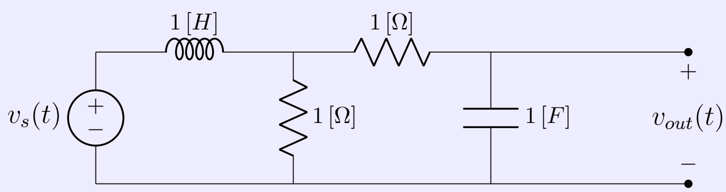
Lastly, taking the solution to the time domain, we get

$$v_{out}(t) = 5e^{-0.5t}u(t) - 5e^{-2t}u(t)\,[V].$$

■

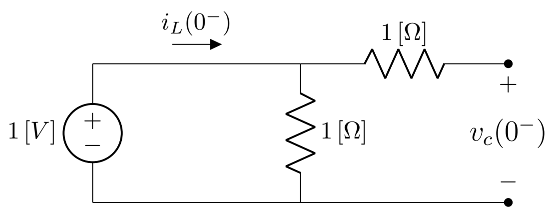
Example 4.7.2. Given the plot of input excitation $v_s(t)$, find the conditional transfer function $H_C(s) = V_{out}(s)/V_s(s)$ of the circuit below for time $t \geq 0$. Then find $v_{out}(t)$ for $t \geq 0$. Note that an ellipsis indicates a continuing pattern on the corresponding end.



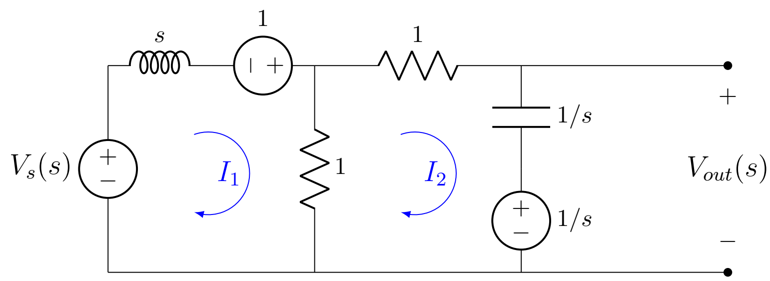


SOLUTION

First, find the initial conditions at time $t = 0^-$. Since $v_s(t) = 1$ for $t < 0$, we can assume the circuit is at steady-state at $t = 0^-$. The steady-state circuit for $t < 0$ is given by



Here we see that $i_L(0^-) = 1\,[V]/1\,[\Omega] = 1\,[A]$ and $v_c(0^-) = 1\,[V]$. Using these initial conditions, we can set up an s-domain equivalent of the circuit for $t \geq 0$.



We can use the mesh current method to solve the circuit.

$$-V_s(s) + sI_1 - 1 + 1\cdot(I_1 - I_2) = 0 \Longrightarrow (s+1)I_1 - I_2 = V_s(s) + 1$$
$$1\cdot(I_2 - I_1) + I_2(1 + 1/s) + 1/s = 0 \Longrightarrow -sI_1 + (2s+1)I_2 = -1$$

In matrix form, the mesh current equations can be written as

$$\begin{bmatrix} s+1 & -1 \\ -s & 2s+1 \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} V_s(s)+1 \\ -1 \end{bmatrix}.$$

From here, we can use Cramer's rule to solve for I_2 which can then be used to solve for $V_{out}(s)$.

$$\Delta = \begin{vmatrix} s+1 & -1 \\ -s & 2s+1 \end{vmatrix} = 2s^2 + 2s + 1$$
$$\Delta_2 = \begin{vmatrix} s+1 & V_s(s)+1 \\ -s & -1 \end{vmatrix} = sV_s(s) - 1$$
$$I_2 = \frac{\Delta_2}{\Delta} = \frac{sV_s(s) - 1}{2s^2 + 2s + 1}$$

The conditional transfer function can now be solved.

$$V_{out}(s) = I_2 \cdot \frac{1}{s} + \frac{1}{s} = \frac{sV_s(s) - 1}{s(2s^2 + 2s + 1)} + \frac{1}{s} = \frac{sV_s(s) + 2s^2 + 2s}{s(2s^2 + 2s + 1)}$$
$$\Longrightarrow H_C(s) = [V_{out}(s)\big|_{V_s(s)=1} = \frac{2s+3}{2s^2 + 2s + 1} = \frac{s + 1.5}{s^2 + s + 0.5}$$

With the Laplace transform of input $v_s(t)$ for $t \geq 0$ given by

$$V_s(s) = \mathcal{L}[v_s(t)] = \mathcal{L}[u(t-1) - u(t-2)] = \frac{1}{s}(e^{-s} - e^{-2s}),$$

the equation for $V_{out}(s)$ is given by

$$\begin{aligned}
V_{out}(s) = H_C(s)V_s(s) &= (e^{-s} - e^{-2s})\left[\frac{s+1.5}{s(s^2+s+0.5)}\right] \\
&= (e^{-s} - e^{-2s})\left[\frac{3}{s} + \frac{-1.5-j0.5}{s+0.5+j0.5} + \frac{-1.5+j0.5}{s+0.5-j0.5}\right] \\
&= (e^{-s} - e^{-2s})\underbrace{\left[\frac{3}{s} + \frac{1.581e^{-j2.82}}{s+0.5+j0.5} + \frac{1.581e^{+j2.82}}{s+0.5-j0.5}\right]}_{G(s)}.
\end{aligned}$$

The inverse Laplace transform of $G(s)$ is

$$g(t) = 3u(t) + 3.162e^{-0.5t}\cos(0.5t + 2.82\ [\text{rad}])u(t)$$

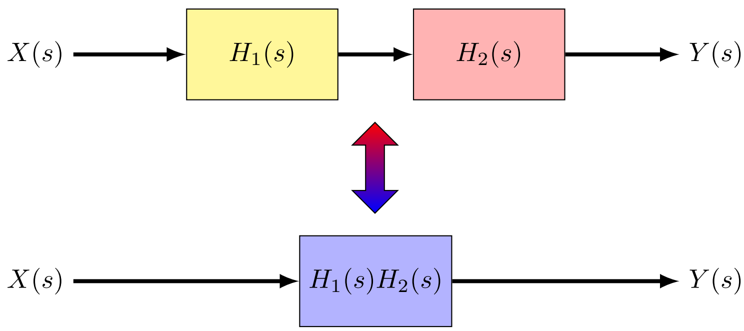
The output voltage $v_{out}(t)$ for $t \geq 0$ (with phase expressed in radians) is then

$$\begin{aligned}
v_{out}(t) &= \mathcal{L}^{-1}[(e^{-s} - e^{-2s})G(s)] = g(t-1) - g(t-2) \\
&= 3u(t-1) + 3.162e^{-0.5(t-1)}\cos(0.5(t-1) + 2.82)u(t-1) \\
&\quad - 3u(t-2) - 3.162e^{-0.5(t-2)}\cos(0.5(t-2) + 2.82)u(t-2).
\end{aligned}$$

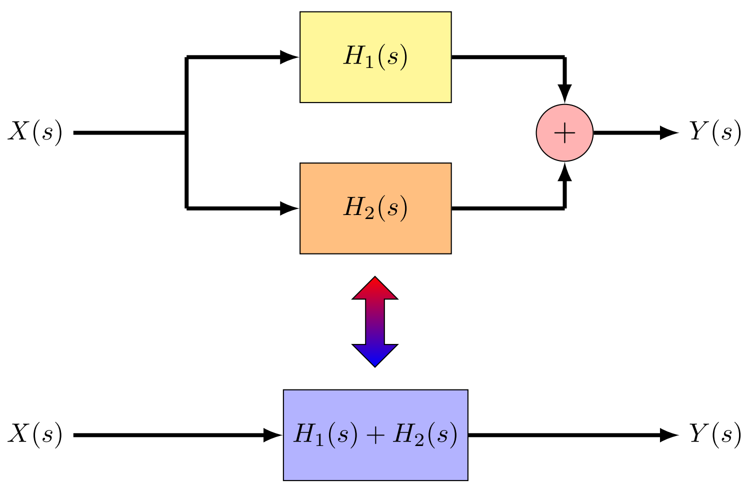
■

4.8 s–Domain Block Diagrams

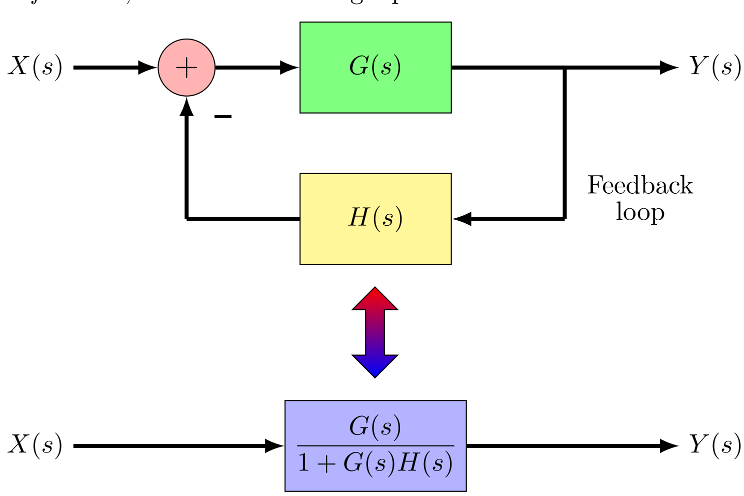
Similar to the interconnection of LTI systems in the time domain, LTI systems can be combined in the s-domain. When LTI systems are connected in series, the interconnection has the following equivalence:



When LTI systems are connected in parallel, the interconnection has the following equivalence:



A minus sign adjacent to an arrowhead at a summation node indicates that the corresponding addend signal is multiplied by -1 before being summed. When an interconnection of LTI systems implement *negative feedback*, it has the following equivalence:



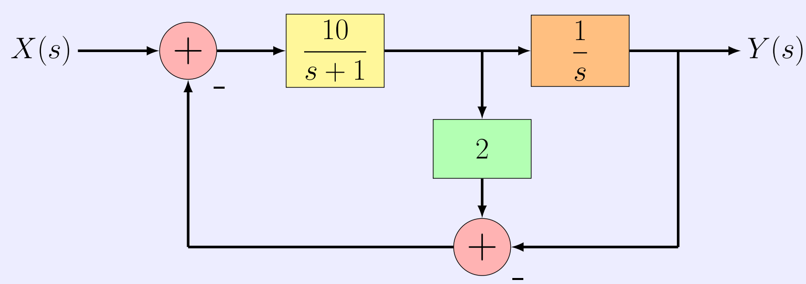
Feedback takes a sample of the output signal and feeds it back to the input signal. A system without feedback is called an *open-loop system*, whereas a system that utilizes feedback to form a *feedback loop* is called a *closed-loop system*.

A closed-loop system has *positive feedback* if the strength of the input signal is increased as a result; the closed-loop system has *negative feedback* if the strength of the input signal is decreased.

An interconnection of LTI systems with feedback loops can be simplified to an equivalent single-transfer-function model.

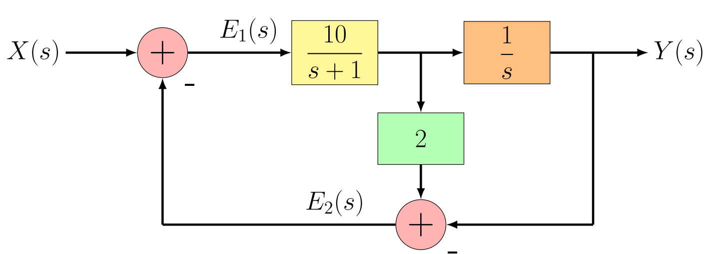
1. Condense any systems in parallel or series first.
2. Count the number of summation nodes left. The number of equations needed is the number of summation nodes plus one, with the one representing the output.
3. Define an intermediary signal $E_n(s)$ after each summation node n.
4. Starting from the output signal $Y(s)$, write an equation for the output. Then working backwards from $Y(s)$, write equations for each summation node.
5. Begin substituting and rearranging equations until a single equation in terms of only $X(s)$ and $Y(s)$ is left.

Example 4.8.1. Find an equivalent transfer function for the following interconnection of LTI systems.



SOLUTION

There are two summation nodes. We will need to define $E_1(s), E_2(s)$ and write three equations.



We can first write an equation for $Y(s)$, following the feedforward path at the top.

$$\begin{aligned} Y(s) &= E_1(s) \cdot \frac{10}{s+1} \cdot \frac{1}{s} \\ &= E_1(s) \left[\frac{10}{s(s+1)} \right] \end{aligned}$$

Now we write an equation for $E_2(s)$ using the rightmost branch.

$$\begin{aligned} E_2(s) &= E_1(s) \cdot \frac{10}{s+1} \cdot 2 - Y(s) \\ &= E_1(s) \cdot \frac{10}{s+1} \cdot 2 - E_1(s) \left[\frac{10}{s(s+1)} \right] \\ &= E_1(s) \cdot \frac{10}{s+1} \left[2 - \frac{1}{s} \right] \\ &= E_1(s) \left[\frac{10(2s-1)}{s(s+1)} \right] \end{aligned}$$

Lastly, we write an equation for $E_1(s)$ using the leftmost branch.

$$\begin{aligned} E_1(s) &= X(s) - E_2(s) = X(s) - E_1(s) \left[\frac{10(2s-1)}{s(s+1)} \right] \\ \implies X(s) &= E_1(s) \left[1 + \frac{10(2s-1)}{s(s+1)} \right] \end{aligned}$$

We can rearrange the equation for $Y(s)$ and substitute into $X(s)$.

$$\begin{aligned} E_1(s) &= Y(s) \left[\frac{s(s+1)}{10} \right] \\ \implies X(s) &= E_1(s) \left[1 + \frac{10(2s-1)}{s(s+1)} \right] \\ &= Y(s) \left[\frac{s(s+1)}{10} \right] \left[1 + \frac{10(2s-1)}{s(s+1)} \right] \\ &= Y(s) \left[\frac{s(s+1)}{10} + (2s-1) \right] \\ &= Y(s) \left[\frac{s^2+21s-10}{10} \right] \end{aligned}$$

Then the equivalent transfer function $H(s)$ is

$$H(s) = \frac{Y(s)}{X(s)} = \frac{10}{s^2+21s-10}.$$

■

Chapter 5

Fourier Series

Previously, we have seen that causal signals and systems can be analyzed using the unilateral Laplace transform. Once transformed to the s-domain, equations can be solved algebraically before transforming back to the time domain.

For LTI systems with everlasting signals, the unilateral Laplace transform is not suitable. While the bilateral Laplace transform could be used, *Fourier analysis* is far more suitable as Fourier analysis techniques transform data to the *frequency domain*, where frequency content can be analyzed.

There are two types of Fourier analysis techniques for continuous-time signals and systems: the Fourier series and the Fourier transform. The Fourier series is used for periodic signals, whereas the Fourier transform can be generalized for any signal regardless of periodicity and causality.

This chapter will first explore the analysis of periodic signals with Fourier series.

5.1 Phasor Domain

Before delving into Fourier series, first we consider periodic signals with only one frequency – that is, sinusoidal signals. It follows that an input sinusoidal signal into an LTI system will generate an output sinusoidal signal with the same frequency.

The relationship between a sinusoid $x(t)$ and a phasor $\underline{X} = |\underline{X}|e^{j\phi}$ with $\phi = \arg(\underline{X})$ can be exploited such that

$$\begin{aligned} x(t) &= \mathrm{Re}[\,\underline{X}e^{j\omega t}] \qquad (5.1) \\ &= \mathrm{Re}[|\underline{X}|e^{j\phi}e^{j\omega t}] \\ &= \mathrm{Re}[|\underline{X}|e^{j(\omega t+\phi)}] \\ &= |\underline{X}|\cos(\omega t+\phi). \qquad (5.2) \end{aligned}$$

The *phasor transform* offers a way to solve signals and systems with only one frequency involved. Because a sinusoid contain only a single frequency, the magnitude and phase of sinusoidal signals can be isolated and analyzed. The phasor transform can be defined as

$$\underline{X} = \mathcal{P}[x(t)], \tag{5.3}$$

where the inverse phasor transform is defined as

$$x(t) = \mathcal{P}^{-1}[\,\underline{X}\,] = \mathrm{Re}[\,\underline{X}e^{j\omega t}]. \tag{5.4}$$

The table of phasor transform pairs is given in Table 5.1.

Table 5.1: Phasor transform pairs.

$x(t) = \mathrm{Re}[\,\underline{X}e^{j\omega t}]$	$\underline{X} = \mathcal{P}[x(t)]$
$A\cos(\omega t)$	A
$A\cos(\omega t + \phi)$	$Ae^{j\phi}$
$-A\cos(\omega t + \phi)$	$Ae^{j(\phi \pm \pi)}$
$A\sin(\omega t)$	$Ae^{-j\pi/2} = -jA$
$A\sin(\omega t + \phi)$	$Ae^{j(\phi - \pi/2)}$
$-A\sin(\omega t + \phi)$	$Ae^{j(\phi + \pi/2)}$
$\frac{d}{dt}\left[A\cos(\omega t + \phi)\right]$	$j\omega Ae^{j\phi}$
$\frac{d^n}{dt^n}\left[A\cos(\omega t + \phi)\right]$	$(j\omega)^n Ae^{j\phi}$
$\int A\cos(\omega\tau + \phi)\,d\tau$	$\frac{1}{j\omega}Ae^{j\phi}$

Example 5.1.1. Find the output $y(t)$ of a system characterized by the differential equation

$$\frac{d^2y(t)}{dt^2} + 300\frac{dy(t)}{dt} + (5 \times 10^4)y(t) = x(t),$$

given input signal $x(t) = 10\sin(100t + 60^\circ)$.

SOLUTION

First, we transform the LCCDE to the phasor domain.

$$
\begin{aligned}
(j\omega)^2\underline{Y} + 300(j\omega)\underline{Y} + (5\times 10^4)\underline{Y} &= 10e^{j(60^\circ-90^\circ)} \\
\underline{Y}[(j\omega)^2 + 300(j\omega) + (5\times 10^4)] &= 10e^{-j30^\circ}
\end{aligned}
$$

We can now rearrange to solve for $\underline{Y}$ and substitute $\omega = 100$.

$$
\begin{aligned}
\underline{Y} &= \frac{10e^{-j30^\circ}}{(j\omega)^2 + 300(j\omega) + (5\times 10^4)} \\
&= \frac{10e^{-j30^\circ}}{(j100)^2 + 300(j100) + (5\times 10^4)} \\
&= (0.2\times 10^{-3})e^{-j66.87^\circ}
\end{aligned}
$$

Transforming back to the time domain, the output is

$$y(t) = \mathrm{Re}[\underline{Y}e^{j\omega t}] = (0.2\times 10^{-3})\cos(100t - 66.87^\circ).$$

■

Example 5.1.2. Find the output $y(t)$ of a system characterized by the differential equation

$$(4\times 10^{-3})\frac{dy(t)}{dt} + 3y(t) = x(t),$$

given input signal $x(t) = 5\cos(\omega_0 t) - 10\cos(2\omega_0 t)$ with $\omega_0 = 10^3$ [rad/s].

SOLUTION

Let $x_1(t) = 5\cos(\omega_0 t)$ and $x_2(t) = -10\cos(2\omega_0 t)$.

First, solving for $\underline{Y_1}$ with $\omega_1 = \omega_0$, we get

$$
\begin{aligned}
&(4\times 10^{-3})\frac{dy_1(t)}{dt} + 3y_1(t) = 5\cos(\omega_1 t) \\
\implies &(4\times 10^{-3})(j\omega_1)\underline{Y_1} + 3\underline{Y_1} = 5 \\
\implies &\underline{Y_1} = \frac{5}{(4\times 10^{-3})j\omega_1 + 3} = \frac{5}{3+j4} = e^{-j53.13^\circ}.
\end{aligned}
$$

Then, solving for $\underline{Y_2}$ with $\omega_2 = 2\omega_0$, we get

$$
\begin{aligned}
&(4\times 10^{-3})\frac{dy_2(t)}{dt} + 3y_2(t) = -10\cos(\omega_2 t) \\
\implies &(4\times 10^{-3})(j\omega_2)\underline{Y_2} + 3\underline{Y_2} = 10e^{j\pi} \\
\implies &\underline{Y_2} = \frac{10e^{j\pi}}{(4\times 10^{-3})j\omega_2 + 3} = \frac{10e^{j\pi}}{3+j8} = 1.17e^{j20.56^\circ}.
\end{aligned}
$$

Since $\underline{Y} = \underline{Y_1} + \underline{Y_2}$ by the superposition principle, it follows that

$$\begin{aligned} y(t) &= \mathrm{Re}[\,\underline{Y_1} e^{j\omega_1 t}] + \mathrm{Re}[\,\underline{Y_2} e^{j\omega_2 t}] \\ &= \mathrm{Re}[e^{-j53.13^\circ} e^{j\omega_0 t}] + \mathrm{Re}[1.17 e^{j20.56^\circ} e^{j2\omega_0 t}] \\ &= \cos(\omega_0 t - 53.13^\circ) + 1.17\cos(2\omega_0 t + 20.56^\circ). \end{aligned}$$

■

5.2 Fourier Series

While the superposition principle can be applied for an input signal comprised of a few sinusoidal addends, using the phasor transform becomes more tedious when more and more sinusoids with different frequencies are introduced. In fact, the procedure becomes impossible when an infinite number of sinusoids are added. Why is this relevant?

According to *Fourier's theorem*, any physically realizable periodic signal can be represented as a sum of sinusoids, whether or not it is a finite sum or an infinite series. While finite sums can be handled using the phasor transform, periodic signals represented by infinite series must be analyzed using a different tool.

Here, the *Fourier series* is introduced, where any physically realizable periodic signal with period T_0 can be broken down into fundamentals (which are terms with fundamental angular frequency $\omega_0 = 2\pi/T_0$) and its n^{th} harmonics (which are terms with frequency $n\omega_0$ for each n).

There are three different representations of the Fourier series: the sine/cosine representation, the amplitude/phase representation, and the exponential representation.

5.2.1 Sine / Cosine Representation

The sine/cosine representation of a physically realizable periodic signal $x(t)$ is given by

$$x(t) = a_0 + \sum_{n=1}^{\infty} [a_n \cos(n\omega_0 t) + b_n \sin(n\omega_0 t)], \tag{5.5}$$

where the *Fourier coefficients* are calculated from the following integrals:

$$a_0 = \frac{1}{T_0} \int_{t_0}^{t_0+T_0} x(t)\, dt \tag{5.6}$$

$$a_n = \frac{2}{T_0} \int_{t_0}^{t_0+T_0} x(t)\cos(n\omega_0 t)\, dt \tag{5.7}$$

$$b_n = \frac{2}{T_0} \int_{t_0}^{t_0+T_0} x(t)\sin(n\omega_0 t)\, dt \tag{5.8}$$

Here, the limits of integration $[t_0, t_0 + T_0]$ can be set to $[0, T_0]$, $[-T_0/2, +T_0/2]$, or a reasonable interval $[t_1, t_2]$ such that $t_2 - t_1 = T_0$. Sometimes, to indicate that the integral is over a period T_0, the integral (not including the integrand) may be written as $\int_{T_0} dt$.

The coefficient a_0 is also referred to as the *DC (Fourier) component* and can be restated as

$$a_0 = \frac{\text{Area of } x(t) \text{ during } T_0}{\text{Period } T_0 \text{ of } x(t)}. \tag{5.9}$$

The Fourier series of a select set of waveforms are listed in Table 5.2.

5.2.2 Amplitude / Phase Representation

The amplitude/phase representation of a physically realizable periodic signal $x(t)$ is most closely related to the phasor transform technique and is given by

$$x(t) = c_0 + \sum_{n=1}^{\infty} c_n \cos(n\omega_0 t + \phi_n), \tag{5.10}$$

where the Fourier coefficients are derived from the relationship $c_n e^{j\phi_n} = a_n - jb_n$ from the sine/cosine representation such that

$$c_0 = a_0 \tag{5.11}$$

$$c_n = \sqrt{a_n + b_n} \tag{5.12}$$

$$\phi_n = -\operatorname{atan2}(b_n, a_n) = \begin{cases} -\arctan(b_n/a_n), & a_n > 0 \\ \pi - \arctan(b_n/a_n), & a_n < 0 \end{cases} \tag{5.13}$$

While the phasor transform isolates an amplitude and phase of a signal for a single frequency, the amplitude/phase representation of the Fourier series isolates the amplitudes c_n and phases ϕ_n of a physically realizable periodic signal across an infinite number of harmonic frequencies.

These amplitudes and phases can be visually depicted on a set of line spectra called the *amplitude spectrum* and the *one-sided phase spectrum*, respectively. The amplitude spectrum plots the amplitudes against the corresponding harmonic frequencies, and the one-sided phase spectrum plots the phases against the corresponding harmonic frequencies. Often the two spectra together are collectively referred to as the *one-sided Fourier spectrum.*

Example 5.2.1. Plot the one-sided Fourier spectrum of the periodic signal $x(t)$ depicted below.

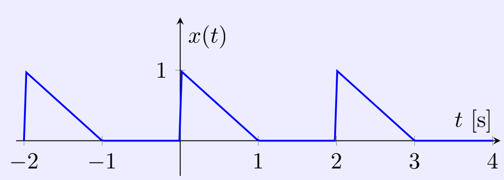
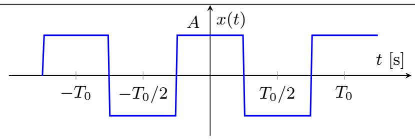
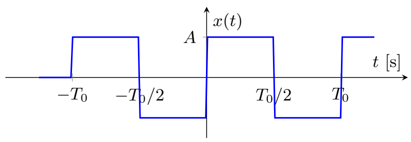
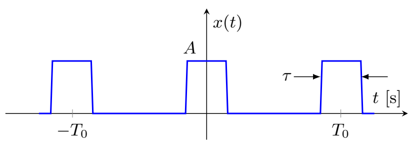
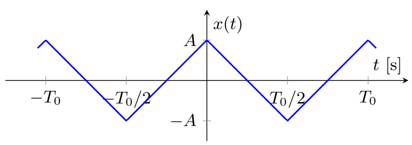
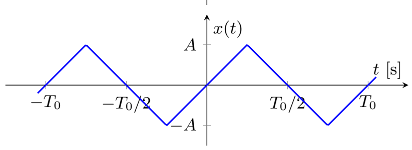
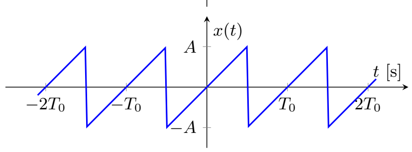
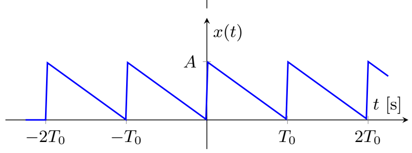
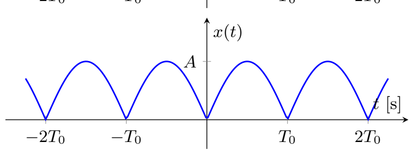
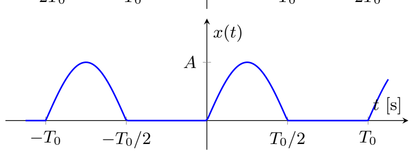


Table 5.2: Select set of periodic waveforms.

	Waveform	Fourier Series
Square wave		$x(t) = \sum_{n=1}^{\infty} \frac{4A}{n\pi} \sin\left(\frac{n\pi}{2}\right) \cos\left(\frac{2n\pi t}{T_0}\right)$
Time-shifted square wave		$x(t) = \sum_{\substack{n=1 \\ n=\text{odd}}}^{\infty} \frac{4A}{n\pi} \sin\left(\frac{2n\pi t}{T_0}\right)$
Pulse train		$x(t) = \frac{A\tau}{T_0} + \sum_{n=1}^{\infty} \frac{2A}{n\pi} \sin\left(\frac{n\pi\tau}{T_0}\right) \cos\left(\frac{2n\pi t}{T_0}\right)$
Triangular wave		$x(t) = \sum_{\substack{n=1 \\ n=\text{odd}}}^{\infty} \frac{8A}{n^2\pi^2} \cos\left(\frac{2n\pi t}{T_0}\right)$
Time-shifted triangular wave		$x(t) = \sum_{\substack{n=1 \\ n=\text{odd}}}^{\infty} \frac{8A}{n^2\pi^2} \sin\left(\frac{n\pi}{2}\right) \sin\left(\frac{2n\pi t}{T_0}\right)$
Sawtooth		$x(t) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{2A}{n\pi} \sin\left(\frac{2n\pi t}{T_0}\right)$
Backward sawtooth		$x(t) = \frac{A}{2} + \sum_{n=1}^{\infty} \frac{A}{n\pi} \sin\left(\frac{2n\pi t}{T_0}\right)$
Full-wave rectified sinusoid		$x(t) = \frac{2A}{\pi} + \sum_{n=1}^{\infty} (-1)^{n+1} \frac{4A}{\pi(1-4n^2)} \cos\left(\frac{2n\pi t}{T_0}\right)$
Half-wave rectified sinusoid		$x(t) = \frac{A}{\pi} + \frac{A}{2} \sin\left(\frac{2\pi t}{T_0}\right) + \sum_{\substack{n=2 \\ n=\text{even}}}^{\infty} \frac{2A}{\pi(1-n^2)} \cos\left(\frac{2n\pi t}{T_0}\right)$

SOLUTION

Visually, we see that the period is $T_0 = 2$. Therefore $\omega_0 = 2\pi/T_0 = \pi$. We can define $x(t)$ during a single period as

$$x(t)\Big|_{t\in(0,T_0)} = \begin{cases} 1-t, & 0<t<1 \\ 0, & 1<t<2 \end{cases}$$

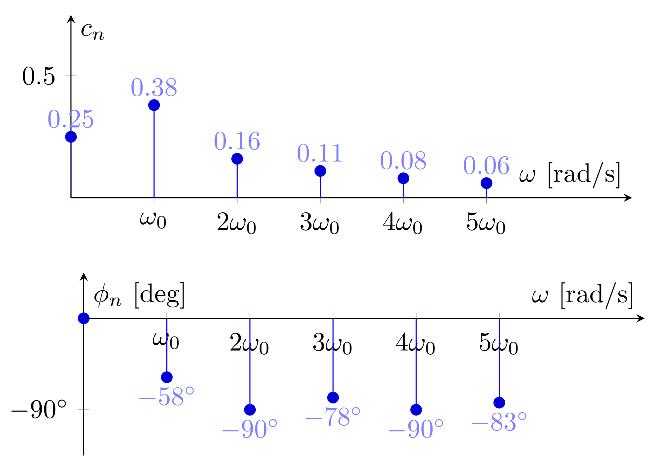
Then the sine/cosine Fourier coefficients are given by

$$\begin{aligned} a_0 &= \frac{\text{Area of } x(t) \text{ during } T_0}{\text{Period } T_0 \text{ of } x(t)} = \frac{0.5}{2} = 0.25 \\ a_n &= \frac{2}{T_0}\int_0^{T_0} x(t)\cos(n\omega_0 t)\,dt = \frac{2}{2}\int_0^1 (1-t)\cos(n\omega_0 t)\,dt = \frac{1-\cos(n\pi)}{(n\pi)^2} \\ b_n &= \frac{2}{T_0}\int_0^{T_0} x(t)\sin(n\omega_0 t)\,dt = \frac{2}{2}\int_0^1 (1-t)\sin(n\omega_0 t)\,dt = \frac{1}{n\pi} \end{aligned}$$

Note that $a_n \geq 0$. Converting the Fourier coefficients to amplitude/phase representation,

$$\begin{aligned} c_0 &= a_0 = 0.25 \\ c_n &= \sqrt{a_n^2 + b_n^2} = \sqrt{\left(\frac{1-\cos(n\pi)}{(n\pi)^2}\right)^2 + \left(\frac{1}{n\pi}\right)^2} \\ \phi_n &= -\arctan(b_n/a_n) = -\arctan\left(\frac{n\pi}{1-\cos(n\pi)}\right) \end{aligned}$$

Plotting the first few values of c_n and ϕ_n against ω, we get



5.2.3 Exponential Representation

Since sinusoids are closely related to complex exponentials with $\cos(\theta) = \frac{1}{2}(e^{+j\theta} + e^{-j\theta})$, it follows that the exponential representation of a physically realizable periodic signal $x(t)$ is given by

$$x(t) = \sum_{n=-\infty}^{+\infty} x_n e^{jn\omega_0 t}, \tag{5.14}$$

where the Fourier coefficients are calculated from the integral:

$$x_n = \frac{1}{T_0} \int_{t_0}^{t_0+T_0} x(t) e^{-jn\omega_0 t}\, dt, \tag{5.15}$$

with conjugate symmetry properties such that for $x_n = |x_n| e^{j\phi_n}$:

$$x_{-n} = x_n^* \tag{5.16}$$
$$\phi_{-n} = -\phi_n \tag{5.17}$$

In fact, ϕ_n here is the same value as ϕ_n in the amplitude/phase representation. Between the two representations, it also follows that

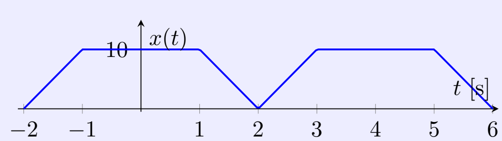
$$x_0 = c_0 \tag{5.18}$$
$$|x_n| = \frac{c_n}{2} \tag{5.19}$$

Similar to the amplitude/phase representation, the magnitudes $|x_n|$ and phases ϕ_n can be collected and plotted against their corresponding harmonic frequencies on a *magnitude spectrum* and a *two-sided phase spectrum*, respectively. Often the two spectra together are collectively referred to as the *two-sided Fourier spectrum*.

Additionally, the relationship between the exponential and sine/cosine representations is given by

$$x_0 = a_0 \tag{5.20}$$
$$x_n = \frac{1}{2}(a_n - jb_n) \tag{5.21}$$

Example 5.2.2. Plot the two-sided Fourier spectrum of the periodic signal $x(t)$ depicted below.



SOLUTION

Visually, we see that the period is $T_0 = 4$. Therefore $\omega_0 = 2\pi/T_0 = \pi/2$. We can define $x(t)$ during a single period from starting point $t_0 = -2$ as

$$x(t)\Big|_{t\in(t_0,t_0+T_0)} = \begin{cases} 10(t+2), & -2 < t < -1 \\ 10, & -1 < t < 1 \\ -10(t-2), & 1 < t < 2 \end{cases}$$

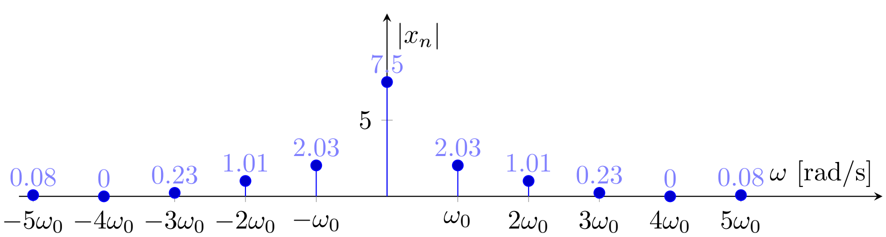
The DC component is given by

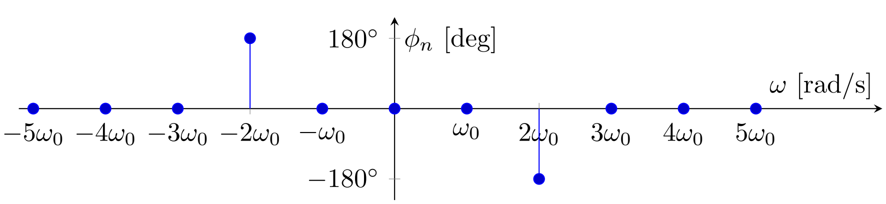
$$x_0 = \frac{\text{Area of } x(t) \text{ during } T_0}{\text{Period } T_0 \text{ of } x(t)} = \frac{30}{4} = 7.5,$$

and the remaining exponential Fourier coefficients are given by

$$\begin{aligned} x_n &= \frac{1}{T_0}\int_{t_0}^{t_0+T_0} x(t)e^{-jn\omega_0 t}\,dt \\ &= \frac{1}{4}\left[\int_{-2}^{-1} 10(t+2)e^{-jn\omega_0 t}\,dt + \int_{-1}^{1} 10e^{-jn\omega_0 t}\,dt + \int_{1}^{2} -10(t-2)e^{-jn\omega_0 t}\,dt\right] \\ &= \left[\frac{5e^{jn\omega_0} - 5e^{j2n\omega_0} + j5n\omega_0 e^{jn\omega_0}}{2n^2\omega_0^2}\right] + \frac{5\sin(n\omega_0)}{n\omega_0} + \left[\frac{5e^{-jn\omega_0} - 5e^{-j2n\omega_0} - j5n\omega_0 e^{-jn\omega_0}}{2n^2\omega_0^2}\right] \\ &= \left[\frac{5\cos(n\omega_0)}{n^2\omega_0^2} - \frac{5\cos(2n\omega_0)}{n^2\omega_0^2} - \frac{5\sin(n\omega_0)}{n\omega_0}\right] + \frac{5\sin(n\omega_0)}{n\omega_0} \\ &= \frac{5[\cos(n\omega_0) - \cos(2n\omega_0)]}{n^2\omega_0^2} \end{aligned}$$

Note that x_n is purely real. Depending on the sign of x_n, the phase ϕ_n takes on values of either $0°$ or $\pm 180°$. Plotting the first few values of $|x_n|$ and ϕ_n against ω and using conjugate symmetry properies, we get





Most of the time, it is easier to find the sine/cosine representation of the Fourier series of a periodic signal first, then convert accordingly to whatever representation is necessary.

However, with the exponential representation of the Fourier series, we observe certain properties (listed in Table 5.3) that will become familiar when comparing to the table of Fourier transform properties in a later chapter.

Table 5.3: Properties of the exponential representation of the Fourier series.

Property	Periodic signal, $x(t)$	Exponential Fourier coefficients, x_n
Superposition	$K_1 x_1(t) + K_2 x_2(t)$	$K_1 (x_1)_n + K_2 (x_2)_n$
Time scaling	$x(at),\ a > 0$	x_n
Time shift	$x(t - t_0),\ t_0 > 0$	$e^{-jn\omega_0 t_0} x_n$
Frequency shift	$e^{-jN\omega_0 t} x(t)$	x_{n+N}
Time reversal	$x(-t)$	x_{-n}
Time derivative	$x'(t) = \dfrac{dx(t)}{dt}$	$jn\omega_0 x_n$
Time integral	$\int_{-\infty}^{t} x(\tau)\, d\tau$	$\dfrac{1}{jn\omega_0} x_n$
Conjugate symmetry	$x(t)$ real	$\begin{cases} x_{-n} = x_n^* \\ \mathrm{Re}(x_n) = \mathrm{Re}(x_{-n}) \\ \mathrm{Im}(x_n) = -\mathrm{Im}(x_{-n}) \\ \lvert x_n \rvert = \lvert x_{-n} \rvert \\ \phi_n = -\phi_{-n} \end{cases}$
Real and even signals	$x(t)$ real and even	x_n purely real and even
Real and odd signals	$x(t)$ real and odd	x_n purely imaginary and odd
Even-odd decomposition of real signals	$\begin{cases} x_e(t) = \frac{1}{2}[x(t) + x(-t)] \\ x_o(t) = \frac{1}{2}[x(t) - x(-t)] \end{cases}$	$\begin{cases} \mathrm{Re}(x_n) \\ j\,\mathrm{Im}(x_n) \end{cases}$

5.3 Symmetry Properties of Fourier Series

From the sine/cosine representation, when computing Fourier coefficients, note that $\cos(n\omega_0 t)$ is an odd function and $\sin(n\omega_0 t)$ is an even function. Recall the following symmetry properties for multiplying symmetric functions:

- (even) $\times$ (even) = even
- (even) $\times$ (odd) = odd
- (odd) $\times$ (odd) = even

Using $[-t_0, +t_0]$ as the limits of integration, it is easier to see that for some odd function $F_{odd}(t)$ and even function $F_{even}(t)$,

$$\int_{-t_0}^{+t_0} F_{odd}(t)\,dt = 2\int_0^{t_0} F_{odd}(t)\,dt \tag{5.22}$$

$$\int_{-t_0}^{+t_0} F_{even}(t)\,dt = 0. \tag{5.23}$$

These integral properties can be applied to the Fourier coefficients from the sine/cosine representation of the Fourier series.

EVEN SYMMETRY. If $x(t)$ is a physically realizable periodic signal with even symmetry $x(t) = x(-t)$, then the sine/cosine Fourier coefficients are

$$a_0 = \frac{2}{T_0}\int_0^{T_0/2} x(t)\,dt \tag{5.24}$$

$$a_n = \frac{4}{T_0}\int_0^{T_0/2} x(t)\cos(n\omega_0 t)\,dt \tag{5.25}$$

$$b_n = 0 \tag{5.26}$$

with the amplitude/phase Fourier coefficients given by

$$c_n = |a_n| \tag{5.27}$$

$$\phi_n = \begin{cases} 0, & a_n > 0 \\ \pi, & a_n < 0 \end{cases} \tag{5.28}$$

ODD SYMMETRY. If $x(t)$ is a physically realizable periodic signal with odd symmetry $x(t) = -x(-t)$, then the sine/cosine Fourier coefficients are

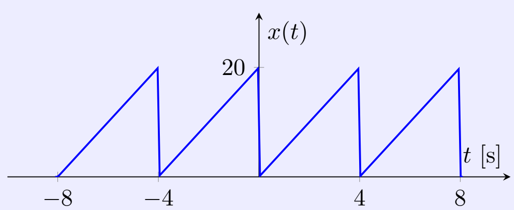
$$a_0 = 0 \tag{5.29}$$
$$a_n = 0 \tag{5.30}$$
$$a_n = \frac{4}{T_0}\int_0^{T_0/2} x(t)\sin(n\omega_0 t)\,dt \tag{5.31}$$

with the amplitude/phase Fourier coefficients given by

$$c_n = |b_n| \tag{5.32}$$
$$\phi_n = \begin{cases} -\pi/2, & b_n > 0 \\ \pi/2, & b_n < 0 \end{cases} \tag{5.33}$$

As a_0 is the DC component, it can be thought of as a DC offset which vertically shifts the plot of some signal. Because of this, it is possible that the signal $f(t) = x(t) - a_0$ (i.e., without the DC offset) has even or odd symmetry, for which we can utilize to our advantage when solving Fourier coefficients. (If so, $x(t)$ is said to have *hidden symmetry*, obscured by DC offset a_0.) For reference, see the flowchart in Figure 5.1.

Example 5.3.1. Find the Fourier series for the following periodic signal $x(t)$.



SOLUTION

Visually, we see that the period is $T_0 = 4$. Therefore, $\omega_0 = 2\pi/T_0 = \pi/2$. First, we calculate the DC offset

$$a_0 = \frac{\text{Area of } x(t) \text{ during } T_0}{\text{Period } T_0 \text{ of } x(t)} = \frac{40}{4} = 10.$$

Removing the DC offset, we get the new signal $f(t) = x(t) - a_0$:

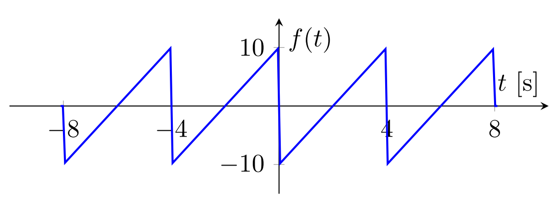
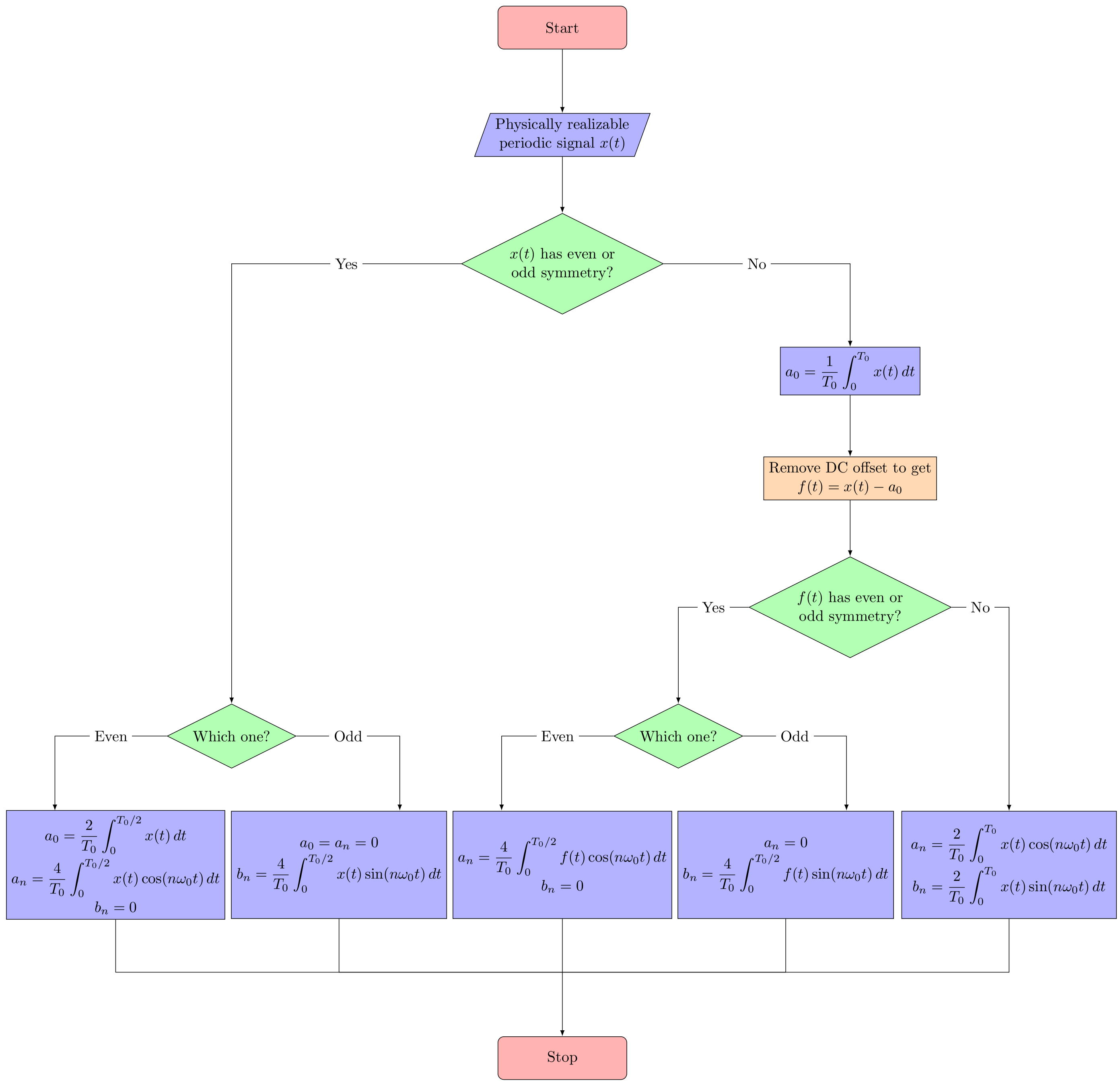


Figure 5.1: Workflow for the sine/cosine representation of the Fourier series.



From here, we can see that $f(t)$ has odd symmetry. We can use odd-symmetry form of the sine/cosine representation of the Fourier series to calculate the Fourier coefficients:

$$\begin{aligned}
a_n &= 0 \\
b_n &= \frac{4}{T_0}\int_0^{T_0/2} f(t)\sin(n\omega_0 t)\,dt = \frac{4}{4}\int_0^2 (5t-10)\sin\left(\frac{n\pi t}{2}\right)\,dt \\
&= \frac{20[\sin(n\pi) - n\pi]}{n^2\pi^2} = \frac{20[(-1)^n - n\pi]}{n^2\pi^2}
\end{aligned}$$

The sine/cosine representation of the Fourier series of $x(t)$ is then

$$x(t) = 10 + \sum_{n=1}^{\infty} \frac{20[(-1)^n - n\pi]}{n^2\pi^2}\sin\left(\frac{n\pi t}{2}\right)$$

■

5.4 Fourier Series of Periodic Extensions

Suppose a function $x(t)$ is defined strictly over interval $t \in (0, T_0)$. Then its *periodic extension* $\tilde{x}(t)$ is defined as

$$\tilde{x}(t + nT_0) = x(t), \text{ for all } t \in (0, T_0) \text{ and } n \in \mathbb{Z}, \tag{5.34}$$

with $\tilde{x}(t)$ having a period of T_0. Essentially, a periodic extension is the result of summing shifted copies of $x(t)$, with the shifts determined by integer amounts of the period, which is defined from the duration of $x(t)$. If the values at the endpoints of $x(t)$ are not equal, then $\tilde{x}(t)$ has jump discontinuities at $t = nT_0$.

While a simple periodic extension can be implemented, one can also apply symmetry in order to generate symmetric signals. While there are multiple types of symmetry, we will briefly cover only five of them:

- even symmetry: $\tilde{x}(t) = \tilde{x}(-t)$
- odd symmetry: $\tilde{x}(t) = -\tilde{x}(-t)$
- half-wave symmetry: $\tilde{x}(t) = -\tilde{x}(t - T_0/2)$
- quarter-wave even symmetry: $\tilde{x}(t)$ has even and half-wave symmetry
- quarter-wave odd symmetry: $\tilde{x}(t)$ has odd and half-wave symmetry

EVEN PERIODIC EXTENSIONS. If $x(t)$ is function defined on interval $t \in (0, L)$, then an even periodic extension $\tilde{x}(t)$ has period $T_0 = 2L$ with Fourier coefficients

$$a_0 = \frac{2}{T_0} \int_0^{T_0/2} x(t)\, dt \tag{5.35}$$

$$a_n = \frac{4}{T_0} \int_0^{T_0/2} x(t) \cos(n\omega_0 t)\, dt \tag{5.36}$$

$$b_n = 0 \tag{5.37}$$

and can be plotted using the following steps:

1. Flip $x(t)$ over the vertical axis.
2. Horizontally shift the overall new plot by nT_0 units for all integers n.

ODD PERIODIC EXTENSIONS. If $x(t)$ is function defined on interval $t \in (0, L)$, then an odd periodic extension $\tilde{x}(t)$ has period $T_0 = 2L$ with Fourier coefficients

$$a_0 = 0 \tag{5.38}$$

$$a_n = 0 \tag{5.39}$$

$$b_n = \frac{4}{T_0} \int_0^{T_0/2} x(t) \sin(n\omega_0 t)\, dt \tag{5.40}$$

and can be plotted using the following steps:

1. Flip $x(t)$ over the vertical axis, then flip the mirrored segment over the time axis.
2. Horizontally shift the overall new plot by nT_0 units for all integers n.

HALF-WAVE SYMMETRY. If $x(t)$ is function defined on interval $t \in (0, L)$, then periodic extension $\tilde{x}(t)$ can have half-wave symmetry for period $T_0 = 2L$ with Fourier coefficients

$$a_0 = 0 \tag{5.41}$$

$$a_n = \begin{cases} 0, & n \text{ even} \\ \frac{4}{T_0} \int_0^{T_0/2} x(t) \cos(n\omega_0 t)\, dt, & n \text{ odd} \end{cases} \tag{5.42}$$

$$b_n = \begin{cases} 0, & n \text{ even} \\ \frac{4}{T_0} \int_0^{T_0/2} x(t) \sin(n\omega_0 t)\, dt, & n \text{ odd} \end{cases} \tag{5.43}$$

and can be plotted using the following steps:

1. Flip $x(t)$ over the time axis, then shift the flipped segment to the right by $T_0/2 = L$ units.
2. Horizontally shift the overall new plot by nT_0 units for all integers n.

QUARTER-WAVE EVEN SYMMETRY. If $x(t)$ is function defined on interval $t \in (0, L)$, then periodic extension $\tilde{x}(t)$ can have even and half-wave symmetry for period $T_0 = 4L$ with Fourier coefficients

$$a_0 = 0 \tag{5.44}$$

$$a_n = \begin{cases} 0, & n \text{ even} \\ \frac{8}{T_0} \int_0^{T_0/4} x(t) \cos(n\omega_0 t)\, dt, & n \text{ odd} \end{cases} \tag{5.45}$$

$$b_n = 0 \tag{5.46}$$

and can be plotted using the following steps:

1. Flip $x(t)$ over the vertical axis. The overall new plot is a half-wave.
2. Flip the half-wave over the time axis and shift it to the right by $T_0/2 = 2L$ units.
3. Horizontally shift the overall new plot by nT_0 units for all integers n.

QUARTER-WAVE ODD SYMMETRY. If $x(t)$ is function defined on interval $t \in (0, L)$, then periodic extension $\tilde{x}(t)$ can have odd and half-wave symmetry for period $T_0 = 4L$ with Fourier coefficients

$$a_0 = 0 \tag{5.47}$$

$$a_n = 0 \tag{5.48}$$

$$b_n = \begin{cases} 0, & n \text{ even} \\ \frac{8}{T_0} \int_0^{T_0/4} x(t) \sin(n\omega_0 t)\, dt, & n \text{ odd} \end{cases} \tag{5.49}$$

and can be plotted using the following steps:

1. Flip $x(t)$ over the vertical axis, then flip the mirrored segment over the time axis. The overall new plot is a half-wave.
2. Flip the half-wave over the time axis and shift it to the right by $T_0/2 = 2L$ units.
3. Horizontally shift the overall new plot by nT_0 units for all integers n.

Example 5.4.1. Plot the following types of periodic extensions of $x(t)$, given that

$$x(t) = (1 - t)[u(t) - u(t - 1)].$$

(a) general periodic extension

(b) even periodic extension

(c) odd periodic extension

(d) half-wave symmetry

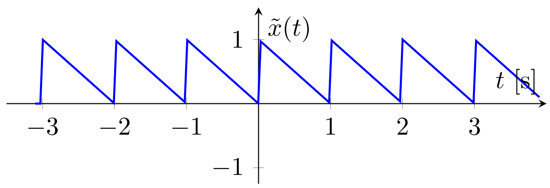
(e) quarter-wave even symmetry

(f) quarter-wave odd symmetry

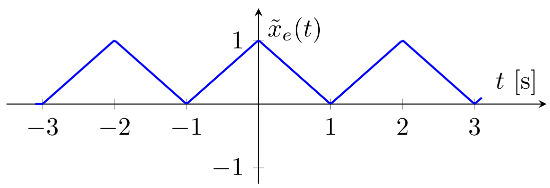
Then find the Fourier series of the quarter-wave odd symmetric periodic extension.

SOLUTION

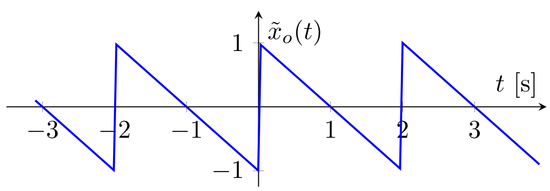
Applying a general periodic extension, we get $\tilde{x}(t)$ with period $T_0 = 1$:



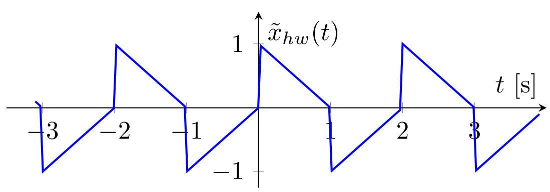
Applying an even periodic extension, we get $\tilde{x}_e(t)$ with period $T_0 = 2$:



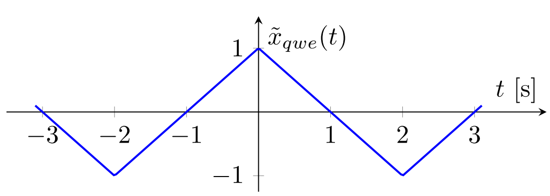
Applying an odd periodic extension, we get $\tilde{x}_o(t)$ with period $T_0 = 2$:



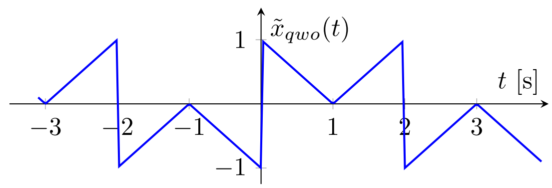
Applying half-wave symmetry, we get $\tilde{x}_{hw}(t)$ with period $T_0 = 2$:



Applying quarter-wave even symmetry, we get $\tilde{x}_{qwe}(t)$ with period $T_0 = 4$:



Applying quarter-wave odd symmetry, we get $\tilde{x}_{qwo}(t)$ with period $T_0 = 4$:



The Fourier coefficients for $\tilde{x}_{qwo}(t)$ is given by

$$\begin{aligned}
a_0 &= 0 \\
a_n &= 0 \\
b_n\big|_{n \text{ even}} &= 0 \\
b_n\big|_{n \text{ odd}} &= \frac{8}{T_0}\int_0^{T_0/4} x(t)\sin(n\omega_0 t)\,dt = \frac{8}{4}\int_0^1 (1-t)\sin\left(\frac{n\pi t}{2}\right)\,dt = \frac{4n\pi - 8\sin\left(\frac{n\pi}{2}\right)}{n^2\pi^2}
\end{aligned}$$

The Fourier series of $\tilde{x}_{qwo}(t)$ is then

$$\tilde{x}_{qwo}(t) = \sum_{\substack{n=1 \\ n=\text{odd}}}^{\infty} \left[\frac{4n\pi - 8\sin\left(\frac{n\pi}{2}\right)}{n^2\pi^2}\right]\sin\left(\frac{n\pi t}{2}\right)$$

5.5 Convergence of the Fourier Series

Earlier, it was prefaced that any physically realizable periodic signal follows Fourier's theorem; we assume that there are no periodic infinite discontinuities as those cannot be physically realized.

In fact, in order for a periodic signal $x(t)$ with period T_0 to follow Fourier's theorem (i.e., be physically realizable), it must satisfy the *Dirichlet conditions*:

1. $x(t)$ has a finite number of finite discontinuities in each period.
2. $x(t)$ has a finite number of maxima and minima in each period.
3. $x(t)$ is absolutely integrable over a period such that $\displaystyle\int_0^{T_0} |x(t)|\,dt < \infty$.

If the Dirichlet conditions are met, then there are two possible functions the Fourier series converges to, depending on the number of jump discontinuities within a period:

- If the periodic signal $x(t)$ is continuous with no jump discontinuities, then the Fourier series converges to $x(t)$.
- If the periodic signal $x(t)$ has nonzero number of jumps within a period, then the Fourier series converges to $x(t)$ everywhere except at the jumps, where the points converge to $\frac{1}{2}[x(t_i^-) + x(t_i^+)]$ for each jump discontinuity at $t = t_i$.

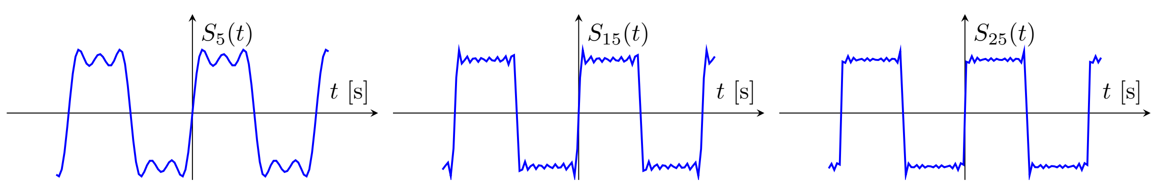
In general, for $S_N(t)$ is the N^{th} partial sum of the Fourier series,

$$\lim_{N\to\infty} S_N(t) = \begin{cases} x(t), & \text{if } x \text{ is continuous at } t \\ \frac{1}{2}[x(t^-) + x(t^+)], & \text{if } x \text{ has a jump at } t \end{cases} \tag{5.50}$$

In fact, while periodic signals with jump discontinuities such as the square wave can be represented as a Fourier series, they cannot be reproduced practically in the real world by generating and summing an infinite number of sine waves.

One could get close with a large N^{th} partial sum of the Fourier series. However, regardless of the value of N, the *Gibbs phenomenon* will be observed, in which oscillatory overshoots occur about the points of discontinuity, as seen in Figure 5.2.

Figure 5.2: Gibbs phenomenon observed in the partial sum $S_N(t)$ of a square wave.



5.6 Parseval's Theorem for Fourier Series

Parseval's theorem for Fourier series is essentially a "conservation of (average) power" theorem. When physically realizable periodic signals are mapped from the continuous-time Fourier series to the discrete-frequency Fourier spectrum, the total signal average power is conserved. Depending on the Fourier series representation used, it follows that the total average power of a physically realizable periodic signal $x(t)$ can be evaluated as:

$$P_x = \frac{1}{T_0}\int_0^{T_0} |x(t)|^2\, dt = a_0^2 + \sum_{n=1}^{\infty}(a_n^2 + b_n^2)/2 \tag{5.51}$$

$$P_x = \frac{1}{T_0}\int_0^{T_0} |x(t)|^2\, dt = c_0^2 + \sum_{n=1}^{\infty} c_n^2/2 \tag{5.52}$$

$$P_x = \frac{1}{T_0}\int_0^{T_0} |x(t)|^2\, dt = \sum_{n=-\infty}^{+\infty} |x_n|^2 \tag{5.53}$$

Here, the *DC power* is given by $a_0^2 = c_0^2 = |x_0|^2$ whereas the rest of the respective summation is the *average AC power*.

The *AC power fraction* is then given by

$$pf = \frac{P_{AC}}{P_{DC} + P_{AC}} \tag{5.54}$$

Additionally, we can define the *one-sided power spectral density* (1-sided PSD) to be

$$PSD_1 = \begin{cases} c_0^2 = |x_0|^2, & n = 0 \\ \frac{c_n^2}{2} = 2|x_n|^2, & n > 0 \end{cases} \tag{5.55}$$

and the *two-sided power spectral density* (2-sided PSD) to be

$$PSD_2 = |x_n|^2. \tag{5.56}$$

Note that the 2-sided PSD is defined for all frequencies, whereas the 1-sided PSD is defined for only nonnegative frequencies. Because of this, the AC component of the 2-sided PSD are half the values of the 1-sided PSD. In signal processing, we tend to be more interested in the 1-sided PSD.

Example 5.6.1. The current flowing through a 12 [Ω] resistor is

$$i(t) = 2 + 4\cos(377t - 30^\circ)\ [\text{A}].$$

Find the average power consumed by the resistor. Then find the AC power fraction.

SOLUTION

The equation for current closely resembles the amplitude/phase representation of the Fourier series, with $c_0 = 2$ and $c_1 = 4$. By Parseval's theorem, the normalized average power is given by

$$P_{av,norm} = c_0^2 + \frac{c_1^2}{2} = 2^2 + \frac{4^2}{2} = 12.$$

Again, note that this is the normalized average power. Since the $P_{av,norm}$ above has units of $[\text{A}^2]$, the actual average power is given by

$$P = P_{av,norm}R,$$

for $R = 12$ [Ω]. Therefore, the actual average power absorbed by the resistor is

$$P_{av} = P_{av,norm}R = 12 \cdot 12 = 144\ [\text{W}].$$

We can isolate the DC power from the average AC power to get

$$P_{DC,norm} = c_0^2 = 2^2 = 4$$

$$P_{AC,norm} = \frac{c_1^2}{2} = 8$$

The AC power fraction is then

$$pf = \frac{P_{AC,norm}}{P_{DC,norm} + P_{AC,norm}} = \frac{8}{4+8} = \frac{2}{3} \Longrightarrow 66.7\%$$

Note that for AC power fraction calculations, using normalized power values should yield the same answer as using actual power values.

5.7 LTI Systems with Fourier Series

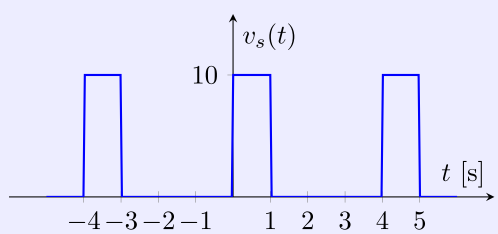
As we saw with one of the phasor transform examples, the superposition principle can be applied to find the output of an LTI system, given that the input signal is a finite sum of sinusoids. Extending the notion to Fourier series, we can represent a physically realizable periodic signal $x(t)$ as a Fourier series using the closely related amplitude/phase representation and use the superposition principle to find the output of an LTI system.

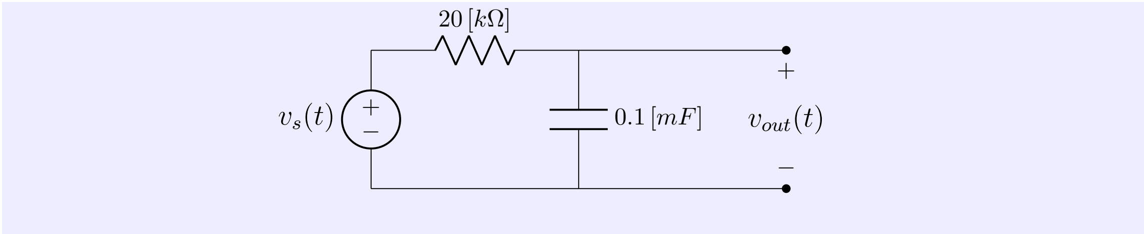
FOURIER SERIES ANALYSIS.

1. Express the input signal $x(t)$ as a Fourier series using the amplitude/phase representation such that $x(t) = c_0 + \sum_{n=1}^{\infty} c_n \cos(n\omega_0 t + \phi_n)$, with phasor $c_n e^{j\phi_n} = a_n - jb_n$.
2. Use the phasor transform to find the generic frequency response function $H(\omega)$ by letting $x(t) = 1 \cdot \cos(\omega t)$ such that $\underline{X} = 1$ and $\underline{Y} = H(\omega)$.
3. Determine the output signal for

$$\begin{aligned} y(t) &= c_0 H(\omega = 0) + \sum_{n=1}^{\infty} c_n \operatorname{Re}[H(\omega = n\omega_0) e^{jn\omega_0 t + \phi_n}] \\ &= c_0 H(\omega = 0) + \sum_{n=1}^{\infty} c_n |H(\omega = n\omega_0)| \cos(n\omega_0 t + \underline{/H(\omega = n\omega_0)} + \phi_n). \end{aligned}$$

Example 5.7.1. Given the plot of periodic input excitation $v_s(t)$, find the output $v_{out}(t)$ of the following RC circuit.





SOLUTION

First, we find a Fourier series expression for $v_s(t)$. From the plot, we can visually see that $T_0 = 4$. Therefore $\omega_0 = 2\pi/T_0 = \pi/2$. During a single period, we can define $v_s(t)$ as

$$v_s(t) = \begin{cases} 10, & 0 < t < 1 \\ 0, & 1 < t < 4 \end{cases}$$

Using the sine/cosine representation, the Fourier coefficients are

$$\begin{aligned} a_0 &= \frac{\text{Area of } x(t) \text{ during } T_0}{\text{Period } T_0 \text{ of } x(t)} = \frac{10}{4} = 2.5 \\ a_n &= \frac{2}{T_0}\int_0^{T_0} v_s(t)\cos(n\omega_0 t)\,dt = \frac{2}{4}\int_0^1 10\cos(n\omega_0 t)\,dt = \frac{10}{n\pi}\sin\left(\frac{n\pi}{2}\right) \\ b_n &= \frac{2}{T_0}\int_0^{T_0} v_s(t)\sin(n\omega_0 t)\,dt = \frac{2}{4}\int_0^1 10\sin(n\omega_0 t)\,dt = \frac{10}{n\pi}\left[1 - \cos\left(\frac{n\pi}{2}\right)\right] \end{aligned}$$

We can then transform the Fourier coefficients to the amplitude/phase representation such that the DC component is given by

$$c_0 = a_0 = 2.5,$$

the amplitudes are given by

$$\begin{aligned} c_n &= \sqrt{a_n + b_n} = \frac{10}{n\pi}\sqrt{\sin^2\left(\frac{n\pi}{2}\right) + \left[1 - \cos\left(\frac{n\pi}{2}\right)\right]^2} = \frac{10}{n\pi}\sqrt{2 - 2\cos\left(\frac{n\pi}{2}\right)} \\ &= \frac{20}{n\pi}\sqrt{\frac{1}{2}\left[1 - \cos\left(\frac{n\pi}{2}\right)\right]} = \frac{20}{n\pi}\sin\left(\frac{n\pi}{4}\right) \end{aligned}$$

and the phases are given by

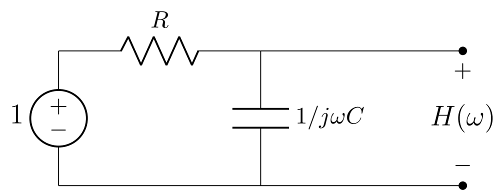
$$\begin{aligned} \phi_n &= \begin{cases} -\arctan(b_n/a_n), & a_n > 0 \\ \pi - \arctan(b_n/a_n), & a_n < 0 \end{cases} \\ &= -\operatorname{atan2}(b_n, a_n) \\ &= -\operatorname{atan2}\left[1 - \cos\left(\frac{n\pi}{2}\right), \sin\left(\frac{n\pi}{2}\right)\right] \\ &= -\frac{(n \bmod 4)\pi}{4} \end{aligned}$$

It just so happens that when plotting the *atan*2 function above, the plot has a period of 4. We can utilize this to apply modulo 4 to the values of n when writing a final expression for ϕ_n.

The input voltage can then be written as

$$v_s(t) = 2.5 + \sum_{n=1}^{\infty} \frac{20}{n\pi} \sin\left(\frac{n\pi}{4}\right) \cos\left(n\omega_0 t - \frac{(n \bmod 4)\pi}{4}\right).$$

Now we convert the circuit to phasor domain with $\underline{V_s} = 1$ and $\underline{V_{out}} = H(\omega)$ such that



The frequency response is given by

$$\begin{aligned} H(\omega) &= \frac{1/j\omega C}{R + 1/j\omega C} = \frac{1}{1 + j\omega RC} \\ |H(\omega)| &= \frac{1}{\sqrt{1 + 4\omega^2}} \\ \underline{/H(\omega)} &= -\arctan(2\omega) \end{aligned}$$

Then for

$$\begin{aligned} H(\omega = 0) &= 1 \\ |H(\omega = n\omega_0)| &= \frac{1}{\sqrt{1 + (n\pi)^2}} \\ \underline{/H(\omega = n\omega_0)} &= -\arctan(n\pi), \end{aligned}$$

the output voltage is given by

$$v_{out}(t) = 2.5 + \sum_{n=1}^{\infty} \left[\frac{20}{n\pi} \sin\left(\frac{n\pi}{4}\right)\right] \left[\frac{1}{\sqrt{1 + (n\pi)^2}}\right] \cos\left(n\omega_0 t - \arctan(n\pi) - \frac{(n \bmod 4)\pi}{4}\right).$$

■

Chapter 6

Fourier Transform

In the previous chapter, we analyzed periodic signals using the Fourier series. However, not all signals are periodic. While the Fourier transform will not be derived here, the Fourier transform is closely related to the Fourier series, except now letting $T \to \infty$ to account for nonperiodic waveforms.

Generally, the Fourier transform can be applied to any signal. However, for best practices, the following table describes what technique is best used for certain signals.

Input $x(t)$		Solution Method	Output $y(t)$
Duration	Waveform		
Everlasting	Sinusoid	Phasor Transform	Steady-State Component
Everlasting	Periodic	Fourier Series and Phasor Transform	Steady-State Component
Causal ($x(t) = 0$ for $t < 0$)	Any	Unilateral Laplace Transform	Complete Solution (transient + steady-state)
Everlasting	Any	Fourier Transform or Bilateral Laplace Transform	Complete Solution (transient + steady-state)

6.1 Fourier Transform

The *Fourier transform* of a signal $x(t)$ is defined as

$$X(\omega) = \mathcal{F}[x(t)] = \int_{-\infty}^{+\infty} x(t)e^{-j\omega t}\, dt. \tag{6.1}$$

Similar to the Laplace transform, the Fourier transform is essentially an operator that maps a signal defined in the *time domain* to another signal defined in the *frequency domain.* $x(t)$ and $X(\omega)$ constitute a unique *Fourier transform pair.* This relationship can be written as

$$x(t) \iff X(\omega). \tag{6.2}$$

Because of this relationship, there exists an *inverse Fourier transform* defined by

$$x(t) = \mathcal{F}^{-1}[X(\omega)] = \frac{1}{2\pi}\int_{-\infty}^{+\infty} X(\omega)e^{+j\omega t}\, d\omega, \tag{6.3}$$

though referencing a table of Fourier transform pairs is a lot simpler than computing the integral. Interestingly, the Fourier transform is related to the bilateral Laplace transform:

$$X(\omega) = X(s)\big|_{s=j\omega} \tag{6.4}$$

The Dirichlet conditions for the Fourier transform is as follows:

1. $x(t)$ has a finite number of finite discontinuities in every finite interval of time.
2. $x(t)$ has a finite number of maxima and minima in every finite interval of time.
3. $x(t)$ is absolutely integrable such that $\int_{-\infty}^{+\infty} |x(t)|\, dt < \infty$.

If the Dirichlet conditions are met, then $x(t)$ is physically realizable and has a Fourier transform. However, these are sufficient but not necessary conditions as some functions like the constant function or the unit step function are not absolutely integrable, yet they have Fourier transforms. (Instead, limit definitions of those functions are evaluated to obtain the Fourier transform.)

Table 6.1 lists the properties of the Fourier transform. From the conjugate symmetry property, there are important implications about the properties of $x(t)$ and its Fourier transform $X(\omega)$:

$x(t)$	$X(\omega)$
Real and even	Real and even
Real and odd	Imaginary and odd
Imaginary and even	Imaginary and even
Imaginary and odd	Real and odd

Example 6.1.1. Find the Fourier transform of the constant function $x(t) = 1$.

SOLUTION

Consider the delta function $\delta(t)$. Its Fourier transform is given by

$$\mathcal{F}[\delta(t)] = \int_{-\infty}^{+\infty} \delta(t) e^{-j\omega t}\, dt = 1.$$

Then by the duality property, the Fourier transform of 1 is

$$\mathcal{F}[1] = 2\pi\delta(-\omega) = 2\pi\delta(\omega).$$

■

Example 6.1.2. Find the Fourier transform of the unit step function $u(t)$.

SOLUTION

From the previous example, the Fourier transform of the delta function $\delta(t)$ is 1. Recall that the unit step function can be defined as

$$u(t) = \int_{-\infty}^{t} \delta(\tau)\, d\tau.$$

Table 6.1: Properties of the Fourier transform.

Property	$x(t)$	$X(\omega) = \mathcal{F}[x(t)]$
Superposition	$K_1 x_1(t) + K_2 x_2(t)$	$K_1 X_1(\omega) + K_2 X_2(\omega)$
Time scaling	$x(at)$	$\frac{1}{\lvert a\rvert} X\left(\frac{\omega}{a}\right)$
Time shift	$x(t - t_0)$	$e^{-j\omega t_0} X(\omega)$
Frequency shift	$e^{j\omega_0 t} x(t)$	$X(\omega - \omega_0)$
Time derivative	$x'(t) = \frac{dx(t)}{dt}$	$j\omega X(\omega)$
Time nth derivative	$x^{(n)}(t) = \frac{d^n x(t)}{dt^n}$	$(j\omega)^n X(\omega)$
Time integral	$\int_{-\infty}^{t} x(\tau)\, d\tau$	$\frac{X(\omega)}{j\omega} + \pi\delta(\omega) \int_{-\infty}^{+\infty} x(t)\, dt$
Frequency derivative	$t^n x(t)$	$(j)^n \frac{d^n X(\omega)}{d\omega^n} = (j)^n X^{(n)}(\omega)$
Convolution	$x_1(t) * x_2(t)$	$X_1(\omega) X_2(\omega)$
Multiplication	$x_1(t) x_2(t)$	$X_1(\omega) * X_2(\omega)$
Modulation	$x(t)\cos(\omega_0 t)$	$\frac{1}{2}[X(\omega - \omega_0) + X(\omega + \omega_0)]$
Duality	$X(t)$	$2\pi x(-\omega)$
Conjugate symmetry	$x(t)$ real	$\begin{cases} X(-\omega) = X^*(\omega) \\ \mathrm{Re}(X(\omega)) = \mathrm{Re}(X(-\omega)) \\ \mathrm{Im}(X(\omega)) = -\mathrm{Im}(X(\omega)) \\ \lvert X(\omega)\rvert = \lvert X(-\omega)\rvert \\ \arg[X(\omega)] = -\arg[X(-\omega)] \end{cases}$
Even-odd decomposition of real signals	$\begin{cases} x_e(t) = \frac{1}{2}[x(t) + x(-t)] \\ x_o(t) = \frac{1}{2}[x(t) - x(-t)] \end{cases}$	$\begin{cases} \mathrm{Re}(X(\omega)) \\ j\,\mathrm{Im}(X(\omega)) \end{cases}$

Then by the time integral property, the Fourier transform of $u(t)$ is

$$\mathcal{F}[u(t)] = \mathcal{F}\left[\int_{-\infty}^{t} \delta(\tau)\, d\tau\right] = \frac{1}{j\omega} + \pi\delta(\omega)\int_{-\infty}^{\infty} \delta(t)\, dt = \frac{1}{j\omega} + \pi\delta(\omega).$$

■

Table 6.2 lists some common Fourier transform pairs. While one could find the inverse Fourier transform using the inverse Fourier integral, it is much simpler to reference a table of transform pairs.

6.2 Fourier Transform of Periodic Signals

Recall that the exponential form of the Fourier series for a physically realizable periodic signal $x(t)$ is given by

$$x(t) = \sum_{n=-\infty}^{+\infty} x_n e^{jn\omega_0 t}, \tag{6.5}$$

with exponential Fourier coefficients

$$x_n = \frac{1}{T_0}\int_{t_0}^{t_0+T_0} x(t)e^{-jn\omega_0 t}\, dt. \tag{6.6}$$

The Fourier transform of periodic signal $x(t)$ with exponential Fourier coefficients x_n is then

$$\mathcal{F}[x(t)] = \mathcal{F}\left[\sum_{n=-\infty}^{+\infty} x_n e^{jn\omega_0 t}\right] = \sum_{n=-\infty}^{+\infty} \mathcal{F}[x_n e^{jn\omega_0 t}] = \sum_{n=-\infty}^{+\infty} x_n 2\pi\delta(\omega - n\omega_0). \tag{6.7}$$

6.3 Parseval's Theorem for Fourier Transforms

Similar to how Parseval's theorem for Fourier series is essentially a "conservation of (average) power" theorem as periodic signals are power signals, Parseval's theorem for Fourier transforms is essentially a "conservation of energy" theorem for any physically realizable signal (aperiodic or periodic). When signals are mapped from the continuous-time function to the continuous-frequency Fourier spectrum, the total signal energy is conserved. It follows that the total energy of a physically realizable signal $x(t)$ can be evaluated as:

$$E_x = \int_{-\infty}^{+\infty} |x(t)|^2\, dt = \frac{1}{2\pi}\int_{-\infty}^{+\infty} |X(\omega)|^2\, d\omega \tag{6.8}$$

Additionally, we can define the *one-sided energy spectral density* (1-sided ESD) to be

$$ESD_1 = \frac{1}{\pi}|X(\omega)|^2, \tag{6.9}$$

and the *two-sided energy spectral density* (2-sided ESD) to be

$$ESD_2 = \frac{1}{2\pi}|X(\omega)|^2. \tag{6.10}$$

Note that the 2-sided ESD is defined for all frequencies, whereas the 1-sided ESD is defined for only nonnegative frequencies. Because of this, the values of the 2-sided ESD are half the values of the 1-sided ESD. In signal processing, we tend to be more interested in the 1-sided ESD.

Table 6.2: Fourier transform pairs.

$x(t)$	$X(\omega) = \mathcal{F}[x(t)]$
$\delta(t)$	1
$\delta(t - t_0)$	$e^{-j\omega t_0}$
1	$2\pi\delta(\omega)$
$u(t)$	$\pi\delta(\omega) + \frac{1}{j\omega}$
$\mathrm{sgn}(t)$	$\frac{2}{j\omega}$
$\mathrm{rect}\left(\frac{t}{\tau}\right)$	$\tau\,\mathrm{Sa}\left(\frac{\omega\tau}{2}\right)$
$\frac{\exp[-t^2/(2\sigma^2)]}{\sqrt{2\pi\sigma^2}}$	$\exp[-\omega^2\sigma^2/2]$
$e^{-at}u(t)$	$\frac{1}{a + j\omega}$
$e^{at}u(-t)$	$\frac{1}{a - j\omega}$
$\cos(\omega_0 t)$	$\pi[\delta(\omega - \omega_0) + \delta(\omega + \omega_0)]$
$\sin(\omega_0 t)$	$\frac{\pi}{j}[\delta(\omega - \omega_0) - \delta(\omega + \omega_0)]$
$e^{j\omega_0 t}$	$2\pi\delta(\omega - \omega_0)$
$te^{-at}u(t)$	$\frac{1}{(a + j\omega)^2}$
$e^{-at}\sin(\omega_0 t)u(t)$	$\frac{\omega_0}{(a + j\omega)^2 + \omega_0^2}$
$\sin(\omega_0 t)u(t)$	$\frac{\pi}{2j}[\delta(\omega - \omega_0) - \delta(\omega + \omega_0)] + \frac{\omega_0}{\omega_0^2 - \omega^2}$
$e^{-at}\cos(\omega_0 t)u(t)$	$\frac{a + j\omega}{(a + j\omega)^2 + \omega_0^2}$
$\cos(\omega_0 t)u(t)$	$\frac{\pi}{2}[\delta(\omega - \omega_0) - \delta(\omega + \omega_0)] + \frac{j\omega}{\omega_0^2 + \omega^2}$

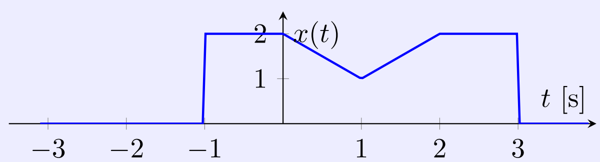
Example 6.3.1. The signal $x(t)$ is plotted below. Use the properties of the Fourier transform to find:

(a) $\angle X(\omega)$

(b) $X(\omega)$ at $\omega = 0$

(c) $\int_{-\infty}^{+\infty} X(\omega)\, d\omega$

(d) $\int_{-\infty}^{+\infty} |X(\omega)|^2\, d\omega$



SOLUTION

For part (a), notice that $x(t)$ has hidden symmetry. If we let $y(t) = x(t+1)$, then $y(t)$ is real and even. Therefore,

$$y(t) \text{ real and even} \iff Y(\omega) \text{ real and even}$$

Since $Y(\omega)$ is purely real, it has zero phase. Then

$$\begin{aligned} y(t) = x(t+1) \iff Y(\omega) &= X(\omega)e^{j\omega} \\ \angle Y(\omega) &= \angle X(\omega) + \omega = 0. \end{aligned}$$

Then the phase of $X(\omega)$ is

$$\angle X(\omega) = -\omega.$$

Substituting $\omega = 0$ into the Fourier transform integral for part (b), we get

$$X(\omega = 0) = \int_{-\infty}^{+\infty} x(t)\, dt = [\text{Area under the curve}] = 7$$

For part (c), we can manipulate the inverse Fourier transform integral such that

$$\begin{aligned} x(t) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} X(\omega) e^{j\omega t}\, d\omega \\ \implies 2\pi x(t=0) &= \int_{-\infty}^{+\infty} X(\omega)\, d\omega = 2\pi \times 2 = 4\pi \end{aligned}$$

Lastly, for part (d), by Parseval's theorem,

$$\int_{-\infty}^{+\infty} |X(\omega)|^2\, d\omega = 2\pi \int_{-\infty}^{+\infty} |x(t)|^2\, dt = \frac{76\pi}{3}$$

6.4 LTI Systems with Fourier Transforms

Similar to the Laplace transform, LTI systems characterized by LCCDEs can be solved by applying the Fourier transform, computing the ω-domain solution, and mapping the solution back to the domain. While the unilateral Laplace transform can only be used on causal signals and systems, the Fourier transform can be used on all physically realizable signals and LTI systems, though everlasting signals are more adequate.

An LCCDE of the form

$$\sum_{\ell=0}^{N} a_{N-\ell} \frac{d^\ell y(t)}{dt^\ell} = \sum_{\ell=0}^{M} b_{M-\ell} \frac{d^\ell x(t)}{dt^\ell} \tag{6.11}$$

can be transformed to the ω-domain to get

$$\left[\sum_{\ell=0}^{N} a_{N-\ell} (j\omega)^\ell\right] Y(\omega) = \left[\sum_{\ell=0}^{M} b_{M-\ell} (j\omega)^\ell\right] X(\omega). \tag{6.12}$$

Example 6.4.1. An LTI system is characterized by the LCCDE

$$\frac{d^2y(t)}{dt^2} + 7\frac{dy(t)}{dt} + 12y(t) = \frac{dx(t)}{dt} + 2x(t).$$

Find the system response to input $x(t) = e^{-2t}u(t)$.

SOLUTION

Transforming the LCCDE to ω-domain, we get

$$\begin{aligned} [(j\omega)^2 + 7j\omega + 12]Y(\omega) &= [j\omega + 2]X(\omega) \\ \Longrightarrow Y(\omega) &= \left[\frac{2 + j\omega}{(j\omega)^2 + 7j\omega + 12}\right] X(\omega) \end{aligned}$$

From the table of Fourier transform pairs, we see that

$$\mathcal{F}[e^{-2t}u(t)] = \frac{1}{2 + j\omega}$$

Therefore, by applying partial fraction expansion with respect to $j\omega$,

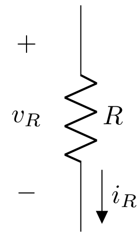
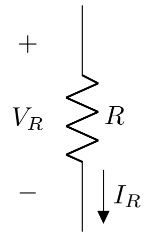
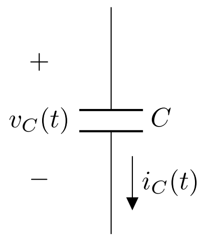
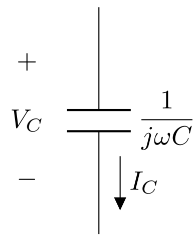
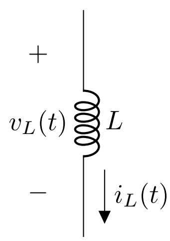
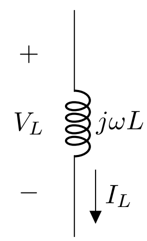
$$Y(\omega) = \left[\frac{2 + j\omega}{(j\omega)^2 + 7j\omega + 12}\right] \frac{1}{2 + j\omega} = \frac{1}{(j\omega)^2 + 7j\omega + 12} = \frac{1}{3 + j\omega} - \frac{1}{4 + j\omega},$$

we get the solution

$$y(t) = e^{-3t}u(t) - e^{-4t}u(t).$$

■

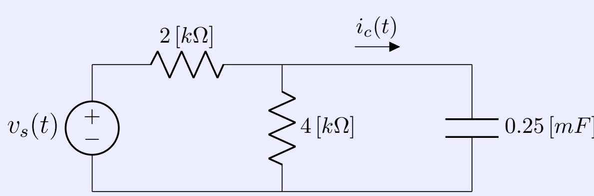
Table 6.3: Circuit models for electrical components in the ω-domain.

Electrical Component	Time Domain	ω-Domain
Resistor		
Capacitor		
Inductor		

Just as differential equations can be solved in the ω-domain, the systems represented by differential equations can be solved in the ω-domain. For instance, an electric circuit can be transformed to an ω-domain equivalent and solved using algebraic equations, with the ω-domain electrical components referred to as impedances. The equivalent ω-domain models can be seen in Table 6.3.

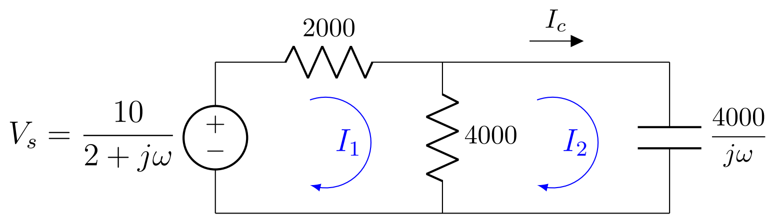
Example 6.4.2. Find an equation for the current flowing through the capacitor $i_c(t)$ in the following circuit, given that the input excitation is

$$v_s(t) = 10e^{-2t}u(t)\ [\mathrm{V}].$$



SOLUTION

First, we transform the circuit to the ω-domain.



Then by using the mesh current method, we get

$$-V_s + 2000I_1 + 4000(I_1 - I_2) = 0 \Longrightarrow I_1 = \frac{V_s + 4000I_2}{6000}$$

$$4000(I_2 - I_1) + \frac{4000}{j\omega} I_2 = 0 \Longrightarrow I_1 = \left(1 + \frac{1}{j\omega}\right) I_2$$

Setting the two equations together and using partial fraction expansion with respect to $j\omega$:

$$I_2 = \frac{V_s}{6000} \cdot \frac{3j\omega}{3+j\omega} = \frac{(5\times 10^{-3})j\omega}{(2+j\omega)(3+j\omega)} = \left(\frac{-10}{2+j\omega} + \frac{15}{3+j\omega}\right) \times 10^{-3}$$

The current running through the capacitor is then given by

$$i_c(t) = \mathcal{F}^{-1}[I_2] = (-10e^{-2t} + 15e^{-3t})u(t) \text{ [mA]}.$$

6.5 Frequency Response

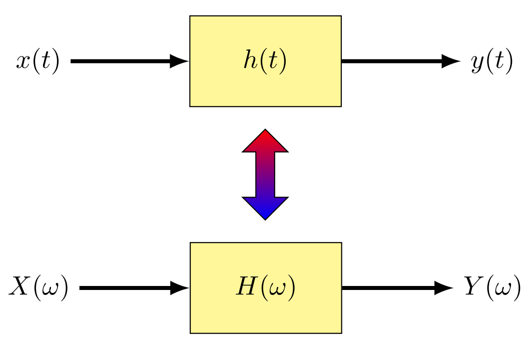
The *frequency response function* is the Fourier transform of the impulse response:

$$H(\omega) = \frac{Y(\omega)}{X(\omega)} = \mathcal{F}[h(t)]. \tag{6.13}$$

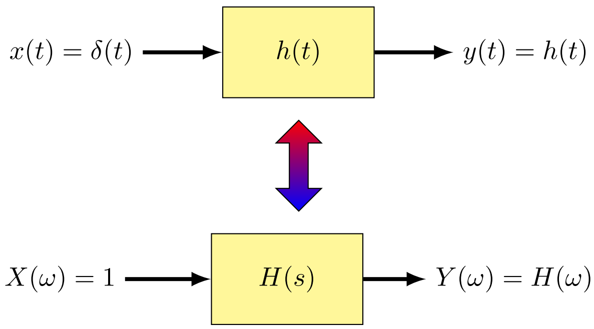
In fact, by the convolution property, it follows that

$$y(t) = x(t) * h(t) \iff Y(\omega) = X(\omega)H(\omega). \tag{6.14}$$

Symbolically,



As the impulse response is the system response to input $x(t) = \delta(t)$, consider the Fourier transform of the impulse signal $\mathcal{F}[\delta(t)] = 1$. Then it also follows that



Interestingly, the frequency response is related to the transfer function:

$$H(\omega) = H(s)\Big|_{s=j\omega} \tag{6.15}$$

As before, there are two approaches to finding the frequency response of an LTI system:

1. Find $X(\omega), Y(\omega)$. Then calculate $H(\omega) = Y(\omega)/X(\omega)$.
2. Find $X(\omega), Y(\omega)$. Then substitute $X(\omega) = 1$ and $Y(\omega) = H(\omega)$.

While the frequency response can be used for determining system stability just like the transfer function, the frequency response has the additional tool of analyzing the gain and phase of an LTI system at particular frequencies of interest, as borrowed from the Fourier series analysis.

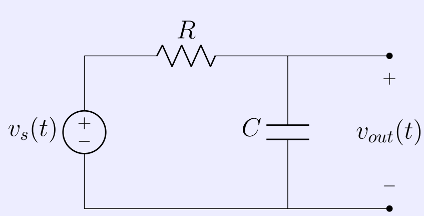
The magnitude $|H(\omega)|$ of the frequency response is called the *magnitude response* and has an associated *phase reponse* $\angle H(\omega)$. The plots of the magnitude response and phase response are two-sided and continuous and are collectively referred to as the *frequency spectrum*.

When expressed as $H(\omega) = \dfrac{N(\omega)}{D(\omega)}$, the magnitude and phase responses can be computed as

$$|H(\omega)| = \frac{|N(\omega)|}{|D(\omega)|} = \frac{\sqrt{[\mathrm{Re}(N(\omega))]^2 + [\mathrm{Im}(N(\omega))]^2}}{\sqrt{[\mathrm{Re}(D(\omega))]^2 + [\mathrm{Im}(D(\omega))]^2}} \tag{6.16}$$

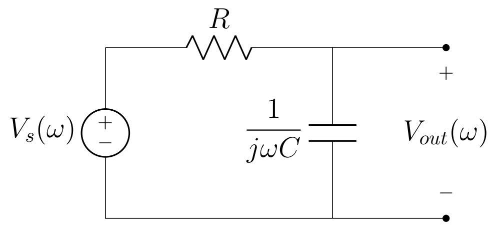
$$\angle H(\omega) = \angle N(\omega) - \angle D(\omega) = \arctan\left(\frac{\mathrm{Im}(N(\omega))}{\mathrm{Re}(N(\omega))}\right) - \arctan\left(\frac{\mathrm{Im}(D(\omega))}{\mathrm{Re}(D(\omega))}\right) \tag{6.17}$$

Example 6.5.1. Find the frequency response of the following RC circuit.



SOLUTION

Mapping to the ω-domain, we get



From here, we can use the voltage divider rule to get

$$V_{out}(\omega) = \left[\frac{1/j\omega C}{R + 1/j\omega C}\right] V_s(\omega)$$

The frequency response is then given by

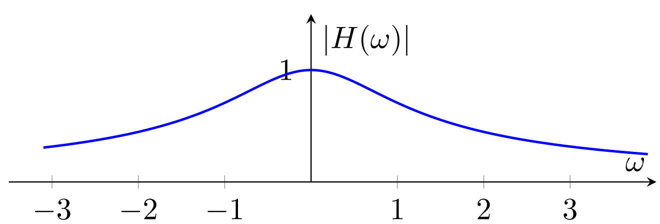
$$H(\omega) = \frac{V_{out}(\omega)}{V_s(\omega)} = \frac{1/j\omega C}{R + 1/j\omega C} = \frac{1}{1 + j\omega RC},$$

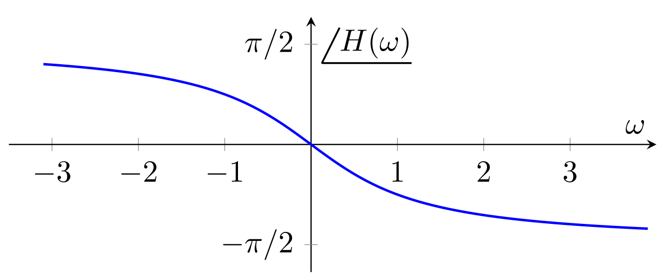
with magnitude and phase responses

$$|H(\omega)| = \frac{1}{\sqrt{1 + (\omega RC)^2}}$$

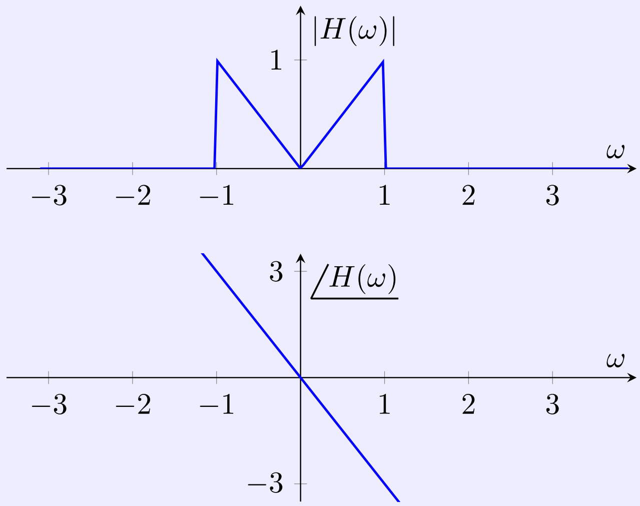
$$\angle H(\omega) = -\arctan(\omega RC)$$

Let $RC = 1$. The frequency spectrum is plotted below.





Example 6.5.2. Given the frequency spectrum of an LTI system plotted below, find the impulse response $h(t)$.



SOLUTION

The phase is given by $\angle H(\omega) = -3\omega$. The frequency response is then

$$H(\omega) = |H(\omega)| \exp(j\angle H(\omega)) = |H(\omega)|e^{-j3\omega}.$$

Let $G(\omega) = |H(\omega)|$. Then by the time shift property,

$$H(\omega) = G(\omega)e^{-j3\omega} \iff h(t) = g(t-3).$$

Solving for $g(t)$ first,

$$\begin{aligned} g(t) = \mathcal{F}^{-1}[G(\omega)] &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} G(\omega)e^{j\omega t}\, d\omega \\ &= \frac{1}{2\pi}\left[\int_{-1}^{0} -\omega e^{j\omega t}\, d\omega + \int_{0}^{1} \omega e^{j\omega t}\, d\omega\right] \\ &= \frac{1}{2\pi}\left[\frac{e^{-jt} + jte^{-jt} - 1}{t^2} + \frac{e^{+jt} - jte^{+jt} - 1}{t^2}\right] \\ &= \frac{1}{\pi}\left[\frac{\cos(t) - 1}{t^2} + \frac{\sin(t)}{t}\right] \end{aligned}$$

Then applying the time shift property,

$$h(t) = g(t-3) = \frac{1}{\pi}\left[\frac{\cos(t-3) - 1}{(t-3)^2} + \frac{\sin(t-3)}{t-3}\right].$$

6.6 Bode Plots

Rather than plotting the frequency spectrum with linear units, an alternative graph of the frequency response to consider is the *Bode plot*, which is comprised of the *Bode magnitude plot* and the *Bode phase plot*. The Bode plot logarithmically scales the ω-axis and additionally converts the magnitude values to decibels. As a result, Bode plots are plotted on semilog templates and are inherently one-sided.

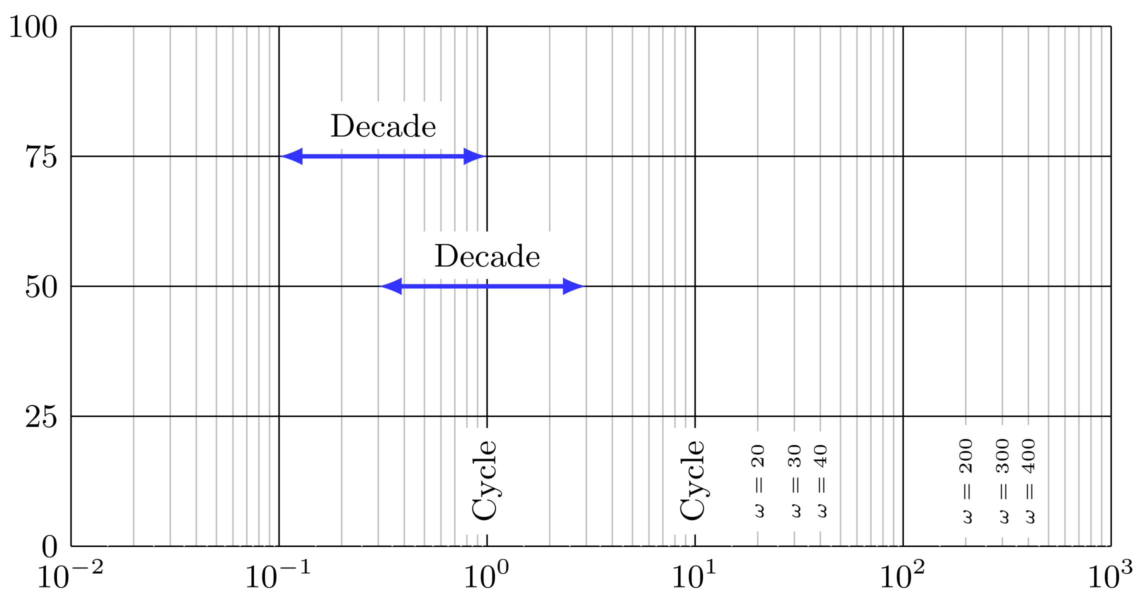
Additionally, Bode plots are convenient for providing straight-line approximations to the magnitude and phase responses. While the Bode plot does not truely give the same response as the frequency response, LTI systems can be modeled such that the asymptotic approximations are close enough to where the errors are minimal.

Figure 6.1 depicts a semilog template, where Bode plots can be drawn on. The main feature to examine is the logarithmically scaled ω-axis. Each prominent solid vertical line that takes on powers of 10 as its angular frequency ω is referred to as a *cycle*. While never depicted on Bode plots, the unseen angular frequency $\omega = 0$ is called the *origin*.

Any solid light tick marks after particular cycle represent multiples of that cycle – at least up until the next cycle. Generally, a *decade* (or *dec* for shorthand) is the horizontal spacing between one cycle and the next; this spacing however can also be shifted to start at any angular frequency.

Usually we choose an appropriate small power of 10 and appropriate large power of 10 for the minimum and maximum when labeling the ω-axis such that any and all behaviors due to the poles and zeros of the system are within scope.

Figure 6.1: Characteristics of a semilog template.



Before delving into producing Bode plots, first we must define the logarithmic unit decibels [dB]. Depending on how the frequency response was obtained, the magnitude response can be converted from unitless to decibels.

Most of the time, we assume that the input and output signals are voltages such that the frequency response is unitless $[V/V]$. Therefore, its magnitude response in decibels is given by

$$|H(\omega)|[dB] = 20\log\left(|H(\omega)|\left[\frac{V}{V}\right]\right)[dB]. \tag{6.18}$$

Table 6.4 lists some examples of decibel values corresponding to a select few magnitudes.

Table 6.4: Examples of corresponding decibel values.

M [V/V]	M [dB]
10^N	$20N$ [dB]
10^3	60 [dB]
100	40 [dB]
10	20 [dB]
4	≈ 12 [dB]
2	≈ 6 [dB]
$\sqrt{2}$	≈ 3 [dB]
1	0 [dB]
$1/\sqrt{2}$	≈ -3 [dB]
0.5	≈ -6 [dB]
0.25	≈ -12 [dB]
0.1	-20 [dB]
10^{-N}	$-20N$ [dB]

Example 6.6.1. Find an expression for $|H(\omega)| = \dfrac{1}{\sqrt{1+(\omega RC)^2}}$ [V/V] in decibels.

SOLUTION

It follows that

$$\begin{aligned}
20\log(|H(\omega)|) &= 20\log\left(\frac{1}{\sqrt{1+(\omega RC)^2}}\right) \\
&= 20\left[\log(1) - \log\left(\sqrt{1+(\omega RC)^2}\right)\right] \\
&= 20\left[\log(1) - \frac{1}{2}\log\left[1+(\omega RC)^2\right]\right] \\
&= -10\log\left[1+(\omega RC)^2\right] \text{ [dB]}
\end{aligned}$$

We can let $\omega_c = 1/(RC)$ such that

$$|H(\omega)| = -10\log\left[1+\left(\frac{\omega}{\omega_c}\right)^2\right] \text{ [dB]}$$

It follows that

$$\text{Low-frequency asymptote: } \left(\frac{\omega}{\omega_c}\right) \to 0 \implies |H(\omega)| \to 0 \text{ [dB]}$$

$$\text{High-frequency asymptote: } \left(\frac{\omega}{\omega_c}\right) \to \infty \implies |H(\omega)| \to -20\log\left(\frac{\omega}{\omega_c}\right) \text{ [dB]}$$

■

The standard form of a generic frequency response is given by

$$H(\omega) = \frac{\overset{\text{constant}}{K} \cdot \overset{\text{zero at origin}}{(j\omega)} \; \overset{\text{real zero}}{\left(1+\frac{j\omega}{\omega_0}\right)}}{\underset{\text{pole at origin}}{(j\omega)} \; \underset{\text{real pole}}{\left(1+\frac{j\omega}{\omega_0}\right)} \quad \underset{\text{multiplicity}}{N}}. \tag{6.19}$$

The corresponding effects of each term on the Bode plots are outlined in Table 6.5.

Note that the Bode magnitude plot of a real zero or real pole has a slope of ± 20 [dB/dec] for only $\omega > \omega_0$, and the Bode phase plot of a real zero or real pole has a slope of ± 45 [deg/dec] for only $0.1\omega_0 < \omega < 10\omega_0$. All other types of terms (i.e., constants and origin terms) have everlasting effects on the Bode plots – at least for the scope of $\omega > 0$.

Additionally, the multiplicity N (i.e., exponent) of a term scales that term's behavior on the Bode plots by a factor of N. For instance, the term $(j\omega)^N$ means it contributes a Bode magnitude slope of $20N$ [dB/dec] and a Bode phase of $90N$ [deg]. On the other hand, the term $\left(1+\frac{j\omega}{\omega_0}\right)^N$ contributes a Bode magnitude slope of $20N$ [dB/dec] for only $\omega > \omega_c$ and a Bode phase slope of $45N$ [deg/dec] for only $0.1\omega_0 < \omega < 10\omega_0$.

Similar to the linear combination of signals, the effects of each individual term of the frequency response can be added. Each term has a corresponding slope change on the Bode plot; these graphical properties are outlined in Table 6.6.

The same tips for waveform synthesis and deconstruction can be applied to Bode plots.

Table 6.5: Straight-line approximations of terms for Bode plots.

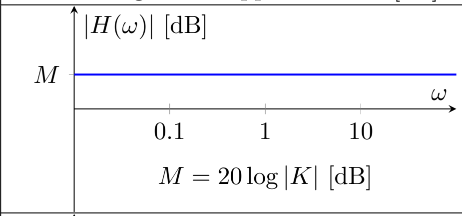
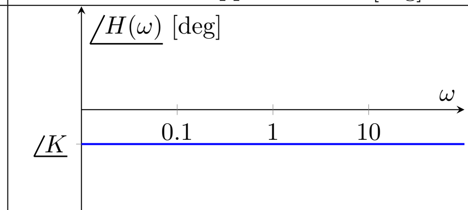
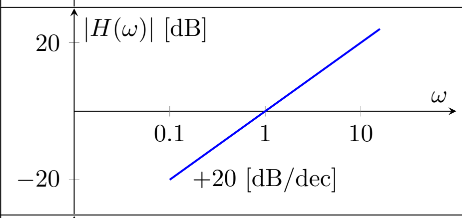
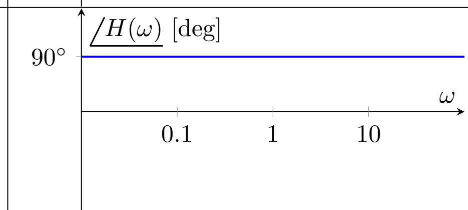
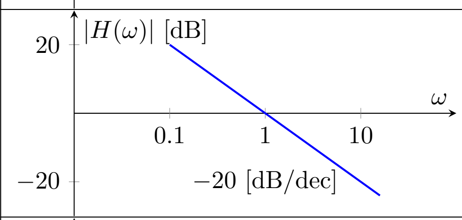
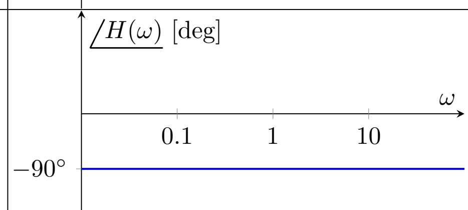
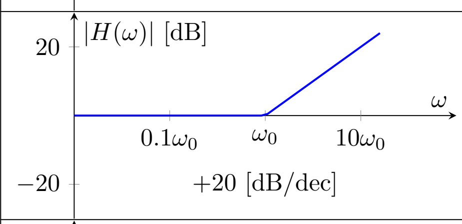
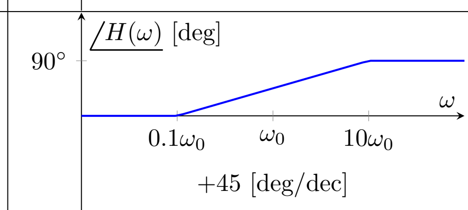
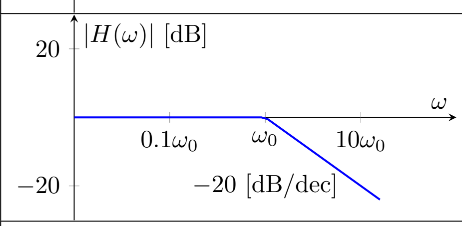
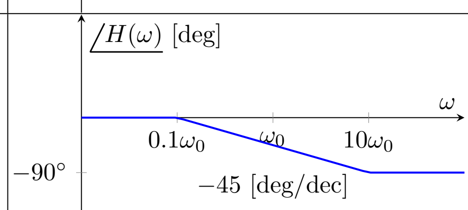
Term	Bode Magnitude Approximation [dB]	Bode Phase Approximation [deg]
Constant $H(\omega) = K$		
Zero at origin $H(\omega) = j\omega$		
Pole at origin $H(\omega) = \frac{1}{j\omega}$		
Real zero $H(\omega) = 1 + \frac{j\omega}{\omega_0}$		
Real pole $H(\omega) = \frac{1}{1+\frac{j\omega}{\omega_0}}$		
Multiplicity, N	Constant: $M \times N$ All zeros/poles: $N\times$(slope)	Constant/origin: $\angle H(\omega) \times N$ Real zeros/poles: $N\times$(slope)

Table 6.6: Graphical properties of frequency response terms on the Bode plot.

	Magnitude Slope Change $\Delta m(\Omega) = m(\Omega^+) - m(\Omega^-)$	Phase Slope Change $\Delta\phi(\Omega) = \phi(\Omega^+) - \phi(\Omega^-)$
Constant	None ($20\log\lvert K\rvert$ everywhere)	None ($\angle K$ everywhere)
Zero at origin	None (+20 [dB/dec] everywhere)	None (+90° everywhere)
Pole at origin	None (-20 [dB/dec] everywhere)	None ($-90°$ everywhere)
Real zero	$\Delta m(\omega_0) = +20$ [dB/dec]	$\Delta\phi(0.1\omega_0) = +45$ [deg/dec], $\Delta\phi(10\omega_0) = -45$ [deg/dec]
Real pole	$\Delta m(\omega_0) = -20$ [dB/dec]	$\Delta\phi(0.1\omega_0) = -45$ [deg/dec], $\Delta\phi(10\omega_0) = +45$ [deg/dec]

Bode plot graphing tips:

- Rewrite the frequency response in standard form
- Mark all starting and ending angular frequencies on the ω-axis
- Determine a starting value from any constant and origin terms (applying any multiplicity if necessary)
- From the starting point, make changes at each marked angular frequency, going from left to right along the ω-axis

Example 6.6.2. Sketch the Bode plots for the frequency response

$$H(\omega) = \frac{48j(20+4j\omega)(1000+2j\omega)}{(4+20j\omega)^2(400+2j\omega)}.$$

SOLUTION

Putting $H(\omega)$ in standard form,

$$\begin{aligned} H(\omega) &= \frac{48j\left[20\left(1+\frac{j\omega}{5}\right)\right]\left[1000\left(1+\frac{j\omega}{500}\right)\right]}{\left[4^2\left(1+\frac{j\omega}{0.2}\right)^2\right]\left[400\left(1+\frac{j\omega}{200}\right)\right]} \\ &= \frac{150j\left(1+\frac{j\omega}{5}\right)\left(1+\frac{j\omega}{500}\right)}{\left(1+\frac{j\omega}{0.2}\right)^2\left(1+\frac{j\omega}{200}\right)} \end{aligned}$$

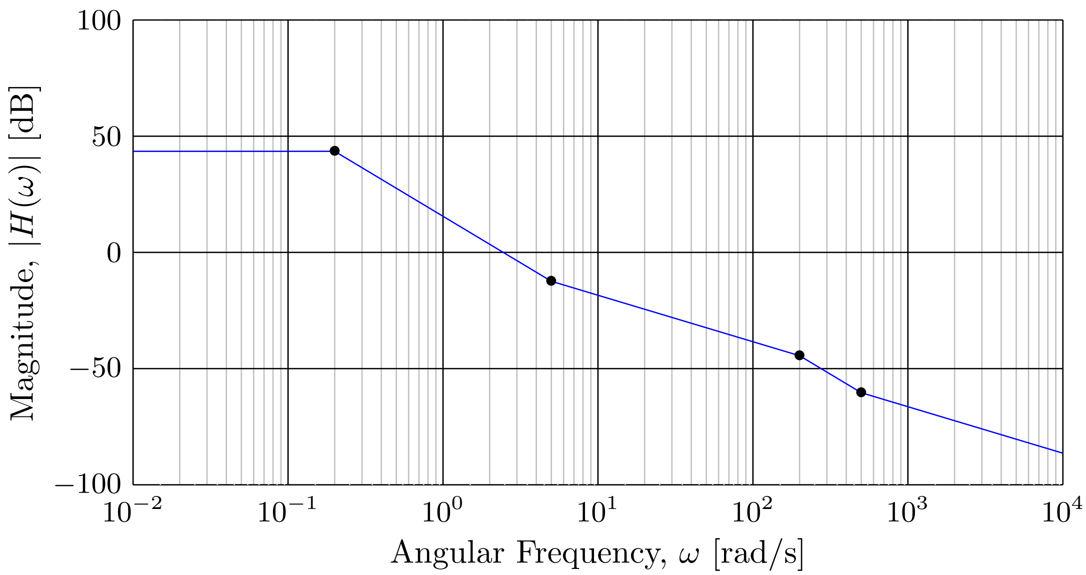
Focusing on just the Bode magnitude plot first, we make a note of all slope changes due to each term at each frequency.

$$\begin{aligned} 150j &\implies 20\log|K| = 43.5 \text{ [dB]} \\ \left(1+\frac{j\omega}{5}\right) &\implies \Delta m = 20 \text{ [dB/dec] at } \omega_0 = 5 \\ \left(1+\frac{j\omega}{500}\right) &\implies \Delta m = 20 \text{ [dB/dec] at } \omega_0 = 500 \\ \frac{1}{\left(1+\frac{j\omega}{0.2}\right)^2} &\implies \Delta m = -40 \text{ [dB/dec] at } \omega_0 = 0.2 \\ \frac{1}{1+\frac{j\omega}{200}} &\implies \Delta m = -20 \text{ [dB/dec] at } \omega_0 = 200 \end{aligned}$$

Collecting the magnitude changes with respect to angular frequency:

$$
\begin{aligned}
\omega = 0 &\Longrightarrow \text{start at } 43.5 \text{ [dB]} \\
\omega = 0.2 &\Longrightarrow \Delta m = -40 \\
\omega = 5 &\Longrightarrow \Delta m = 20 \\
\omega = 200 &\Longrightarrow \Delta m = -20 \\
\omega = 500 &\Longrightarrow \Delta m = 20
\end{aligned}
$$

Using the collection of magnitude changes, plot the Bode magnitude plot.



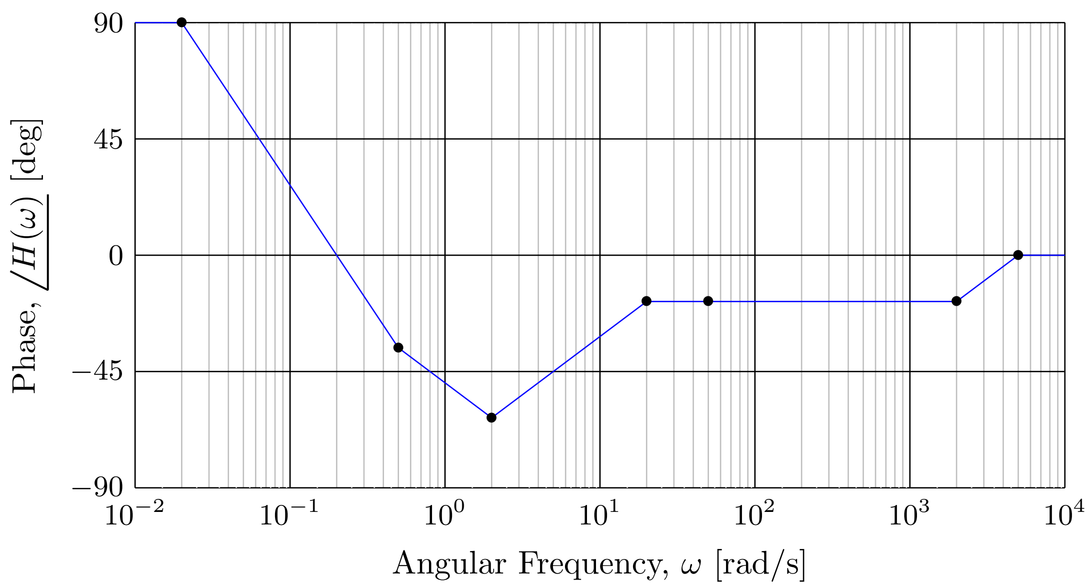
Similarly for the Bode phase plot, we make a note of all slope changes.

$$
\begin{aligned}
150j &\Longrightarrow \underline{/K} = 90^\circ \\
\left(1 + \frac{j\omega}{5}\right) &\Longrightarrow \Delta m = +45 \text{ [deg/dec] at } 0.1\omega_0 = 0.5, \Delta m = -45 \text{ [deg/dec] at } 10\omega_0 = 50 \\
\left(1 + \frac{j\omega}{500}\right) &\Longrightarrow \Delta m = +45 \text{ [deg/dec] at } 0.1\omega_0 = 50, \Delta m = -45 \text{ [deg/dec] at } 10\omega_0 = 5000 \\
\frac{1}{\left(1 + \frac{j\omega}{0.2}\right)^2} &\Longrightarrow \Delta m = -90 \text{ [deg/dec] at } 0.1\omega_0 = 0.02, \Delta m = +90 \text{ [deg/dec] at } 10\omega_0 = 2 \\
\frac{1}{1 + \frac{j\omega}{200}} &\Longrightarrow \Delta m = -45 \text{ [deg/dec] at } 0.1\omega_0 = 20, \Delta m = +45 \text{ [deg/dec] at } 10\omega_0 = 2000
\end{aligned}
$$

Collecting the phase changes with respect to angular frequency:

$$\begin{aligned}
\omega = 0 &\Longrightarrow \text{start at } 90^\circ \\
\omega = 0.02 &\Longrightarrow \Delta m = -90 \\
\omega = 0.5 &\Longrightarrow \Delta m = +45 \\
\omega = 2 &\Longrightarrow \Delta m = +90 \\
\omega = 20 &\Longrightarrow \Delta = -45 \\
\omega = 50 &\Longrightarrow \Delta m = -45 + 45 = 0 \\
\omega = 2000 &\Longrightarrow \Delta m = +45 \\
\omega = 5000 &\Longrightarrow \Delta m = -45
\end{aligned}$$

Using the collection of phase changes, plot the Bode phase plot.



Chapter 7

Continuous-Time Applications

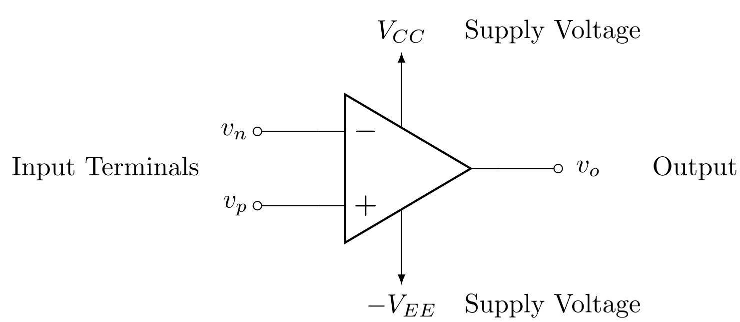
7.1 Introduction to Operational Amplifiers

Electronic components are called *active components* since active components can actively control the behavior of the circuit, whether that is voltage/power gain (such as operational amplifiers and transistors) or current flow (such as diodes and transistors again). Active components introduce nonlinear operating characteristics to electronic components.

On the other hand, electric components such as resistors, inductors, and capacitors are called *passive components* and have no control over voltage or current; the behavior of a passive component is predetermined by its impedance. *Passive circuits* contain strictly passive components, whereas *active circuits* contain at least one active component.

As prefaced, *operational amplifiers* (or *op-amps* for short) are active components with versatile functionality that can be used in circuits for various purposes such as gain, filtering, and hysteresis. As seen in Figure 7.1, an op-amp consists of two input terminals – an *inverting input terminal* (or negative terminal) with voltage v_n and a *noninverting input terminal* (or positive terminal) with voltage v_p – and an output terminal with voltage v_o. However, in order for op-amps to work, two *DC power supply voltages* (also called voltage rails) must be connected, labeled V_{CC} and $-V_{EE}$.

Figure 7.1: Op-amp terminals.

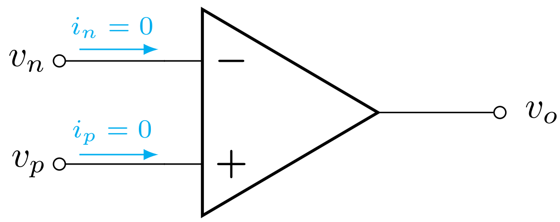


In fact, real op-amps have the constraint that $-V_{EE} \leq v_o \leq V_{CC}$. Note that the power supplies do not necessarily have to be equal in magnitude. While not depicted, all labeled voltages in Figure 7.1 are with respect to an implicit ground node. The lack of a ground node depiction is for simplistic circuit diagram purposes.

7.1.1 Ideal Op-Amps

Ideal op-amps are op-amp models that generate desired responses without component limitations such as voltage saturation. As such, ideal op-amps can be drawn just like in Figure 7.2.

Figure 7.2: Ideal op-amp.



The characteristics of an *ideal op-amp* are introduced:

- Infinite input impedance
- Zero output impedance
- Zero common-mode gain (or equivalently, infinite common-mode rejection)
- Infinite open-loop gain
- Infinite bandwidth

A more detailed explanation of the ideal op-amp characteristics are as follows:

- An ideal op-amp has infinite input impedance such that the currents "drawn" into both input terminals are zero, i.e., $i_p = i_n = 0$.
- An ideal op-amp has zero output impedance such that the output $v_o = A(v_p - v_n)$ without an op-amp output impedance load to drop the output voltage.
- An ideal op-amp has zero common-mode gain such that if $v_p - v_n = 0$ is the difference signal, then $v_o = 0$.
- An ideal op-amp has infinite open loop gain such that $A = \dfrac{v_o}{v_p - v_n}$ can be very large and ideally infinite.
- An ideal op-amp has infinite bandwidth such that it can operate over any angular frequencies $\omega \in [0, \infty)$.

Notice that voltage rails are not drawn nor discussed. Ideal op-amps assume that $|V_{CC}|$ and $|-V_{EE}|$ are large enough such that the output voltage will always be well within range of $-V_{EE} \leq v_o \leq V_{CC}$ without saturation.

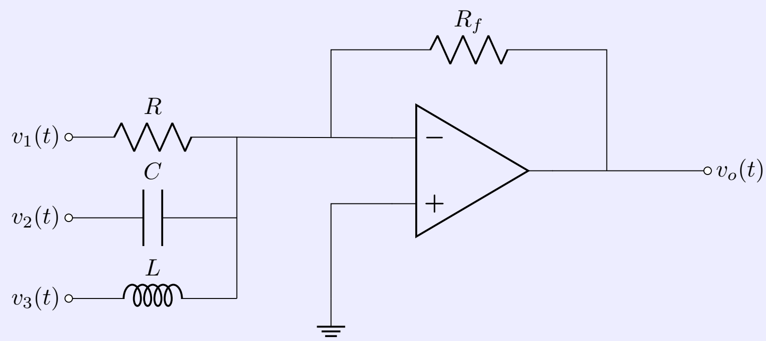
What do ideal op-amps offer then? When ideal op-amps are used in active circuits, the internal device characteristics of the op-amp do not affect the circuit, and only the external loads connected to the ideal op-amp matter.

In other words, the ideal op-amp can be thought of as a black box with some transfer function

$H(s)$ or frequency response $H(\omega)$ with respect to some input voltage(s) connected to the ideal op-amp and the output voltage of the op-amp. If the output terminal is externally connected to the noninverting input terminal, the system is said to have *positive feedback*; similarly, if the output terminal is externally connected to the inverting input terminal, then the circuit has *negative feedback*. As with systems, once feedback is introduced, the circuit has some closed-loop gain.

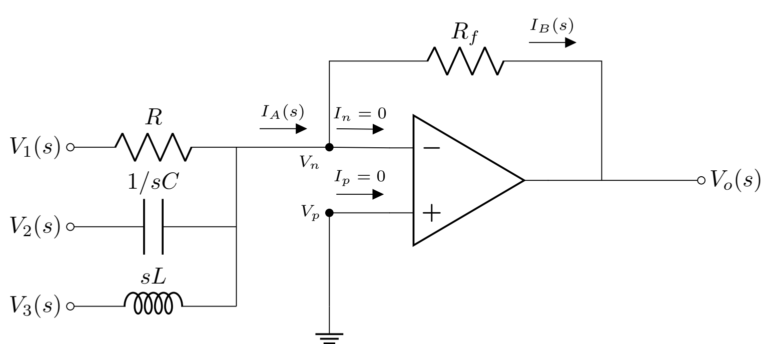
When necessary, op-amp circuits can be transformed into the s-domain or ω-domain for analysis. From then, the transfer function or frequency response can be solved by using the ideal op-amp characteristics.

Example 7.1.1. Find the s-domain output of the following circuit. Assume ideal op-amps.



SOLUTION

The circuit is already depicted as if the node voltage method (NVM) will be used. We will use this fact when calculating currents. First, convert to the s-domain.



Using Kirchhoff's current law (KCL) and NVM, it follows that

$$I_A(s) = \frac{V_1(s) - V_n}{R} + \frac{V_2(s) - V_n}{1/sC} + \frac{V_3(s) - V_n}{sL}$$
$$I_B(s) = \frac{V_n - V_o(s)}{R_f}$$

Using ideal op-amp characteristics, it follows that $I_n = I_p = 0$ and $V_n = V_p$. Since $V_p = 0$, then $V_n = 0$ as well. It also follows that by KCL, $I_A(s) = I_B(s)$. Therefore, the new equation with criteria $V_n = V_p = 0$ and $I_A(s) = I_B(s)$ is:

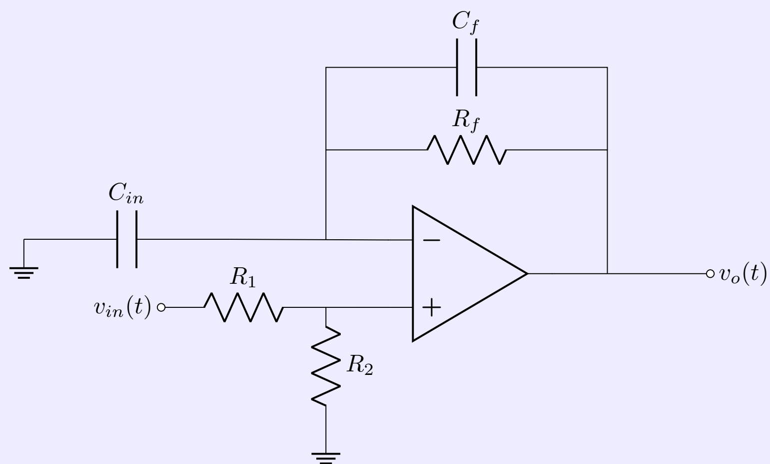
$$\frac{V_1(s)}{R} + \frac{V_2(s)}{1/sC} + \frac{V_3(s)}{sL} = \frac{-V_o(s)}{R_f}$$

Therefore, the s-domain output is given by

$$V_o(s) = -R_f \left[\frac{V_1(s)}{R} + \frac{V_2(s)}{1/sC} + \frac{V_3(s)}{sL} \right].$$

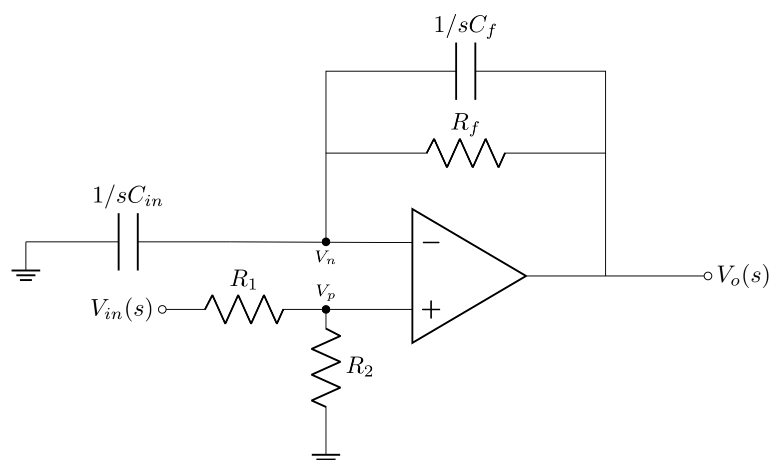
■

Example 7.1.2. Find the transfer function of the following circuit. Then plot the Bode magnitude plot, assuming measurements of unity. Assume ideal op-amps.



SOLUTION

First, convert to the s-domain.



Using the voltage divider rule and ideal op-amp characteristics, it follows that

$$V_n = V_p = V_{in}(s)\left[\frac{R_2}{R_1+R_2}\right]$$

Additionally, the equivalent impedance in the negative feedback loop is given by

$$Z_{eq} = R_f \,||\, (1/sC_f) = \frac{R_f/sC_f}{R_f + 1/sC_f} = \frac{R_f}{1+sR_fC_f}$$

Then from Kirchhoff's current law (KCL), it follows that

$$\frac{0-V_n}{1/sC_{in}} = \frac{V_n - V_o}{Z_{eq}}$$
$$\Longrightarrow V_o = V_n\left[1+\frac{Z_{eq}}{1/sC_{in}}\right] = V_{in}(s)\left[\frac{R_2}{R_1+R_2}\right]\left[1+\frac{sR_fC_{in}}{1+sR_fC_f}\right]$$

The transfer function is then

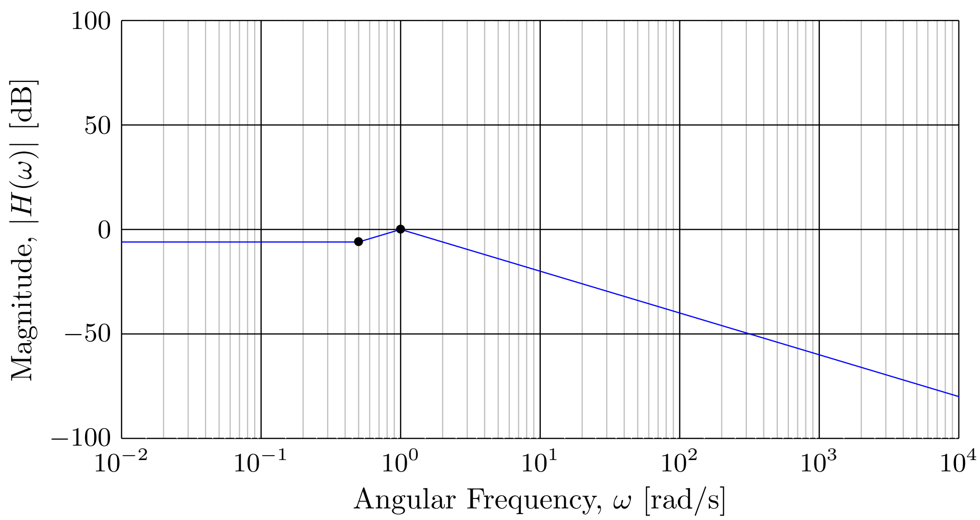
$$H(s) = \frac{V_o(s)}{V_{in}(s)} = \left[\frac{R_2}{R_1+R_2}\right]\left[1+\frac{sR_fC_{in}}{1+sR_fC_f}\right] = \left[\frac{R_2}{R_1+R_2}\right]\left[\frac{1+sR_f(C_{in}+C_f)}{1+sR_fC_f}\right]$$

Finally, letting $s = j\omega$ and assuming measurements of unity, we get the frequency response

$$H(\omega) = \frac{1}{2}\left[\frac{1+2j\omega}{1+j\omega}\right] = \frac{0.5\left(1+\frac{j\omega}{0.5}\right)}{1+j\omega}.$$

Since we only care about the Bode magnitude plot, the changes are as follows:

$$0.5 \Longrightarrow 20\log|K| = -6 \text{ [dB]}$$
$$\left(1+\frac{j\omega}{0.5}\right) \Longrightarrow \Delta m = 20 \text{ [dB/dec] at } \omega_0 = 0.5$$
$$\frac{1}{1+j\omega} \Longrightarrow \Delta m = -20 \text{ [dB/dec] at } \omega_0 = 1$$



7.2 Introduction to Analog Filters

Filters are special types of LTI systems that modify the frequency spectrum of an input signal to produce a desired output signal. Filters are most commonly used to remove unwanted components or features such as noise from a signal. Depending on which set of frequencies to retain or remove, there are four types of filters:

- lowpass filter (LPF)
- highpass filter (HPF)
- bandpass filter (BPF)
- bandreject filter (BRF)

Filters are best characterized by their frequency responses, with the *order* of the filter determined by $\deg[D(\omega)]$ for $H(\omega) = \dfrac{N(\omega)}{D(\omega)}$.

7.2.1 Ideal Filters

Before going in depth, the notion of an *ideal filter* needs introducing. Also called a *brickwall filter* due to the rectangular shape of its magnitude response, the ideal filter perfectly removes a set of unwanted frequencies (called the *stopband*) while retaining the remaining frequency content (called the *passband*) without loss of information. The *cutoff frequency* $\omega = \omega_c$ of an ideal filter is the frequency that sharply separates the passband from the stopband.

The passband and stopband regions of each type of filter are defined in Table 7.1.

Table 7.1: Types of filters for $\omega_c > 0$.

Filter	Passband	Stopband
Lowpass filter (LPF)	$\lvert\omega\rvert < \omega_c$	$\lvert\omega\rvert > \omega_c$
Highpass filter (HPF)	$\lvert\omega\rvert > \omega_c$	$\lvert\omega\rvert < \omega_c$
Bandpass filter (BPF)	$\lvert\omega\rvert \in (\omega_{c_1}, \omega_{c_2})$	$\lvert\omega\rvert \notin (\omega_{c_1}, \omega_{c_2})$
Bandreject filter (BRF)	$\lvert\omega\rvert \notin (\omega_{c_1}, \omega_{c_2})$	$\lvert\omega\rvert \in (\omega_{c_1}, \omega_{c_2})$

Note that the passband and stopband regions are swapped between the lowpass filter and the highpass filter; the same is true between the bandpass filter and the bandreject filter. Assuming the same cutoff frequencies, it follows that

$$H_{LP}(\omega) = \begin{cases} 1, & |\omega| < \omega_c \\ 0, & otherwise \end{cases} \tag{7.1}$$

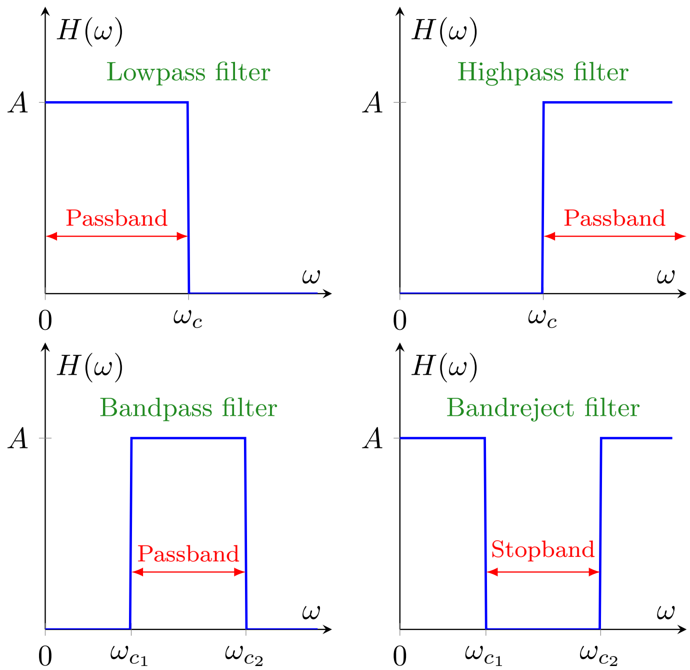
$$H_{HP}(\omega) = 1 - H_{LP}(\omega) \tag{7.2}$$

$$H_{BP}(\omega) = H_{HP,\omega_{c_1}}(s) \times H_{LP,\omega_{c_2}}(s) = H_{LP,\omega_{c_2}}(\omega) - H_{LP,\omega_{c_1}}(\omega) \tag{7.3}$$

$$H_{BR}(\omega) = 1 - H_{BP}(\omega) \tag{7.4}$$

While frequency responses are two-sided, generally only the *one-sided filter response* is of interest, which is essentially just limiting the range of the frequency response to $\omega \geq 0$, without altering the magnitude values. Alternatively, the Bode plots can be used as approximate graphs due to their one-sided nature. The ideal filter response of the four filter types are seen in Figure 7.3.

Figure 7.3: Ideal filter responses for $A = \max[|H(\omega)|]$.



However, by taking the inverse Fourier transforms of the ideal filter responses, the impulse responses of the ideal filters are

$$h_{LP}(t) = \frac{\sin(\omega_c t)}{\pi t} \tag{7.5}$$

$$h_{HP}(t) = \delta(t) - h_{LP}(t) \tag{7.6}$$

$$h_{BP}(t) = \frac{\sin(\omega_{c_2} t)}{\pi t} - \frac{\sin(\omega_{c_1} t)}{\pi t} \tag{7.7}$$

$$h_{BR}(t) = \delta(t) - h_{BP}(t) \tag{7.8}$$

Modified versions of the sinc function in the time domain, these filters are not physically realizable since the impulse responses are noncausal and everlasting.

7.2.2 Passive vs Active Filters

While ideal filters cannot be physically implemented, the frequency response can be approximately emulated such that the errors are minimal. This can be done physically or digitally. As with LTI systems, filters can either be analog (implemented via circuits) or digital (implemented via computers). This section will only cover analog filters.

As with circuits, there are two types of analog filters: passive and active. *Passive filters* contain only passive circuit components such as resistors, capacitors, and inductors, whereas *active filters* contain at least one active component such as the op-amp. Passive filters tend to be designed for high frequencies, whereas active filters tend to be designed for low frequencies. These differences are outlined in Table 7.2.

Table 7.2: Comparison of passive and active filters.

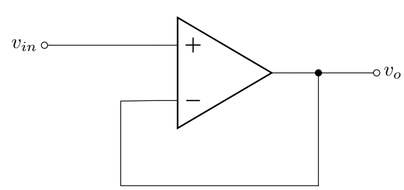
Passive Filters	Active Filters
• Uses only resistors, capacitors, inductors • No external power source needed • Adaptable to large signals (if an inductor is present; good for power systems) • Adept for high-frequency signals (if an inductor is present) • Limited to gain of 1 or less • Loading effect observed in multi-stage filters (each added stage draws more current away from the previous stages, changing the behavior of the previous stages)	• Omits inductors and uses op-amps • Requires DC power supply for op-amps • Response limited by DC power supply for op-amps (good for electronics circuits) • Limited frequency range (favors low-frequency signals) • Capable of signal gain greater than 1 • Performance independence between cascaded stages (as long as there exists voltage buffers between stages)

Note that inductors are generally avoided in low-frequency circuits as they are bulky in size/weight, costly, and possess greater nonideal behavior than other components; as a result, inductors are usually reserved for high-frequency passive circuits such as electric power transmission systems.

Multi-stage filters describe N^{th} order filters (for $N \geq 2$) in which individual filters are cascaded (i.e., the output of a filter is fed into the input node of the next filter).

- While multi-stage passive filters do have a loading effect, this can be minimized by making sure the impedance of each successive stage is 10 times greater than the impedance of its previous stage, such that the loading effect becomes less than 10%, and the overall frequency response function could be approximated as if there were performance independence between the cascaded stages.
- Of course, with active filters, there is no loading effect, as long as voltage buffers with gain 1 (like the ones shown in Figure 7.4) are placed between stages.

Figure 7.4: Voltage buffer with unity gain.



Lastly, recall that the power equation for a component or system is given by

$$P = \frac{V^2}{Z_{eq}}. \tag{7.9}$$

Since the frequency response assumes that input and output signals are voltage signals, when designing real filters, for

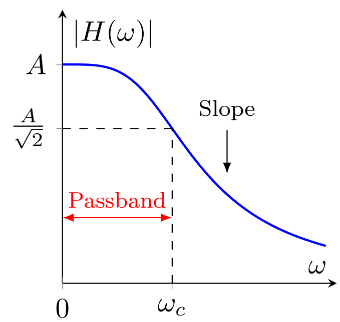
$$A = \max[|H(\omega)|], \tag{7.10}$$

the objective is to find an appropriate cutoff frequency ω_c, which is when the power dissipated is half. This occurs when the voltage is $(1/\sqrt{2}) \times V$. Therefore, for analog filter design, the cutoff frequency ω_c is the frequency at which

$$|H(\omega_c)| = \frac{A}{\sqrt{2}} = 0.707A. \tag{7.11}$$

Because of this, the cutoff frequency ω_c is also called the *half-power frequency* in the context of filter design. It is also called the *3-dB frequency* since it is the point where the magnitude response drops 3 [dB] below the ideal passband level A [V/V]. Note that there exists some slope (also called *roll-off rate*) after the cutoff frequency ($\omega > \omega_c$) in the filter magnitude response. The steeper the slope, the closer the response is to ideal.

Figure 7.5: Example of a real lowpass filter magnitude response.



It is convenient then that the Bode approximations already account for cutoff frequencies when used on filters.

Example 7.2.1. Find the cutoff frequency of the transfer function $H(\omega) = \dfrac{1}{1 + j\omega RC}$.

SOLUTION

Note that the frequency response is passive. The cutoff frequency happens when the magnitude response is attenuated by a factor of $1/\sqrt{2}$.

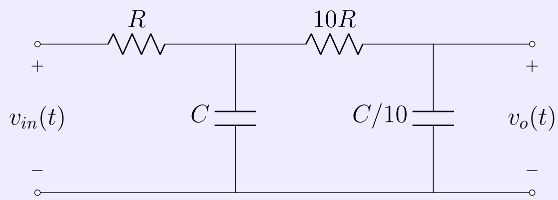
$$|H(\omega)| = \left|\frac{1}{1 + j\omega RC}\right| = \frac{1}{\sqrt{1 + (\omega RC)^2}} = \frac{1}{\sqrt{2}}$$
$$\implies 1 + (\omega RC)^2 = 2$$
$$\implies \omega = \frac{1}{RC} \triangleq \omega_c$$

Interestingly, when rewriting $H(\omega)$ in Bode standard form, it follows that

$$H(\omega) = \frac{1}{1 + j\omega RC} = \frac{1}{1 + \frac{j\omega}{1/RC}} = \frac{1}{1 + \frac{j\omega}{\omega_c}}$$

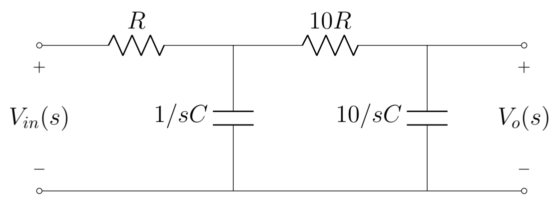
■

Example 7.2.2. Describe the following filter to the fullest extent.



SOLUTION

From observation, the circuit is a two-stage passive filter. Converting to the s-domain, we get



Since the second stage has an impedance of ten times that of the first stage, we can approximate the overall transfer function by multiplying the transfer functions of the two individual stages.

$$
\begin{aligned}
H_1(s) &= \frac{1/sC}{R + 1/sC} = \frac{1}{1 + sRC} \\
H_2(s) &= \frac{10/sC}{10R + 10/sC} = \frac{1}{1 + sRC} \\
\Longrightarrow H(s) &\approx H_1(s)H_2(s) = \frac{1}{(1 + sRC)^2}
\end{aligned}
$$

The frequency response is given by

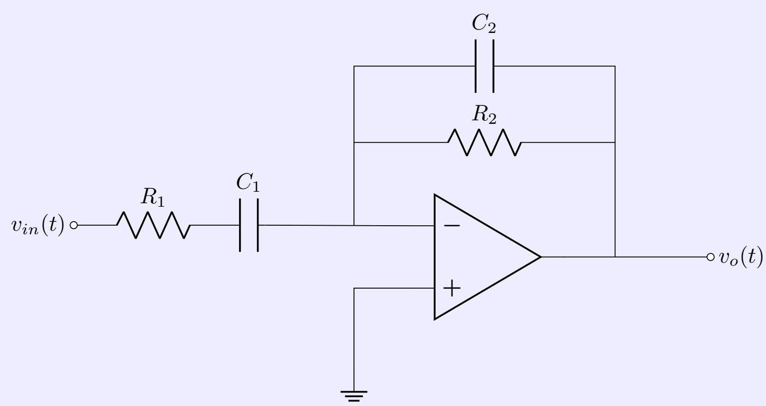
$$
H(\omega) = \frac{1}{(1 + j\omega RC)^2} = \frac{1}{(1 + \frac{j\omega}{1/RC})^2} = \frac{1}{(1 + \frac{j\omega}{\omega_c})^2},
$$

for cutoff frequency $\omega_c = 1/RC$. Since there is a real pole at ω_c with multiplicity 2, the roll-off rate is then -40 [dB/dec].

Overall the filter:

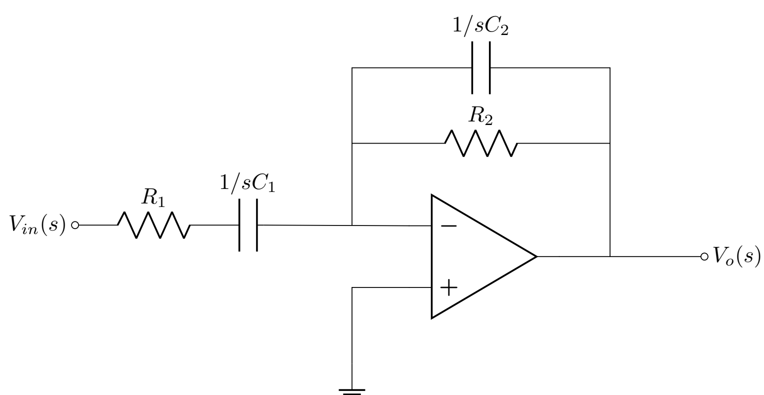
- is a two-stage, second-order passive lowpass filter,
- has cutoff frequency $\omega_c = 1/RC$,
- and has roll-off rate -40 [dB/dec].

Example 7.2.3. Describe the following filter to the fullest extent. Assume ideal op-amps.



SOLUTION

From observation, the circuit is a single-stage active filter. Converting to the s-domain, we get



The equivalent impedance in the negative feedback loop is given by

$$Z_{eq} = R_2 \,||\, (1/sC_2) = \frac{R_2}{1 + sR_2C_2}.$$

Then it follows that

$$\frac{V_{in}(s)}{R_1 + 1/sC_1} = \frac{-V_o(s)}{Z_{eq}}$$

such that the transfer function is

$$H(s) = \frac{V_o(s)}{V_{in}(s)} = -\frac{Z_{eq}}{R_1 + 1/sC_1} = -\frac{sR_2C_1}{(1 + sR_1C_1)(1 + sR_2C_2)}.$$

The frequency response is then

$$H(\omega) = -\frac{j\omega R_2 C_1}{\left(1 + \frac{j\omega}{1/R_1C_1}\right)\left(1 + \frac{j\omega}{1/R_2C_2}\right)} = -\frac{j\omega R_2 C_1}{\left(1 + \frac{j\omega}{\omega_{c_1}}\right)\left(1 + \frac{j\omega}{\omega_{c_2}}\right)},$$

for $\omega_{c_1} = 1/R_1C_1$ and $\omega_{c_2} = 1/R_2C_2$. Assume $\omega_{c_1} < \omega_{c_2}$. Then the asymptotic magnitude changes are as follows:

$$\begin{aligned} j\omega &\Longrightarrow \Delta m = +20 \text{ [dB/dec] at } \omega = 0 \\ \frac{1}{1 + \frac{j\omega}{\omega_{c_1}}} &\Longrightarrow \Delta m = -20 \text{ [dB/dec] at } \omega = \omega_{c_1} \\ \frac{1}{1 + \frac{j\omega}{\omega_{c_2}}} &\Longrightarrow \Delta m = -20 \text{ [dB/dec] at } \omega = \omega_{c_2} \end{aligned}$$

Even without sketching the Bode magnitude plot, we can see that the filter is a bandpass filter.

Overall the filter:

- is a single-stage, second-order active bandpass filter,
- has a low cutoff frequency $\omega_{c_1} = 1/R_1C_1$,
- has a high cutoff frequency $\omega_{c_2} = 1/R_2C_2$,
- and has respective roll-off rates of ±20 [dB/dec].

■

7.3 Butterworth Filters

Butterworth filters are LTI systems designed to have maximally flat magnitude responses that closely approximates the ideal filter responses. Initially designed to approximate the ideal lowpass filter, the Butterworth filter has since been extended to other filter types by modifying the Butterworth lowpass filter to achieve such means.

As with any filter, Butterworth filters can be designed physically or digitally. Real analog Butterworth filters can be designed using either the order-N passive Cauer LC filter or the order-$2N$ active Sallen–Key filter; neither methods will be discussed here. Instead, a theoretical design approach of the Butterworth filters will be explored.

Figure 7.6: Second-order passive Cauer lowpass filter.

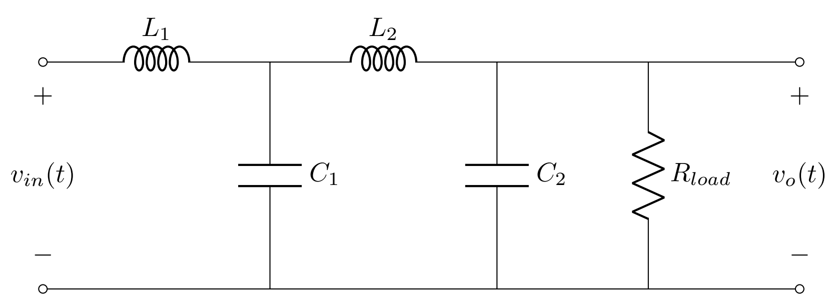
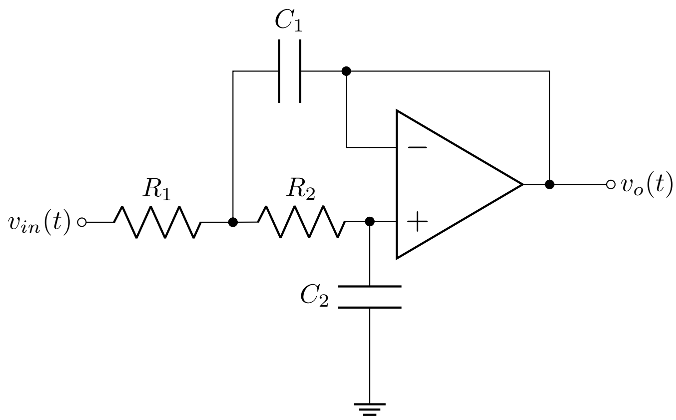


Figure 7.7: Second-order active Sallen–Key lowpass filter

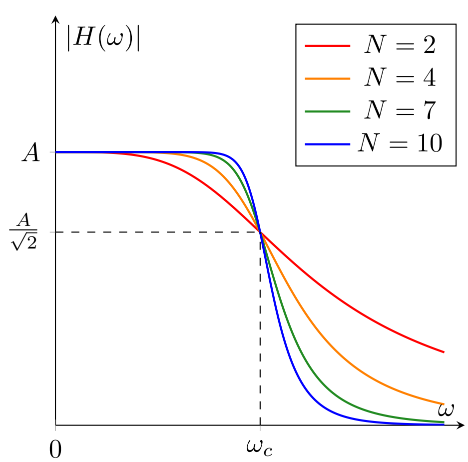


For a desired cutoff frequency ω_c and filter order N, the *magnitude-squared response* of the Butterworth lowpass filter is given by

$$|H_{LP}(\omega)|^2 = \frac{A^2}{1 + \left(\frac{j\omega}{j\omega_c}\right)^{2N}} \tag{7.12}$$

where $A = \max[|H_{LP}(\omega)|]$. Interestingly, as the filter order N increases, the roll-off (slope) in the magnitude response becomes steeper and closer to the ideal sharp transition, as seen in Figure 7.8.

Figure 7.8: N^{th} order Butterworth lowpass filter magnitude response.



Assume $A = 1$ such that the *passive magnitude-squared response* of the Butterworth lowpass filter is

$$|H_{LP}(\omega)|^2 = \frac{1}{1 + \left(\frac{j\omega}{j\omega_c}\right)^{2N}} \tag{7.13}$$

The passive magnitude-squared response can be converted to the s-domain such that the s-domain passive magnitude-squared response is

$$|H_{LP}(s)|^2 = H_{LP}(s)H^*_{LP}(s) = \frac{1}{1 + \left(\frac{s}{j\omega_c}\right)^{2N}} \tag{7.14}$$

and the passive lowpass transfer function can be more succinctly defined as

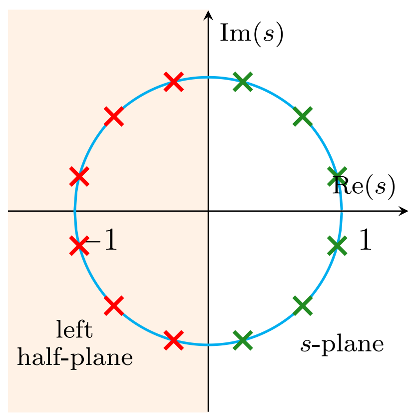
$$H_{LP}(s) = \prod_{k=1}^{N} \frac{\omega_c}{s - \omega_c \exp(j\theta_k)} = \prod_{k=1}^{N} \frac{1}{s_c - \exp(j\theta_k)}, \quad \text{for } s_c = \frac{s}{\omega_c} \tag{7.15}$$

$$\theta_k = \left[\frac{2k + N - 1}{2N}\right]\pi, \quad \text{for } k = 1, ..., N. \tag{7.16}$$

For filter order N, the squared transfer function $|H_{LP}(s)|^2$ produces $2N$ poles that can be equidistantly plotted along the unit circle on the s-plane; from there, only the N poles from the open left half-plane (OLHP) are selected as poles of $H_{LP}(s)$, which are given by $\exp(j\theta_k)$ for $k = 1, ..., N$. This is to ensure that the LTI system characterized by the filter is BIBO stable.

An example of the pole placement for filter order $N = 6$ can be found in Figure 7.9, where all crosses represent the poles of $|H_{LP}(s)|^2$, and the red crosses represent the selected poles for $H_{LP}(s)$.

Figure 7.9: Pole placement of a 6th order Butterworth lowpass filter on the s-plane.



As every Butterworth filter relates back to the Butterworth lowpass filter, the passive lowpass transfer function can be written as

$$H_{LP}(s) = \frac{1}{B_N(s_c)}, \tag{7.17}$$

where

$$B_N(s_c) = 1 + \sum_{k=1}^{N} a_k [s_c]^k \tag{7.18}$$

is the N^{th} order *(normalized) Butterworth polynomial* with respect to s_c. The *Butterworth coefficients* a_k for the first few values of N are outlined in Table 7.3.

Table 7.3: Rounded Butterworth coefficients for $B_N(s_c)$ with $s_c = s/\omega_c$ (lowpass only).

N	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8	a_9	a_{10}
1	1									
2	1.41	1								
3	2	2	1							
4	2.61	3.41	2.61	1						
5	3.24	5.24	5.24	3.24	1					
6	3.87	7.46	9.14	7.46	3.87	1				
7	4.49	10.10	14.59	14.59	10.10	4.49	1			
8	5.13	13.14	21.85	25.69	21.85	13.14	5.13	1		
9	5.76	16.58	31.16	41.99	41.99	31.16	16.58	5.76	1	
10	6.39	20.43	42.80	64.88	74.23	64.88	42.80	20.43	6.39	1

While only the passive lowpass transfer function was considered, the active transfer function $H_a(s)$ of some filter type can be derived from the passive transfer function $H_p(s)$ of the same filter type:

$$H_a(s) = A \cdot H_p(s) \tag{7.19}$$

Example 7.3.1. Find the transfer function of a third-order Butterworth lowpass filter with cutoff frequency $\omega_c = 10^3$ [rad/s] and a DC gain of 10.

SOLUTION

First, find the passive transfer function. Using the Butterworth polynomial in $s_c = s/\omega_c$, it follows that

$$H_{LP}(s) = \frac{1}{B_3(s_c)} = \frac{1}{s_c^3 + 2s_c^2 + 2s_c + 1}.$$

The passive transfer function simplifies to

$$\begin{aligned} H_{LP}(s) &= \frac{1}{\left(\frac{s}{\omega_c}\right)^3 + 2\left(\frac{s}{\omega_c}\right)^2 + 2\left(\frac{s}{\omega_c}\right) + 1} = \frac{\omega_c^3}{s^3 + 2\omega_c s^2 + 2\omega_c^2 s + \omega_c^3} \\ &= \frac{10^9}{s^3 + (2 \times 10^3)s^2 + (2 \times 10^6)s + 10^9} \end{aligned}$$

Lastly, we find the active transfer function $H(s)$. Since the DC gain is 10 and the filter is lowpass, it follows that $A = 10$ and

$$\begin{aligned} H(s) = A \cdot H_{LP}(s) &= 10 \times \frac{10^9}{s^3 + (2 \times 10^3)s^2 + (2 \times 10^6)s + 10^9} \\ &= \frac{10^{10}}{s^3 + (2 \times 10^3)s^2 + (2 \times 10^6)s + 10^9} \end{aligned}$$

■

The passive transfer function of the Butterworth highpass filter $H_{HP}(s)$ can be derived from the passive transfer function of the Butterworth lowpass filter $H_{LP}(s)$ and is given by

$$H_{HP}(s) = H_{LP}(s)\Big|_{s_c \leftarrow (1/s_c)} \quad \text{for } s_c = \frac{s}{\omega_c} \tag{7.20}$$

Example 7.3.2. Find the transfer function of a third-order Butterworth highpass filter with cutoff frequency $\omega_c = 10^3$ [rad/s] and a high frequency gain of 10.

SOLUTION

From an earlier example, the passive lowpass transfer function is

$$H_{LP}(s) = \frac{1}{s_c^3 + 2s_c^2 + 2s_c + 1}$$

Then the passive highpass transfer function is

$$\begin{aligned}
H_{HP}(s) = H_{LP}(s)\Big|_{s_c \leftarrow (1/s_c)} &= \frac{1}{\left(\frac{1}{s_c}\right)^3 + 2\left(\frac{1}{s_c}\right)^2 + 2\left(\frac{1}{s_c}\right) + 1} \\
&= \frac{s_c^3}{1 + 2s_c + 2s_c^2 + s_c^3} \\
&= \frac{\left(\frac{s}{\omega_c}\right)^3}{1 + 2\left(\frac{s}{\omega_c}\right) + 2\left(\frac{s}{\omega_c}\right)^2 + \left(\frac{s}{\omega_c}\right)^3} \\
&= \frac{s^3}{s^3 + 2\omega_c s^2 + 2\omega_c^2 s + \omega_c^3} \\
&= \frac{s^3}{s^3 + (2 \times 10^3)s^2 + (2 \times 10^6)s + 10^9}
\end{aligned}$$

Lastly, we find the active transfer function $H(s)$. Since the high frequency gain is 10 and the filter is highpass, it follows that $A = 10$ and

$$\begin{aligned}
H(s) = A \cdot H_{HP}(s) &= 10 \times \frac{s^3}{s^3 + (2 \times 10^3)s^2 + (2 \times 10^6)s + 10^9} \\
&= \frac{10s^3}{s^3 + (2 \times 10^3)s^2 + (2 \times 10^6)s + 10^9}
\end{aligned}$$

■

The Butterworth bandpass filter can also be derived from the lowpass and highpass filters such that for $\omega_{c_1} < \omega_{c_2}$,

$$H_{BP}(s) = H_{HP,\omega_{c_1}}(s) \times H_{LP,\omega_{c_2}}(s) \tag{7.21}$$

Alternatively, the passive transfer function of the Butterworth bandpass filter can be derived solely from the passive lowpass filter such that

$$H_{BP}(s) = H_{LP}(s)\Big|_{s_c \leftarrow s'_c} \tag{7.22}$$

$$\text{for } s'_c = \frac{1}{\omega_{c_2} - \omega_{c_1}} \left[\frac{s^2 + \omega_{c_1}\omega_{c_2}}{s}\right] \tag{7.23}$$

Similarly, the passive transfer function of the Butterworth bandreject filter can be derived from the passive lowpass filter such that

$$H_{BR}(s) = H_{LP}(s)\Big|_{s_c \leftarrow (1/s'_c)} \tag{7.24}$$

$$\text{for } \frac{1}{s'_c} = (\omega_{c_2} - \omega_{c_1}) \left[\frac{s}{s^2 + \omega_{c_1}\omega_{c_2}}\right] \tag{7.25}$$

Note that both the bandpass and bandreject filters can only have even filter orders.

Example 7.3.3. Find the transfer function of a second-order Butterworth bandpass filter with passband $\omega \in (10, 40)$ [rad/s], assuming a maximum gain of unity.

SOLUTION

Since it is a bandpass filter with order 2, let $N = 1$ for the passive lowpass transfer function:

$$H_{LP}(s) = \frac{1}{s_c + 1}$$

Then it follows that

$$s'_c = \frac{1}{\omega_{c_2} - \omega_{c_1}} \left[\frac{s^2 + \omega_{c_1}\omega_{c_2}}{s}\right] = \frac{1}{40 - 10} \left[\frac{s^2 + 10 \times 40}{s}\right] = \frac{s^2 + 400}{30s},$$

and the passive bandpass transfer function is

$$H_{BP}(s) = H_{LP}(s)\Big|_{s_c \leftarrow s'_c} = \frac{1}{\left(\frac{s^2+400}{30s}\right) + 1} = \frac{30s}{s^2 + 30s + 400}.$$

■

Example 7.3.4. Find the transfer function of a second-order Butterworth bandreject filter with stopband $\omega \in (10, 40)$ [rad/s], assuming a maximum gain of unity.

SOLUTION

Since it is a bandreject filter with order 2, let $N = 1$ for the passive lowpass transfer function:

$$H_{LP}(s) = \frac{1}{s_c + 1}$$

Then it follows that

$$\frac{1}{s_c'} = (\omega_{c_2} - \omega_{c_1}) \left[\frac{s}{s^2 + \omega_{c_1}\omega_{c_2}} \right] = \frac{30s}{s^2+400},$$

and the passive bandreject transfer function is

$$H_{BR}(s) = H_{LP}(s)\Big|_{s_c \leftarrow 1/s_c'} = \frac{1}{\left(\frac{30s}{s^2+400}\right)+1} = \frac{s^2+400}{s^2+30s+400}.$$

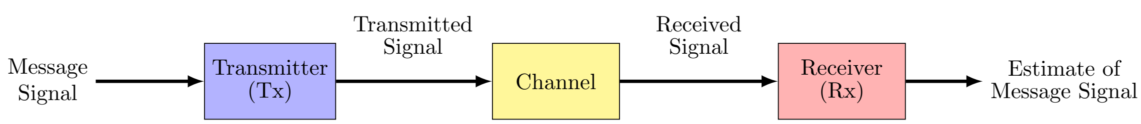
■

7.4 Introduction to Communications Systems

In the broadest sense, *communication* is the exchange of information. Before modern times, messages were carried by foot (postmen, runners), carrier pigeons, light (Morse code), and fires (smoke signals). In the present day, most parts of the world now use *(electrical) communications systems* (also called *telecommunications*) instead, which can transmit signals over longer distances at the speed of light.

The block diagram of a typical communications system is shown in Figure 7.10. Here, both the message signal and the output signal are electrical signals. The *source* originates a message (i.e., a

Figure 7.10: Communications system.



human voice, TV picture, e-mail message, etc.). If the message is not an electrical signal, it must first be converted into one via an *input inducer* (i.e., a microphone, camera, computer keyboard, etc.); once converted, this electrical signal is called a *message signal*.

The message signal is inputted into a *transmitter*, which modifies the message signal for efficient transmission. It may be comprised of a combination of systems such as an A/D converter, an encoder, and a modulator. The modified signal then gets transmitted through a *(communications) channel* (also called *transmission medium*); examples include copper wires, coaxial cables, optical fiber, and a radio link. In the channel, unwanted distortion and noise is introduced to the transmitted signal.

The received signal with distortion and noise is then inputted into a *receiver*, which reprocesses the received signal by reversing the transmitter operations (i.e., D/A converter, decoder, demodulator) and removes the unwanted distortion and noise (as best as possible) from the received signal.

The receiver output signal is then fed into an *output transducer*, which converts the electrical signal back to a message (i.e., the same form as the message at the source). The message is then finally delivered to the *destination*.

7.4.1 Classification of Communications Systems

7.4.1.1 Analog vs Digital Communications

Communications systems can be classified by the type of signals that are being transmitted and received. In *analog communications systems*, analog signals are used for information transmission. In *digital communications systems* (also called *data communications systems*), digital signals are used for information transmission.

7.4.1.2 Baseband vs Passband Communications

Along the same line of signal-based classification, communications systems can also be classified by how the message signals are modified before transmission. If the signal is transmitted with *modulation*, then the signal is called a *passband signal* (or simply a *modulated signal*), and the communications system is classified as a *passband communications system* (also called a *carrier communications system*). Otherwise, if the message signal is not modulated, then the signal is called a *baseband signal*, and the communications system is classified as a *baseband communications system*. Typically, baseband transmission is more suitable for short distance communication, whereas passband transmission is required for long distance communication.

Additionally, passband signals can be further classified as either *narrowband*, *broadband* (also called *wideband*), or *ultra-wideband* (UWB), depending on the limited bandwidth allowed that a passband signal can take up in the frequency domain. The passband communications system that transmits a particular type of passband signal can also then take on the very same descriptor (i.e., broadband communications systems transmit broadband signals). However, in a special type of communications systems called *spread spectrum communications systems*, a narrowband signal can be spread over a broadband frequency band.

7.4.1.3 Wired vs Wireless Communications

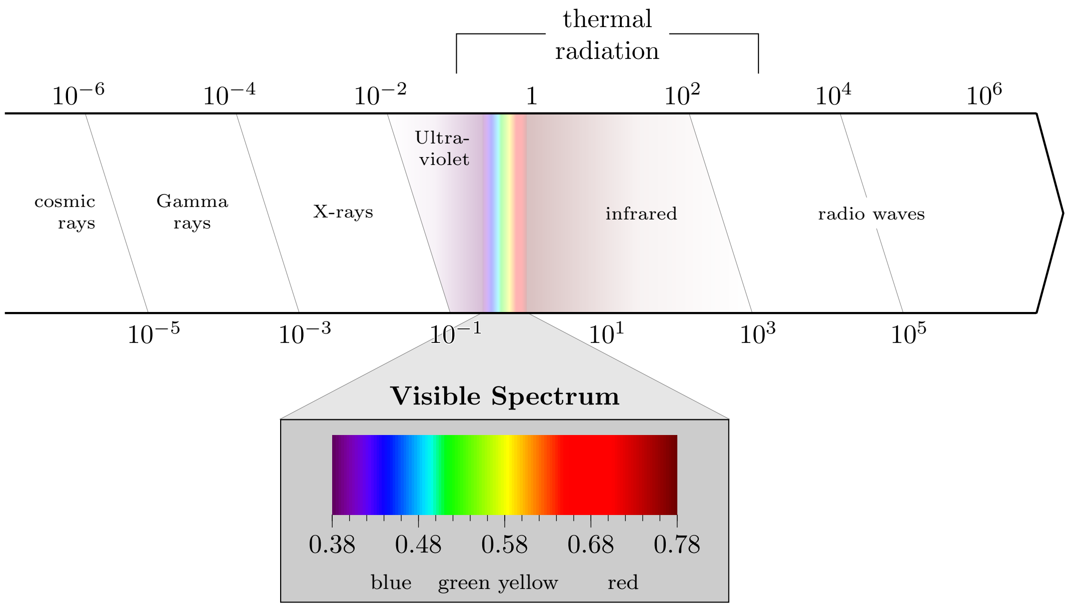
Communications systems can be classified by the type of transmission medium used for the communications channel. *Guided transmission media* (also called *wired* or *bounded*) are physical (and tangible) links over which signals are sent from transmitter to receiver; communications systems with guided transmission media are called *wired communications systems* (also called *wireline communications systems*). Wired communications systems can further be classified by the type of guided media used, such as fiber-optic communications systems or cable networks.

In contrast, *unguided transmission media* (also called *wireless* or *unbounded*) describes free space (i.e., over-the-air) over which signals are being transmitted and received; communications systems with unguided transmission media are called *wireless communications systems*. Wireless communications systems can further be classified by the type of electromagnetic wave that is carrying the message signal, such as RF communications systems (radio waves), microwave links, and optical wireless communications (visible, infrared, or ultraviolet waves). In addition, wireless communications can also be described by the number of antennas on the transmitter and receiver ends (SISO, SIMO, MISO, MIMO).

7.4.1.4 Short-Range vs Long-Range Communications

As implied in the name, *short-range communications systems* can only carry a message signal over short distances, whereas *long-range communications systems* can carry a message signal over long distances. Examples of short-range communications systems include Wi-Fi, Bluetooth, and RFID. Examples of long-range communications systems include 4G/5G cellular networks and low-power wide area networks.

Figure 7.11: Electromagnetic spectrum with linear frequencies [Hz].



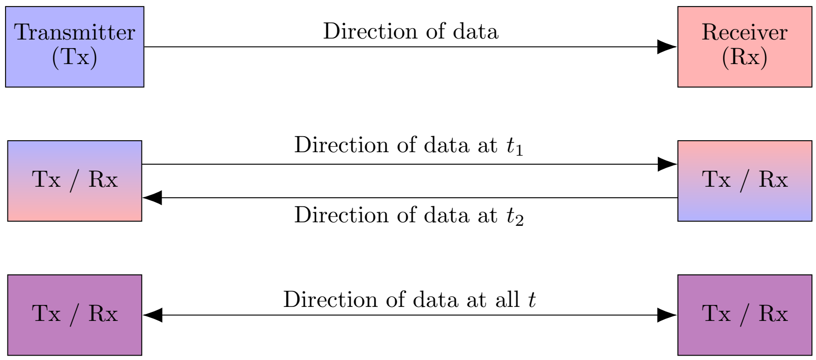
7.4.1.5 Simplex vs Duplex Communications

Communications systems can be classified by the directionality of the exchange of information over a channel, whether that is unidirectional or bidirectional. In a *simplex communications system* with two parties, only one party can be a transmitter while the other is strictly a receiver such that the flow of information is unidirectional from transmitter to receiver. Examples include TV broadcasting systems or PC-to-keyboard connections.

In a *half-duplex communications system* with two parties, either party can take on the role of transmitter or receiver at any time, as long as signal transmission happens one at a time. That is, the communication is technically bidirectional but only one direction can be provided at a time (not simultaneously). This is usually set up to conserve bandwidth, though collisions can occur if both parties attempt to transmit at the same time. Examples include walkie-talkies, USB data transfers, and Internet access.

In a *full-duplex communications system* with two parties, both parties can take on the role of either transmitter or receiver at any time and can transmit and receive at any time; that is, communication is bidirectional and can be simultaneous. Examples include phone networks, video conferencing, and live chats.

Figure 7.12: Simplex vs half-duplex vs full-duplex communications systems.



7.4.2 Signal Transmission Through LTI Systems

For *distortionless transmission* through a system, the exact input signal shape must be reproduced at the output. That is,

$$y(t) = K \cdot x(t - t_0), \tag{7.26}$$

for K is the *gain constant* and t_0 is the *time delay*. The frequency response function is given by

$$H(\omega) = \frac{Y(\omega)}{X(\omega)} = Ke^{-j\omega t_0}. \tag{7.27}$$

Distortionless transmission can only happen if the amplitude $|H(\omega)| = K$ is constant over the entire frequency range and if the phase $\angle H(\omega) = -t_0\omega$ is linear with frequency.

When $|H(\omega)|$ is not constant, the frequency components of the input signal are transmitted with different amounts of gain or attenuation, resulting in a different waveform due to *amplitude distortion*. When $\angle H(\omega)$ is not linear, the frequency components of the input signal pass through the system at different delays, resulting in a different waveform due to *phase distortion*.

7.4.3 Fourier Transform Using Linear Frequency

Since the field of telecommunications frequently uses linear frequency f [Hz] and since it follows that $\omega = 2\pi f$, it is only appropriate to rewrite the Fourier transform using linear frequency.

$$X(f) = \mathcal{F}[x(t)] = \int_{-\infty}^{+\infty} x(t)e^{-j2\pi ft}\,dt. \tag{7.28}$$

The inverse Fourier transform is given by

$$x(t) = \mathcal{F}^{-1}[X(f)] = \int_{-\infty}^{+\infty} X(f)e^{+j2\pi ft}\,df. \tag{7.29}$$

Using the scaling property of delta functions

$$\delta(\omega) = \delta(2\pi f) = \frac{1}{2\pi}\delta(f) \tag{7.30}$$

and the convolution with delta function

$$X(\omega) * \delta(\omega) = X(\omega) \tag{7.31}$$

$$\Longrightarrow X(2\pi f) * \delta(2\pi f) = X(f) * \frac{1}{2\pi}\delta(f) = \frac{1}{2\pi}X(f), \tag{7.32}$$

the revised Fourier transform tables are provided in Tables 7.4 and 7.5.

Table 7.4: Properties of the Fourier transform using linear frequency.

Property	$x(t)$	$X(f) = \mathcal{F}[x(t)]$
Superposition	$K_1 x_1(t) + K_2 x_2(t)$	$K_1 X_1(f) + K_2 X_2(f)$
Time scaling	$x(at)$	$\frac{1}{\lvert a \rvert} X\left(\frac{f}{a}\right)$
Time shift	$x(t - t_0)$	$e^{-j2\pi f t_0} X(f)$
Frequency shift	$e^{+j2\pi f_0 t} x(t)$	$X(f - f_0)$
Time nth derivative	$x^{(n)}(t) = \frac{d^n x(t)}{dt^n}$	$(j2\pi f)^n X(f)$
Time integral	$\int_{-\infty}^{t} x(\tau)\, d\tau$	$\frac{X(f)}{j2\pi f} + \frac{1}{2}\delta(f) \int_{-\infty}^{+\infty} x(t)\, dt$
Frequency derivative	$t^n x(t)$	$\left(\frac{j}{2\pi}\right)^n \frac{d^n X(f)}{df^n}$
Convolution	$x_1(t) * x_2(t)$	$X_1(f) X_2(f)$
Multiplication	$x_1(t) x_2(t)$	$X_1(f) * X_2(f)$
Modulation	$x(t)\cos(2\pi f_0 t)$	$\frac{1}{2}[X(f - f_0) + X(f + f_0)]$
Duality	$X(t)$	$x(-f)$

Table 7.5: Fourier transform pairs using linear frequency.

$x(t)$	$X(f) = \mathcal{F}[x(t)]$
$\delta(t)$	1
$\delta(t - t_0)$	$e^{-j2\pi f t_0}$
1	$\delta(f)$
$u(t)$	$\frac{1}{2}\delta(f) + \frac{1}{j2\pi f}$
$\text{sgn}(t)$	$\frac{1}{j\pi f}$
$\text{rect}\left(\frac{t}{\tau}\right)$	$\tau\,\text{Sa}(\pi\tau f) = \tau\,\text{sinc}(\tau f)$
$\text{tri}\left(\frac{t}{\tau}\right)$	$\tau\,\text{Sa}^2(\pi\tau f) = \tau\,\text{sinc}^2(\tau f)$
$e^{j2\pi f_0 t}$	$\delta(f - f_0)$
$\cos(2\pi f_0 t)$	$\frac{1}{2}[\delta(f - f_0) + \delta(f + f_0)]$
$\sin(2\pi f_0 t)$	$\frac{1}{2j}[\delta(f - f_0) - \delta(f + f_0)]$
$e^{-at}u(t)$	$\frac{1}{a + j2\pi f}$
$te^{-at}u(t)$	$\frac{1}{(a + j2\pi f)^2}$
$\frac{1}{\pi t}$	$-j\,\text{sgn}(f)$
$\text{Sa}(\pi\tau f) = \text{sinc}(\tau f)$	$\frac{1}{\tau}\,\text{rect}\left(\frac{f}{\tau}\right)$

Additionally, Parseval's theorem can be rewritten with linear frequency such that the total energy of a physically realizable signal $x(t)$ is

$$E_x = \int_{-\infty}^{+\infty} |x(t)|^2\,dt = \int_{-\infty}^{+\infty} |X(f)|^2\,df, \tag{7.33}$$

with the 1-sided and 2-sided energy spectral densities defined as

$$ESD_1 = 2\,|X(f)|^2$$
$$ESD_2 = |X(f)|^2$$

7.4.4 Hilbert Transform

The *Hilbert transform* of a signal $x(t)$ is defined as

$$\hat{x}(t) = \mathcal{H}[x(t)] = x(t) * \frac{1}{\pi t}. \tag{7.34}$$

The *inverse Hilbert transform* is defined as

$$x(t) = \mathcal{H}^{-1}[\hat{x}(t)] = -\mathcal{H}[\hat{x}(t)] = -\hat{x}(t) * \frac{1}{\pi t}. \tag{7.35}$$

It follows that if the Fourier transform exists for some signal, then the signal must also have an existing Hilbert transform. Essentially, the Hilbert transform is the system response to an LTI system with impulse response

$$h_Q(t) = \frac{1}{\pi t} \tag{7.36}$$

and frequency response

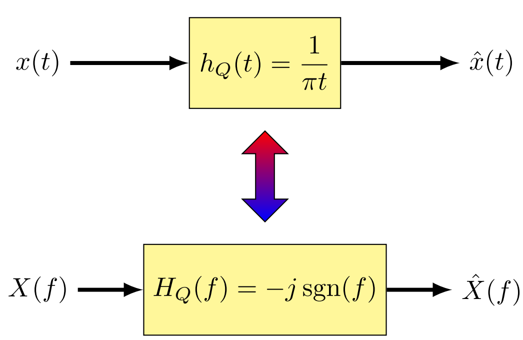
$$H_Q(f) = -j\,\mathrm{sgn}(f). \tag{7.37}$$

From the frequency response, an *allpass filter* response is observed:

$$|H_Q(f)| = 1 \tag{7.38}$$

$$\angle H_Q(f) = \begin{cases} -\pi/2, & f > 0 \\ +\pi/2, & f < 0 \end{cases} \tag{7.39}$$

The Hilbert transform can then be seen as the system response to an ideal phase shifter which shifts positive-frequency components by $-\pi/2$ and negative-frequency components by $+\pi/2$, called the *quadrature filter*. The ideal quadrature filter is characterized by impulse response $h_Q(t) = 1/\pi t$.



Since the impulse response of the quadrature filter is noncausal, the filter is not physically realizable.

However, the Hilbert transform can still be used to create *analytic signals* from real signals. An analytic signal is a complex-valued signal with no negative frequency components and is often used in SSB modulation which will be explored later. The *analytic representation* of a real signal $x(t)$ is

$$z(t) = x(t) + j\hat{x}(t) = x(t) + j\mathcal{H}[x(t)], \tag{7.40}$$

where $x(t) = \mathrm{Re}(z(t))$ is the *in-phase component* and $\hat{x}(t) = \mathrm{Im}(z(t))$ is the *quadrature component*; from the nomenclature, analytic signals are also called *I/Q signals*.

Table 7.6: Hilbert transform pairs.

$x(t)$	$\hat{x}(t) = \mathcal{H}[x(t)]$
$\delta(t)$	$\dfrac{1}{\pi t}$
$u(t-T)$	$\dfrac{1}{\pi}\ln\lvert t-T\rvert$
$u(t-T_1)-u(t-T_2)$	$\dfrac{1}{\pi}\ln\left\lvert\dfrac{t-T_1}{t-T_2}\right\rvert$
$e^{+j2\pi f_0 t}$, $f_0>0$	$-je^{+j2\pi f_0 t}$
$e^{-j2\pi f_0 t}$, $f_0>0$	$je^{-j2\pi f_0 t}$
$\cos(2\pi f_0 t)$, $f_0>0$	$\sin(2\pi f_0 t)$
$\sin(2\pi f_0 t)$, $f_0>0$	$-\cos(2\pi f_0 t)$
$\mathrm{sinc}(t)$	$\dfrac{\pi t}{2}\,\mathrm{sinc}^2\left(\dfrac{t}{2}\right)$
$\dfrac{1}{1+t^2}$	$\dfrac{t}{1+t^2}$
$\dfrac{a}{a+t^2}$, $\mathrm{Re}(a)>0$	$\dfrac{t}{a^2+t^2}$
$\dfrac{\sin(2\pi f_0 t)}{t}$, $f_0>0$	$\dfrac{1-\cos(2\pi f_0 t)}{t}$

Example 7.4.1. Suppose $x(t)=\sin(3t)$. Find the corresponding analytic signal, $z(t)$.

SOLUTION

Using the Hilbert transform chart, it follows that

$$\begin{aligned} z(t) = x(t) + j\mathcal{H}[x(t)] &= \sin(3t) + j\mathcal{H}[\sin(3t)] \\ &= \sin(3t) - j\cos(3t) \\ &= -j[\cos(3t) + j\sin(3t)] \\ &= -je^{j3t} \\ &= e^{-j\pi/2}e^{j3t} \\ &= \exp[j(3t-\pi/2)]. \end{aligned}$$

■

7.5 Introduction to Modulation

Modulation is the process of imposing a low-frequency message signal onto a very high-frequency *carrier wave* to create a passband signal to be transmitted. In doing so, the shape of the carrier wave is modified so that information from the message signal is somehow encoded. The reasons for modulation are as follows:

- Practical antenna length: for efficient signal transmission, a minimum antenna length of $\lambda/4$ is needed. Since low-frequency message signals have long wavelengths λ, the required antenna length would have to be on the order of kilometers. By modulating with a high-frequency carrier wave (with short wavelengths), the minimum antenna length is at a reasonable length.
- Frequency-division multiplexing: multiple message signals with similar frequency ranges can only be sent over the same channel if the corresponding carrier wave for each message signal are unique in frequency such that the message signals can be distinguished by the carrier wave frequency (i.e., allocating different parts of the frequency spectrum to each message signal).
- Longer range: while a message alone does not travel far (i.e., human speech), modulating the message signal with a carrier wave allows for longer communications range. While some waves travel further than others, a wave regardless can still carry a message over a longer distance than the physical limitations of the source.

Example 7.5.1. For frequency-division multiplexing, *guard bands* separate the spectra of adjacent signals to avoid interference. Suppose a set of passband signals is designed to occupy 3.5 [kHz] widebands across an available transmission bandwidth of 10 [MHz]. If a guard band of 0.5 [kHz] is allowed, determine the number of signals that can be transmitted at a time via multiplexing.

SOLUTION

It follows that

$$n = \frac{10 \text{ [MHz]}}{(3.5 + 0.5) \text{ [kHz]}} = 2500 \text{ signals}$$

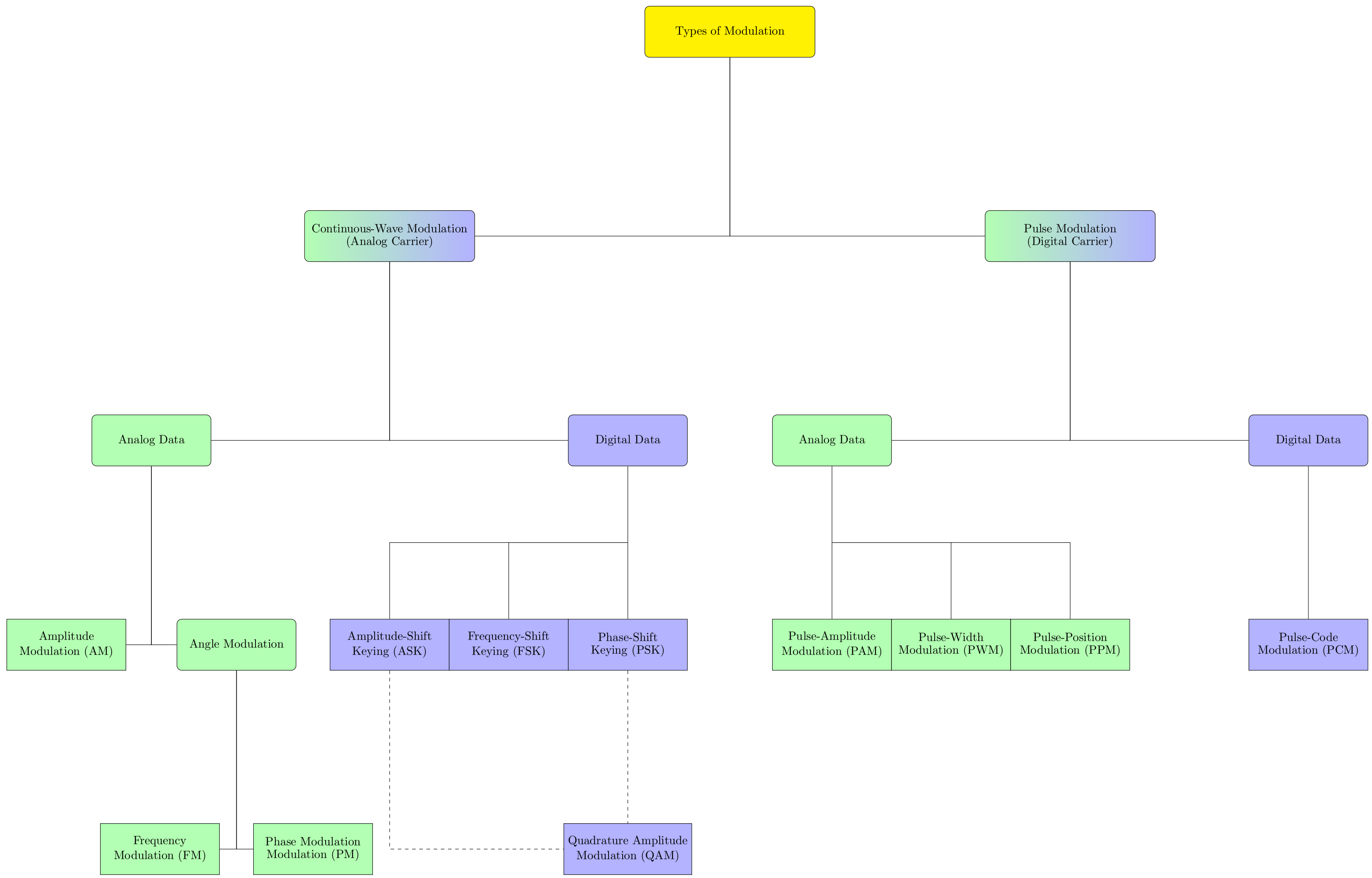
■

In passband communications systems, the transmitter includes a *modulator* which performs the modulation, and the receiver includes a *demodulator* which undoes the modulation (called *demodulation*).

Broadly speaking, *analog modulation* is the modulation of analog message signals, whereas *digital modulation* is the modulation of digital (or digitized) message signals; for both definitions, the classification of the carrier wave does not matter. The flowchart in Figure 7.13 shows the different types of modulation.

The rest of this section will briefly cover an analog modulation technique called amplitude modulation, as both the message signal and the carrier wave are analog in nature.

Figure 7.13: Types of modulation for passband communications systems.



7.6 Amplitude Modulation

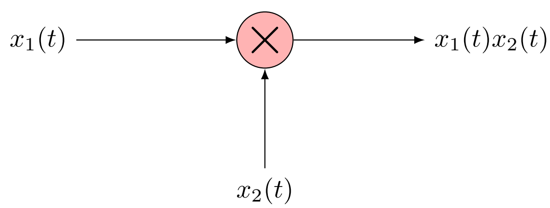
Amplitude modulation (AM) is an analog modulation technique used in passband communications systems, where the amplitude of the carrier wave varies proportionally with the amplitude of the message signal. There are four types of amplitude modulation schemes:

- Double-sideband suppressed carrier (DSB-SC)
- Double-sideband large carrier (DSB-LC)
- Single-sideband (SSB, or SSB-SC)
- Vestigial sideband (VSB)

Since DSB-LC is the most common AM scheme, sometimes DSB-LC is synonomous with the umbrella term AM. For the context of this section, DSB-LC is a specific type of AM. This section will only briefly cover DSB-SC, DSB-LC, and SSB with noiseless channels.

Before going in depth, a new block diagram element is introduced in Figure 7.14: the *multiplier*. While multipliers are actual circuits that can be built, for the rest of the text, multipliers will be analyzed as conceptual building blocks for time-domain block diagrams.

Figure 7.14: Time-domain multiplier.



7.6.1 Double-Sideband Suppressed Carrier (DSB-SC)

7.6.1.1 DSB-SC Modulation

The *double-sideband suppressed carrier* (DSB-SC) modulation scheme is the result of directly interpreting the modulation property of the Fourier transform. Given that a message signal $m(t)$ is modulated with a carrier wave $\cos(2\pi f_c t)$, it follows that the modulated (passband) signal $y_m(t)$ is

$$y_m(t) = m(t)\cos(2\pi f_c t) \tag{7.41}$$

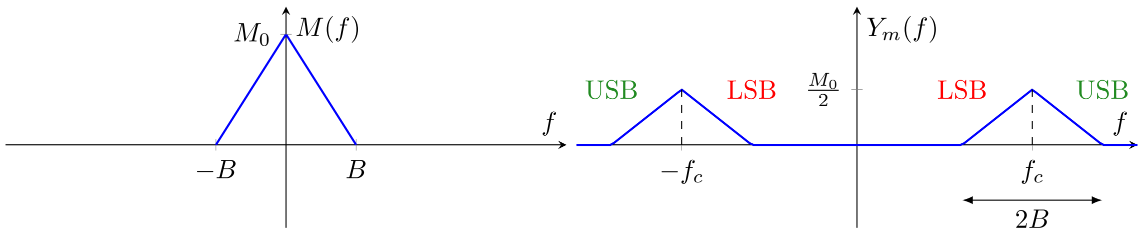
$$\Updownarrow$$

$$Y_m(f) = \frac{1}{2}[M(f + f_c) + M(f - f_c)] \tag{7.42}$$

As seen in Figure 7.15, the passband signal spectrum $Y_m(f)$ contains copies of the baseband spectrum $M(f)$ centered at $\pm f_c$. Furthermore, the modulated signal spectrum can be split into two portions: the outer portion described by $|f| > f_c$ is called the *upper sideband* (USB), whereas the inner portion $|f| < f_c$ is called the *lower sideband* (LSB). It then follows that if the message signal $m(t)$ has a bandwidth B, then the passband signal $y_m(t)$ has a bandwidth $2B$.

Additionally, note that $Y_m(f)$ does not explicitly contain the Fourier transform of a sinusoid. This means that DSB-SC does not introduce a sinusoid at f_c, hence the term *suppressed carrier*.

Figure 7.15: Sample spectra of baseband and DSB-SC passband signals.



7.6.1.2 DSB-SC Demodulation

Interestingly, the same carrier wave can be used to demodulate the passband signal $y_m(t)$ to get $y_d(t)$. It follows that

$$y_d(t) = y_m(t)\cos(2\pi f_c t) = m(t)\cos^2(2\pi f_c t) = \frac{m(t)}{2} + \frac{m(t)\cos(2\pi \cdot 2 f_c t)}{2} \tag{7.43}$$

$$\Updownarrow$$

$$Y_d(f) = \frac{1}{2}M(f) + \frac{1}{4}[M(f+2f_c) + M(f-2f_c)] \tag{7.44}$$

Therefore, a lowpass filter (LPF) can be applied to get the output $\frac{1}{2}m(t)$, with the condition that

$$2f_c - B > B \Longrightarrow f_c > B \tag{7.45}$$

Figure 7.16: Ideal lowpass filtering of a demodulated DSB-SC signal.

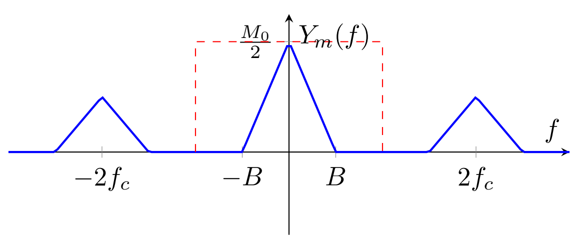
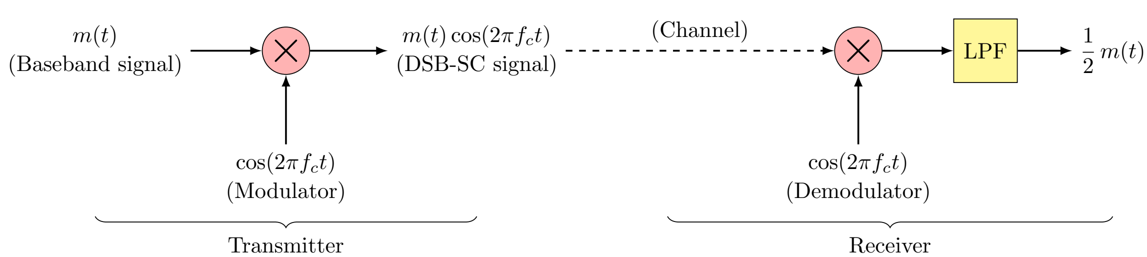


Figure 7.17: DSB-SC block diagram.



However, for DSB-SC to function properly, both the modulating carrier wave and the demodulating carrier wave must be in sync. This is not simple to achieve over a large distance.

7.6.2 Double-Sideband Large Carrier (DSB-LC)

7.6.2.1 DSB-LC Modulation

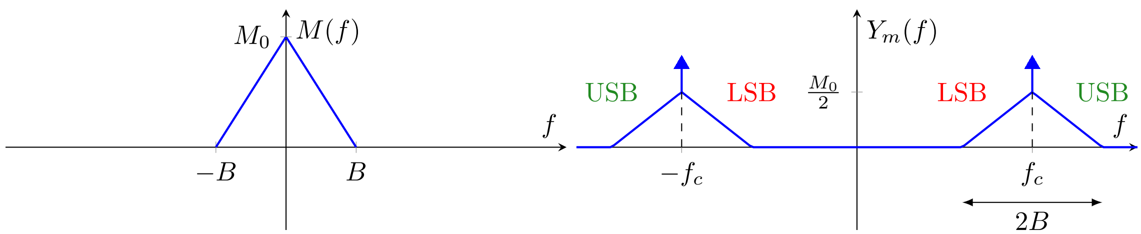
The *double-sideband large carrier* (DSB-LC) is an alternative modulation scheme to DSB-SC that removes the need to use a carrier wave at the receiver side. DSB-LC uses a DC bias such that the carrier wave $\cos(2\pi f_c t)$ does indeed get transmitted and appears in the passband spectrum. It follows that

$$y_m(t) = [A + m(t)]\cos(2\pi f_c t) \tag{7.46}$$

$$\Updownarrow$$

$$Y_m(f) = \frac{A}{2}[\delta(f+f_c) + \delta(f-f_c)] + \frac{1}{2}[M(f+f_c) + M(f-f_c)] \tag{7.47}$$

Figure 7.18: Sample spectra of baseband and DSB-LC passband signals.



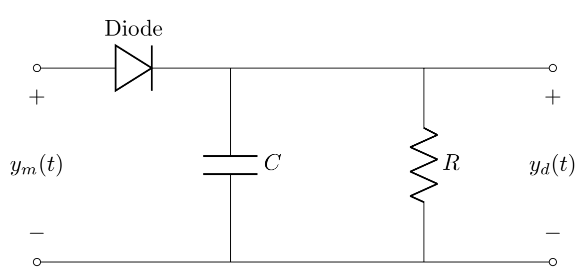
7.6.2.2 DSB-LC Demodulation

An envelope detector is a circuit that can be used as a demodulator for DSB-LC instead of using a carrier oscillator, given the conditions that

$$A + m(t) \geq 0, \text{ for all } t \tag{7.48}$$

$$f_c \gg B \tag{7.49}$$

Figure 7.19: Envelope detector.



An additional design parameter for the envelope detector is that

$$2\pi B < \frac{1}{RC} \ll 2\pi f_c \tag{7.50}$$

A huge drawback of DSB-LC however is the power efficiency since the DC bias does not contain any useful information. From

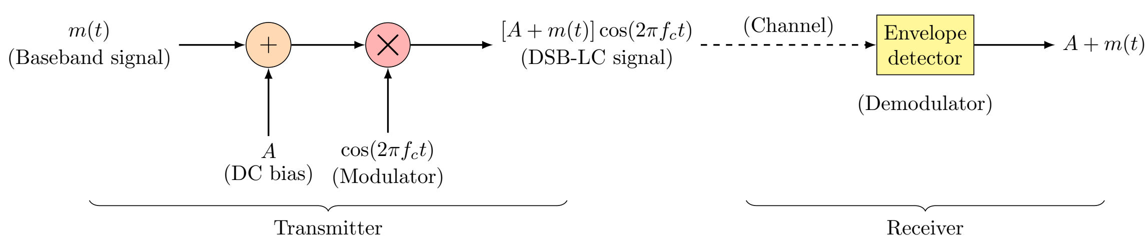
$$y_m(t) = \underbrace{A\cos(2\pi f_c t)}_{\text{carrier}} + \underbrace{m(t)\cos(2\pi f_c t)}_{\text{sidebands}} \tag{7.51}$$

the carrier power P_c is wasteful, and the sideband power P_s is the only useful power. The power efficiency is then defined as

$$\eta = \frac{\text{useful power}}{\text{total power}} = \frac{P_s}{P_c + P_s} = \frac{P_{av}[m(t)]/2}{A^2/2 + P_{av}[m(t)]/2} = \frac{P_{av}[m(t)]}{A^2 + P_{av}[m(t)]}, \tag{7.52}$$

where $P_{av}[m(t)]$ is the average power of the message signal $m(t)$.

Figure 7.20: DSB-LC block diagram.



Example 7.6.1. Suppose a tone $m(t) = \cos(2\pi f t)$ with frequency $f = 440$ [Hz] is transmitted over a channel using DSB-LC. Determine the optimal power efficiency.

SOLUTION

First, since

$$A + m(t) \geq 0,$$

the optimal DC bias is then

$$A = -\min[m(t)] = -(-1) = 1.$$

Additionally, since the tone is sinusoidal, the average power of $m(t)$ is

$$P_{av}[m(t)] = \frac{1}{2}.$$

Therefore, the optimal power efficiency is

$$\eta = \frac{P_s}{P_c + P_s} = \frac{P_{av}[m(t)]}{A^2 + P_{av}[m(t)]} = \frac{1/2}{1/2 + 1/2} = 0.5 \Longrightarrow 50\%$$

■

7.6.3 Single-Sideband (SSB)

7.6.3.1 SSB Modulation

The pitfall of both DSB-SC and DSB-LC is that both schemes occupy twice the signal bandwidth when modulated and are inefficient in their usage of the frequency spectrum. *Single-sideband* (SSB, or SSB-SC) modulation offers an alternative solution to the bandwidth issue by removing the lower sidebands.

SSB modulation makes use of $h_Q(t)$, previously introduced as the quadrature filter (also called the Hilbert transformer or the 90°–phase shifter). As seen in Figure 7.21, the system at the transmitter end outputs the signal

$$y_m(t) = m(t)\cos(2\pi f_c t) - \hat{m}(t)\sin(2\pi f_c t), \tag{7.53}$$

where $\hat{m}(t)$ is the Hilbert transform of $m(t)$. The frequency response is given by

$$Y_m(f) = \begin{cases} M(f + f_c), & f < -f_c \\ 0, & |f| < f_c \\ M(f - f_c), & f > f_c \end{cases} \tag{7.54}$$

Figure 7.21: SSB block diagram.

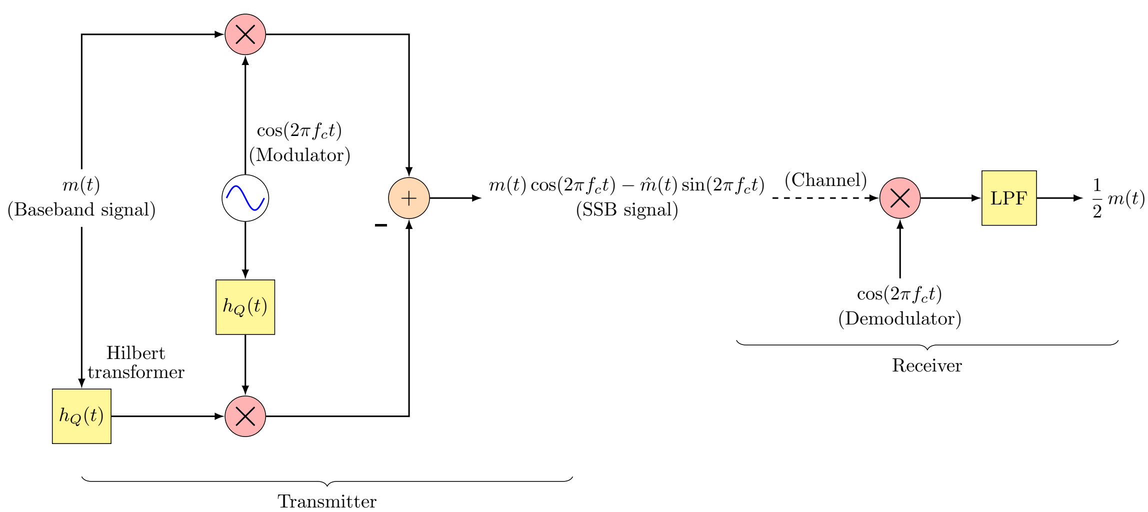
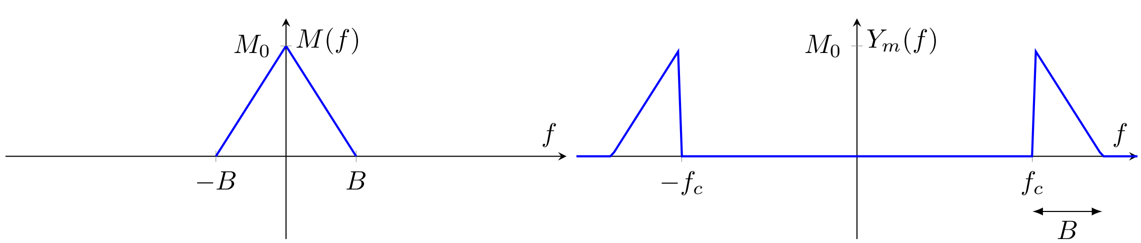


Figure 7.22: Sample spectra of baseband and SSB passband signals.



7.6.3.2 SSB Demodulation

SSB demodulation is the same as DSB-SC demodulation, using the same carrier oscillator as the transmitter and lowpass filtering afterwards. It follows that

$$y_d(t) = y_m(t)\cos(2\pi f_c t) = m(t)\cos^2(2\pi f_c t) - \hat{m}(t)\cos(2\pi f_c t)\sin(2\pi f_c t) \tag{7.55}$$

$$= \frac{m(t)}{2} + \frac{m(t)\cos(2\pi \cdot 2f_c t)}{2} - \frac{\hat{m}(t)\sin(2\pi \cdot 2f_c t)}{2} \tag{7.56}$$

Assuming $f_c > B$, the signal $y_d(t)$ can be inputted through a lowpass filter such that the output at the receiver end is $\frac{1}{2}m(t)$.

7.6.4 Quadrature Amplitude Modulation

Interestingly, the same transmitter block diagram for SSB modulation can similarly be applied to the analog version of *quadrature amplitude modulation* (QAM), which is reserved for I/Q signals (also called analytic signals).

When isolating the real and imaginary parts to get the in-phase component $m_I(t)$ and the quadrature component $m_Q(t)$, the components can be reconstructed at the receiver end by applying the same transmitter circuit on the receiver end, then lowpass filtering each component.

Figure 7.23: Analog QAM block diagram.