

# Chapter 3

## Basics of Systems

### 3.1 What are Systems?

As previously mentioned, a signal represents some underlying physical variable of interest. As an abstraction, we consider a signal as simply a real-valued (or sometimes complex-valued) function defined on some domain. For example, we might represent a signal as just a real-valued function of time  $x(\cdot)$  where  $x(t)$  represents the value of the signal at time  $t$ . Similarly, we will consider an abstraction of the general system of Figure 3.1 from Chapter 1.

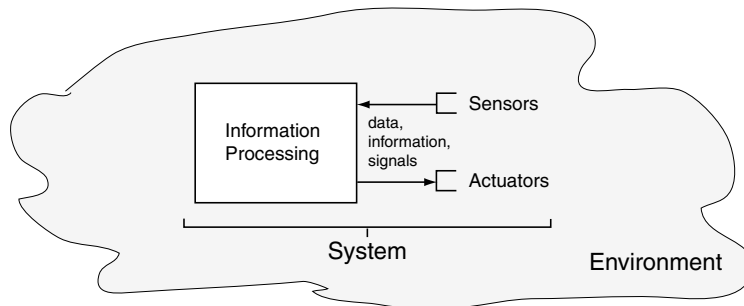


Figure 3.1: A general system immersed in its environment.

In this figure, we think of a system as some part of the physical world that interacts with its environment and is designed to carry out some task. The box labeled “information processing” receives input signals from various sensors and produces output signals for various actuators. Thus, we think of the system as transforming input signals into output signals. A number of subtasks within

---

\*©1999-2002 by Sanjeev R. Kulkarni. All rights reserved. Please do not distribute without permission of the author.

<sup>†</sup>Lecture Notes for ELE201 Introduction to Electrical Signals and Systems.

<sup>‡</sup>Thanks to Rich Radke and Sean McCormick for producing the figures.

“information processing” themselves perform operations on a signal to produce another signal — i.e., they transform one function  $x(t)$  to produce a new function  $y(t)$ . The term *system* is used in this abstract and technical sense to refer to such mappings that take a signal as input and produce another signal as output. As we’ll see, by making this abstraction and imposing additional assumptions, we’ll be able to study special types of systems in a precise way that leads to useful insights and results.

As with signals, we can have various types of systems such as continuous-time, discrete-time, analog, digital, etc. We will let  $x(t)$  or  $x[n]$  denote the input signal for continuous-time and discrete-time systems, respectively, and let  $y(t)$  and  $y[n]$  denote the corresponding the output signals.  $\mathcal{H} = \mathcal{H}[\cdot]$  will denote the system, so that

$$y(t) = \mathcal{H}[x(t)]$$

and similarly in the discrete-time case with “(t)” replaced by “[n]”. It is standard to represent this relationship for the system  $\mathcal{H}$  with a schematic diagram as shown in Figure 3.2.

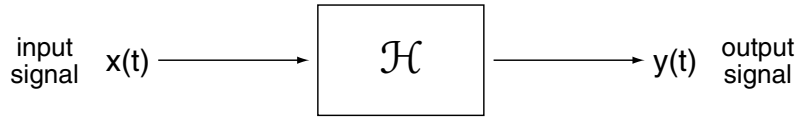


Figure 3.2: Abstraction of a general system.

The range of possible systems is so wide that it is difficult to say much that applies to completely general systems. However, by restricting attention to special classes of systems a set of tools can be developed that helps in the modeling, analysis, and design of systems. Two particularly useful properties are linearity and time-invariance, which we describe next. A rich theory has been developed for this class of systems. Although real physical systems are rarely exactly linear or time-invariant, results for the idealized assumptions still provide useful models for understanding more complex systems.

## 3.2 Linearity

One extremely useful class of systems is the class of linear systems. Roughly, a system is linear if adding or scaling inputs to the system simply results in adding or scaling the corresponding outputs. More precisely, a system  $\mathcal{H}$  is *linear* if the following two conditions are satisfied:

1.

$$\mathcal{H}[a x(t)] = a \mathcal{H}[x(t)] \quad (3.1)$$

for all constants  $a$  and input signals  $x(t)$ ; and

2.

$$\mathcal{H}[x_1(t) + x_2(t)] = \mathcal{H}[x_1(t)] + \mathcal{H}[x_2(t)] \quad (3.2)$$

for all input signals  $x_1(t), x_2(t)$ .

The first condition implies, for example, that once we know the output of the system due to  $x(t)$ , then the output due to  $2x(t)$  will be twice the original output. The second condition implies that to find the output due to the sum of two signals  $x_1(t) + x_2(t)$ , we can first find the outputs due to each signal separately and then just sum these outputs. The two conditions (3.1) and (3.2) can be combined into the single condition

$$\mathcal{H}[a_1x_1(t) + a_2x_2(t)] = a_1\mathcal{H}[x_1(t)] + a_2\mathcal{H}[x_2(t)] \quad (3.3)$$

for all constants  $a_1, a_2$  and input signals  $x_1(t), x_2(t)$ . This property is often referred to as *superposition*. The definition of linearity can also be written for discrete-time systems in the obvious way.

The beauty of linear systems arises from the superposition property, which lets us analyze outputs due to very complicated inputs by first just considering some standard set of simple inputs, and then considering additive combinations of these.

### 3.3 Time-Invariance

Time-invariance is another useful concept to consider in studying systems. Time-invariance means that whatever behavior a system has, if the inputs were delayed by some amount of time, then the behavior of the system would be unchanged other than just being delayed by that same amount of time. More precisely, suppose  $y(t)$  denotes the output of the system due to  $x(t)$ , so that as before  $y(t) = \mathcal{H}[x(t)]$ . Then  $\mathcal{H}$  is *time-invariant* if

$$\mathcal{H}[x(t - t_0)] = y(t - t_0) \quad (3.4)$$

for all times  $t_0$  and input signals  $x(t)$ . If a system is not time-invariant then it is said to be *time-varying*.

In this condition,  $t_0$  denotes the delay, so that  $x(t - t_0)$  is the signal  $x(t)$  delayed by time  $t_0$ . Note that  $t_0$  is allowed to be negative, so that the condition also applies to starting the input early as well as delaying it.

The left hand side of (3.4) denotes the output of the system if we apply the delayed input  $x(t - t_0)$ . The right hand side is the output  $y(t)$  delayed by time  $t_0$ . That is, the left hand side first delays the input to get  $x(t - t_0)$  and then applies this to the system, while the right hand side first puts  $x(t)$  into the system and then delays the resulting output. Time-invariance requires these to give the same final result.

The notion of time-invariance can be defined in a similar way for discrete-time systems. Of course, in this case, since the time variable  $n$  only takes on integer values, we only need to consider integer shifts in the equation for time-invariance.

Like linearity, the power of time-invariance lies in the nice behavioral properties this condition imposes on the system. It allows us to use knowledge of

the system behavior on certain inputs to determine what the system would do on other (namely shifted) inputs. The two conditions together (linearity and time-invariance) are particularly powerful, as we'll see in the next few sections. In fact, with these two conditions the behavior of a system to *any* input can be determined once we know the behavior due to just one special input – namely the delta function.

### 3.4 Impulse Response and Output of LTI Systems

We now focus on the class of linear time-invariant systems (also known as LTI systems). With these two properties (i.e., linearity and time-invariance) a rich theory can be developed. Here we will touch on just some of the basic tools for describing and analyzing such systems, but these tools form the basis for much of the field of signals and systems.

We will first consider the case of discrete-time systems. Suppose we apply the discrete-time delta function as the input to the system, namely we let  $x[n] = \delta[n]$ . The output will be some signal which we will denote as  $h[n]$ , so that

$$h[n] = \mathcal{H}[\delta[n]]$$

The function  $h[n]$  is called the *impulse response* of the system. It is the output (response) of the system when the input is a delta function (impulse).

For the moment, let's not be concerned with what particular values we observe for the impulse response, but only that we observe some output that we will record and manipulate in various ways. Our focus is on seeing whether we can use our knowledge that the output is  $h[n]$  when the input is  $\delta[n]$  to predict what output we will get for some other inputs.

Without some restrictions or assumptions on the system, we have no hope of predicting the output on some new inputs. However, recall that we have assumed that the system is LTI, which are very powerful conditions indeed, as we shall now see.

First, since the system is time-invariant, we can easily determine the input to the shifted delta function  $\delta[n - n_0]$ . This input is just a delayed version of the delta function  $\delta[n]$  that takes the value 1 at time  $n = n_0$  and is 0 elsewhere. Time-invariance tells us that the output due to this delayed delta function is just  $h[n - n_0]$ , namely the output due to  $\delta[n]$  but just shifted by the same amount. In equation form,

$$\mathcal{H}[\delta[n - n_0]] = h[n - n_0]$$

Now, we can use linearity to determine the output due to many other inputs. For example, what is the output due to applying  $2\delta[n]$ ? Well, we already know that the output  $h[n]$  results from the input  $\delta[n]$ . So by the scaling property of linearity, we must get the output  $2h[n]$  when we apply  $2\delta[n]$ .

What about the output due to  $\delta[n - 1]$ ? Time-invariance tells us that the output must be the same as the output due to  $\delta[n]$  but just shifted by one unit.

I.e., the output is  $h[n - 1]$ . Likewise, if we apply  $\delta[n] + \delta[n - 1]$ , the additive property of linearity together with time-invariance tells us that the output must be  $h[n] + h[n - 1]$ .

Now, we can use both the scaling and additive properties of linearity together with time-invariance, to determine that if we apply the input  $2\delta[n] + 3\delta[n - 1]$ , the output will be  $2h[n] + 3h[n - 1]$ . Thus, by just knowing the output due to  $\delta[n]$  and using linearity and time-invariance, we can determine the output due to some pretty complicated inputs by just “building up” these inputs in terms of shifted and scaled delta functions, and then combining the correspondingly shifted and scaled versions of the impulse response.

In fact, by carrying this idea further, we see that by appropriate combinations of scaled and shifted delta functions, we can in fact synthesize *any* desired input function. Suppose we want to create an arbitrary input  $x[n]$ . We can start by making sure we create the proper value at time 0, namely  $x[0]$ . To do this we include  $\delta[n]$  scaled by  $x[0]$ , i.e.,  $x[0]\delta[n]$ . Next, to get the desired value at time 1,  $x[1]$ , we simply add  $x[1]\delta[n - 1]$ , which is  $\delta[n - 1]$  scaled by  $x[1]$ . Then to get the desired value at time  $-1$ , we add  $x[-1]\delta[n + 1]$ . Notice that each shifted delta function is non-zero only at one time, so that adding new terms won't ruin the values we've already created. Thus, continuing in this manner, we can write the input signal  $x[n]$  as

$$\begin{aligned} x[n] = & \cdots + x[-2]\delta[n + 2] + x[-1]\delta[n + 1] + x[0]\delta[n] \\ & + x[1]\delta[n - 1] + x[2]\delta[n - 2] + \cdots \end{aligned}$$

or using summation notation

$$x[n] = \sum_{i=-\infty}^{\infty} x[i]\delta[n - i] \quad (3.5)$$

This is actually just a version of the sifting property of the delta function that we encountered in Section XX.

Although expressing  $x[n]$  in this way might seem degenerate or not particularly useful, we have actually achieved a great deal. Initially, we had no idea what output we would get if the input  $x[n]$  was applied to the system. But now, we can use linearity and time-invariance as we did before to see that the output will just be a combination of a bunch (albeit, potentially an infinite number) of shifted and scaled versions of the impulse response  $h[n]$ . The fact that the scaling terms are the values of the input signal itself shouldn't bother or surprise us. On the contrary, that is exactly what we did above in finding the output due to  $2\delta[n] + 3\delta[n - 1]$ , the only difference being that we used specific numbers (2 and 3) for values of the input signal. In fact, the output of the system generally *should* depend on the input signal, otherwise it would be a dull system.

Formally carrying out the manipulations, we get that the output  $y[n]$  due to the input  $x[n]$  is given by

$$y[n] = \mathcal{H}[x[n]]$$

$$\begin{aligned}
&= \mathcal{H} \left[ \sum_{i=-\infty}^{\infty} x[i] \delta[n-i] \right] \\
&= \sum_{i=-\infty}^{\infty} x[i] \mathcal{H}[\delta[n-i]] \\
&= \sum_{i=-\infty}^{\infty} x[i] h[n-i]
\end{aligned}$$

The first equality is the definition of  $y[n]$ . The second uses the expression (3.5) for  $x[n]$  obtained above, which is really the sifting property of the delta function. The third equality follows from linearity, where we have used both the additive property to bring  $\mathcal{H}$  inside the sum, and the scaling property to bring  $x[i]$  out of  $\mathcal{H}[\cdot]$ . The fourth equality follows from time-invariance and the definition of the impulse response  $h[n]$ .

This final result is important enough to repeat as a separate equation. Namely, if  $h[n]$  is the impulse response of a system, and we apply an input  $x[n]$  then the output  $y[n]$  is given by

$$y[n] = \sum_{i=-\infty}^{\infty} x[i] h[n-i] \quad (3.6)$$

A few comments on this important result are in order. The somewhat complicated/convoluted operation on the right hand side of Equation (3.6) combining  $x[n]$  and  $h[n]$  is appropriately called the *convolution* of  $x[n]$  and  $h[n]$ . It arises so often that it is given the special notation  $(x * h)[n]$ . This operation will be discussed in more detail in the next section.

Although the expression for the convolution looks complicated, the remarkable thing is that this result implies that once we know the impulse response of a system (i.e., the output due to the single input  $\delta[n]$ ), then we can find the output due to *any* input! In fact, the output is obtained by just convolving the input with the impulse response. We see that linearity and time-invariance impose so much structure on the system that knowing how the system behaves for just the one special input  $\delta[n]$  determines the behavior of the system for all inputs. Thus, as far as just the input/output behavior of a system is concerned, the system is completely described by its impulse response. In principle, if we have a black box and know nothing about its operation except that it happens to be linear and time-invariant, we can find out everything about its behavior by doing one simple experiment. We apply the delta function  $\delta[n]$  to the input and measure the output signal that results. We then know exactly how the black box will work under other inputs. Of course, in reality things aren't quite so simple for a variety of reasons, including the fact that most real systems are neither exactly linear nor time-invariant. Nevertheless, the result is extremely useful as we'll see on many occasions.

Finally, we should mention that a result analogous to the discrete-time case can also be obtained for continuous-time systems, as long as the continuous-

time system satisfies certain mild continuity conditions that in practical cases are always satisfied. In particular, we let  $h(t)$  denote the impulse response of the system. That is,  $h(t)$  is the output of the system due to the input  $\delta(t)$ , so that

$$h(t) = \mathcal{H}[\delta(t)]$$

Then (under suitable continuity conditions) the output  $y(t)$  due to *any* input  $x(t)$  is given by

$$y(t) = \int_{-\infty}^{\infty} x(\tau) h(t - \tau) d\tau \quad (3.7)$$

A rough derivation of this result follows very similar lines as in the discrete-time case. First, by the sifting property of  $\delta(t)$ , we can write

$$x(t) = \int_{-\infty}^{\infty} x(\tau) \delta(t - \tau) d\tau$$

Then, using linearity and time-invariance we get

$$\begin{aligned} y(t) &= \mathcal{H}[x(t)] \\ &= \mathcal{H}\left[\int_{-\infty}^{\infty} x(\tau)\delta(t - \tau) d\tau\right] \\ &= \int_{-\infty}^{\infty} x(\tau)\mathcal{H}[\delta(t - \tau)] d\tau \\ &= \int_{-\infty}^{\infty} x(\tau)h(t - \tau) d\tau \end{aligned}$$

The third equality uses linearity, and this is where the continuity conditions are needed in a rigorous derivation. The last equality uses time-invariance.

As in the discrete-time case, the right-hand-side of Equation (3.7) is called the (continuous-time) *convolution* of  $x(t)$  and  $h(t)$ , and denoted  $(x * h)(t)$ . The result again implies that once we know the impulse response, we can determine the output of the system due to any input via a convolution. As a practical matter, finding the impulse response of a continuous-time black-box is more challenging since we are generally unable to produce a continuous-time delta function. However, approximations to the delta function can be produced and are useful, and in any case the result is extremely useful as a modeling and analysis tool.

## 3.5 Convolution

In the previous section, we introduced the notion of the convolution of two signals. We repeat the definition here for convenience. In discrete-time, the convolution of  $x[n]$  and  $h[n]$  is defined by

$$(x * h)[n] = \sum_{i=-\infty}^{\infty} x[i]h[n - i].$$

In continuous-time, the convolution  $x(t)$  and  $h(t)$  is defined by

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

We have seen that the convolution of two signals arose naturally as a way to understand the behavior of LTI systems. Specifically, the output of an LTI system is simply the input convolved with the impulse response. However, in addition to this interpretation, the convolution operation can be motivated by simply thinking of it as a rather general and useful way to transform one signal into another signal. If  $y(t) = (x * h)(t)$ , we think of  $x(t)$  as the original signal,  $y(t)$  as the new signal, and  $h(t)$  as representing the particular type of transformation we wish to perform. With this perspective, it is immaterial that the convolution operation can be viewed as the output of an LTI system. In any case, regardless of the particular view or motivation of convolution, it is an operation that occurs so frequently in both the theory and practice of signals and systems that it well worth spending time to understand the convolution operation.

Let's consider the discrete-time case first. Let

$$y[n] = (x * h)[n] = \sum_{i=-\infty}^{\infty} x[i]h[n - i].$$

First notice that the convolution of two signals results in another signal, not just a single number. Thus to understand what the convolution does, we need to understand what it does at each time. Notice that for any fixed  $n$ , the expression on the right hand side is the sum of a bunch of (infinitely many) terms. Each term is the value of  $x[\cdot]$  at some time and the value of  $h[\cdot]$  at some other time. Thus the value of  $y[n]$  at some time  $n$  is really just a weighted sum of the signal  $x[\cdot]$  where the weights depend on the signal  $h[\cdot]$ . At some different time, the convolution is still just a weighted sum where again the weights depend on  $h[\cdot]$  but now different weights are assigned to values of  $x[\cdot]$ .

Notice that part of the expression for the convolution sum, namely the  $x[i]$  term, does not depend on  $n$ . The summation index  $i$  is just a dummy variable to carry out the weighted sum. So we can think of the convolution as follows (and as illustrated in Figure 3.3):

1. Draw the signal  $x[\cdot]$  where we use the dummy variable  $i$ , since we will be taking a weighted sum.
2. Flip the signal  $h[i]$  about the vertical (range) axis to get  $h[-i]$  and then shift by  $n$  to get  $h[n - i]$ .
3. Multiply the corresponding values of  $x[i]$  and  $h[n - i]$  and add up to find  $y[n]$ .

In continuous-time, the convolution

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



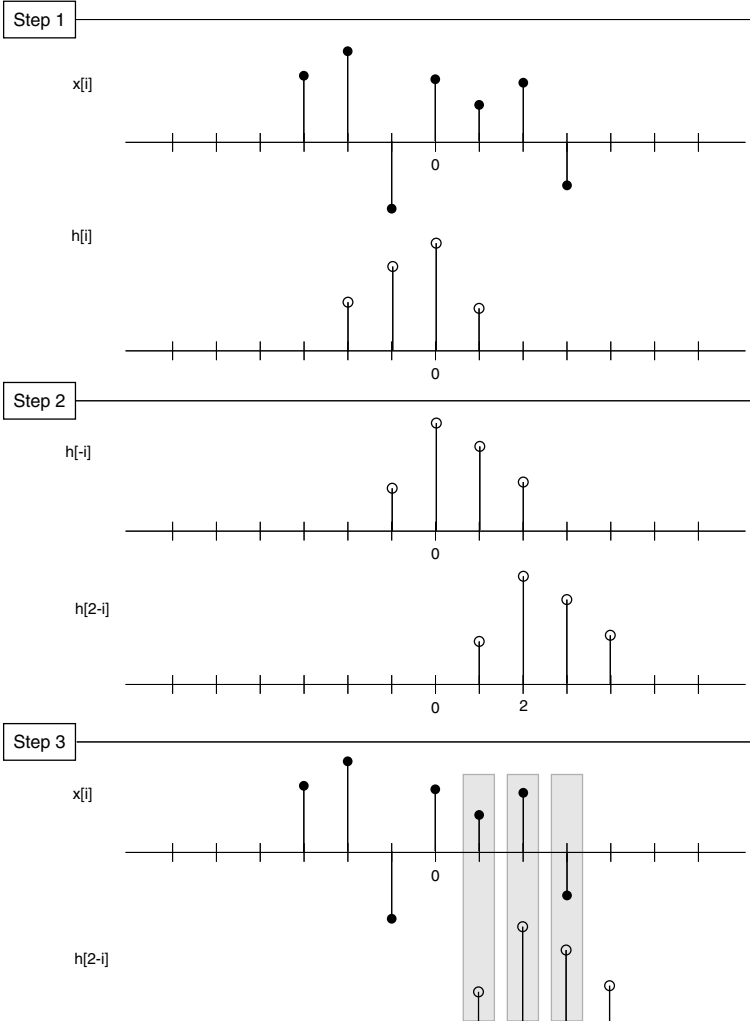


Figure 3.3: Schematic convolution of  $x[n]$  with  $h[n]$

has a similar interpretation, but with the sum replaced by an integral. Namely, the convolution at time  $t$  is obtained as follows:

1. Consider  $x(\tau)$  where we now use  $\tau$  as the dummy variable for the weighted integral.
2. Flip the signal  $h(\tau)$  about the vertical (range) axis to get  $h(-\tau)$  and then shift by  $t$  to get  $h(t - \tau)$ .
3. Multiply the corresponding values of  $x(\tau)$  and  $h(t - \tau)$  and integrate to find  $y(t)$ .

### 3.6 Frequency Response of LTI Systems

In this section, we consider the response of an LTI system to a sinusoidal input. If the input to the system is a sinusoid at a particular frequency, we would like to know what can be said about the output of the system.

It turns out that the behavior of LTI systems to sinusoids is simple. All that an LTI system can do to a sinusoidal input is to change its amplitude and phase. That is, if the input to an LTI system is a sinusoid at some frequency  $\omega$  then the output is also a sinusoid at frequency  $\omega$ , but possibly with a different amplitude and phase.

The theory is cleaner if we allow complex sinusoids that were discussed in Chapter XX. Namely, we consider the input

$$x(t) = e^{j\omega t} = \cos \omega t + j \sin \omega t$$

Now, let's see what happens if we apply this input to an LTI system with impulse response  $h(t)$ .

Recall that the output  $y(t)$  is just the convolution of  $x(t)$  and  $h(t)$ , i.e.,

$$y(t) = (x * h)(t)$$

Since we have seen that convolution is commutative, we can write this as

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

So far, this is completely general, i.e., we haven't yet used the fact that  $x(t)$  is sinusoidal. We'll now use this fact by substituting  $x(t) = e^{j\omega t}$  into the above expression to get

$$\begin{aligned} y(t) &= \int_{-\infty}^{\infty} h(\tau)e^{j\omega(t-\tau)} d\tau \\ &= \int_{-\infty}^{\infty} h(\tau)e^{j\omega t}e^{-j\omega\tau} d\tau \\ &= e^{j\omega t} \int_{-\infty}^{\infty} h(\tau)e^{-j\omega\tau} d\tau \end{aligned}$$

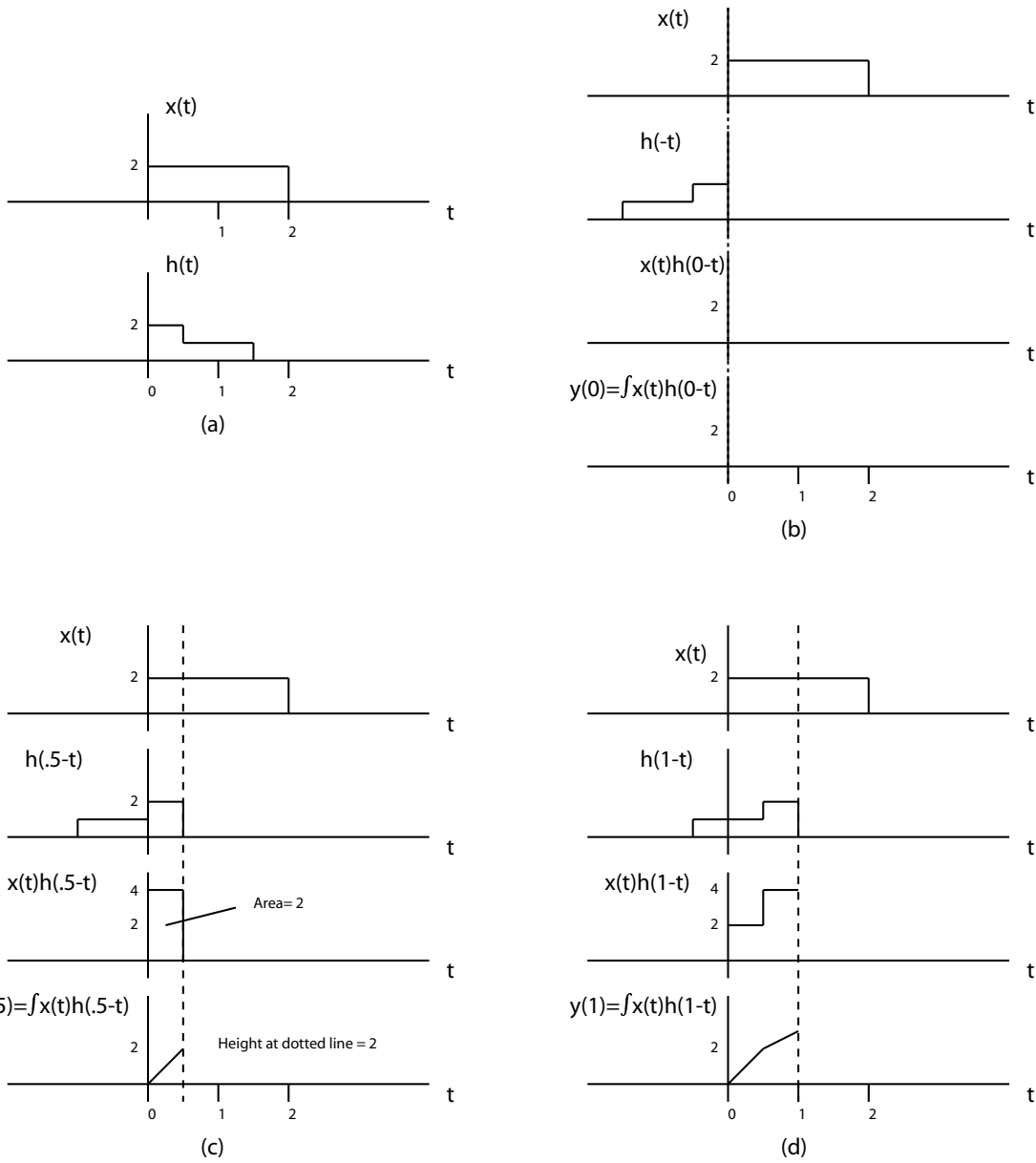


Figure 3.4: (a.) The signal,  $x(t)$  top and system response  $h(t)$  bottom. (b.)  $x(t)$  top.  $h(t)$  flipped about vertical axis middle top. Product of  $x(\tau)$  and  $h(t-\tau)$  middle bottom. Integral of product bottom. (c. - h.) Same as (b.) but further along in time.

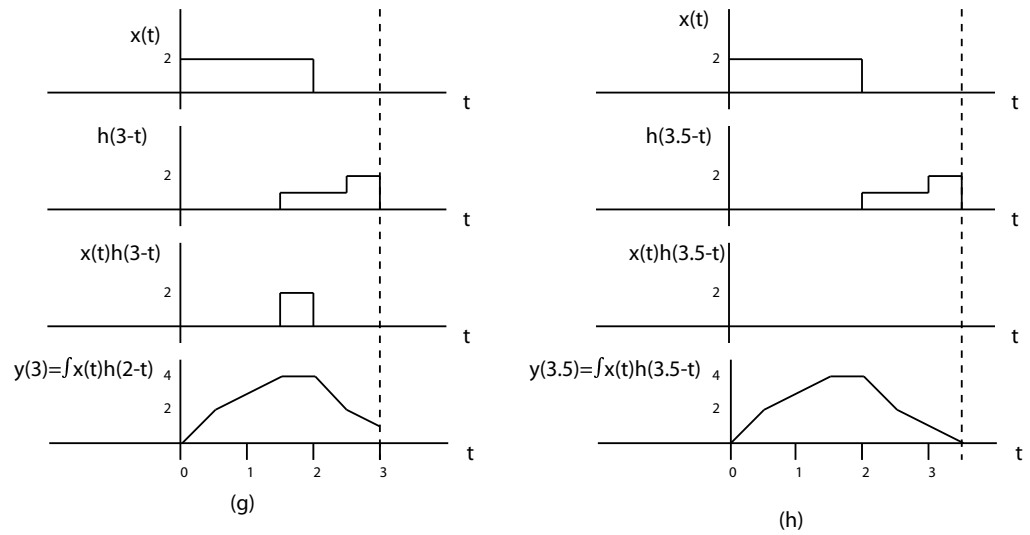
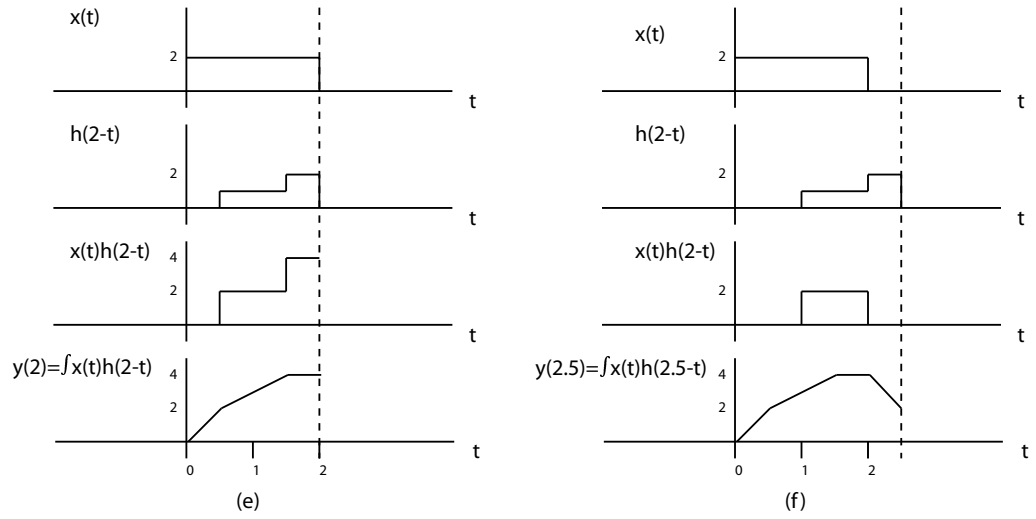


Figure 3.5:

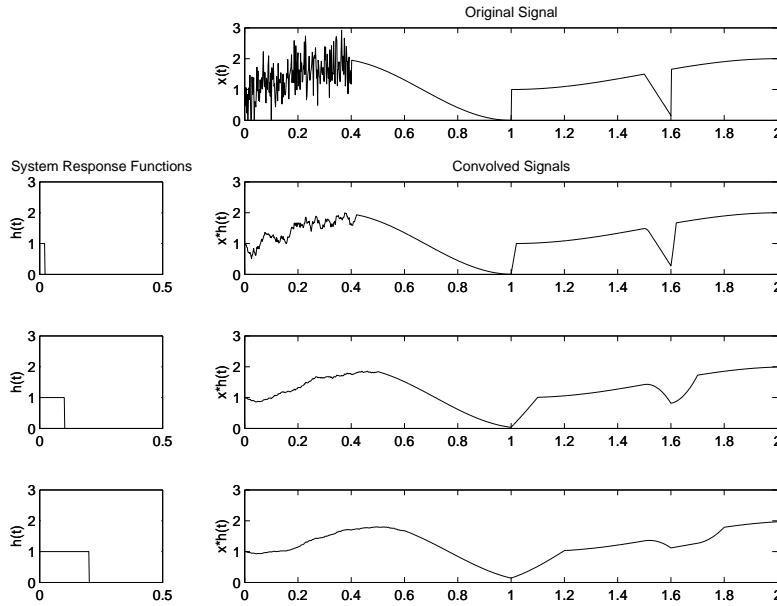


Figure 3.6: Convolving with rect functions of different lengths

Notice that on the right hand side we have the original complex sinusoid  $e^{j\omega t}$  times a term that does not depend on  $t$ . Thus, we have shown the desired result.

Of course, the term multiplying  $e^{j\omega t}$  depends on  $\omega$ . In fact, this term is called the *frequency response* of the system, since it tells us what the system does to sinusoidal inputs at each frequency. We will let  $H(\omega)$  denote the frequency response, so that

$$H(\omega) = \int_{-\infty}^{\infty} h(\tau)e^{-j\omega\tau} d\tau \tag{3.8}$$

and then we can write

$$y(t) = H(\omega)e^{j\omega t}$$

Once we know the impulse response  $h(t)$ , then in principle we can compute the frequency response  $H(\omega)$  by simply carrying out the integral in Equation (3.8).

The fact that the frequency response can be computed from the impulse response should no longer be surprising to us since we have previously seen that the impulse response actually determines the output for *any* input. So, in particular, it determines the output for a sinusoid at frequency  $\omega$ . What may be surprising is that the behavior of LTI systems on sinusoids is so simple. What may be more surprising is that not only does the impulse response determine the frequency response, but the reverse is also true. That is, if we know the frequency response for all  $\omega$  then we can determine the impulse response, and hence we can determine the output due to any input. Thus, once we know what the system does to sinusoids at each frequency, we know how the system

responds to all inputs. This result is closely connected to extremely useful concepts of frequency domain representations of signals and Fourier transforms, which is the topic we will consider in the next chapter. In fact, as we'll soon see, the frequency response  $H(\omega)$  is simply the Fourier transform (i.e., frequency domain representation) of the impulse response.

We finish this section by showing that similar ideas hold in the case of discrete-time systems. For the discrete-time case, we consider the discrete-time complex sinusoid

$$x[n] = e^{j\omega n}$$

As in the continuous-time case, the output  $y[n]$  of an LTI system with impulse response  $h[n]$  is given by

$$\begin{aligned} y[n] &= (x * h)[n] \\ &= (h * x)[n] \\ &= \sum_{k=-\infty}^{\infty} h[k]x[n-k] \end{aligned}$$

Now, substituting the expression for  $x[n]$  we get

$$\begin{aligned} y[n] &= \sum_{k=-\infty}^{\infty} h[k]e^{j\omega(n-k)} \\ &= \sum_{k=-\infty}^{\infty} h[k]e^{j\omega n}e^{-j\omega k} \\ &= e^{j\omega n} \sum_{k=-\infty}^{\infty} h[k]e^{-j\omega k} \end{aligned}$$

Thus, in the discrete-time case the frequency response  $H(\omega)$  is given by

$$H(\omega) = \sum_{k=-\infty}^{\infty} h[k]e^{-j\omega k} \quad (3.9)$$

The output can then be written as

$$y[n] = H(\omega)e^{j\omega n}$$

so that in the case of a discrete-time LTI system we also get a sinusoidal output at the same frequency as the input.