This chapter is optional but recommended.

CHAPTER 13

Continuous Signal Processing

Continuous signal processing is a parallel field to DSP, and most of the techniques are nearly identical. For example, both DSP and continuous signal processing are based on linearity, decomposition, convolution and Fourier analysis. Since continuous signals cannot be directly represented in digital computers, don't expect to find computer programs in this chapter. Continuous signal processing is based on *mathematics*; signals are represented as equations, and systems change one equation into another. Just as the *digital computer* is the primary tool used in DSP, *calculus* is the primary tool used in continuous signal processing. These techniques have been used for centuries, long before computers were developed.

The Delta Function

Continuous signals can be decomposed into scaled and shifted *delta functions*, just as done with discrete signals. The difference is that the continuous delta function is much more complicated and mathematically abstract than its discrete counterpart. Instead of defining the continuous delta function by what it *is*, we will define it by the *characteristics it has*.

A thought experiment will show how this works. Imagine an electronic circuit composed of linear components, such as resistors, capacitors and inductors. Connected to the input is a signal generator that produces various shapes of short *pulses*. The output of the circuit is connected to an oscilloscope, displaying the waveform produced by the circuit in response to each input pulse. The question we want to answer is: *how is the shape of the output pulse related to the characteristics of the input pulse*? To simplify the investigation, we will only use input pulses that are much shorter than the output. For instance, if the system responds in milliseconds, we might use input pulses only a few microseconds in length.

After taking many measurement, we come to three conclusions: First, the *shape* of the input pulse does not affect the shape of the output signal. This is illustrated in Fig. 13-1, where various shapes of short input pulses produce exactly the same shape of output pulse. Second, the shape of the output waveform is totally determined by the characteristics of the system, i.e., the value and configuration of the resistors, capacitors and inductors. Third, the *amplitude* of the output pulse is directly proportional to the *area* of the input pulse. For example, the output will have the same amplitude for inputs of: 1 volt for 1 microsecond, 10 volts for 0.1 microseconds, 1,000 volts for 1 nanosecond, etc. This relationship also allows for input pulses with *negative* areas. For instance, imagine the combination of a 2 volt pulse lasting 2 microseconds being quickly followed by a -1 volt pulse lasting 4 microseconds. The total area of the input signal is *zero*, resulting in the output doing *nothing*.

An action potential is an impulse from the perspective of GCaMP, but not a genetically-encoded voltage indicator.

Input signals that are brief enough to have these three properties are called **impulses**. In other words, an impulse is any signal that is entirely zero except for a short *blip* of arbitrary shape. For example, an impulse to a microwave transmitter may have to be in the *picosecond* range because the electronics responds in *nanoseconds*. In comparison, a volcano that erupts for *years* may be a perfectly good impulse to geological changes that take *millennia*.

Mathematicians don't like to be limited by any particular system, and commonly use the term *impulse* to mean a signal that is short enough to be an impulse to *any possible* system. That is, a signal that is *infinitesimally* narrow. The **continuous delta function** is a normalized version of this type of impulse. Specifically, the continuous delta function is mathematically defined by three idealized characteristics: (1) the signal must be infinitesimally brief, (2) the pulse must occur at time zero, and (3) the pulse must have an area of one.

Since the delta function is defined to be infinitesimally narrow *and* have a fixed area, the amplitude is implied to be *infinite*. Don't let this bother you; it is completely unimportant. Since the amplitude is part of the *shape* of the impulse, you will never encounter a problem where the amplitude makes any difference, infinite or not. The delta function is a mathematical construct, not a real world signal. Signals in the real world that *act* as delta functions will always have a finite duration and amplitude.

Just as in the discrete case, the continuous delta function is given the mathematical symbol: $\delta(\ )$. Likewise, the output of a continuous system in response to a delta function is called the **impulse response**, and is often denoted by: $h(\ )$. Notice that parentheses, ( ), are used to denote continuous signals, as compared to brackets, [ ], for discrete signals. This notation is used in this book and elsewhere in DSP, but isn't universal. Impulses are displayed in graphs as vertical arrows (see Fig. 13-1d), with the *length* of the arrow indicating the *area* of the impulse.

To better understand real world impulses, look into the night sky at a *planet* and a *star*, for instance, Mars and Sirius. Both appear about the same brightness and size to the unaided eye. The reason for this similarity is not

This could be an illustration of different light pulses being delivered to ChR2 (except ChR2's impulse response function has a monoexponential decay).

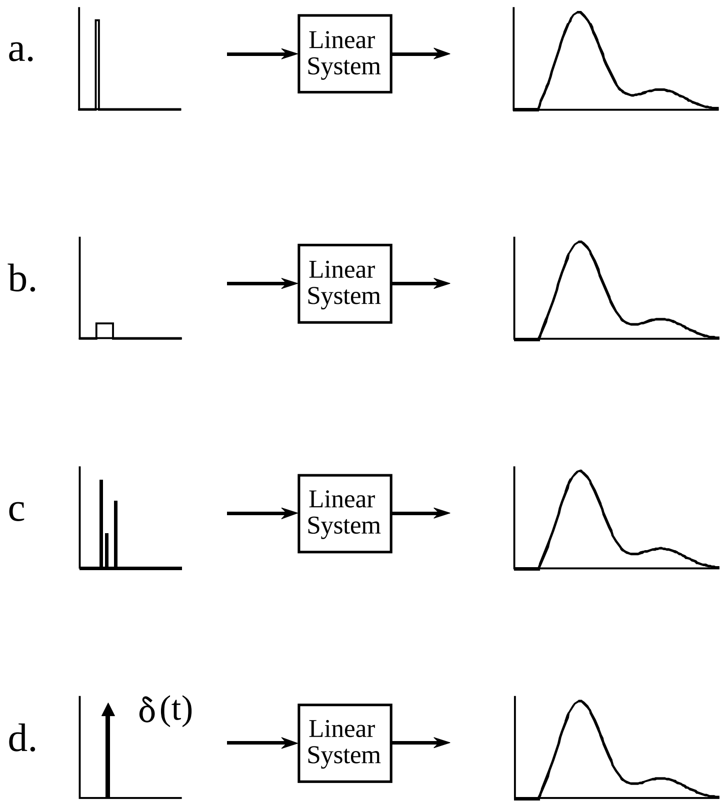


FIGURE 13-1
The continuous delta function. If the input to a linear system is brief compared to the resulting output, the shape of the output depends only on the characteristics of the system, and not the shape of the input. Such short input signals are called *impulses*. Figures a,b & c illustrate example input signals that are impulses for this particular system. The term *delta function* is used to describe a normalized impulse, i.e., one that occurs at $t = 0$ and has an area of one. The mathematical symbols for the delta function are shown in (d), a vertical arrow and $\delta(t)$.

obvious, since the viewing geometry is drastically different. Mars is about 6000 kilometers in diameter and 60 million kilometers from earth. In comparison, Sirius is about 300 times larger and over one-million times farther away. These dimensions should make Mars appear more than *three-thousand* times larger than Sirius. How is it possible that they look alike?

These objects look the same because they are small enough to be *impulses* to the human visual system. The perceived shape is the impulse response of the eye, not the actual image of the star or planet. This becomes obvious when the two objects are viewed through a small telescope; Mars appears as a dim disk, while Sirius still appears as a bright impulse. This is also the reason that stars twinkle while planets do not. The image of a star is small enough that it can be briefly blocked by particles or turbulence in the atmosphere, whereas the larger image of the planet is much less affected.

Convolution

Just as with discrete signals, the convolution of continuous signals can be viewed from the *input signal*, or the *output signal*. The input side viewpoint is the best *conceptual* description of how convolution operates. In comparison, the output side viewpoint describes the *mathematics* that must be used. These descriptions are virtually identical to those presented in Chapter 6 for discrete signals.

Figure 13-2 shows how convolution is viewed from the input side. An input signal, $x(t)$, is passed through a system characterized by an impulse response, $h(t)$, to produce an output signal, $y(t)$. This can be written in the familiar mathematical equation, $y(t) = x(t) * h(t)$. The input signal is divided into narrow columns, each short enough to act as an *impulse* to the system. In other words, the input signal is decomposed into an infinite number of scaled and shifted delta functions. Each of these impulses produces a scaled and shifted version of the impulse response in the output signal. The final output signal is then equal to the combined effect, i.e., the sum of all of the individual responses.

For this scheme to work, the width of the columns must be much shorter than the response of the system. Of course, mathematicians take this to the extreme by making the input segments *infinitesimally* narrow, turning the situation into a calculus problem. In this manner, the input viewpoint describes how a single point (or narrow region) in the input signal affects a larger portion of output signal.

In comparison, the output viewpoint examines how a single point in the output signal is determined by the various values from the input signal. Just as with discrete signals, each instantaneous value in the output signal is affected by a section of the input signal, weighted by the impulse response flipped left-for-right. In the discrete case, the signals are multiplied and *summed*. In the continuous case, the signals are multiplied and *integrated*. In equation form:

EQUATION 13-1
The convolution integral. This equation defines the meaning of: $y(t) = x(t) * h(t)$.

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

This equation is called the convolution integral, and is the twin of the convolution sum (Eq. 6-1) used with discrete signals. Figure 13-3 shows how this equation can be understood. The goal is to find an expression for calculating the value of the output signal at an arbitrary time, *t*. The first step is to change the independent variable used to move through the input signal and the impulse response. That is, we replace *t* with τ (a lower case

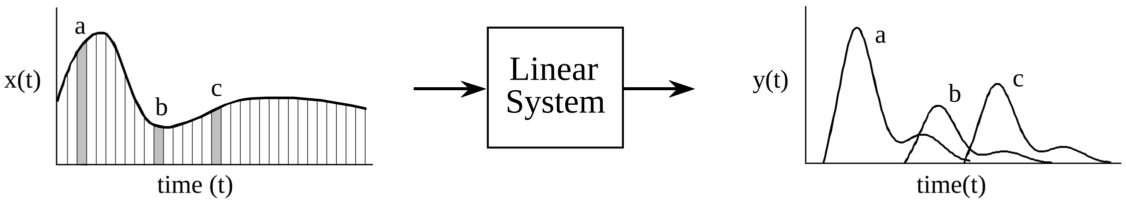


FIGURE 13-2
Convolution viewed from the input side. The input signal, $x(t)$, is divided into narrow segments, each acting as an impulse to the system. The output signal, $y(t)$, is the sum of the resulting scaled and shifted impulse responses. This illustration shows how three points in the input signal contribute to the output signal.

Greek tau). This makes $x(t)$ and $h(t)$ become $x(\tau)$ and $h(\tau)$, respectively. This change of variable names is needed because t is already being used to represent the point in the output signal being calculated. The next step is to flip the impulse response left-for-right, turning it into $h(-\tau)$. Shifting the flipped impulse response to the location t, results in the expression becoming $h(t-\tau)$. The input signal is then weighted by the flipped and shifted impulse response by multiplying the two, i.e., $x(\tau)\, h(t-\tau)$. The value of the output signal is then found by integrating this weighted input signal from negative to positive infinity, as described by Eq. 13-1.

If you have trouble understanding how this works, go back and review the same concepts for discrete signals in Chapter 6. Figure 13-3 is just another way of describing the convolution machine in Fig. 6-8. The only difference is that integrals are being used instead of summations. Treat this as an extension of what you already know, not something new.

An example will illustrate how continuous convolution is used in real world problems and the mathematics required. Figure 13-4 shows a simple continuous linear system: an electronic low-pass filter composed of a single resistor and a single capacitor. As shown in the figure, an impulse entering this system produces an output that quickly jumps to some value, and then exponentially decays toward zero. In other words, the impulse response of this simple electronic circuit is a *one-sided exponential*. Mathematically, the

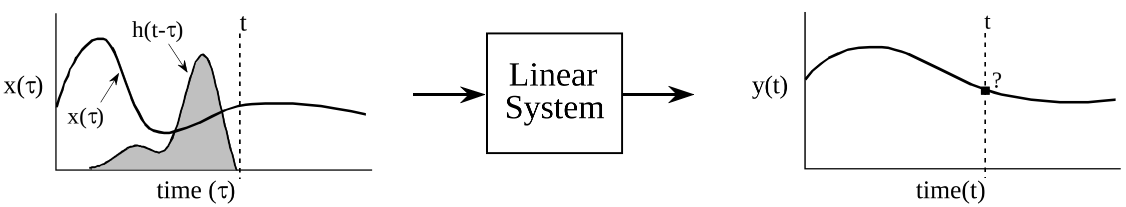


FIGURE 13-3
Convolution viewed from the output side. Each value in the output signal is influenced by many points from the input signal. In this figure, the output signal at time t is being calculated. The input signal, $x(\tau)$, is *weighted* (multiplied) by the flipped and shifted impulse response, given by $h(t-\tau)$. Integrating the weighted input signal produces the value of the output point, $y(t)$

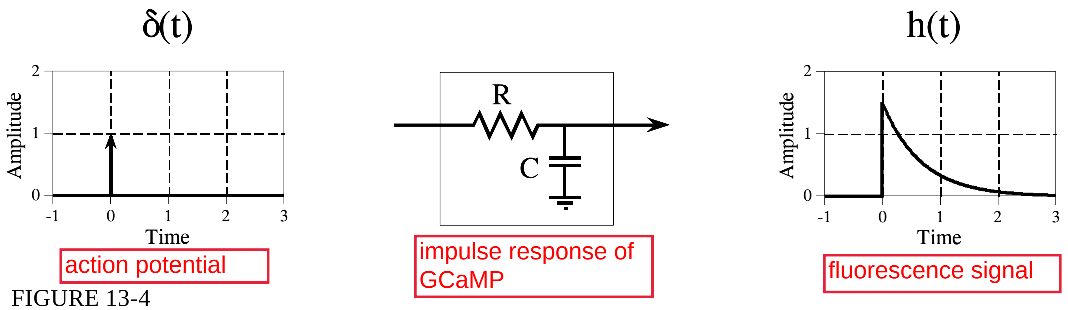


FIGURE 13-4
Example of a continuous linear system. This electronic circuit is a low-pass filter composed of a single resistor and capacitor. The impulse response of this system is a one-sided exponential.

impulse response of this system is broken into two sections, each represented by an equation:

$$h(t) = 0 \qquad \text{for } t < 0$$

$$h(t) = \alpha e^{-\alpha t} \qquad \text{for } t \geq 0$$

where $\alpha = 1/RC$ (R is in ohms, C is in farads, and t is in seconds). Just as in the discrete case, the continuous impulse response contains complete information about the system, that is, how it will react to all possible signals. To pursue this example further, Fig. 13-5 shows a square pulse entering the system, mathematically expressed by:

$$x(t) = 1 \qquad \text{for } 0 \leq t \leq 1$$

$$x(t) = 0 \qquad \text{otherwise}$$

Since both the input signal and the impulse response are completely known as mathematical expressions, the output signal, $y(t)$, can be calculated by evaluating the convolution integral of Eq. 13-1. This is complicated by the fact that both signals are defined by *regions* rather than a single

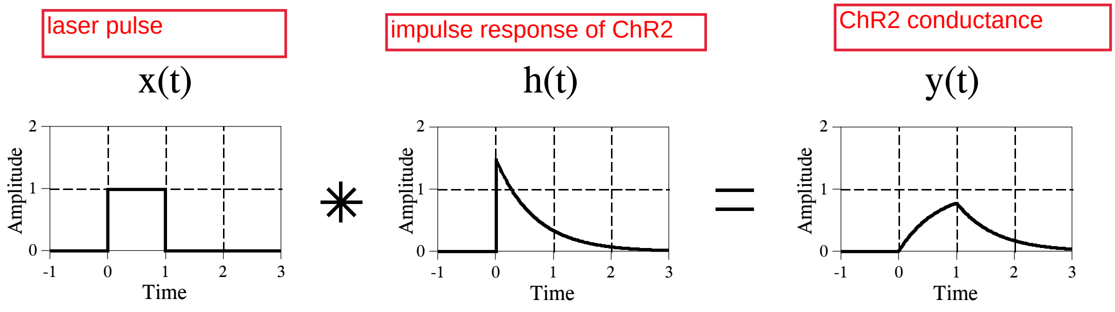


FIGURE 13-5
Example of continuous convolution. This figure illustrates a square pulse entering an RC low-pass filter (Fig. 13-4). The square pulse is convolved with the system's impulse response to produce the output.

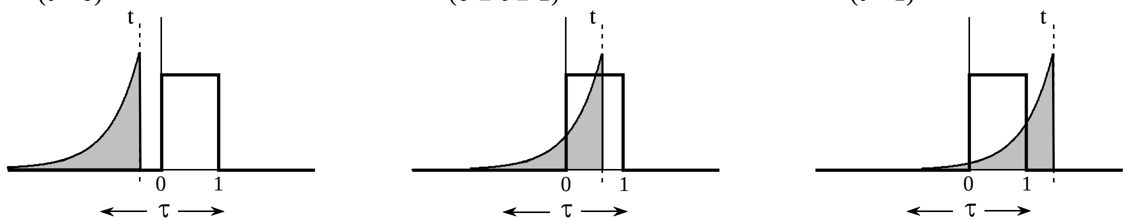


FIGURE 13-6
Calculating a convolution by segments. Since many continuous signals are defined by *regions*, the convolution calculation must be performed region-by-region. In this example, calculation of the output signal is broken into three sections: (a) no overlap, (b) partial overlap, and (c) total overlap, of the input signal and the shifted-flipped impulse response.

mathematical expression. This is very common in continuous signal processing. It is usually essential to draw a picture of how the two signals shift over each other for various values of t. In this example, Fig. 13-6a shows that the two signals do not overlap at all for $t < 0$. This means that the product of the two signals is zero at all locations along the τ axis, and the resulting output signal is:

$$y(t) = 0 \qquad \text{for } t < 0$$

A second case is illustrated in (b), where t is between 0 and 1. Here the two signals partially overlap, resulting in their product having nonzero values between $\tau = 0$ and $\tau = t$. Since this is the only nonzero region, it is the only section where the integral needs to be evaluated. This provides the output signal for $0 \le t \le 1$, given by:

$$y(t) = \int_{-\infty}^{\infty} x(\tau)\, h(t-\tau)\, d\tau \qquad \text{(start with Eq. 13-1)}$$

$$y(t) = \int_{0}^{t} 1 \cdot \alpha e^{-\alpha(t-\tau)}\, d\tau \qquad \text{(plug in the signals)}$$

$$y(t) = e^{-\alpha t}\, [\, e^{\alpha\tau} \,] \Big|_{0}^{t} \qquad \text{(evaluate the integral)}$$

$$y(t) = e^{-\alpha t}\, [\, e^{\alpha t} - 1 \,] \qquad \text{(reduce)}$$

$$y(t) = 1 - e^{-\alpha t} \qquad \text{for } 0 \le t \le 1$$

Figure (c) shows the calculation for the third section of the output signal, where t > 1. Here the overlap occurs between $\tau = 0$ and $\tau = 1$, making the calculation the same as for the second segment, except a change to the limits of integration:

$$y(t) = \int_0^1 1 \cdot \alpha e^{-\alpha(t-\tau)} \, d\tau \qquad \text{(plug into Eq. 13-1)}$$

$$y(t) = e^{-\alpha t} \left[e^{\alpha\tau} \right] \Big|_0^1 \qquad \text{(evaluate the integral)}$$

$$y(t) = \left[e^{\alpha} - 1 \right] e^{-\alpha t} \qquad \text{for } t > 1$$

The waveform in each of these three segments should agree with your knowledge of electronics: (1) The output signal must be zero until the input signal becomes nonzero. That is, the first segment is given by $y(t) = 0$ for $t < 0$. (2) When the step occurs, the RC circuit exponentially increases to match the input, according to the equation: $y(t) = 1 - e^{-\alpha t}$. (3) When the input is returned to zero, the output exponentially decays toward zero, given by the equation: $y(t) = ke^{-\alpha t}$ (where $k = e^{\alpha} - 1$, the voltage on the capacitor just before the discharge was started).

More intricate waveforms can be handled in the same way, although the mathematical complexity can rapidly become unmanageable. When faced with a nasty continuous convolution problem, you need to spend significant time evaluating *strategies* for solving the problem. If you start blindly evaluating integrals you are likely to end up with a mathematical mess. A common strategy is to break one of the signals into simpler additive components that can be *individually* convolved. Using the principles of linearity, the resulting waveforms can be added to find the answer to the original problem.

Figure 13-7 shows another strategy: modify one of the signals in some linear way, perform the convolution, and then undo the original modification. In this example the modification is the *derivative*, and it is undone by taking the *integral*. The derivative of a unit amplitude square pulse is two *impulses*, the first with an area of one, and the second with an area of negative one. To understand this, think about the opposite process of taking the integral of the two impulses. As you integrate past the first impulse, the integral rapidly increases from zero to one, i.e., a step function. After passing the negative impulse, the integral of the signal rapidly returns from one back to zero, completing the square pulse.

Taking the derivative simplifies this problem because convolution is easy when one of the signals is composed of impulses. Each of the two impulses in $x'(t)$ contributes a scaled and shifted version of the impulse response to

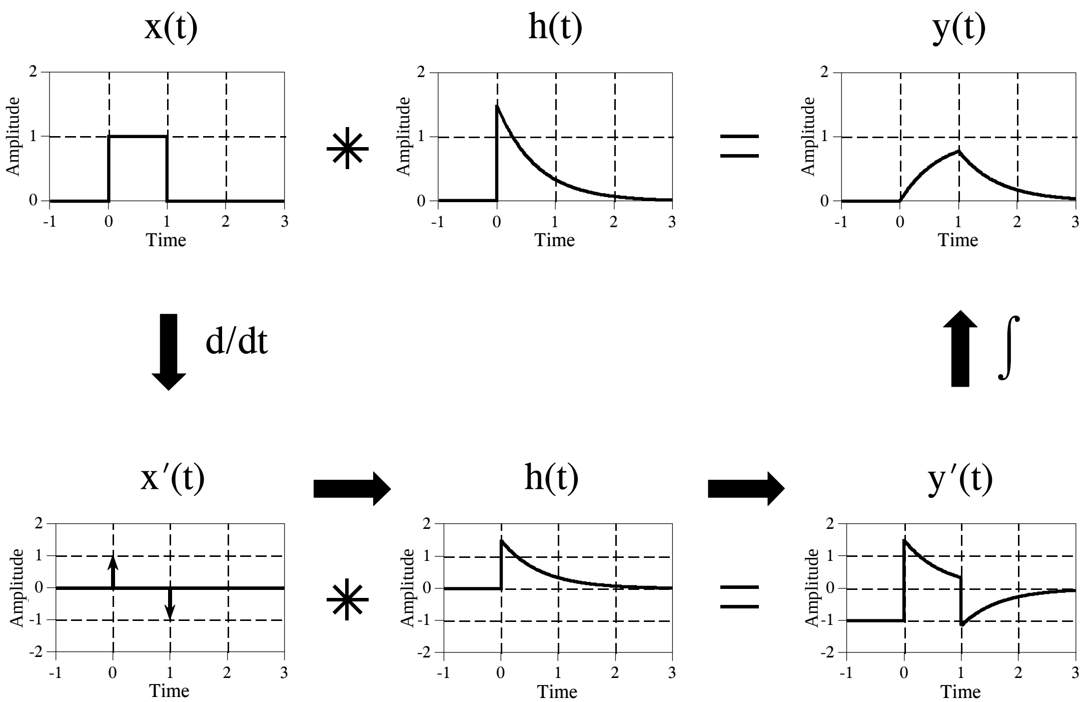


FIGURE 13-7
A strategy for convolving signals. Convolution problems can often be simplified by clever use of the rules governing linear systems. In this example, the convolution of two signals is simplified by taking the derivative of one of them. After performing the convolution, the derivative is undone by taking the integral.

the derivative of the output signal, $y'(t)$. That is, by inspection it is known that: $y'(t) = h(t) - h(t-1)$. The output signal, $y(t)$, can then be found by plugging in the exact equation for $h(t)$, and integrating the expression.

A slight nuisance in this procedure is that the DC value of the input signal is lost when the derivative is taken. This can result in an error in the DC value of the calculated output signal. The mathematics reflects this as the arbitrary constant that can be added during the integration. There is no systematic way of identifying this error, but it can usually be corrected by inspection of the problem. For instance, there is no DC error in the example of Fig. 13-7. This is known because the calculated output signal has the correct DC value when t becomes very large. If an error is present in a particular problem, an appropriate DC term is manually added to the output signal to complete the calculation.

This method also works for signals that can be reduced to impulses by taking the derivative *multiple* times. In the jargon of the field, these signals are called *piecewise polynomials*. After the convolution, the initial operation of multiple derivatives is undone by taking multiple integrals. The only catch is that the lost DC value must be found at each stage by finding the correct constant of integration.

Before starting a difficult continuous convolution problem, there is another approach that you should consider. Ask yourself the question: *Is a mathematical expression really needed for the output signal, or is a graph of the waveform sufficient?* If a graph is adequate, you may be better off to handle the problem with *discrete* techniques. That is, approximate the continuous signals by samples that can be directly convolved by a computer program. While not as mathematically pure, it can be much easier.

The Fourier Transform

Fourier transforms are very important, but we won't have time to cover them in this course. This part is not required, but I recommend reading it.

The Fourier Transform for continuous signals is divided into two categories, one for signals that are *periodic*, and one for signals that are *aperiodic*. Periodic signals use a version of the Fourier Transform called the **Fourier Series**, and are discussed in the next section. The Fourier Transform used with aperiodic signals is simply called the **Fourier Transform**. This chapter describes these Fourier techniques using only *real* mathematics, just as the last several chapters have done for discrete signals. The more powerful use of *complex* mathematics will be reserved for Chapter 31.

Figure 13-8 shows an example of a continuous aperiodic signal and its frequency spectrum. The time domain signal extends from negative infinity to positive infinity, while each of the frequency domain signals extends from zero to positive infinity. This frequency spectrum is shown in rectangular form (real and imaginary parts); however, the polar form (magnitude and phase) is also used with continuous signals. Just as in the discrete case, the **synthesis equation** describes a recipe for constructing the time domain signal using the data in the frequency domain. In mathematical form:

$$x(t) = \frac{1}{\pi} \int_{0}^{+\infty} Re\,X(\omega) \cos(\omega t) - Im\,X(\omega) \sin(\omega t) \, d\omega$$

EQUATION 13-2
The Fourier transform synthesis equation. In this equation, $x(t)$ is the time domain signal being synthesized, and $Re\,X(\omega)$ & $Im\,X(\omega)$ are the real and imaginary parts of the frequency spectrum, respectively.

In words, the time domain signal is formed by adding (with the use of an integral) an infinite number of scaled sine and cosine waves. The real part of the frequency domain consists of the scaling factors for the cosine waves, while the imaginary part consists of the scaling factors for the sine waves. Just as with discrete signals, the synthesis equation is usually written with *negative* sine waves. Although the negative sign has no significance in this discussion, it is necessary to make the notation compatible with the complex mathematics described in Chapter 29. The key point to remember is that some authors put this negative sign in the equation, while others do not. Also notice that frequency is represented by the symbol, ω, a lower case

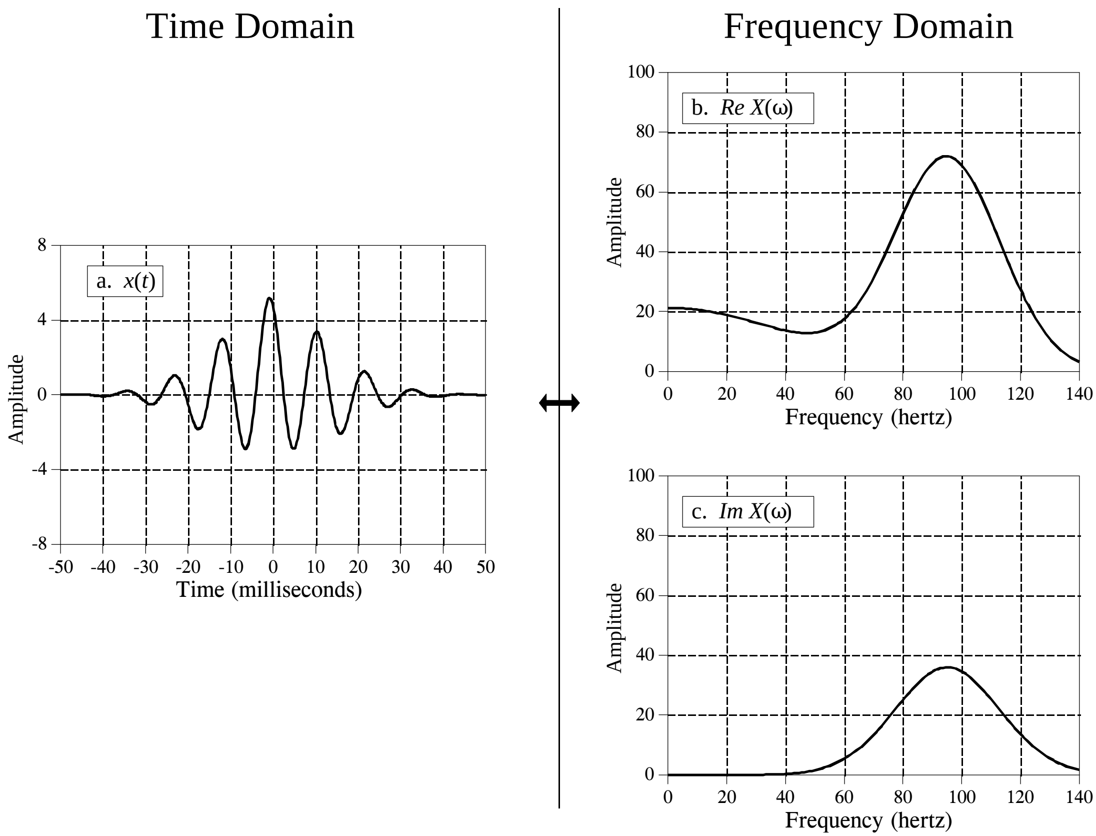


FIGURE 13-8
Example of the Fourier Transform. The time domain signal, $x(t)$, extends from negative to positive infinity. The frequency domain is composed of a real part, $Re\,X(\omega)$, and an imaginary part, $Im\,X(\omega)$, each extending from zero to positive infinity. The frequency axis in this illustration is labeled in *cycles per second* (hertz). To convert to natural frequency, multiply the numbers on the frequency axis by 2π.

Greek omega. As you recall, this notation is called the **natural frequency**, and has the units of radians per second. That is, $\omega = 2\pi f$, where f is the frequency in cycles per second (hertz). The natural frequency notation is favored by mathematicians and others doing signal processing by *solving equations*, because there are usually fewer symbols to write.

The **analysis equations** for continuous signals follow the same strategy as the discrete case: *correlation* with sine and cosine waves. The equations are:

EQUATION 13-3
The Fourier transform analysis equations. In this equation, $Re\,X(\omega)$ & $Im\,X(\omega)$ are the real and imaginary parts of the frequency spectrum, respectively, and $x(t)$ is the time domain signal being analyzed.

$$Re\,X(\omega) = \int_{-\infty}^{+\infty} x(t)\,\cos(\omega t)\,dt$$

$$Im\,X(\omega) = -\int_{-\infty}^{+\infty} x(t)\,\sin(\omega t)\,dt$$

As an example of using the analysis equations, we will find the frequency response of the RC low-pass filter. This is done by taking the Fourier transform of its impulse response, previously shown in Fig. 13-4, and described by:

$$h(t) = 0 \qquad \text{for } t < 0$$

$$h(t) = \alpha e^{-\alpha t} \qquad \text{for } t \geq 0$$

The frequency response is found by plugging the impulse response into the analysis equations. First, the real part:

$$Re\,H(\omega) = \int_{-\infty}^{+\infty} h(t)\cos(\omega t)\,dt \qquad \text{(start with Eq. 13-3)}$$

$$Re\,H(\omega) = \int_{0}^{+\infty} \alpha e^{-\alpha t}\cos(\omega t)\,dt \qquad \text{(plug in the signal)}$$

$$Re\,H(\omega) = \frac{\alpha e^{-\alpha t}}{\alpha^2 + \omega^2}\left[-\alpha\cos(\omega t) + \omega\sin(\omega t)\right]\Big|_{0}^{+\infty} \qquad \text{(evaluate)}$$

$$Re\,H(\omega) = \frac{\alpha^2}{\alpha^2 + \omega^2}$$

Using this same approach, the imaginary part of the frequency response is calculated to be:

$$Im\,H(\omega) = \frac{-\omega\alpha}{\alpha^2 + \omega^2}$$

Just as with discrete signals, the rectangular representation of the frequency domain is great for mathematical manipulation, but difficult for human understanding. The situation can be remedied by converting into polar notation with the standard relations: $Mag\,H(\omega) = [Re\,H(\omega)^2 + Im\,H(\omega)^2]^{½}$ and $Phase\,H(\omega) = \arctan[Re\,H(\omega) / Im\,H(\omega)]$. Working through the algebra provides the frequency response of the RC low-pass filter as magnitude and phase (i.e., polar form):

$$Mag\ H(\omega) = \frac{\alpha}{[\alpha^2 + \omega^2]^{1/2}}$$

$$Phase\ H(\omega) = \arctan\left[-\frac{\omega}{\alpha}\right]$$

Figure 13-9 shows graphs of these curves for a cutoff frequency of 1000 hertz (i.e., $\alpha = 2\pi 1000$).

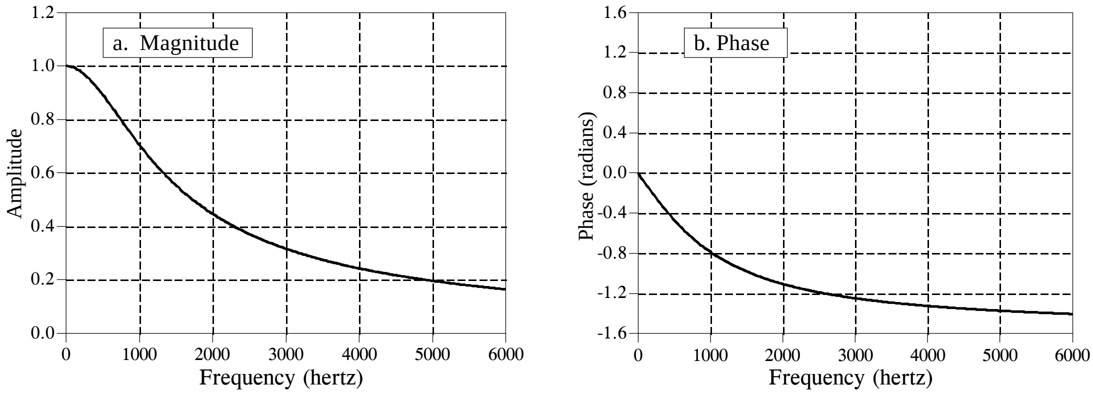


FIGURE 13-9
Frequency response of an RC low-pass filter. These curves were derived by calculating the Fourier transform of the impulse response, and then converting to polar form.

The Fourier Series

This brings us to the last member of the Fourier transform family: the *Fourier series*. The time domain signal used in the Fourier series is *periodic* and *continuous*. Figure 13-10 shows several examples of continuous waveforms that repeat themselves from negative to positive infinity. Chapter 11 showed that periodic signals have a frequency spectrum consisting of **harmonics**. For instance, if the time domain repeats at 1000 hertz (a period of 1 millisecond), the frequency spectrum will contain a first harmonic at 1000 hertz, a second harmonic at 2000 hertz, a third harmonic at 3000 hertz, and so forth. The first harmonic, i.e., the frequency that the time domain repeats itself, is also called the **fundamental frequency.** This means that the frequency spectrum can be viewed in two ways: (1) the frequency spectrum is *continuous*, but zero at all frequencies except the harmonics, or (2) the frequency spectrum is *discrete*, and only *defined* at the harmonic frequencies. In other words, the frequencies between the harmonics can be thought of as having a value of zero, or simply not existing. The important point is that they do not contribute to forming the time domain signal.

The Fourier series **synthesis equation** creates a continuous periodic signal with a fundamental frequency, f, by adding scaled cosine and sine waves with frequencies: f, $2f$, $3f$, $4f$, etc. The amplitudes of the cosine waves are held in the variables: a_1, a_2, a_3, a_4, etc., while the amplitudes of the sine waves are held in: b_1, b_2, b_3, b_4, and so on. In other words, the "a" and "b" coefficients are the real and imaginary parts of the frequency spectrum, respectively. In addition, the coefficient a_0 is used to hold the DC value of the time domain waveform. This can be viewed as the amplitude of a cosine wave with zero frequency (a constant value). Sometimes a_0 is grouped with the other "a" coefficients, but it is often handled separately because it requires special calculations. There is no b_0 coefficient since a sine wave of zero frequency has a constant value of zero, and would be quite useless. The synthesis equation is written:

$$x(t) = a_0 + \sum_{n=1}^{\infty} a_n \cos(2\pi f t n) - \sum_{n=1}^{\infty} b_n \sin(2\pi f t n)$$

EQUATION 13-4
The Fourier series synthesis equation. Any periodic signal, $x(t)$, can be reconstructed from sine and cosine waves with frequencies that are multiples of the fundamental, f. The a_n and b_n coefficients hold the amplitudes of the cosine and sine waves, respectively.

The corresponding **analysis equations** for the Fourier series are usually written in terms of the *period* of the waveform, denoted by T, rather than the fundamental frequency, f (where $f = 1/T$). Since the time domain signal is periodic, the sine and cosine wave correlation only needs to be evaluated over a single period, i.e., $-T/2$ to $T/2$, 0 to T, $-T$ to 0, etc. Selecting different limits makes the mathematics different, but the final answer is always the same. The Fourier series analysis equations are:

$$a_0 = \frac{1}{T} \int_{-T/2}^{T/2} x(t)\, dt \qquad a_n = \frac{2}{T} \int_{-T/2}^{T/2} x(t) \cos\left(\frac{2\pi t n}{T}\right) dt$$

$$b_n = \frac{-2}{T} \int_{-T/2}^{T/2} x(t) \sin\left(\frac{2\pi t n}{T}\right) dt$$

EQUATION 13-5
Fourier series analysis equations. In these equations, $x(t)$ is the time domain signal being decomposed, a_0 is the DC component, a_n & b_n hold the amplitudes of the cosine and sine waves, respectively, and T is the period of the signal, i.e., the reciprocal of the fundamental frequency.

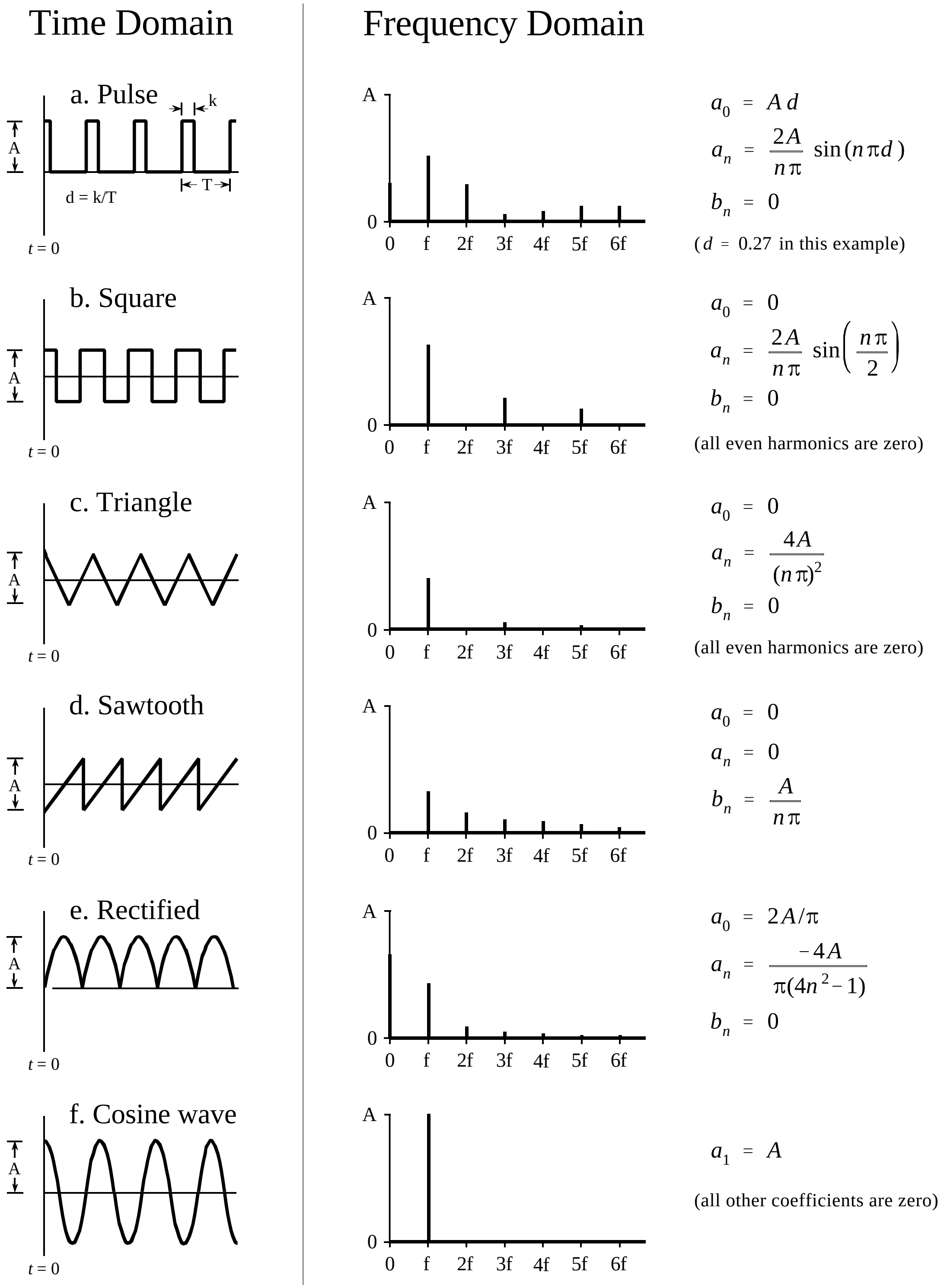


FIGURE 13-10
Examples of the Fourier series. Six common time domain waveforms are shown, along with the equations to calculate their "*a*" and "*b*" coefficients.

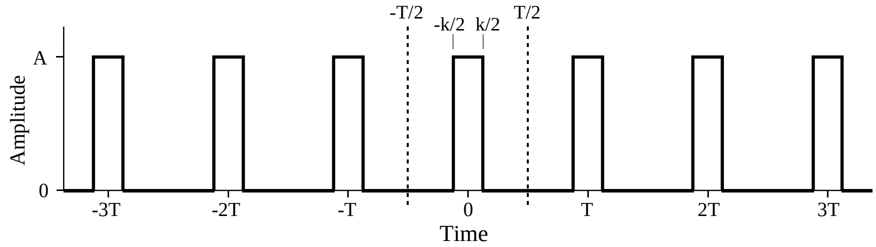


FIGURE 13-11
Example of calculating a Fourier series. This is a pulse train with a duty cycle of $d = k/T$. The Fourier series coefficients are calculated by correlating the waveform with cosine and sine waves over any full period. In this example, the period from $-T/2$ to $T/2$ is used.

Figure 13-11 shows an example of calculating a Fourier series using these equations. The time domain signal being analyzed is a *pulse train*, a square wave with unequal high and low durations. Over a single period from $-T/2$ to $T/2$, the waveform is given by:

$$x(t) = A \qquad \text{for } -k/2 \le t \le k/2$$

$$x(t) = 0 \qquad \text{otherwise}$$

The *duty cycle* of the waveform (the fraction of time that the pulse is "high") is thus given by $d = k/T$. The Fourier series coefficients can be found by evaluating Eq. 13-5. First, we will find the DC component, a_0:

$$a_0 = \frac{1}{T}\int_{-T/2}^{T/2} x(t)\,dt \qquad \text{(start with Eq. 13-5)}$$

$$a_0 = \frac{1}{T}\int_{-k/2}^{k/2} A\,dt \qquad \text{(plug in the signal)}$$

$$a_0 = \frac{A\,k}{T} \qquad \text{(evaluate the integral)}$$

$$a_0 = A\,d \qquad \text{(substitute: } d = k/T\text{)}$$

This result should make intuitive sense; the DC component is simply the average value of the signal. A similar analysis provides the "*a*" coefficients:

$$a_n = \frac{2}{T} \int_{-T/2}^{T/2} x(t) \cos\left(\frac{2\pi t n}{T}\right) dt \qquad \text{(start with Eq. 13-4)}$$

$$a_n = \frac{2}{T} \int_{-k/2}^{k/2} A \cos\left(\frac{2\pi t n}{T}\right) dt \qquad \text{(plug in the signal)}$$

$$a_n = \frac{2A}{T} \left[\frac{T}{2\pi n} \sin\left(\frac{2\pi t n}{T}\right)\right] \Bigg|_{-k/2}^{k/2} \qquad \text{(evaluate the integral)}$$

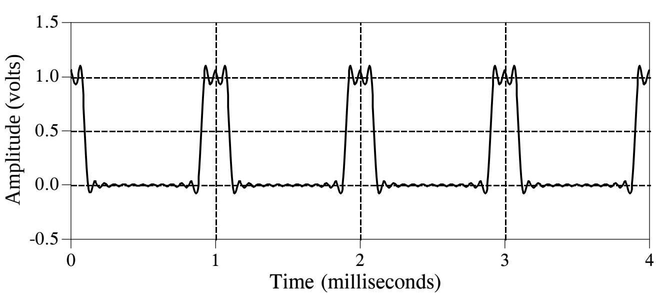
$$a_n = \frac{2A}{n\pi} \sin(\pi n d) \qquad \text{(reduce)}$$

The "*b*" coefficients are calculated in this same way; however, they all turn out to be *zero*. In other words, this waveform can be constructed using only cosine waves, with no sine waves being needed.

The "*a*" and "*b*" coefficients will change if the time domain waveform is shifted left or right. For instance, the "*b*" coefficients in this example will be zero *only* if one of the pulses is centered on $t = 0$. Think about it this way. If the waveform is *even* (i.e., symmetrical around $t = 0$), it will be composed solely of *even* sinusoids, that is, cosine waves. This makes all of the "*b*" coefficients equal to zero. If the waveform if *odd* (i.e., symmetrical but opposite in sign around $t = 0$), it will be composed of *odd* sinusoids, i.e., sine waves. This results in the "*a*" coefficients being zero. If the coefficients are converted to polar notation (say, M_n and θ_n coefficients), a shift in the time domain leaves the magnitude unchanged, but adds a linear component to the phase.

To complete this example, imagine a pulse train existing in an electronic circuit, with a frequency of 1 kHz, an amplitude of one volt, and a duty cycle of 0.2. The table in Fig. 13-12 provides the amplitude of each harmonic contained in this waveform. Figure 13-12 also shows the synthesis of the waveform using only the *first fourteen* of these harmonics. Even with this number of harmonics, the reconstruction is not very good. In mathematical jargon, the Fourier series *converges* very *slowly*. This is just another way of saying that sharp edges in the time domain waveform results in very high frequencies in the spectrum. Lastly, be sure and notice the overshoot at the sharp edges, i.e., the Gibbs effect discussed in Chapter 11.

An important application of the Fourier series is electronic **frequency multiplication**. Suppose you want to construct a very stable sine wave oscillator at 150 MHz. This might be needed, for example, in a radio transmitter operating at this frequency. High stability calls for the circuit to be *crystal controlled*. That is, the frequency of the oscillator is determined by a resonating quartz crystal that is a part of the circuit. The problem is, quartz crystals only work to about 10 MHz. The solution is to build a crystal controlled oscillator operating somewhere between 1 and 10 MHz, and then *multiply* the frequency to whatever you need. This is accomplished by *distorting* the sine wave, such as by clipping the peaks with a diode, or running the waveform through a squaring circuit. The harmonics in the distorted waveform are then isolated with band-pass filters. This allows the frequency to be doubled, tripled, or multiplied by even higher integers numbers. The most common technique is to use sequential stages of doublers and triplers to generate the required frequency multiplication, rather than just a single stage. The Fourier series is important to this type of design because it describes the *amplitude* of the multiplied signal, depending on the type of distortion and harmonic selected.



frequency	amplitude (volts)
DC	0.20000
1 kHz	0.37420
2 kHz	0.30273
3 kHz	0.20182
4 kHz	0.09355
5 kHz	0.00000
6 kHz	-0.06237
7 kHz	-0.08649
8 kHz	-0.07568
9 kHz	-0.04158
10 kHz	0.00000
11 kHz	0.03402
12 kHz	0.05046
⋮	
123 kHz	0.00492
124 kHz	0.00302
125 kHz	0.00000
126 kHz	-0.00297
⋮	
803 kHz	0.00075
804 kHz	0.00046
805 kHz	0.00000
806 kHz	-0.00046

FIGURE 13-12
Example of Fourier series synthesis. The waveform being constructed is a pulse train at 1 kHz, an amplitude of one volt, and a duty cycle of 0.2 (as illustrated in Fig. 13-11). This table shows the amplitude of the harmonics, while the graph shows the reconstructed waveform using only the first fourteen harmonics.