Digital Signal Processing: A Computer Science Perspective
Jonathan Y. Stein

Print ISBN 0-471-29546-9 Online ISBN 0-471-20059-X

3

The Spectrum of Periodic Signals

Signals dwell both in the time and frequency domains; we can equally accurately think of them as values changing in time (time domain), or as blendings of fundamental frequencies (spectral domain). The method for determining these fundamental frequencies from the time variations is called *Fourier* or *spectral* analysis. Similar techniques allow returning to the time domain representation from the frequency domain description.

It is hard to believe that 300 years ago the very idea of *spectrum* didn't even exist, that less than 200 years ago the basic mechanism for its calculation was still controversial, and that as recently as 1965 the algorithm that made its digital computation practical almost went unpublished due to lack of interest. Fourier analysis is used so widely today that even passing mention of its most important applications is a lengthy endeavor. Fourier analysis is used in quantum physics to uncover the structure of matter on the smallest of scales, and in cosmology to study the universe as a whole. Spectroscopy and X-ray crystallography rely on Fourier analysis to analyze the chemical composition and physical structure from minute quantities of materials, and spectral analysis of light from stars tells us of the composition and temperature of bodies separated from us by light years. Engineers routinely compute Fourier transforms in the analysis of mechanical vibrations, in the acoustical design of concert halls, and in the building of aircraft and bridges. In medicine Fourier techniques are called upon to reconstruct body organs from CAT scans and MRI, to detect heart malfunctions and sleep disorders. Watson and Crick discovered the double-helix nature of DNA from data obtained using Fourier analysis. Fourier techniques can help us differentiate musical instruments made by masters from inferior copies, can assist in bringing back to life deteriorated audio recordings of great vocalists, and can help in verifying a speaker's true identity.

In this chapter we focus on the concepts of spectrum and frequency, but only for periodic signals where they are easiest to grasp. We feel that several brief historical accounts will assist in placing the basic ideas in proper context. We derive the Fourier series (FS) of a periodic signal, find the FS for various signals, and see how it can be utilized in radar signal processing. We briefly discuss its convergence and properties, as well as its major drawback, the Gibbs phenomenon. We also introduce a new notation that uses complex numbers and negative frequencies, in order to set the stage for the use of Fourier techniques in the analysis of nonperiodic signals in the next chapter.

3.1 Newton's Discovery

Isaac Newton went over to the window and shuttered it, completely darkening the room. He returned to his lab bench, eager to get on with the experiment. Although he was completely sure of the outcome, he had been waiting to complete this experiment for a long time.

The year was 1669 and Newton had just taken over the prestigious Lucasian chair at Cambridge. He had decided that the first subject of his researches and lectures would be optics, postponing his further development of the theory of fluxions (which we now call the differential calculus). During the years 1665 and 1666 Newton had been forced to live at his family's farm in Lincolnshire for months at time, due to the College being closed on account of the plague. While at home he had worked out his theory of fluxions, but he had also done something else. He had perfected a new method of grinding lenses.

While working with these lenses he had found that when white light passed through lenses it always produced colors. He finally gave up on trying to eliminate this 'chromatic aberration' and concluded (incorrectly) that the only way to make a truly good telescope was with a parabolic mirror instead of a lens. He had just built what we now call a Newtonian reflector telescope proving his theory. However, he was not pleased with the theoretical aspects of the problem. He had managed to avoid the chromatic aberration, but had not yet explained the source of the problem. Where did the colors come from?

His own theory was that white light was actually composed of all possible colors mixed together. The lenses were not creating the colors, they were simply decomposing the light into its constituents. His critics on this matter were many, and he could not risk publishing this result without iron clad proof; and this present experiment would vindicate his ideas.

He looked over the experimental setup. There were two prisms, one to break the white light into its constituent colors, and one that would hopefully combine those colors back into white light again. He had worked hard in polishing these prisms, knowing that if the experiment failed it would be because of imperfections in the glass. He carefully lit up his light source and positioned the prisms. After a little experimentation he saw what he had expected; in between the prisms was a rainbow of colors, but after the second prism the light was perfectly white. He tried blocking off various colors and observed the recomposed light's color, putting back more and more colors until the light was white again. Yes, even his most vehement detractors at the Royal society would not be able to argue with this proof.

Newton realized that the white light had all the colors in it. He thought of these colors as *ghosts* which could not normally be seen, and in his Latin write-up he actually used the word *specter*. Later generations would adopt this word into other languages as *spectrum*, meaning all of the colors of the rainbow.

Newton's next step in understanding these components of white light should have been the realization that the different colors he observed corresponded to different frequencies of radiation. Unfortunately, Newton, the greatest scientist of his era, could not make that step, due to his firm belief that light was not composed of waves. His years of experimentation with lenses led him to refute such a wave theory as proposed by others, and to assert a corpuscular theory, that light was composed of small particles. Only in the twentieth century was more of the truth finally known; light is both waves and particles, combined in a way that seventeenth-century science could not have imagined. Thus, paradoxically, Newton discovered the spectrum of light, without being able to admit that *frequency* was involved.

EXERCISES

3.1.1 Each of the colors of the rainbow is characterized by a single frequency, while artists and computer screens combine *three* basic colors. Reconcile the one-dimensional *physical* concept of frequency with the three-dimensional *psychological* concept of color.

3.1.2 Wavepackets are *particle-like* waves, that is, waves that are localized in space. For example, you can create a wavepacket by multiplying a sine wave by a Gaussian

$$s(t) = e^{\frac{(x-\mu)^2}{2\sigma^2}} \sin(\omega t)$$

where μ is the approximate location. Plot the signal in space for a given time, and in time for a given location. What is the uncertainty in the location of the 'particle'? If one wishes the 'particle' to travel at a speed v, one can substitute $\mu = vt$. What happens to the space plot now? How accurately can the velocity be measured?

3.2 Frequency Components

Consider a simple analog sinusoid. This signal may represent monochromatic light (despite Newton's prejudices), or a single tone of sound, or a simple radio wave. This signal is obviously periodic, and its *basic period* T is the time it takes to complete one cycle. The reciprocal of the basic period, $f = \frac{1}{T}$, the number of cycles it completes in a second, is called the *frequency*. Periods are usually measured in seconds per cycle and frequencies in cycles per second, or Hertz (Hz). When the period is a millisecond the frequency is a kilohertz (KHz) and a microsecond leads to a megahertz (MHz).

Why did we need the qualifier *basic* in 'basic period'? Well, a signal which is periodic with basic period T, is necessarily also periodic with period $2T$, $3T$, and all other multiples of the basic period. All we need for periodicity with period P is for $s(t + P)$ to equal $s(t)$ for all t, and this is obviously the case for periods P which contain any whole number of cycles. Hence if a sinusoid of frequency f is periodic with period P, the sinusoid with double that frequency is also periodic with period P. In general, sinusoids with period nf (where n is any integer) will all be periodic with period P. Frequencies that are related in this fashion are called *harmonics*.

A pure sine is completely specified by its frequency (or basic period), its amplitude, and its phase at time $t = 0$. For more complex periodic signals the frequency alone does not completely specify the signal; one has to specify the content of each cycle as well. There are several ways of doing this. The most straightforward would seem to require full specification of the *waveform*, that is the values of the signal in the basic period. This is feasible for digital signals, while for analog signals this would require an infinite number of values to be specified. A more sophisticated way is to recognize that complex periodic signals have, in addition to their main frequency, many other component frequencies. Specification of the contributions of all these components determines the signal. This specification is called the signal's *spectrum*.

What do we mean by frequency components? Note the following facts.

- The multiplication of a periodic signal by a number, and the addition of a constant signal, do not affect the periodicity.
- Sinusoids with period nf (where n is any integer) are all periodic with period $P = \frac{1}{f}$. These are harmonics of the basic frequency sinusoid.
- The sum of any number of signals all of which are periodic with period T, is also periodic with the same period.

From all of these facts together we can conclude that a signal that results from weighted summing of sinusoidal signals with frequencies nf, and possibly addition of a constant signal, is itself periodic with period $P = \frac{1}{f}$. Such a trigonometric series is no longer sinusoidal, indeed it can look like just about anything, but it *is* periodic. You can think of the spectrum as a recipe for preparing an arbitrary signal; the frequencies needed are the ingredients, and the weights indicate how much of each ingredient is required.

The wealth of waveforms that can be created in this fashion can be demonstrated with a few examples. In Figure 3.1 we start with a simple sine, and progressively add harmonics, each with decreased amplitude (the sine of frequency kf having amplitude $\frac{1}{k}$). On the left side we see the harmonics themselves, while the partial sums of all harmonics up to that point appear on the right. It would seem that the sum tends to a periodic sawtooth signal,

$$\sum_{k=0}^{K} \frac{\sin(k\omega t)}{k} \xrightarrow{K\to\infty} -\mathcal{T}(t) \tag{3.1}$$

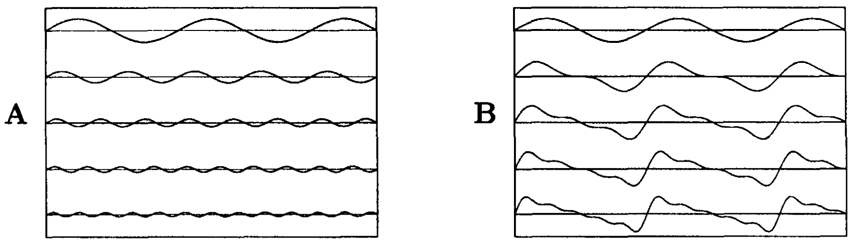

Figure 3.1: Building up a periodic sawtooth signal $-\mathcal{T}(t)$ from a sine and its harmonics. In (A) are the component sinusoids, and in (B) the composite signal.

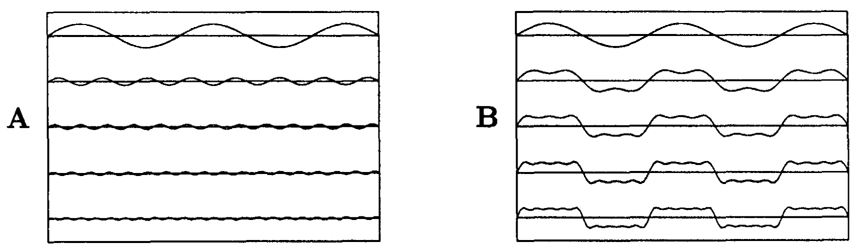

Figure 3.2: Building up a periodic square wave signal from a sine and its odd harmonics. In (A) are the component sinusoids, and in (B) the composite signal.

and this feeling is strengthened when the summation is carried out to higher harmonics. Surprisingly, when we repeat this feat with odd harmonics only we get a square wave

$$\sum_{k=0}^{K} \frac{1}{2k+1} \sin\Big((2k+1)\omega t\Big) \longrightarrow \square(t) \tag{3.2}$$

as can be seen in Figure 3.2.

The signal $f(t) = \sin(\omega t)$ is an *odd* function of t, that is $f(-t) = -f(t)$. Since the sum of odd functions is odd, all signals generated by summing only harmonically related sines will be odd as well. If our problem requires an *even* function, one for which $f(-t) = f(t)$, we could sum cosines in a similar way. In order to produce a signal that is neither odd nor even, we need to sum harmonically related sines *and* cosines, which from here on we shall call **H**armonically **R**elated **S**inusoids (HRSs). In this way we can produce a huge array of general periodic signals, since any combination of sines and cosines with frequencies all multiples of some basic frequency will be periodic with that frequency.

In fact, just about anything, as long as it *is* periodic, can be represented as a trigonometric series involving harmonically related sinusoids. Just about anything, as long as it *is* periodic, can be broken down into the weighted sum of sinusoidal signals with frequencies nf, and possibly a constant signal. When first discovered, this statement surprised even the greatest of mathematicians.

EXERCISES

3.2.1 In the text we considered summing all harmonics and all odd harmonics with amplitude decreasing as $\frac{1}{n}$. Why didn't we consider all even harmonics?

3.2.2 When two sinusoids with close frequencies are added *beats* with two observable frequencies result. Explain this in terms of the arguments of this section.

3.2.3 To what waveforms do the following converge?

1. $\frac{1}{2} - \frac{4}{\pi^2}\left(\frac{\cos(x)}{1^2} + \frac{\cos(3x)}{3^2} + \frac{\cos(5x)}{5^2} + \cdots\right)$
2. $\frac{2}{\pi} - \frac{4}{\pi}\left(\frac{\cos(2x)}{1\cdot 3} + \frac{\cos(4x)}{3\cdot 5} + \frac{\cos(6x)}{5\cdot 7} + \cdots\right)$
3. $\frac{1}{\pi} + \frac{1}{2}\sin(x) - \frac{2}{\pi}\left(\frac{\cos(2x)}{1\cdot 3} + \frac{\cos(4x)}{3\cdot 5} + \frac{\cos(6x)}{5\cdot 7} + \cdots\right)$
4. $\frac{1}{3} - \frac{4}{\pi^2}\left(\frac{\cos(x)}{1^2} - \frac{\cos(2x)}{2^2} + \frac{\cos(3x)}{3^2} + \cdots\right)$

3.3 Fourier's Discovery

The idea of constructing complex periodic functions by summing trigonometric functions is very old; indeed it is probable that the ancient Babylonians and Egyptians used it to predict astronomical events. In the mid-eighteenth century this idea engendered a great deal of excitement due to its possible application to the description of vibrating strings (such as violin strings). The great eighteenth-century Swiss mathematician Leonard Euler realized that the equations for the deflection of a freely vibrating string admit sinusoidal solutions. That is, if we freeze the string's motion, we may observe a sinusoidal pattern. If the string's ends are fixed, the boundary conditions of nondeflecting endpoints requires that there be an even number of half wavelengths, as depicted in Figure 3.3. These different *modes* are accordingly harmonically related. The lowest spatial frequency has one half-wavelength in the string's length L, and so is of spatial frequency $\frac{1}{2L}$ cycles per unit length. The next completes a single cycle in L, and so is of frequency $\frac{1}{L}$. This is followed by three half cycles giving frequency $\frac{3}{2L}$, and so on. The boundary conditions ensure that all sinusoidal deflection patterns have spatial frequency that is a multiple of $\frac{1}{2L}$.

However, since the equations for the deflection of the string are linear, any linear combination of sinusoids that satisfy the boundary conditions is also a possible oscillation pattern. Consequently, a more general transverse deflection trace will be the sum of the basic modes (the sum of HRSs). The

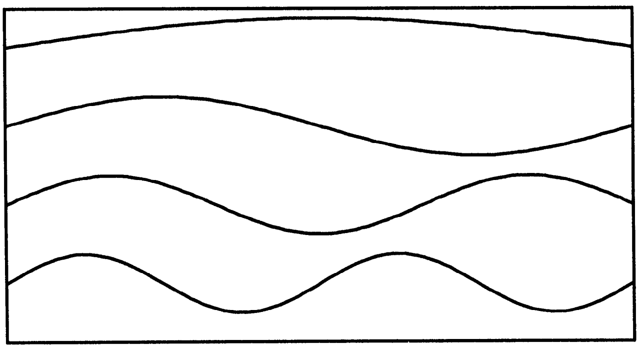

Figure 3.3: The instantaneous deflection of a vibrating string may be sinusoidal, and the boundary conditions restrict the possible frequencies of these sines. The top string contains only half of its wavelength between the string's supports; the next contains a full wavelength, the third three-quarters, etc.

question is whether this is the *most* general pattern of deflection. In the eighteenth and nineteenth century there were good reasons for suspecting the answer to be negative. Not having the benefit of the computer-generated plots of sums of HRSs presented in the previous section, even such great mathematicians as Lagrange believed that all such sums would yield *smooth* curves. However, it was easy to deform the string such that its shape would be noncontinuous (e.g., by pulling it up at its middle point forcing a triangular shape). What would happen the moment such a plucked string was released? Since the initial state was supposedly not representable in terms of the basic sinusoidal modes, there must be other, nonsinusoidal, solutions. This was considered to be a fatal blow to the utility of the theory of trigonometric series. It caused all of the mathematicians of the day to lose interest in them; all except Jean Baptiste Joseph Fourier. In his honor we are more apt today to say 'Fourier series' than 'trigonometric series'.

Although mathematics was Fourier's true interest, his training was for the military and clergy. He was sorely vexed upon reaching his twenty-first birthday without attaining the stature of Newton, but his aspirations had to wait for some time due to his involvement in the French revolution. Fourier (foolishly) openly criticized corrupt practices of officials of Robespierre's government, an act that led to his arrest and incarceration. He would have gone to the guillotine were it not for Robespierre himself having met that fate first. Fourier returned to mathematics for a time, studying at the Ecole Normal in Paris under the greatest mathematicians of the era, Lagrange and Laplace. After that school closed, he began teaching mathematics at the Ecole Polytechnique, and later succeeded Lagrange to the chair of mathematical analysis. He was considered a gifted lecturer, but as yet had made no outstanding contributions to science or mathematics.

Fourier then once again left his dreams of mathematics in order to join Napoleon's army in its invasion of Egypt. After Napoleon's loss to Nelson in the Battle of the Nile, the French troops were trapped in Egypt, and Fourier's responsibilities in the French administration in Cairo included founding of the Institut d'Egypte (of which he was secretary and member of the mathematics division), the overseeing of archaeological explorations, and the cataloging of their finds. When he finally returned to France, he resumed his post as Professor of Analysis at the Ecole Polytechnique, but Napoleon, recalling his administrative abilities, snatched him once again from the university, sending him to Grenoble as Prefect. Although Fourier was a most active Prefect, directing a number of major public works, he neglected neither his Egyptological writing nor his scientific research. His contributions to Egyptology won him election to the French Academy and to the Royal

Society in London. His most significant mathematical work is also from this period. This scientific research eventually led to his being named perpetual secretary of the Paris Academy of Sciences.

Fourier was very interested in the problem of heat propagation in solids, and in his studies derived the partial differential equation

$$\frac{\partial v}{\partial t} = K \frac{\partial^2 v}{\partial x^2}$$

now commonly known as the *diffusion equation*. The solution to such an equation is, in general, difficult, but Fourier noticed that there were solutions of the form $f(t)g(x)$, where $f(t)$ were decreasing exponentials and $g(x)$ were either $\sin(nx)$ or $\cos(nx)$. Fourier claimed that the most general $g(x)$ would therefore be a linear combination of such sinusoids

$$g(x) = \sum_{k=0}^{\infty} \Big(a_k \sin(kx) + b_k \cos(kx) \Big) \tag{3.3}$$

the expansion known today as the Fourier series. This expansion is more general than that of Euler, allowing both sines and cosines to appear simultaneously. Basically Fourier was claiming that arbitrary functions could be written as weighted sums of the sinusoids $\sin(nx)$ and $\cos(nx)$, a result we now call Fourier's theorem.

Fourier presented his theorem to the Paris Institute in 1807, but his old mentors Lagrange and Laplace criticized it and blocked its publication. Lagrange once again brought up his old arguments based on the inability of producing nonsmooth curves by trigonometric series. Fourier eventually had to write an entire book to answer the criticisms, and only this work was ever published. However, even this book fell short of complete rigorous refutation of Lagrange's claims. The full proof of validity of Fourier's ideas was only established later by the works of mathematicians such as Dirichlet, Riemann, and Lebesgue. Today we know that all functions that obey certain conditions (known as the Dirichlet conditions), even if they have discontinuous derivatives or even if they are themselves discontinuous, have Fourier expansions.

EXERCISES

3.3.1 Consider functions $f(t)$ defined on the interval $-1 \leq t \leq 1$ that are defined by finite weighted sums of the form $\sum_k f_k \cos(\pi k t)$, where k is an integer. What do all these functions have in common? What about weighted sums of $\sin(\pi k t)$?

3.3.2 Show that any function $f(t)$ defined on the interval $-1 \le t \le 1$ can be written as the sum of an even function $f_e(t)$ $(f_e(-t) = f_e(-t))$ and an odd function $(f_o(-t) = -f_o(-t))$.

3.3.3 Assume that all even functions can be represented as weighted sums of cosines as in the first exercise, and that all odd functions can be similarly represented as weighted sums of sines. Explain how Fourier came to propose equation (3.3).

3.3.4 How significant is the difference between a parabola and half a period of a sinusoid? To find out, approximate $x(t) = \cos(t)$ for $-\frac{\pi}{2} \le t \le \frac{\pi}{2}$ by $y(t) = at^2 + bt + c$. Find the coefficients by requiring $y(-t) = y(t)$, $y(0) = 1$ and $y(\pm\frac{\pi}{2}) = 0$. Plot the cosine and its approximation. What is the maximal error? The cosine has slope 1 at the ends of the interval; what is the slope of the approximation? In order to match the slope at $t = \pm\frac{\pi}{2}$ as well, we need more degrees of freedom, so we can try $y(t) = at^4 + bt^2 + c$. Find the coefficients and the maximum error.

3.4 Representation by Fourier Series

In this section we extend our discussion of the mathematics behind the Fourier series. We will not dwell upon formal issues such as conditions for convergence of the series. Rather, we have two related tasks to perform. First, we must convince ourselves that Fourier was right, that indeed *any* function (including nonsmooth ones) can be uniquely expanded in a Fourier Series (FS). This will demonstrate that the sinusoids, like the SUIs of Section 2.5, form a basis for the vector space of periodic signals with period T. The second task is a practical one. In Section 3.2 we posited a series and graphically determined the periodic signal it represented. Our second task is to find a way to do the converse operation—given the periodic signal to find the series.

In Section 2.5 we saw that *any* digital signal could be expanded in the set of all SUIs. It was left as exercises there to show that the same is true for the analog domain, and in particular for periodic analog signals. The set of all shifted analog impulses (Dirac delta functions) $\delta(t - \tau)$ forms a basis in which all analog signals may be expanded. Now, since we are dealing with periodic signals let us focus on the signal's values in the time interval between time zero and time T. It is clear that it is sufficient to employ shifted impulses for times from zero to T to recreate any waveform in this time interval.

The desired proof of a similar claim for HRSs can rest on our showing that any shifted analog impulse in the required time interval can be built up from such sinusoids. Due to the HRS's periodicity in T, the shifted impulse will automatically be replicated in time to become a periodic 'impulse train'. Consequently the following algorithm finds the HRS expansion of any function of period T.

```
focus on the interval of time from t = 0 to t = T
expand the desired signal in this interval in shifted impulses
for each impulse substitute its HRS expansion
rearrange and sort the HRS terms
consider this to be the desired expansion for all t
```

All that remains is to figure out how to represent an impulse in terms of HRSs. In Section 3.2 we experimented with adding together an infinite number of HRSs, but always with amplitudes that decreased with increasing frequency. What would happen if we used all harmonics equally?

$$b_0 + \cos(t) + \cos(2t) + \cos(3t) + \cos(4t) + \ldots \tag{3.4}$$

At time zero all the terms contribute unity and so the infinite sum diverges. At all other values the oscillations cancel themselves out. We demonstrate graphically in Figure 3.4 that this sum converges to an impulse centered at time zero. We could similarly make an impulse centered at any desired time by using combinations of sin and cos terms. This completes the demonstration that any analog impulse centered in the basic period, and thus any periodic signal, can be expanded in the infinite set of HRSs.

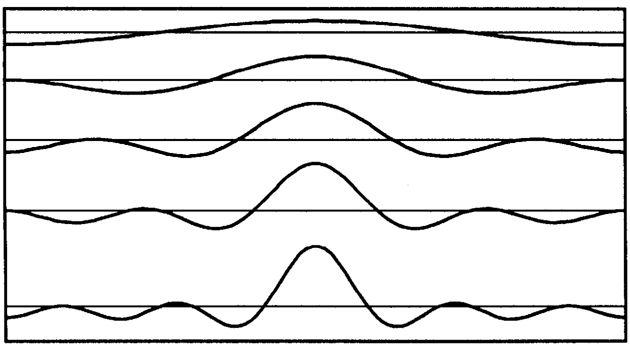

Figure 3.4: Building up an impulse from a cosine and its harmonics.

We are almost done. We have just shown that the HRSs span the vector space of periodic analog signals. In order for this set to be a basis the expansions must be unique. The usual method of proving uniqueness involves showing that there are no extraneous signals in the set, i.e., by showing that the HRSs are linearly independent. Here, however, there is a short-cut; we can show that the HRSs comprise an orthonormal set, and we know from Appendix A.14 that all orthonormal sets are linearly independent.

In Section 2.5 the dot product was shown to be a valid scalar multiplication operation for the vector space of analog signals. For periodic analog signals we needn't integrate over all times, rather the product given by

$$r = x \cdot y \qquad \text{means} \qquad r = \int_0^T x(t)\, y(t)\, dt \tag{3.5}$$

(where the integration can actually be performed over any whole period) should be as good. Actually it is strictly better since the product over all times of finite-valued periodic signals may be infinite, while the present product always finite. Now it will be useful to try out the dot product on sinusoids.

We will need to know only a few definite integrals, all of which are derivable from equation A.34. First, the integral of any sinusoid over any number of whole periods gives zero

$$\int_0^T \sin\left(\frac{2\pi}{T}t\right)\, dt = 0 \tag{3.6}$$

since $\sin(-x) = -\sin(x)$, and so for every positive contribution to the integral there is an equal and opposite negative contribution. Second, the integral of $\sin^2$ (or $\cos^2$) over a single period is

$$\int_0^T \sin^2\left(\frac{2\pi}{T}t\right)\, dt = \frac{T}{2} \tag{3.7}$$

which can be derived by realizing that symmetry dictates

$$I = \int_0^T \sin^2\left(\frac{2\pi}{T}t\right)\, dt = \int_0^T \cos^2\left(\frac{2\pi}{T}t\right)\, dt$$

and so

$$2I = \int_0^T \left(\sin^2\left(\frac{2\pi}{T}t\right) + \cos^2\left(\frac{2\pi}{T}t\right)\right)\, dt = \int_0^T 1\, dt = T$$

by identity (A.20). Somewhat harder to guess is the fact that the integral of the product of different harmonics is always zero, i.e.

$$\begin{aligned}
\int_0^T \sin\left(\frac{2\pi n}{T}t\right)\cos\left(\frac{2\pi m}{T}t\right)dt &= 0 \qquad \forall n,m>0 \\
\int_0^T \sin\left(\frac{2\pi n}{T}t\right)\sin\left(\frac{2\pi m}{T}t\right)dt &= \delta_{n,m}\frac{T}{2} \\
\int_0^T \cos\left(\frac{2\pi n}{T}t\right)\cos\left(\frac{2\pi m}{T}t\right)dt &= \delta_{n,m}\frac{T}{2}
\end{aligned} \tag{3.8}$$

the proof of which is left as an exercise.

These relations tell us that the set of normalized signals $\{v_k\}_{k=1}^{\infty}$ defined by

$$\begin{aligned}
v_0(t) &= \sqrt{\frac{1}{T}} \\
v_{2k+1}(t) &= \sqrt{\frac{2}{T}}\cos\left(\frac{2\pi k}{T}t\right) \qquad \forall k>0 \\
v_{2k}(t) &= \sqrt{\frac{2}{T}}\sin\left(\frac{2\pi k}{T}t\right) \qquad \forall k>0
\end{aligned}$$

forms an orthonormal set of signals. Since we have proven that *any* signal of period T can be expanded in these signals, they are an orthonormal set of signals that span the space of periodic signals, and so an orthonormal basis. The $\{v_k\}$ are precisely the HRSs to within unimportant multiplicative constants, and hence the HRSs are an orthogonal basis of the periodic signals. The Fourier series takes on a new meaning. It is the expansion of an arbitrary periodic signal in terms of the orthogonal basis of HRSs.

We now return to our second task—given a periodic signal $s(t)$, we now know there is an expansion:

$$s(t) = \sum_{k=1}^{\infty} c_k\, v_k(t)$$

How do we find the expansion coefficients c_k? This task is simple due to the basis $\{v_k\}$ being orthonormal. From equation A.85 we know that for an orthonormal basis we need only to *project* the given signal onto each basis signal (using the dot product we defined above).

$$c_k = s \cdot v = \int_0^T s(t)\, v_k(t)\, dt$$

This will give us the coefficients for the normalized basis. To return to the usual HRSs

$$\begin{aligned} s(t) &= \sum_{k=1}^{\infty} a_k \sin\left(\frac{2\pi k}{T}t\right) + \sum_{k=0}^{\infty} b_k \cos\left(\frac{2\pi k}{T}t\right) \\ &= \sum_{k=1}^{\infty} a_k \sin\left(\frac{2\pi k}{T}t\right) + b_0 + \sum_{k=1}^{\infty} b_k \cos\left(\frac{2\pi k}{T}t\right) \end{aligned} \tag{3.9}$$

is not difficult.

$$\begin{aligned} a_k &= \frac{2}{T}\int_0^T s(t)\sin\left(\frac{2\pi k}{T}t\right)\,dt \\ b_0 &= \frac{1}{T}\int_0^T s(t)\,dt \\ b_k &= \frac{2}{T}\int_0^T s(t)\cos\left(\frac{2\pi k}{T}t\right)\,dt \end{aligned} \tag{3.10}$$

This result is most fortunate; were the sinusoids not orthogonal, finding the appropriate coefficients would require solving 'normal equations' (see Appendix A.14). When there are a finite number N of basis functions, this is a set of N equations in N variables; if the basis is infinite we are not even able to write down the equations!

These expressions for the FS coefficients might seem a bit abstract, so let's see how they really work. First let's start with a simple sinusoid $s(t) = A\sin(\omega t) + B$. The basic period is $T = \frac{2\pi}{\omega}$ and so the expansion can contain only sines and cosines with periods that divide this T. The DC term is, using equations (3.6) and (3.7),

$$b_0 = \frac{1}{T}\int_0^T s(t)\,dt = \frac{1}{T}\int_0^T \left(A\sin\left(\frac{2\pi}{T}t\right) + B\right)dt = \frac{1}{T}BT = B$$

as expected, while from equations (3.8) all other terms are zero except for one.

$$\begin{aligned} a_1 &= \frac{2}{T}\int_0^T s(t)\sin\left(\frac{2\pi k}{T}t\right)dt \\ &= \frac{2}{T}\int_0^T \left(A\sin\left(\frac{2\pi}{T}t\right) + B\right)\sin\left(\frac{2\pi}{T}t\right)dt = \frac{2}{T}A\frac{T}{2} = A \end{aligned}$$

This result doesn't surprise us since the expansion of one of basis signals must be exactly that signal!

Slightly more interesting is the case of the square wave $\square(t/T)$. There will be no DC term nor any cosine terms, as can be seen by direct symmetry. To show this mathematically we can exploit a fact we have previously mentioned, that the domain of integration can be over any whole period. In this case it is advantageous to use the interval from $-T/2$ to $T/2$. Since $\square(t/T)$ is an odd function, i.e., $\square(-t/T) = -\square(t/T)$, the contribution from the left half interval exactly cancels out the contribution of the right half interval. This is a manifestation of a general principle; odd functions have only sine terms, while even functions have only DC and cosine term contributions. The main contribution for $\square(t/T)$ will be from the sine of period T, with coefficient

$$\begin{aligned} a_1 &= \frac{2}{T}\int_0^T s(t)\sin\left(\frac{2\pi}{T}t\right)dt \\ &= \frac{2}{T}\int_0^{\frac{T}{2}} \sin\left(\frac{2\pi}{T}t\right)dt - \frac{2}{T}\int_{\frac{T}{2}}^T \sin\left(\frac{2\pi}{T}t\right)dt \\ &= 2\frac{2}{T}\int_0^{\frac{T}{2}} \sin\left(\frac{2\pi}{T}t\right)dt = \frac{4}{\pi} \end{aligned}$$

while the sine of double this frequency

$$a_2 = \frac{2}{T}\int_0^T s(t)\sin\left(\frac{4\pi}{T}t\right)dt = \frac{2}{T}\int_0^{\frac{T}{2}} \sin\left(\frac{4\pi}{T}t\right)dt - \frac{2}{T}\int_{\frac{T}{2}}^T \sin\left(\frac{4\pi}{T}t\right)dt = 0$$

cannot contribute because of the odd problem once again. Therefore only odd harmonic sinusoids can appear, and for them

$$\begin{aligned} a_k &= \frac{2}{T}\int_0^T s(t)\sin\left(\frac{2\pi k}{T}t\right)dt \\ &= \frac{2}{T}\int_0^{\frac{T}{2}} \sin\left(\frac{2\pi k}{T}t\right)dt - \frac{2}{T}\int_{\frac{T}{2}}^T \sin\left(\frac{2\pi k}{T}t\right)dt \\ &= 2\frac{2}{T}\int_0^{\frac{T}{2}} \sin\left(\frac{2\pi k}{T}t\right)dt = \frac{4}{\pi k} \end{aligned}$$

which is exactly equation (3.2).

EXERCISES

3.4.1 Our proof that the HRSs span the space of periodic signals required the HRSs to be able to reproduce all SUIs, while Figure 3.4 reproduced only an impulse centered at zero. Show how to generate arbitrary SUIs (use a trigonometric sum formula).

3.4.2 Observe the sidelobes in Figure 3.4. What should the constant term b_0 be for the sidelobes to oscillate around zero? In the figure each increase in the number of cosines seems to add another half cycle of oscillation. Research numerically the number and amplitude of these oscillations by plotting the sums of larger numbers of cosines. Do they ever disappear?

3.4.3 Reproduce a graph similar to Figure 3.4 but using sines instead of cosines. Explain the results (remember that sine is an odd function). Why isn't the result simply a shifted version of cosine case?

3.4.4 Find the Fourier series coefficients for the following periodic signals. In order to check your results plot the original signal and the partial sums.

1. Sum of two sines $a_1 \sin(\omega t) + a_2 \sin(2\omega t)$
2. Triangular wave
3. Fully rectified sine $|\sin(x)|$
4. Half wave rectified sine $\sin(x)u(\sin(x))$

3.4.5 We can consider the signal $s(t) = A\sin(\omega t) + B$ to be periodic with period $T = \frac{4\pi}{\omega}$. What is the expansion now? Is there really a difference?

3.4.6 For the two-dimensional plane consider the basis made up of unit vectors along the x axis $A_1 = (1, 0)$ and along the 45° diagonal $A = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$. The unit vector of slope $\frac{1}{2}$ is $Y = (\frac{2}{\sqrt{5}}, \frac{1}{\sqrt{5}})$. Find the coefficients of the expansion $Y = \alpha_1 A_1 + \alpha_2 A_2$ by projecting Y on both A_1 and A_2 and solving the resulting equations.

3.4.7 Find explicitly the normal equations for a set of basis signals $A_k(t)$ and estimate the computational complexity of solving these equations.

3.5 Gibbs Phenomenon

Albert Abraham Michelson was the first American to receive a Nobel prize in the sciences. He is justly famous for his measurement of the speed of light and for his part in the 1887 Michelson-Morley experiment that led to the birth of the special theory of relativity. He invented the *interferometer* which allows measurement of extremely small time differences by allowing two light waves to interfere with each other. What is perhaps less known is that just after the Michelson-Morley experiment he built a practical Fourier analysis device providing a sort of physical proof of Fourier's mathematical claims regarding representation of periodic signals in terms of sinusoids. He was quite surprised when he found that the Fourier series for the square wave

$\square(t)$ didn't converge very well. In fact there was significant 'ringing', bothersome oscillations that wouldn't go away with increasing number of terms. Unsure whether he had discovered a new mathematical phenomenon or simply a bug in his analyzer he turned to the eminent American theoretical physicist of the time, Josiah Willard Gibbs. Gibbs realized that the problem was caused by discontinuities. Dirichlet had shown that the Fourier series converged to the midpoint at discontinuities, and that as long as there were a finite number of such discontinuities the series would globally converge; but no one had previously asked what happened near a discontinuity for a finite number of terms. In 1899 Gibbs published in *Nature* his explanation of what has become known as the Gibbs phenomenon.

In Section 3.3 we mentioned the Dirichlet conditions for convergence of the Fourier series.

Theorem: Dirichlet's Convergence Conditions
Given a periodic signal $s(t)$, if

1. $s(t)$ is absolutely integratable, i.e., $\int |s(t)|dt < \infty$, where the integral is over one period,
2. $s(t)$ has at most a finite number of extrema, and
3. $s(t)$ has at most a finite number of finite discontinuities,

then the Fourier series converges for every time. At discontinuities the series converges to the midpoint. ∎

To rigorously prove Dirichlet's theorem would take us too far afield so we will just give a taste of the mathematics one would need to employ. What is necessary is an analytical expression for the partial sums $S_K(t)$ of the first K terms of the Fourier series. It is useful to define the following sum

$$D_K(t) = \tfrac{1}{2} + \cos(t) + \cos(2t) + \ldots + \cos(Kt) = \tfrac{1}{2} + \sum_{k=1}^{K} \cos(kt) \qquad (3.11)$$

and to find for it an explicit expression by using trigonometric identities.

$$D_K(t) = \frac{\sin\left((K + \frac{1}{2})t\right)}{2\sin(\frac{1}{2}t)} \qquad (3.12)$$

It can then be shown that for any signal $s(t)$ the partial sums equal

$$S_K(t) = \frac{2}{T} \int s(t+\tau)\, D_K\left(\frac{2\pi}{T}\tau\right) d\tau \qquad (3.13)$$

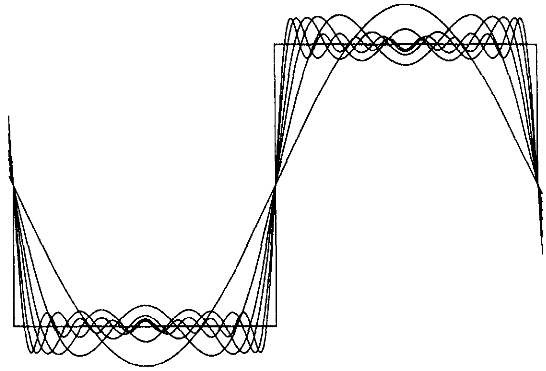

Figure 3.5: Partial sums of the Fourier series of a periodic square wave signal $\square(t)$ for $K = 0, 1, 2, 3, 5$ and 7. Note that although far from the discontinuity the series converges to the square wave, near it the overshoot remains.

(the integration being over one period of duration T) from which Dirichlet's convergence results emerge.

Now you may believe, as everyone did before Gibbs, that Dirichlet's theorem implies that amplitude of the oscillations around the true values decreases as we increase the number of terms in the series. This is the case *except* for the vicinity of a discontinuity, as can be seen in Figure 3.5. We see that close to a discontinuity the partial sums always overshoot their target, and that while the time from the discontinuity to the maximum overshoot decreases with increasing K, the overshoot amplitude does not decrease very much. This behavior does not contradict Dirichlet's theorem since although points close to jump discontinuities may initially be affected by the overshoot, after enough terms have been summed the overshoot will pass them and the error will decay.

For concreteness think of the square wave $\square(t)$. For positive times close to the discontinuity at $t = 0$ equation (3.13) can be approximated by

$$S_K(t) = \frac{2}{\pi} \operatorname{sgn}(t) \operatorname{Sinc}(4\pi K|t|) \tag{3.14}$$

as depicted in Figure 3.6. Sinc is the sine integral.

$$\operatorname{Sinc}(t) = \int_0^t \operatorname{sinc}(\tau)\, d\tau$$

Sinc approaches $\frac{\pi}{2}$ for large arguments, and thus $S_K(t)$ *does* approach unity for large K and/or t. The maximum amplitude of Sinc occurs when its derivative (sinc) is zero, i.e., when its argument is π. It is not hard to find

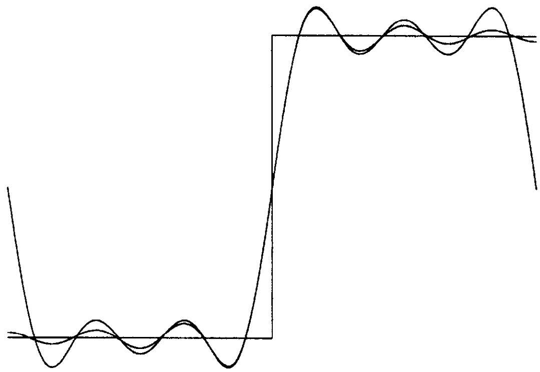

Figure 3.6: Gibbs phenomenon for the discontinuity of the square wave at $t = 0$. Plotted are the square wave, the partial sum with $K = 3$ terms, and the approximation using the sine integral.

numerically that for large K this leads to an overshoot of approximately 0.18, or a little less than 9% of the height of the jump. Also, the sine integral decays to its limiting value like $\frac{1}{t}$; hence with every doubling of distance from the discontinuity the amplitude of the oscillation is halved. We derived these results for the step function, but it is easy to see that they carry over to a general jump discontinuity.

That's what the mathematics says, but what does it mean? The oscillations themselves are not surprising, this is the best way to smoothly approximate a signal—sometimes too high, sometimes too low. As long as these oscillations rapidly die out with increasing number of terms the approximation can be considered good. What do we expect to happen near a discontinuity? The more rapid a change in the signal in the time domain is, the wider the bandwidth will be in the frequency domain. In fact the uncertainty theorem (to be discussed in Section 4.4) tells us that the required bandwidth is inversely proportional to the transition time. A discontinuous jump requires an infinite bandwidth and thus no combination of a finite number of frequencies, no matter how many frequencies are included, can do it justice. Of course the coefficients of the frequency components of the square wave *do* decrease very rapidly with increasing frequency. Hence by including more and more components, that is, by using higher and higher bandwidth, signal values closer and closer to the discontinuity, approach their proper values. However, when we approximate a discontinuity using bandwidth BW, within about $1/BW$ of the discontinuity the approximation cannot possibly approach the true signal.

We can now summarize the Gibbs phenomenon. Whenever a signal has a jump discontinuity its Fourier series converges at the jump time to the midpoint of the jump. The partial sums display oscillations before and after the jump, the number of cycles of oscillation being equal to the number of terms taken in the series. The size of the overshoot decreases somewhat with the number of terms, approaching about 9% of the size of the jump. The amplitude of the oscillations decreases as one moves away from the discontinuity, halving in amplitude with every doubling of distance.

EXERCISES

3.5.1 Numerically integrate $\mathrm{sinc}(t)$ and plot $\mathrm{Sinc}(t)$. Show that it approaches $\pm\frac{\pi}{2}$ for large absolute values. Find the maximum amplitude. Where does it occur? Verify that the asymptotic behavior of the amplitude is $\frac{1}{t}$.

3.5.2 The following exercises are for the mathematically inclined. Prove equation (3.12) by term-by-term multiplication of the sum in the definition of $D_K(t)$ by $\sin\left(\frac{t}{2}\right)$ and using trigonometric identity (A.32).

3.5.3 Prove equation (3.13) and show Dirichlet's convergence results.

3.5.4 Prove the approximation (3.14).

3.5.5 Lanczos proposed suppressing the Gibbs phenomenon in the partial sum S_K by multiplying the k^{th} Fourier coefficient (except the DC) by $\mathrm{sinc}\left(\frac{\pi k}{2K}\right)$. Try this for the square wave. How much does it help? Why does it help?

3.5.6 We concentrated on the Gibbs phenomenon for the square wave. How do we know that other periodic signals with discontinuities act similarly? (Hint: Consider the Fourier series for $s(t) + a\square(t)$ where $s(t)$ is a continuous signal and a a constant.)

3.6 Complex FS and Negative Frequencies

The good news about the Fourier series as we have developed it is that its basis signals are the familiar *sine* and *cosine* functions. The bad news is that its basis signals are the familiar sine *and* cosine functions. The fact that there are two different kinds of basis functions, and that the DC term is somewhat special, makes the FS as we have presented it somewhat clumsy to use. Unfortunately, sines alone span only the subspace composed of all odd signals, while cosines alone span only the subspace of all even signals.

Signals which are neither odd nor even. truly require combinations of both Since the FS in equation (3.9) includes for every frequency both a sine and cosine function (which differ by 90° or a quarter cycle), it is said to be in quadrature form.

The first signal space basis we studied, the SUI basis, required only one functional form. Is there a single set of sinusoidal signals, all of the same type, that forms a basis for the space of periodic signals? Well, for each frequency component ω the FS consists of the sum of two terms $a\cos(\omega t) + b\sin(\omega t)$. Such a sum produces a pure sinusoid of the same frequency, but with some phase offset $d\sin(\omega t + \varphi)$. In fact, it is easy to show that

$$a_k \sin(\omega t) + b_k \cos(\omega t) = d_k \sin(\omega t + \varphi_k) \tag{3.15}$$

as long as

$$d_k = \sqrt{a_k^2 + b_k^2} \qquad \varphi_k = \tan^{-1}(b_k, a_k) \tag{3.16}$$

where the arctangent is the full four-quadrant function, and

$$a_k = d_k \cos\varphi_k \qquad b_k = d_k \sin\varphi_k \tag{3.17}$$

in the other direction.

As a result we *can* expand periodic signals $s(t)$ as

$$s(t) = d_0 + \sum_{k=0}^{\infty} d_k \sin\left(\frac{2\pi k}{T}t + \varphi_k\right) \tag{3.18}$$

with both amplitudes and phases being parameters to be determined.

The amplitude and phase form is intellectually more satisfying than the quadrature one. It represents every periodic signal in terms of harmonic frequency components, each with characteristic amplitude and phase. This is more comprehensible than representing a signal in terms of pairs of sinusoids in quadrature. Also, we are often only interested in the *power spectrum*, which is the amount of energy in each harmonic frequency. This is given by $|d_k|^2$ with the phases ignored.

There *are* drawbacks to the amplitude and phase representation. Chief among them are the lack of symmetry between d_k and φ_k and the lack of simple formulas for these coefficients. In fact, the standard method to calculate d_k and φ_k is to find a_k and b_k and use equations (3.16)!

We therefore return to our original question: Is there a single set of sinusoidal signals, all of the same type, that forms a basis for the space of periodic signals and that can be calculated quickly and with resort to the quadrature representation? The answer turns out to be affirmative.

To find this new representation recall the connection between sinusoids and complex exponentials of equation (A.8).

$$\cos(\omega t) = \frac{1}{2}\left(e^{i\omega t} + e^{-i\omega t}\right) \qquad \sin(\omega t) = \frac{1}{2i}\left(e^{i\omega t} - e^{-i\omega t}\right) \tag{3.19}$$

We can *think of the exponents with* positive $e^{i\omega t}$ and *negative* $e^{-i\omega t}$ exponents as a single type of exponential $e^{i\omega t}$ with positive and negative frequencies ω. Using only such complex exponentials, although of both positive and negative frequencies, we can produce both the sine and cosine signals of the quadrature representation, and accordingly represent any periodic signal.

$$s(t) = \sum_{k=-\infty 1}^{\infty} c_k e^{i\frac{2\pi k}{T}t} \tag{3.20}$$

We *could* once again derive the expression for the coefficients c_k from those for the quadrature representation, but it is simple enough to derive them from scratch. We need to know only a single integral.

$$\int_0^T e^{i\frac{2\pi n}{T}t}\, e^{-i\frac{2\pi m}{T}t}\, dt = \delta_{m,n} T \tag{3.21}$$

This shows that the complex exponentials are orthogonal with respect to the dot product for complex signals

$$s_1 \cdot s_2 = \int_0^T s_1(t)\, s_2^*(t)\, dt \tag{3.22}$$

and that

$$v_k(t) = \frac{1}{\sqrt{T}}\, e^{i\frac{2\pi k}{T}t} \tag{3.23}$$

form a (complex) orthonormal set. From this it is easy to see that

$$c_k(t) = \frac{1}{T}\int_0^T s(t)\, e^{-i\frac{2\pi k}{T}t}\, dt \tag{3.24}$$

with a minus sign appearing in the exponent. Thus Fourier's theorem can be stated in a new form: All periodic functions (which obey certain conditions) can be written as weighted sums of complex exponentials.

The complex exponential form of the FS is mathematically the simplest possible. There is only one type of function, one kind of coefficient, and there is strong symmetry between equations (3.20) and (3.24) that makes them easier to remember. The price to pay has been the introduction of mysterious negative frequencies. What do we mean by -100 Hz? How can something cycle *minus 100* times per second?

Physically, negative frequency signals are almost identical to their positive counterparts, since only the real part of a complex signal counts. Recall the pen-flashlight experiment that you were requested to perform in exercise 2.2.6. The complex exponential corresponds to observing the flashlight head-on, while the real sinusoid is observing it from the side. Rotation of the light in clockwise or counterclockwise (corresponding to positive or negative frequencies) produces the same effect on an observer who perceives just the vertical (real) component; only an observer with a full view notices the difference. However, it would be foolhardy to conclude that negative frequencies are of no importance; when more than one signal is present the relative phases are crucial.

We conclude this section with the computation of a simple complex exponential FS—that of a real sinusoid. Let $s(t) = A\cos(\frac{2\pi k}{T}t)$. The period is of course T, and

$$c_k = \frac{1}{T}\int_0^T A\cos(\frac{2\pi}{T}t)\, e^{-\mathrm{i}\frac{2\pi k}{T}t}\, dt = \frac{A}{T}\int_0^T \tfrac{1}{2}\left(e^{\mathrm{i}\frac{2\pi}{T}t} + e^{-\mathrm{i}\frac{2\pi}{T}t}\right)\, e^{-\mathrm{i}\frac{2\pi k}{T}t}\, dt$$

which after using the orthogonality relation (3.21) leaves two terms.

$$c_k = \frac{A}{2T}\,\delta_{k,-1} \;+\; \frac{A}{2T}\,\delta_{k,+1}$$

This is exactly what we expected considering equation (3.19). Had we chosen $s(t) = A\sin(\frac{2\pi k}{T}t)$ we would have still found two terms with identical k and amplitudes but with phases shifted by 90°. This is hardly surprising; indeed it is easy to see that all $s(t) = A\cos(\frac{2\pi k}{T}t + \varphi)$ will have the same FS except for phase shifts of φ. Such constant phase shifts are meaningless, there being no meaning to absolute phase, only to changes in phase.

EXERCISES

3.6.1 Plot $\sin(x) + \sin(2x + \varphi)$ with $\varphi = 0, \frac{\pi}{2}, \pi, \frac{2\pi}{2}$. What can you say about the effect of phase? Change the phases in the Fourier series for a square wave. What signals can you make?

3.6.2 Derive all the relations between coefficients of the quadrature, amplitude and phase, and complex exponential representations. In other words, show how to obtain a_k and b_k from c_k and vice versa; a_k and b_k from d_k and vice versa; c_k from d_k and vice versa. In your proofs use only trigonometric identities and equation (A.7).

3.6.3 Prove equation (3.21).

3.6.4 Calculate the complex exponential FS of $s(t) = A\sin(\frac{2\pi k}{T}t)$. How does it differ from that of the cosine?

3.6.5 Consistency requires that substituting equation (3.20) for the FS into equation (3.24) for c_k should bring us to an identity. Show this using (3.21). What new expression for the delta function is implied by the reverse consistency argument?

3.6.6 What transformations can be performed on a signal without effecting its power spectrum $|c_k|^2$? What is the physical meaning of such transformations?

3.7 Properties of Fourier Series

In this section we continue our study of Fourier series. We will exclusively use the complex exponential representation of the FS since it is simplest, and in any case we can always convert to other representations if the need arises.

The first property, which is obvious from the expression for c_k, is linearity. Assume $s_1(t)$ has FS coefficients c_k^1 and $s_2(t)$ has coefficients c_k^2, then $s(t) = As_1(t) + Bs_2(t)$ has as its coefficients $c_k = Ac_k^1 + Bc_k^2$. This property is often useful in simplifying calculations, and indeed we already implicitly used it in our calculation of the FS of $\cos(\omega t) = \frac{1}{2}e^{i\omega t} + \frac{1}{2}e^{-i\omega t}$. As a further example, suppose that we need to find the FS of a constant (DC) term plus a sinusoid. We can immediately conclude that there will be exactly three nonzero c_k terms, c_{-l}, c_0, and c_{+l}.

In addition to its being used as a purely computational ploy, the linearity of c_k has theoretic significance. The world would be a completely different place were the FS not to be linear. Were the FS of $As(t)$ not to be Ac_k then simple amplification would change the observed harmonic content of a signal. Linear operators have various other desirable features. For example, small changes to the input of a linear operator can only cause bounded changes to the output. In our case this means that were one to slightly perturb a signal with known FS, there is a limit to how much c_k can change.

The next property of interest is the effect of time shifts on the FS. By time shift we mean replacing t by $t - \tau$, which is equivalent to resetting our clock to read zero at time τ. Since the time we start our clock is arbitrary such time shifts cannot alter any physical aspects of the signal being studied. Once again going back to the expression for c_k we find that the FS of $s(t - \tau)$

is $e^{-i\frac{2\pi k}{T}} c_k$. The coefficients magnitudes are unchanged, but the phases have been linearly shifted. As we know from exercise 3.6.6 such phase shifts do not change the power spectrum but still may be significant. We see here that phase shifts that are linear in frequency correspond to time shifts.

When a transformation leaves a signal unchanged or changes it in some simple way we call that transformation a *symmetry*. Time shift is one interesting symmetry, and another is time reversal Rev s. Although the import of the latter is less compelling than the former many physical operations are unchanged by time reversal. It is not difficult to show that the effect of time reversal is to reverse the FS to c_{-k}.

The next property of importance was discovered by Parseval and tells us how the energy can be recovered from the FS coefficients.

$$E = \frac{1}{T}\int_0^T |s(t)|^2\, dt = \sum_{k=-\infty}^{\infty} |c_k|^2 \qquad (3.25)$$

What does Parseval's relation mean? The left hand side is the power computed over a single period of the periodic signal. The power of the sum of two signals equals the sum of the powers if and only if the signals are orthogonal.

$$\begin{aligned} \frac{1}{T}\int_0^T |x(t)+y(t)|^2\, dt &= \frac{1}{T}\int_0^T \left(x(t)+y(t)\right)^* \left(x(t)+y(t)\right)\, dt \\ &= \frac{1}{T}\int_0^T |x(t)|^2 + |y(t)|^2 + 2\,\Re\left(x^*(t)y(t)\right)\, dt \end{aligned}$$

Since any two different sinusoids are uncorrelated, their powers add, and this can be generalized to the sum of any number of sinusoids. So Parseval's relation is another consequence of the fact that sinusoids are orthogonal.

For complex valued signals $s(t)$ there is a relation between the FS of the signal and that of its complex conjugate $s^*(t)$. The FS of the complex conjugate is c^*_{-k}. For real signals this implies a symmetry of c_k (i.e., $c_{-k} = c^*_k$), which means $|c_{-k}| = |c_k|$ and $\Re(c_{-k}) = \Re(c_k)$ but $\Im(c_{-k}) = -\Im(c_k)$.

There are many more symmetries and relations that can be derived for the FS, e.g., the relationship between the FS of a signal and those of its derivative and integral. There is also an important rule for the FS of the product of two signals, which the reader is not yet ready to digest.

EXERCISES

3.7.1 Show that adding to the argument of a sinusoid a phase that varies linearly with time shifts its frequency by a constant. Relate this to the time shift property of the FS.

3.7.2 Plot the sum of several sinusoids with various phases. Demonstrate that a linear phase shift causes a time shift. Can you tell that all these signals have the same power spectrum?

3.7.3 How does change of time scale $s(\alpha t)$ affect c_k? Prove that the effect of time reversal is to reverse the FS.

3.7.4 Derive Parseval's relation for the FS.

3.7.5 Show that if a signal is symmetric(antisymmetric), i.e., if $s(t + \frac{T}{2}) = \pm s(t)$, then its FS contains only even (odd) harmonics.

3.7.6 The FS of s is c_k; what is the FS of its derivative? Its integral?

3.8 The Fourier Series of Rectangular Wave

Since we have decided to use the complex exponential representation almost exclusively, we really should try it out. First, we want to introduce a slightly different notation. When we are dealing with several signals at a time, say $q(t)$, $r(t)$, and $s(t)$, using c_k for the FS coefficients of all of them, would be confusing to say the least. Since the Fourier coefficients contain exactly the same information as the periodic signal, using the name of the signal, as in q_k, r_k, or S_k, would be justified. There won't be any confusion since $s(t)$ is continuous and S_k is discrete; however, later we will deal with continuous spectra where it wouldn't be clear. So most people prefer to capitalize the Fourier coefficients, i.e., to use Q_k, R_k, and S_K, in order to emphasize the distinction between time and frequency domains. Hence from now on we shall use

$$S_k = \frac{1}{T} \int s(t)\, e^{-i\frac{2\pi k}{T}t}\, dt \tag{3.26}$$

(with the integration over any full period) to go from a signal $s(t)$ to its FS $\{S_k\}_{k=-\infty}^{\infty}$, and

$$s(t) = \sum_{k=-\infty}^{\infty} S_k e^{i\frac{2\pi k}{T}t} \tag{3.27}$$

to get back again.

Now to work. We have already derived the FS of a square wave, at least in the quadrature representation. Here we wish to extend this result to the slightly more general case of a rectangular wave, i.e., a periodic signal that does not necessarily spend half of its time at each level. The fraction of time

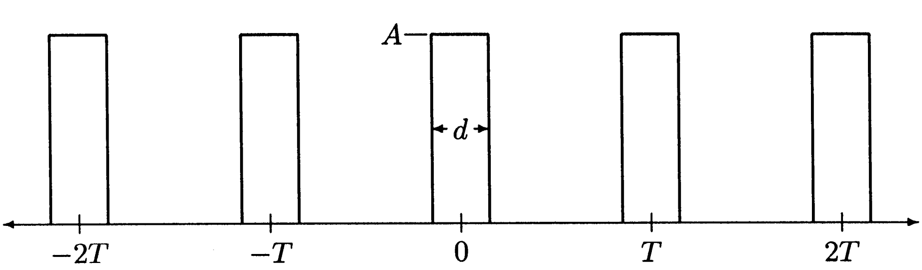

Figure 3.7: The rectangular signal with amplitude A, period T, and duty cycle $\delta = \frac{d}{T}$.

a rectangular wave spends in the higher of its levels is called its *duty cycle* $\delta = \frac{d}{T}$, and a rectangular wave with $\delta = \frac{1}{2}$ duty cycle is a square wave. We also wish to make the amplitude and period explicit, and to have the signal more symmetric in the time domain; we accordingly introduce A, T, and $d = \delta T$, and require the signal to be high from $-\frac{d}{2}$ to $\frac{d}{2}$. Unlike the square wave, a non-50% duty cycle rectangular signal will always have a DC component. There is consequently no reason for keeping the levels symmetric around zero, and we will use 0 and A rather than $\pm A$.

Thus we will study

$$s(t) = A \begin{cases} 1 & |\mathrm{frac}(\frac{t}{T})| < \frac{d}{2} \\ 0 & \frac{d}{2} < |\mathrm{frac}(\frac{t}{T})| < T - \frac{d}{2} \\ 1 & T - \frac{d}{2} < |\mathrm{frac}(\frac{t}{T})| < T \end{cases} \tag{3.28}$$

(where $\mathrm{frac}(x)$ is the fractional part of x) as depicted in Figure 3.7.

The period is T and therefore the angular frequencies in the Fourier series will all be of the form $\omega_k = \frac{2\pi}{T}k$. We can choose the interval of integration in equation (3.24) as we desire, as long as it encompasses a complete period. The most symmetric choice here is from $-\frac{T}{2}$ to $\frac{T}{2}$, since the signal then becomes simply

$$s(t) = A \begin{cases} 1 & |(\frac{t}{T})| < \frac{d}{2} \\ 0 & \text{else} \end{cases} \tag{3.29}$$

and as a consequence

$$\begin{aligned} S_k &= \frac{1}{T} \int_{-\frac{T}{2}}^{\frac{T}{2}} s(t)\, e^{-\mathrm{i}\frac{2\pi k}{T}t}\, dt \\ &= \frac{A}{T} \int_{-\frac{d}{2}}^{\frac{d}{2}} e^{-\mathrm{i}\omega_k t}\, dt \end{aligned}$$

which after change of variable and use of equation (A.8) becomes

$$S_k = A\frac{\sin(\frac{\omega_k d}{2})}{\frac{\omega_k d}{2}} = A\,\mathrm{sinc}\left(\frac{\omega_k d}{2}\right) = A\,\mathrm{sinc}\left(\frac{\pi k d}{T}\right) = A\,\mathrm{sinc}(\pi k \delta) \qquad (3.30)$$

where we have recognized our old friend sinc. The FS is dependent only on the duty cycle, not directly on T. Of course this does not mean that the Fourier series is not dependent on T! The coefficient S_k multiplies the term containing $\omega_k = \frac{2\pi k}{T}$, and consequently the distribution on the frequency axis indeed changes. Taking into account this meaning of S_k we see that the spectral envelope is influenced by the pulse width but not the period.

The main lobe of the sinc function is between $-\pi$ and π, which here means between $\delta k = -1$ and $\delta k = 1$. Hence most of the energy is between $\omega_k = \frac{2\pi k}{T} = \pm\frac{2\pi}{\delta T}$, or otherwise stated, the frequency spread is $\Delta\omega = \frac{4\pi}{\delta T}$. The minimum spacing between two points in time that represent the same point on the periodic signal is obviously $\Delta t = T$. The relationship between the time and frequency spreads can therefore be expressed as

$$\Delta\omega\Delta t = \frac{4\pi}{\delta} \qquad (3.31)$$

which is called the 'time-frequency uncertainty product'. The effect of varying the duty cycle δ at constant period T is demonstrated in Figure 3.8. As δ is decreased the width of the spectrum increases (i.e., the spectral amplitudes become more constant) until finally at zero duty cycle (the signal being a periodic train of impulses) all the amplitudes are equal. If the duty cycle is increased to one (the signal becoming a constant $s(t) = A$), only the DC component remains nonzero.

What happens when the period T is increased, with δ constant? We know that the wider the spacing in the time domain, the narrower the spacing of the frequency components will be. The constancy of the time-frequency uncertainty product tells us that the extent of the sinc function on the frequency axis doesn't change, just the frequency resolution. This is demonstrated in Figure 3.9.

These characteristics of the FS of a rectangular wave are important in the design of pulse radar systems. We will discuss radar in more detail in Section 5.3, for now it is sufficient to assume the following simplistic model. The radar transmits a periodic train of short duration pulses, the period of which is called the **P**ulse **R**epetition **I**nterval (PRI); the reciprocal of the PRI is called the **P**ulse **R**epetition **F**requency (PRF).

This transmitted radar signal is reflected by a target and received back at the radar at this same PRI but offset by the round-trip time. Dividing

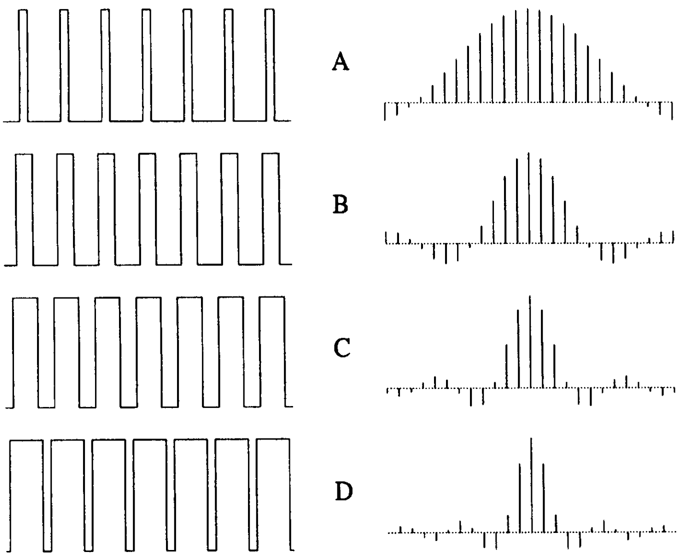

Figure 3.8: The effect of changing the duty cycle at constant period. In these figures we see on the left a periodic rectangular signal, and on the right the absolute squares of its FS amplitudes represented as vertical bars placed at the appropriate frequencies. (A) represents a duty cycle of 20%, (B) 40%, (C) 60% and (D) 80%. Note that when the duty cycle vanishes all amplitudes become equal, while when the signal becomes a constant, only the DC term remains.

the time offset by two and multiplying by the speed of radar waves (the speed of light c) we obtain the distance from radar to target. The round-trip time should be kept lower than the PRI; and echo returning after precisely the PRI is not received since the radar receiver is 'blanked' during transmission; if the round-trip time exceeds the PRI we get aliasing, just as in sampling analog signals. Hence we generally strive to use long PRIs so that the distance to even remote targets can be unambiguously determined. More sophisticated radars vary the PRI from pulse to pulse in order to disambiguate the range while keeping the echo from returning precisely when the next pulse is to be transmitted.

Due to the Doppler effect, the PRF of the reflection from target moving at velocity v is shifted from its nominal value.

$$\Delta\mathrm{PRF} = \mathrm{PRF}\,\frac{v}{c} \tag{3.32}$$

An approaching target is observed with PRF *higher* than that transmitted, while a receding target has a *lower* PRF. The PRF is conveniently found

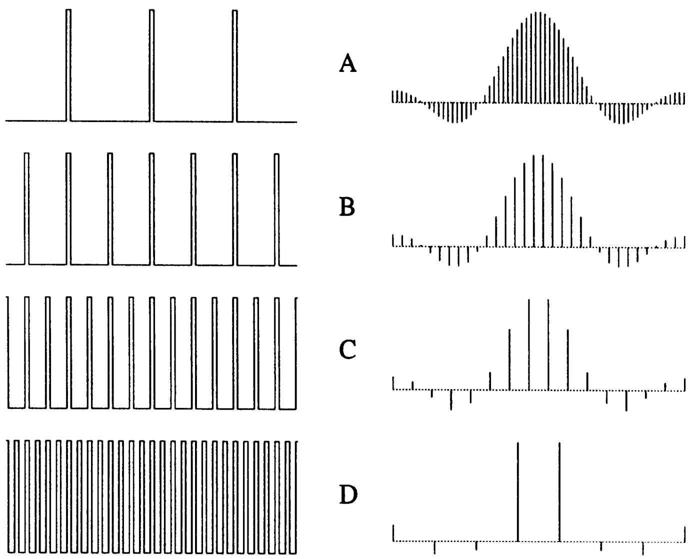

Figure 3.9: The effect of changing the period at constant duty cycle. In these figures we see on the left a periodic rectangular signal, and on the right the absolute squares of its FS amplitudes represented as vertical bars placed at the appropriate frequencies. As we progress from (A) through (D) the period is halved each time. Note that as the period is decreased with constant pulse width the frequency resolution decreases but the underlying sinc is unchanged.

using Fourier analysis techniques, with precise frequency determination favoring high PRF. Since the requirements of unambiguous range (high PRI) and precise velocity (high PRF) are mutually incompatible, simple pulse radars can not provide both simultaneously.

The radar signal is roughly a low duty cycle rectangular wave, and so its FS is approximately that of Figures 3.8 and 3.9. In order to maximize the probability of detecting the echo, we endeavor to transmit as much energy as possible, and thus desire wider pulses and higher duty cycles. Higher duty cycles entail both longer receiver blanking times and narrower sinc functions in the frequency domain. The former problem is easily understood but the latter may be more damaging. In the presence of interfering signals, such as reflections from 'clutter', intentional jamming, and coincidental use of the same spectral region by other services, the loss of significant spectral lines results in reduced target detection capability.

EXERCISES

3.8.1 Show how to regain the Fourier series of the square wave (equation (3.2)), from (3.30) by taking a 50% duty cycle.

3.8.2 We assumed that A in equation (3.28) was constant, independent of T and d. Alternative choices are also of interest. One could demand that the basic rectangle be of unit area $A = \frac{1}{d}$, or of unit energy $A = \frac{1}{\sqrt{d}}$, or that the power (energy per time) be unity $A = \frac{T}{\sqrt{d}}$. Explain the effect of the different choices on the signal and its FS when δ and T are varied.

3.8.3 Show that the FS of a train of impulses $s(t) = \sum \delta(t - kT)$ is a train of impulses in the frequency domain. How does this relate to the calculations of this section? To which choice of A does this correspond?

3.8.4 One technique that radar designers use to disambiguate longer ranges is PRI staggering. Staggering involves alternating between several PRIs. How does staggering help disambiguate? How should the PRIs be chosen to maximize the range? (Hint: Use the Chinese remainder theorem.)

3.8.5 What is the FS of a rectangular wave with stagger two (i.e., alternation between two periods T_1 and T_2)?

Bibliographical Notes

For historical background to the development of the concept of frequency consult [223]. Newton's account of the breaking up of white light into a spectrum of colors can be read in his book *Opticks* [179]. For more information on the colorful life of Fourier consult [83]. Incidentally, Marc Antoine Parseval was a royalist, who had to flee France for a while to avoid arrest by Napoleon. Lord Rayleigh, in his influential 1877 book on the theory of sound (started interestingly enough on a vacation to Egypt where Fourier lived eighty years earlier), was perhaps the first to call the trigonometric series by the name 'Fourier series'. Gibbs' presentation of his phenomenon is [74].

There are many books devoted entirely to Fourier series and transforms. To get more practice in the mechanics of Fourier analysis try [104]. In-depth discussion of the Dirichlet conditions can be found in the mathematical literature on Fourier analysis.