8 Fourier Series

8.1 Overview

Fourier or **harmonic analysis** involves the "decomposition" of a time series into a sum of sinusoidal components (sines and cosines). The decomposition is nothing more than fitting the n values of a time series with n sinusoids of specific ("Fourier") frequencies. Thus, Fourier analysis is simply a specific type of *interpolation*, but one that only holds for time series sampled at even intervals of the independent variable, giving special properties to the sinusoidal interpolant. The fitted function, given this particular basis, is known as a **Fourier series**, and because of the special properties of the sinusoids, their fit provides unique insights to the underlying structure and composition of the data series being analyzed. Sinusoids are nature's natural cycles, being nothing more than the motion of a spot on the side of rotating circle as it rolls along at a steady speed.[1]

When the data points of the time series are separated by even intervals of the independent variable (i.e., the data points are "***evenly spaced***"), the properties of the sinusoids (they become "orthogonal") allow the interpolation solution procedure to be greatly simplified. The simplified procedure is called the **discrete Fourier transform**, and it provides an extremely efficient method (applicable to any orthogonal basis, not just sinusoids) with which to perform the interpolation and determine the coefficient values multiplying the sinusoids, providing the interpolated fit. Because sinusoids are **periodic** functions, their actual characteristics needed to interpolate the time series are often of interest, and their coefficients are usually presented in the form of a **spectrum**.

As a problem in interpolation, Fourier analysis is a means for analyzing deterministic data; namely, the exact fit to the data, not taking into account the uncertainties and problems associated with noise in the data. When applying the tools and concepts of Fourier analysis to stochastic process data – that is recognizing that the data contain noise – additional considerations are adopted, as was the case in curve fitting: a best fit must be determined and the best-fit coefficients must be evaluated, as must the uncertainty in the fit of the sines and cosines. This involves **spectral analysis**. In other words

[1] You can draw a sinusoid by attaching a pencil to a wheel, pointing outwards: as the wheel is moved at a steady pace by an axle in its center, the movement of the wheel will create a perfect sinusoid. Or jam an indelible marker, drawing end out, into the outer edge of a tire on a car and drive at a constant speed along a wall (then please send photos: one of the sinusoid, and another of the car's owner).

(as used here), Fourier analysis represents the interpolation problem useful for deterministic series, whereas spectral analysis represents the curve-fitting problem useful for series containing noise.

Fourier sines and cosines are the natural eigen functions for complex natural systems,[2] which makes them ideal for decomposing such systems. They follow a number of rules that make analysis and interpretation of your time series exceptionally insightful. Fourier series are used in the solution of certain ordinary differential equations and have been studied and used in innumerable ways, making them well understood and a cornerstone of analysis.

In practice, Fourier and spectral analysis represent some of the most powerful and often used methods in data analysis in all fields of science, engineering, and mathematics. Also, the Fourier transform itself, like convolution and many of the other methodological tools forming the base of our analysis techniques, is invaluable in analytic problem-solving as well as data analysis.

8.2 Introduction

Fourier series involve a set of particular sines and cosines that combine to interpolate a data set or time series. We have already devoted time to solving the general interpolation problem, so you might question the need for spending additional time in developing this special case of interpolation. There are a couple of reasons: this special case is later shown to be applicable to interpolation (and least squares problems) using *any* orthogonal basis function for fitting the data – not just for particular sines and cosines. Therefore, the development of the Fourier series is in some respects a natural extension of the more general discussion of interpolation provided in Chapter 4. However, because of the considerable attention, significant insights, and intuitive comprehension provided by the sine and cosine functions, it makes sense to initiate this general development through the particular case of Fourier series and the Fourier transform.

Before developing the special form of solution appropriate for this interpolant using the Fourier transform, an overview of basic properties and definitions of periodic functions is required. Most fundamental is the concept of a periodic function.

8.3 Periodic Functions

8.3.1 Definitions and Concepts

A **periodic function** is one that repeats in time or length (i.e., its independent variable), so

$$f(t + T) = f(T + nT) = f(T) \tag{8.1}$$

where: T = the **period** of the function if it repeats in *time* (or it is the **wavelength** if it repeats in *space*) and

[2] This is explained in the chapter on Empirical Orthogonal Function analysis (Chapter 15).

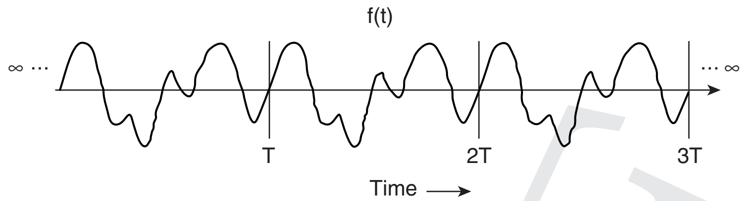

Figure 8.1 Example of a periodic time series that repeats every T units of time.

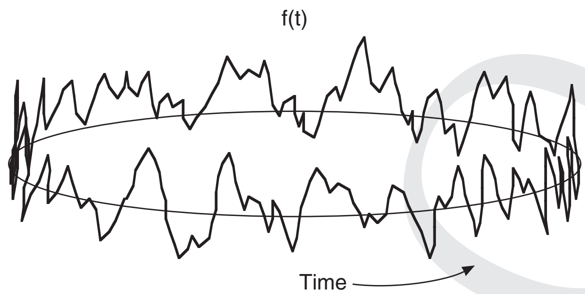

Figure 8.2 One complete period of a periodic time series, wrapped into a closed loop, so that the time series is infinitely long in both directions.

n = an integer, 0,1,2, . . .

The period, T, represents the length of the independent variable after which the time series is repeated. So, for example, Figure 8.1 shows a periodic time series that repeats every T units of time.

The first point in this series is repeated every T units of time. This is true for every point in the time series, so it satisfies the relationship given by (8.1) – an important consequence being: *in order to satisfy* (8.1) *exactly, the time series, f(t), must be infinitely long*. While it may seem esoteric, this generates an unfortunate, troublesome issue that we must later address.

As a general rule, when working with periodic time series, the domain of the independent variable is limited as

$$-\frac{T}{2} \le t < \frac{T}{2}. \tag{8.2}$$

In other words, the time series is typically presented so that the last point in the series is the point just before the first point repeats in the series, due to the periodicity over T. In this manner, each unique data point is represented just once within the domain represented by (8.2). Any point outside the domain, including the point at T/2, is already represented within the domain at some integral multiple of T (so the value at T/2 is already represented by the point at T/2 − T = −T/2).

Another way to look at periodicity, so that this restriction makes conceptual sense, is by considering that the data lie on a *closed loop*, where the circumference of the loop represents one complete period of the series. In other words, imagine a time series in a circular form, such as is shown in Figure 8.2.

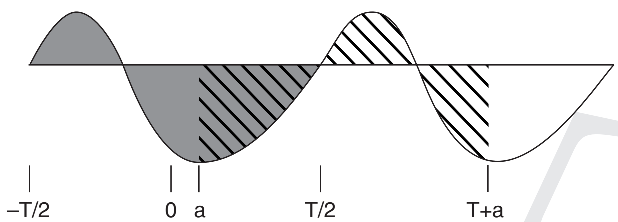

Figure 8.3 Integration over one period of a periodic time series. Both solid-filled and diagonal-line-filled regions give the same integral, since both are over a full period, though they start and stop at different points in the series.

In this respect, it is clear why one would not want the data series to include points at both −T/2 and at +T/2, since this would represent plotting 2 points at the same position on the loop. In other words, the last point in the series is always followed by the first point in the series, regardless of where you make the first point in the series. ***This cyclic representation should always be kept in mind when dealing with periodic data or with analyses that assume that the data are periodic (as discrete Fourier analysis does).*** To put this in standard linear form, you simply disconnect the loop at any point and straighten it out. In that form it would represent just one period of the series. This one period, however, completely represents the unique information of the entire (infinitely long) periodic time series.

For periodic time series, the integral over one full period is independent of where in the series the period begins. That is,

$$\int_{-T/2}^{T/2} \mathrm{f(t)dt} = \int_{-T/2+a}^{T/2+a} \mathrm{f(t)dt},$$

where *a* is a constant. This is shown graphically in Figure 8.3. The shaded region and hatched regions represent the same area because they both encompass one full period of the data.

Also,

$$\int_{a}^{b} \mathrm{f(t)dt} = \int_{a+\mathrm{T}}^{b+\mathrm{T}} \mathrm{f(t)dt}.$$

Sines and Cosines

The simplest periodic functions are the sine and cosine functions (Figure 8.4). These trigonometric functions can be written in a variety of forms. For example,

$$\mathrm{Asin}2\pi\mathrm{ft} \text{ or } \mathrm{Asin}(2\pi\mathrm{ft}+\varphi) \text{ or } \mathrm{Asin}[2\pi\mathrm{f}(\mathrm{t}+\mathrm{t}_{\varphi})] \tag{8.4a}$$

$$\mathrm{Acos}2\pi\mathrm{ft} \text{ or } \mathrm{Acos}(2\pi\mathrm{ft}+\varphi) \text{ or } \mathrm{Acos}[2\pi\mathrm{f}(\mathrm{t}+\mathrm{t}_{\varphi})] \tag{8.4b}$$

where: A = **amplitude** (height from 0) of the sine or cosine;

f = **frequency** (1/T; T = period or wavelength).

Sometimes, the phrase **rotational frequency** is used for this term to avoid confusion with *angular frequency*:

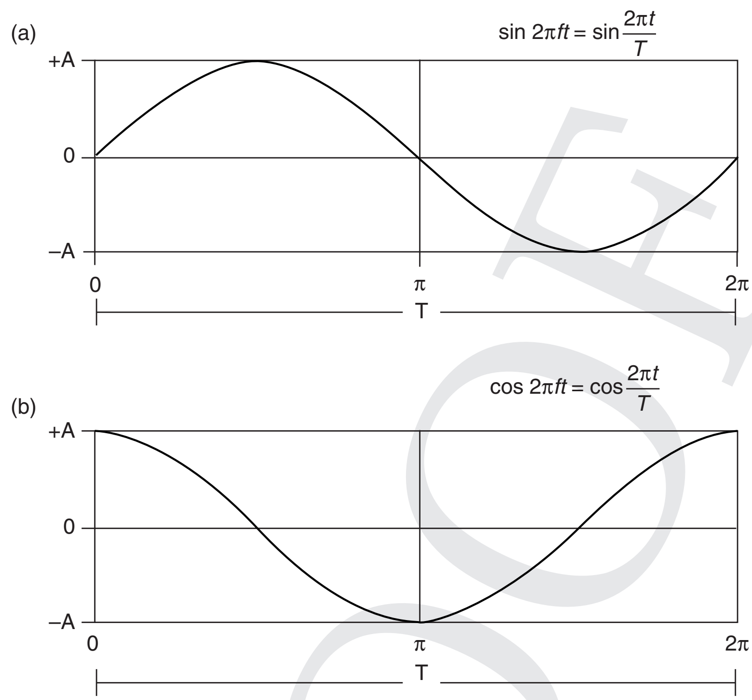

Figure 8.4 (**A**) A generic sine, with zero phase (or a cosine with phase angle $-\pi/2$), has amplitude, A = 1, and period, T. (**B**) A generic cosine, with zero phase and amplitude A = 1.

$\varphi = 2\pi f t_\varphi$ = **phase angle** or phase displacement;
t_φ = translation of the origin (constant displacement in t).
In addition, two additional convenient forms include

$$2\pi \mathrm{f} \;= \begin{cases} \omega \ (= \textbf{angular frequenc}\text{y if T is a period in T}\textit{ime}) \\ k \ (= \textbf{wavenumber} \text{ if T is a wavelength in } \textit{space}) \end{cases}$$

So, (8.4a) can be rewritten in terms of angular frequency or wavenumber as

$$\mathrm{A}\sin\omega \mathrm{t} \text{ or } \mathrm{A}\sin(\omega \mathrm{t} + \varphi) \text{ or } \mathrm{A}\sin[\omega(\mathrm{t} + \mathrm{t}_\varphi)] \tag{8.4c}$$

$$\mathrm{A}\sin k\mathrm{t} \text{ or } \mathrm{A}\sin(k\mathrm{t} + \varphi) \text{ or } \mathrm{A}\sin[k(\mathrm{t} + \mathrm{t}_\varphi)] \tag{8.4d}$$

In either case, ω or *k* is in radians, though it can be given in degrees or grads.

Note that the relationship between angular and rotational frequency, $\omega = 2\pi f$, is similar to that between natural log (ln) and log base 10 (log) where ln = 2.303log. In both cases (frequency and logs), the two forms are related by a constant and thus represent two slightly different means for presenting the same quantity. Which form is used is often simply a matter of convenience and comfort.

The specific position (time or length; ωt or κt) in the waveform is the **phase** (most common when dealing with propagating waves). Though related, do not confuse *phase* with *phase angle*, φ. The latter represents a displacement of the origin, while the former

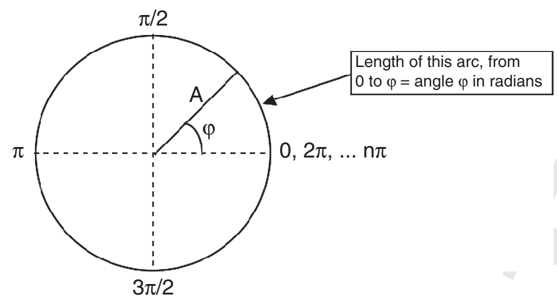

Figure 8.5 Polar form of a sinusoid, giving primary characteristics in terms of amplitude (= radius of circle), and phase angle as length of arc.

represents the location in the waveform (e.g., at its peak or trough). The two are related, since the phase angle is the phase at time 0. Some people do not differentiate between phase and phase angle, so make sure you are clear on how the word is being used. Here, both words are used interchangeably to mean phase angle.

Since the amplitude and phase determine the form of a sinusoid of any particular frequency, waveforms are often represented in **polar form**, in terms of amplitude (A) and phase (φ), as seen in Figure 8.5. Here, the radius of the circle is equal to the amplitude of the sinusoid and the phase is given by the arc length along the circumference of the circle. In this form, the drawn radius rotates around the circle at a rate of ω (or κ). Therefore, a sinusoid is a convenient representation of constant rotational motion, where the rotational rate is given by the angular frequency or wavenumber. The polar form readily conveys this interpretation.

Regardless of how they are presented, sines and cosines represent fundamental physical processes and mathematical functions, and, as such, they naturally arise in a variety of situations.

Even and Odd Functions

Cosines and sines represent simple examples of even and odd functions, respectively. An even function (e.g., cosine) is one that is symmetrical about the ordinate, while an odd function (e.g., sine) is antisymmetrical about the ordinate. Therefore,

$$\mathrm{f_e}(-x) = \mathrm{f_e}(x) \tag{8.5a}$$

$$\mathrm{f_o}(-x) = -\mathrm{f_o}(x), \tag{8.5b}$$

where $\mathrm{f_e}(\mathrm{x})$ is an even function and $\mathrm{f_o}(\mathrm{x})$ is an odd function.

Knowledge of the symmetry properties of a function often proves helpful when manipulating the function. Another useful property involves the integration of even and odd functions.[3] Specifically,

[3] For discrete time series, the integration properties also hold for summation ***only*** if the series is sampled at even increments of the independent variable.

$$\int_{-a}^{a} f_e(x)dx = 2\int_{0}^{a} f_e(x)dx \tag{8.6a}$$

$$\int_{-a}^{a} f_o(x)dx = 0. \tag{8.6b}$$

These properties are apparent upon simple inspection of cosine and sine functions drawn from ±2π.

In general, any function f(t), whether it displays any symmetry or not, can be decomposed into a sum of an even, $f_e(t)$, and odd, $f_o(t)$, function as follows:

$$\begin{aligned} f(t) &= \frac{1}{2}[f(t) + f(t) + f(-t) - f(-t)] \\ &= \underbrace{\frac{1}{2}[f(T) + f(-t)]}_{\text{even function}} + \underbrace{\frac{1}{2}[f(t) - f(-t)]}_{\text{odd function}} \\ &= \quad f_e(t) \quad + \quad f_o(t). \end{aligned} \tag{8.7}$$

It is easy to prove that (1/2)[f(t) + f(−t)] is an even function and (1/2)[f(t) − f(−t)] is an odd function: substitute −t for t in each. For the even function, the result is identical whether we use t or −t, since [f(t) + f(−t)] = [f(−t) + f(t)]; for the odd function [f(t) − f(−t)] = [f(−t) − f(t)] = −[f(t) − f(−t)]. Thus, $f_e(t) = f_e(-t)$ and $f_o(t) = -f_o(-t)$.

Basic Trigonometric Identities

The following identities are very useful for Fourier analysis and for understanding the nature of periodic time series:

$$\sin(x + y) = \cos(y)\sin(x) + \sin(y)\cos(x) \tag{8.8a}$$

$$\cos(x + y) = \cos(y)\cos(x) - \sin(y)\sin(x) \tag{8.8b}$$

$$\cos(x)\cos(y) = \frac{1}{2}[\cos(x + y) + \cos(x - y)] \tag{8.8c}$$

$$\cos(x)\sin(y) = \frac{1}{2}[\sin(x + y) + \sin(x - y)] \tag{8.8d}$$

$$\sin(x)\sin(y) = \frac{1}{2}[\cos(x - y) - \cos(x + y)] \tag{8.8e}$$

Because the cosine is an even function and the sine an odd function, the expansions of identities (8.8a) and (8.8b), for −y, as opposed to +y, are easily given as

$$\begin{aligned} \sin(x - y) &= \cos(-y)\sin(x) + \sin(-y)\cos(x) \\ &= \cos(y)\sin(x) - \sin(y)\cos(x) \end{aligned} \tag{8.9a}$$

and

$$\begin{aligned} \cos(x - y) &= \cos(-y)\cos(x) - \sin(-y)\sin(x) \\ &= \cos(y)\cos(x) + \sin(y)\sin(x) \end{aligned} \tag{8.9b}$$

Consider the standard cosine with angular frequency ω and phase angle, φ. From (8.8b),

$$\mathrm{A}\cos(\omega t+\varphi)=\mathrm{A}\cos(\varphi)\cos(\omega t)-\mathrm{A}\sin(\varphi)\sin(\omega t), \tag{8.10}$$

but since φ is a fixed angle (the initial phase of the sinusoid at time 0), Acos(φ) is a constant, as is Asin(φ), so (8.8a) can be rewritten as

$$\mathrm{A}\cos(\omega t+\varphi)=\mathrm{a}\cos(\omega t)-\mathrm{b}\sin(\omega t), \tag{8.11a}$$

where:

$$\mathrm{a} = \mathrm{A}\cos(\varphi) \tag{8.11b}$$

$$\mathrm{b} = \mathrm{A}\sin(\varphi) \tag{8.11c}$$

and, by inverting (8.11b,c),

$$\varphi = tan^{-1}\frac{\mathrm{b}}{\mathrm{a}} \tag{8.11d}$$

$$\mathrm{A} = (a^2+\mathrm{b}^2)^{1/2} \tag{8.11e}$$

Equations (8.11d,e) are easily obtained by combining (8.11b,c). Specifically, solve for the phase angle (8.11d) by dividing (8.11c) by (8.11b):

$$\frac{\sin(\varphi)}{\cos(\varphi)} = tan(\varphi) = \frac{\mathrm{b}}{\mathrm{a}} \tag{8.12a}$$

so

$$\varphi = tan^{-1}\frac{\mathrm{b}}{\mathrm{a}}.\ [4] \tag{8.12b}$$

Similarly, we can square (8.11b) and (8.11c) and add them to solve for the A,

$$\mathrm{a}^2+\mathrm{b}^2=\mathrm{A}^2[\cos^2(\varphi)\ +\sin^2(\varphi)], \tag{8.13a}$$

and, recalling that $\sin^2 x + \cos^2 x = 1$, yields

$$\mathrm{A} = (a^2+\mathrm{b}^2)^{1/2}. \tag{8.13b}$$

Therefore, you can easily switch between a single sinusoid with phase φ and amplitude A, on the one hand, and the sum of a pure sine and a pure cosine (i.e., sines and cosines without a phase shift) on the other.[5] This means that any single sinusoid can be decomposed as in Figure 8.6.

[4] The computation of the phase angle from the arc-tangent requires care on a computer, since the sign of a and b determine which quadrant the phase angle lies within, and some forms of the arc-tangent function do not differentiate this. For example, $\tan^{-1}$(-b/a) is the same as $\tan^{-1}$(b/-a), yet these two pairs of a and b represent uniquely different phase angles.

[5] Here *sinusoid* is used generically – that is, it represents either a sine or a cosine function.

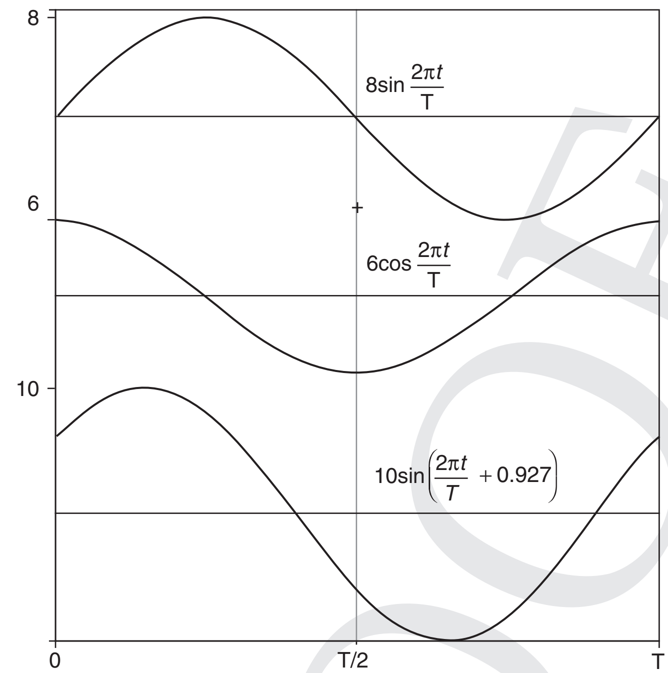

Figure 8.6 Adding the pure sine (amplitude 8) and pure cosine (amplitude 6), both having the same frequency and no phase angle, sum to give the lower sinusoid (thick line) of amplitude 10 and phase angle 0.927. This follows directly from equation (8.11a).

In Figure 8.6 the bottom (bold-line) sinusoid 10sin[(2πt/T) + 0.927] is synthesized (from 8.8a) as the sum of a pure cosine of amplitude a = 10sin(0.927) = 6, and a sine of amplitude b = 10cos(0.927) = 8. So, from (8.11a),

$$10\sin\left(\frac{2\pi T}{T}+0.927\right)=\underbrace{10\sin(0.927)}_{a}\cos\left(\frac{2\pi T}{T}\right)+\underbrace{10\cos(0.927)}_{b}\sin\left(\frac{2\pi T}{T}\right)$$
$$=6\cos\left(\frac{2\pi t}{T}\right)+8\sin\left(\frac{2\pi t}{T}\right).$$

Now consider adding two cosines of the same frequency but different amplitudes and phases. In that case, the trigonometric identity (8.8b) can be used in combination with (8.11) as

$$\begin{aligned}A_1\cos(\omega t+\varphi_1)+A_2\cos(\omega t+\varphi_2)&=A_1\cos(\varphi_1)\cos(\omega t)-A_1\sin(\varphi_1)\sin(\omega t)\\&\quad+(A_2\cos(\varphi_2)\cos(\omega t)-A_2\sin(\varphi_2)\sin(\omega t)\\&=(A_{1c}+A_{2c})\cos(\omega t)-(A_{1s}+A_{2s})\sin(\omega t)\\&=a\cos(\omega t)-b\sin(\omega t)\\&=A\cos(\omega t+\varphi)\end{aligned}\tag{8.15}$$

where:
$a = A_{1c}+A_{2c}, b = A_{1s}+A_{2s}$, $A_{1c} = A_1\cos(\varphi_1)$, $A_{2c} = A_2\cos(\varphi_2), A_{1s} = A_1\sin(\varphi_1)$, $A_{2s} = A_2\sin(\varphi_2)A_{2s} = A_2\sin(\varphi_2)$, and the amplitude and initial phase angle of the resulting cosine is $A = \left[(A_{1c}+A_{2c})^2 + (A_{1s}+A_{2s})^2\right]^{1/2}$, and $\varphi = \mathit{tan}^{-1}[(A_{1s}+A_{2s})/(A_{1c}+A_{2c})]$.

Therefore, the sum of the two cosines of the same frequency produces a new single cosine of that same frequency, ω, but of a different amplitude and phase relative to the two cosines that went into the sum. Similarly, from (8.11), a sine and cosine of the same frequency sum to produce a new single cosine of the same frequency but with a nonzero phase angle.

Pure Sinusoid as Stochastic Process

The assumption of stationarity has considerable implications in spectral analysis, which have led to the analysis focusing on the amplitude or power spectrum, with little emphasis (or attention at all) to the phase of the series. Specifically, consider a process that generates a time series with a perfect or near-perfect sinusoidal component in it. Your initial impression may be that the process is not stationary, since the mean of the series clearly changes with time in a predictable manner, save for some pesky noise. However, while it is true that the single realization of the process contains a predictable sinusoidal component, the process is still stationary, with little ability to predict the value of the time series at a particular time in another realization of the process.

For example, consider a rotating tank of water (rotation makes it behave a bit like the real world oceans on a rotating Earth) where a sensor reports the height of the fluid at a particular location each second of the experiment, which starts after we drop a large stone into the tank next to the tank wall. The signal is dominated by a wavelike pattern as the primary motion. For any particular experiment, the time series shows a rather regular and predictable pattern, giving the impression that the mean is changing with time according to the dominant waveform. It is true that for the one realization, you can gain an excellent understanding and characterization of the realization and make predictions of future times for it.[6]

However, when considering the ensemble average of all realizations that characterize the process itself, the mean height at the location does not change in time, but instead is zero for all times. Likewise, we do not gain an ability to predict the height at any particular time for future experiments in the tank, even if each experiment is dominated by the same general waveform. The reason is that in each experiment the phase of the dominant waveform is different, depending on subtle differences in the initial conditions (e.g., exactly where the object was dropped into the water relative to the wall), spin-up characteristics (time required to spin to a stable state), how synchronous the start time is with the impact of the object, etc. Consequently, if we average together all of the different waveforms from the different experiments, we find that the changing phase leads to cancellation of the different peaks and troughs leading to a zero mean – that is, the *process* is stationary and *stochastic*.

[6] In general, spectra provide a very poor means for prediction, except when the time series is dominated by perfect sinusoidal components (with stable phase through time), in which case prediction can be quite good.

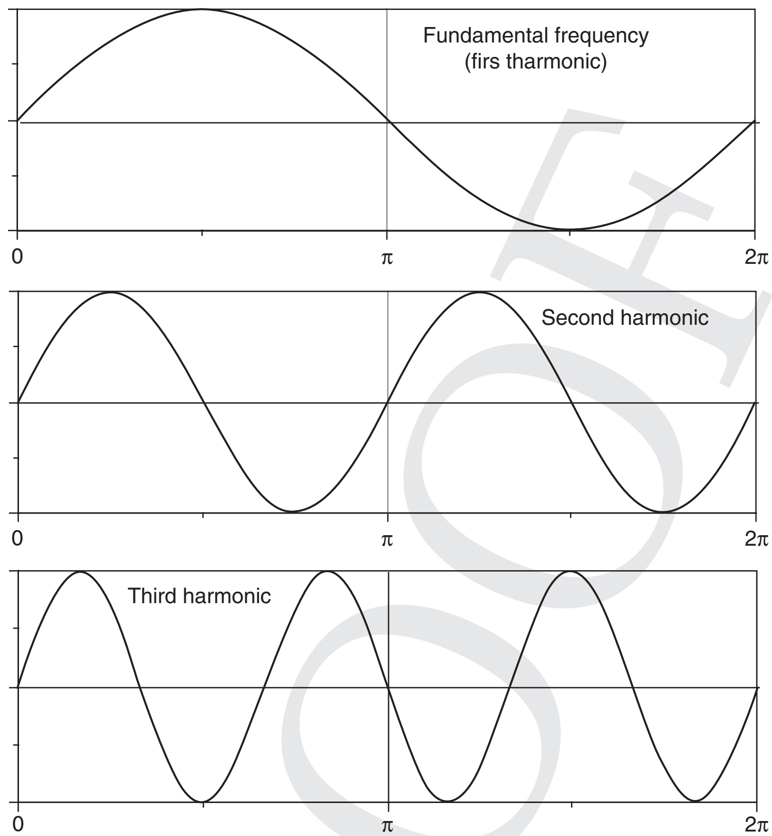

Figure 8.7 Example of first three harmonics for a series of the same length. Independent of initial phase, the harmonics make one, two, and three full cycles over the length of the series.

Harmonics

Harmonics or **overtones** are sines and cosines with frequencies that are integer multiples of a fundamental frequency.

The **fundamental frequency** (or **first harmonic**) is given by $f_1 = 1/T$. This waveform completes one full cycle over the period T, and since this is independent of phase, it is irrelevant whether it is a sine or cosine. In Figure 8.7 this is demonstrated by a sine (we could have an initial phase shift and the relationship would still be the same).

The **second harmonic**, f_2, is the sinusoid that makes two complete cycles or oscillations over period T. Its frequency is twice that of the first harmonic, since it makes twice as many complete oscillations in the same period, so $f_2 = 2f_1$.

The **third harmonic**, f_3, makes three complete oscillations over the same period T, so $f_3 = 3f_1$.

Continuing in this same manner (I'm hoping you see the pattern here), the ***n*th harmonic**, f_n, makes n complete oscillations over the period T, and $f_n = nf_1$.

Therefore,

$$
\begin{aligned}
f_1 &= f_F = \frac{1}{T};\ \omega_1 = 2\pi f_1 \\
f_2 &= 2f_1 = \frac{2}{T};\ \omega_2 = 4\pi f_1 \\
f_3 &= 3f_1 = \frac{1}{T};\ \omega_3 = 6\pi f_1 \\
&\vdots \\
f_n &= nf_1 = \frac{n}{T};\ \omega_n = 2n\pi f_1.^{7}
\end{aligned}
\tag{8.16}
$$

The linear superposition of harmonics (i.e., waveforms that are an integer multiple of the fundamental frequency 1/T) always produce a function that is periodic over T.

Modulation

When two sinusoids (sines or cosines) of different frequencies are linearly combined, they produce a composite signal that has a periodic pinching and swelling of the entire time series, known as a **beat**. If the frequencies of the sinusoids being combined are similar, this "beating" can be considerable.

Specifically, consider

$$y(t) = \cos\omega_1 t + \cos\omega_2 t, \tag{8.17}$$

and rewrite the two angular frequencies as

$$\omega_1 = \overline{\omega} + \delta\omega \tag{8.18a}$$

$$\omega_2 = \overline{\omega} - \delta\omega \tag{8.18b}$$

Then

$$\delta\omega = \frac{\omega_1 - \omega_2}{2} \tag{8.19a}$$

$$\overline{\omega} = \frac{\omega_1 + \omega_2}{2} \tag{8.19b}$$

Substituting (8.18) into (8.17) gives

$$y(t) = \cos(\overline{\omega}t + \delta\omega t) + \cos(\overline{\omega}t - \delta\omega t), \tag{8.20}$$

and applying (8.8c) gives

$$y(t) = 2\cos(\delta\omega t)\cos(\overline{\omega}t). \tag{8.21}$$

[7] Note that some people (acousticians, and Courant, for example) consider that the harmonics are the higher frequency forms of the fundamental frequency. This is consistent with the physics of harmonics in instruments. Given this definition, the fundamental frequency is not considered to be a harmonic at all, and consequently they call $2f_F$ the first harmonic; $3f_F$ the second harmonic, etc. This notation is thus offset from the present (mathematically more common) notation by one. Therefore, you must be careful to clearly indicate how you are defining the terms "first harmonic," etc., and be equally careful in making sure of how others define it (this can get very confusing sometimes, when people start talking about the first harmonic without differentiating whether they are referring to the fundamental frequency or $2f_F$). Unfortunately, both forms have a certain intuitive appeal. I will consistently use the notation as defined here in (8.16).

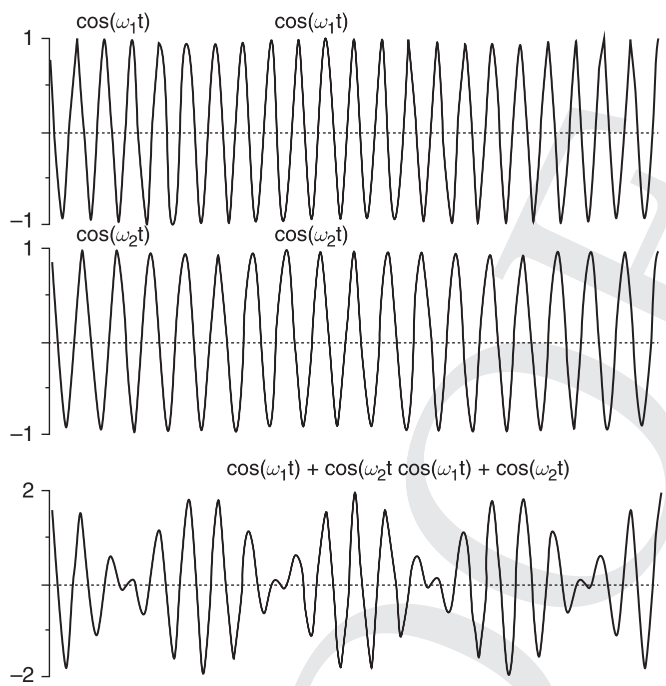

Figure 8.8 The addition of two cosines of similar frequency, giving a modulated cosine with the average frequency.

Examination of this equation shows that the second cosine term, $\cos(\varpi t)$, with period T = $2\pi/\varpi$ has a *time-varying amplitude* given by cos(δωt) which itself has a period of T = 2π/δω.[8] When one function multiplies another, it is said to **modulate** the function it multiplies. That is, in the example just given, the function 2cos(δωt) modulates the function $\cos(\varpi t)$. Therefore, 2cos(δωt) is the **modulation function**, and it essentially forms an **envelope** within which the modulated function is contained. In the case of sinusoids, since the modulation function is periodic it is called a beat, and its frequency of modulation is called the **beat frequency**.

For example, consider two cosines with frequencies given by $\omega_1 = 11\pi/5$ and $\omega_2 = 9\pi/5$. The addition of these two cosines is shown in Figure 8.8. Alternatively, according to (8.21), this is equivalent to multiplying two cosines of the form $2\cos(\delta\omega t)\cos(\varpi t)$, where $\delta\omega = (\omega_1 - \omega_2)/2$ and $\varpi = (\omega_1 + \omega_2)/2$ that product is shown in Figure 8.9.

In those figures, the concept of modulation is clearly seen. The product of the cosines "modulates," as shown by the bold lines forming the envelope of the modulated function, $\cos(\varpi t)$, in Figure 8.9. The period of oscillation (or modulation) of the envelope, or beat frequency, is given as $T_{beat} = \pi/\delta\omega$, while the period of the modulated series is $T = 2\pi/\varpi$. The beat frequency is one-half the frequency of the modulation function. That is, the $\cos(\varpi t)$ is modulated by a function 2cos(δωt), which has a frequency of 2π/δω. The beat frequency is one-half of this modulation frequency, or π/δω. This reflects that fact that the modulation makes complete cycles of the modulated

[8] The period is determined by rewriting 2cos(δωt) in the standard angular frequency form as 2cos(2πt/T), so 2πt/T = δωt and T = 2π/δω.

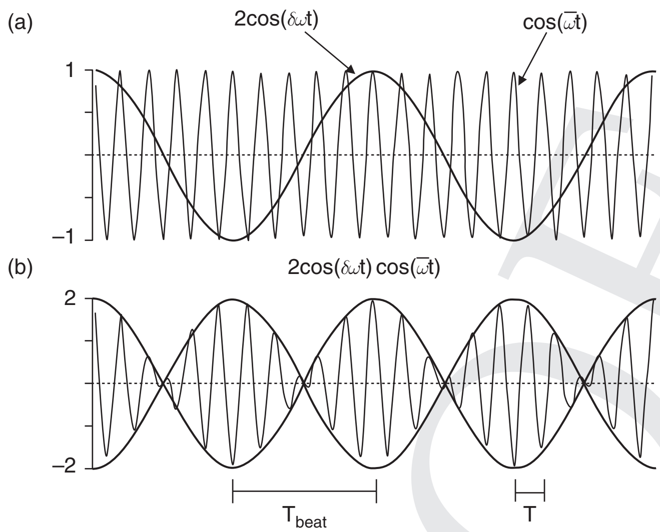

Figure 8.9 (**A**) Average frequency cosine, $\cos(\overline{\omega}t)$, and difference frequencies (the modulating function, $2\cos(\delta\omega t)$, bold line superimposed on the average frequency series) of Figure 8.8 are multiplied (as in 8.21). This multiplication gives the same modulated series as in Figure 8.8, where the two original cosines are added together. (**B**) The envelope is clearly outlined here as the solid bold line in the modulated series.

waveform over each half cycle, because it is symmetrical about the x-axis (abscissa), as seen in the figure.

Also, the relative spacing between peaks in the modulated packets is not regular. Careful examination of the spacing reveals a switch in phase at the period of the beat. Specifically, the peaks are periodic between the zero nodes of the envelope, with the period of the modulated series $T = 2\pi/\overline{\omega}$. But when crossing the zero nodes of the envelope, there is a 180° phase shift. Therefore, each "packet" of the modulated waveform is essentially an inverted form of the preceding (and following) packets. This is due to the x-axis symmetry of the modulating function (i.e., the modulating cosine oscillates symmetrically about zero, flipping the function it modulates when it goes negative).

As $\delta\omega$ gets smaller (i.e., as the two frequencies get closer together), the period of the beat curve gets longer and the phenomena becomes more noticeable. In acoustics, often two frequencies ω_1 and ω_2 are too high to hear, but the beat is within an audible range.

8.4 Fourier Series

8.4.1 Interpolation with Fourier Sines and Cosines

A **Fourier series** is that sequence of sines and cosines that interpolates the specific time series.

Consider fitting a series of *n* data points, y_i, with a set of sines and cosines that are periodic over the length of the time series, that length given by T. This is a standard continuous (global) interpolation problem, as we've previously examined, and it is given as

$$a_0\cos 0\omega t_i + b_0 \sin 0\omega t_i + a_1 \cos 1\omega t_i + b_1 \sin 1\omega t_i + a_2 \cos 2\omega t_i + b_2 \sin 2\omega t_i + \ldots = y_i, \quad (8.22a)$$

[9]

where ω = 2π/T and n terms are included in the sum, so there are as many terms of the basis as there are data points.

The first two terms reduce to a_0 (since cos0 = 1) and 0 (since sin0 = 0), so b_0 drops out completely, requiring that an additional term be added at the end of the series in order to have n terms for n data points.[10] Thus,

$$a + a_1 \cos 1\omega t_i + b_1 \sin 1\omega t_i + a_2 \cos 2\omega t_i + b_2 \sin 2\omega t_i + \ldots = y_i. \quad (8.22b)$$

Written as a system,

$$\begin{aligned} a + a_1 \cos 1\omega t_1 + b_1 \sin 1\omega t_1 + a_2 \cos 2\omega t_1 + b_2 \sin 2\omega t_1 + \ldots &= y_1 \\ a + a_1 \cos 1\omega t_2 + b_1 \sin 1\omega t_2 + a_2 \cos 2\omega t_2 + b_2 \sin 2\omega t_2 + \ldots &= y_2 \\ &\vdots \\ a + a_1 \cos 1\omega t_n + b_1 \sin 1\omega t_n + a_2 \cos 2\omega t_n + b_2 \sin 2\omega t_n + \ldots &= y_n \end{aligned} \quad (8.23a)$$

or

$$a_0 + \sum_{j=1}^{\le n/2} [a_j \cos(j\omega t_i) + b_j \sin(j\omega t_i)] = y_i. \quad (8.23b)$$

This sum of sines and cosines of increasing harmonic terms is called a **Fourier series**.[11] The a_j and b_j coefficients are called the **Fourier coefficients**. Once we solve for the Fourier coefficients, the sines and cosines just discussed superimpose to interpolate the time series in the standard way.

The Fourier series gives n equations in n unknowns and can be written in standard matrix form:

$$\begin{bmatrix} 1 & \cos\omega t_1 & \sin\omega t_1 & \cos 2\omega t_1 & \sin 2\omega t_1 & \cos 3\omega t_1 & \sin 3\omega t_1 & \ldots \\ 1 & \cos\omega t_2 & \sin\omega t_2 & \cos 2\omega t_2 & \sin 2\omega t_2 & \cos 3\omega t_2 & \sin 3\omega t_2 & \\ & & & \vdots & & & & \\ 1 & \cos\omega t_n & \sin\omega t_n & \cos 2\omega t_n & \sin 2\omega t_n & \cos 3\omega t_n & \sin 3\omega t_n & \end{bmatrix} \begin{bmatrix} a_0 \\ a_1 \\ b_1 \\ a_2 \\ b_2 \\ b_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

[9] This series can be written as

$$A_0\cos(0\omega t_i + \varphi) + A_1\cos(1\omega t_i + \varphi) + A_2\cos(2\omega t_i + \varphi) + \ldots = y_i \quad (F8.1)$$

However, recall that for the interpolation problem, we wish to establish a system of the form **Ax** = **b** – that is, one that is linear in the unknown coefficients. That linear form is easily achieved by employing the trigonometric identity, (8.8b), which allows expansion into the pure sines and cosines given in (8.22).

[10] A more thorough discussion regarding these first two terms appears later.

[11] The limit of the sum, as given in (8.23b), is discussed in more detail and given more precisely in the next section.

This system, $\mathbf{A}x = b$, is solved for the n coefficients of this sine and cosine series in the usual (matrix) way for interpolation ($x = \mathbf{A}^{-1}b$).

8.4.2 Interpreting the Fourier Series

Assume for the moment that we have solved the system given in (8.24) for the a_j and b_j coefficients that cause the sum of sines and cosines given in (8.22) to interpolate the n data points. We can then construct the pure sines and cosines with these amplitudes. Alternatively, we can combine them into cosines (dropping the sines) with amplitude and phase according to (8.11a), where the amplitude of each cosine harmonic, i, is given as $(a_i + b_i)^{1/2}$ and its phase $tan^{-1}(b_i/a_i)$.

Since each sine and cosine varies about zero, the series has a zero mean, unless the $a_0 \neq 0$. That coefficient is the mean of the final interpolated series – all of the sines and cosines vary about this value.[12]

The first cosine and sine terms represent the fundamental frequency, or first harmonic. Therefore, they both complete one full cycle over the length of the time series at amplitudes determined by the values of a_1 and b_1 (and with no phase displacement).

For the example of Figure 8.10 where $n = 33$ data points, we have a mean (a_0 coefficient) and 16 harmonic terms, each with an amplitude and phase, the combination of the a_j and b_j coefficients, that together sum to exactly fit the data. The a_0 coefficient, representing the mean of the series is not shown in the figure.

This example shows that the irregularly shaped time series shown at the bottom of the Figure 8.10 (and discretely sampled at 33 positions) is completely decomposed into a series of periodic functions (cosines). If you are interested in examining whether the time series represents a periodic process to some degree, then examination of these components for particularly large amplitude harmonic terms may be an indication of underlying periodic components.

Amplitude Spectrum

In the previous figure, the components of the interpolant (i.e., all of the individual cosines in the fitted curve) were explicitly plotted to show the characteristics of each cosine. While that is a perfectly acceptable manner in which to view the components of the interpolant, that same information is more readily displayed by plotting the amplitude of each harmonic term (i.e., the A_j for each cosine, recalling $A_i = (a_i + b_i)^{1/2}$) as a function of the particular harmonic, j. Such a plot is called the **spectrum,** or, more explicitly, the **amplitude spectrum**.[13]

For this example, the spectrum is shown in Figure 8.11.

[12] This is not true for discrete data that are separated at irregular intervals of time or space.

[13] There are actually many forms with which this graph can be displayed. For example, the amplitude might be presented as the A_j values, as the squared values, or as the squared values divided by the total variance of the time series. The harmonic may be presented as the frequency, period, or angular frequency. Also, the axes of the graph may be linear, log, log-linear, etc. Each of these subtly different forms may go by a different name and emphasize a different aspect of the fitted cosines, a matter that is discussed in context later. However, regardless of the specific form of the display, the general information content is the same – the amplitude of the various cosines that interpolate the time series are displayed as a function of the different harmonics.

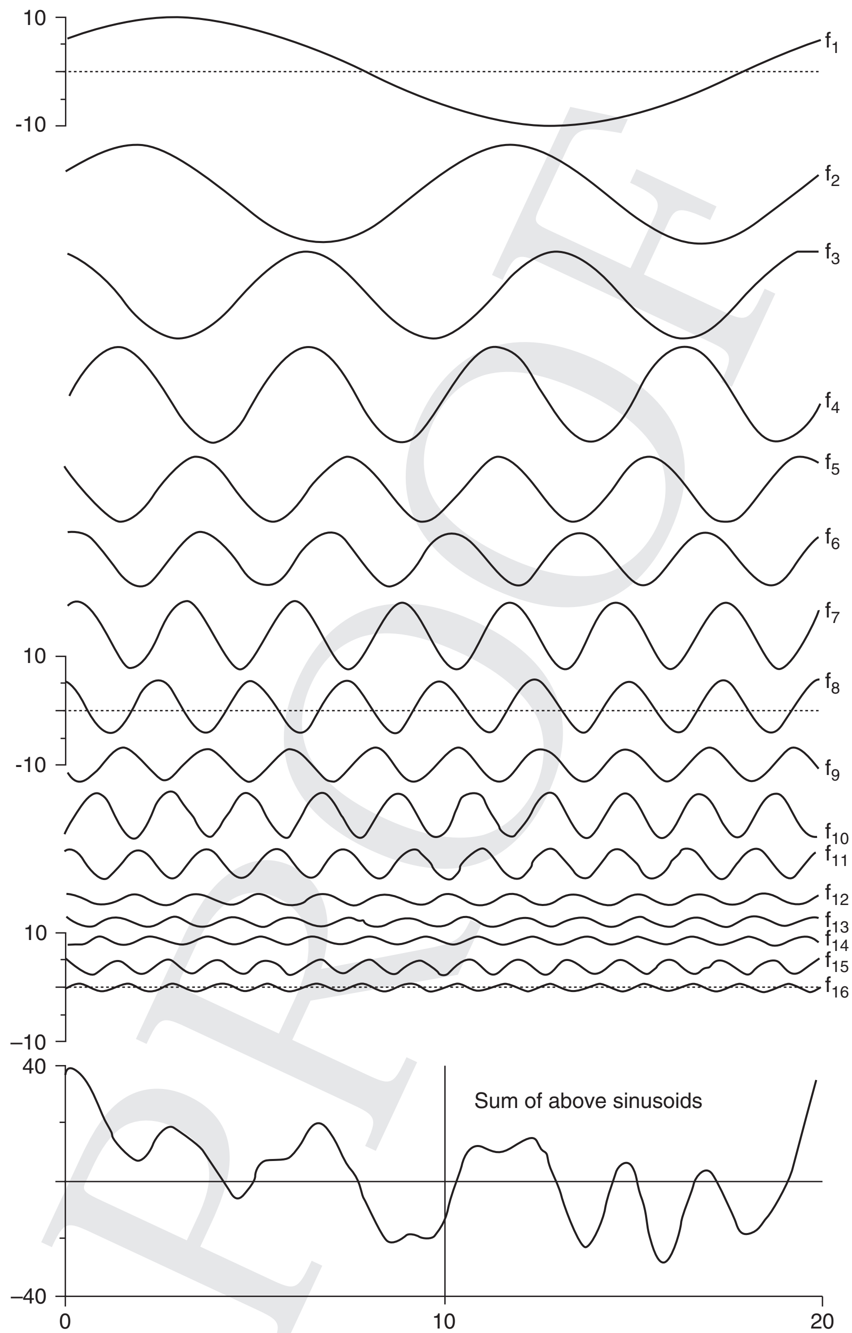

Figure 8.10 Sixteen individual cosines (one for each harmonic), with amplitude and phase, sum to give the more complicated time series at the base of the cosines.

This shows the amplitudes of the various cosines added in Figure 8.10. The harmonic number simply represents the harmonic to which the associated amplitude corresponds. So, for example, harmonic 1 signifies the first (or fundamental) harmonic, represented in

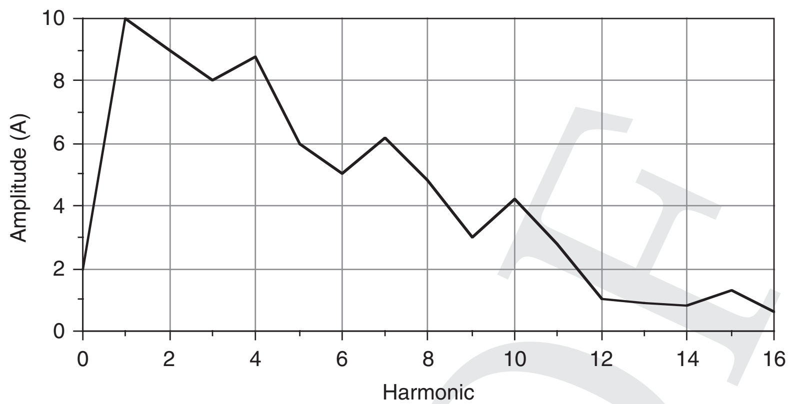

Figure 8.11 Plot of the amplitude of each of the 16 fitted cosines in the example of Figure 8.10.

Figure 8.10 by f_1 (which has an amplitude of 10). The harmonic labeled 0 corresponds to the a_0 coefficient, which represents the mean, as previously discussed.

The higher harmonic terms in this example have relatively small amplitudes, and there is no single harmonic that overwhelmingly dominates. Spectra that show a general dominance of the lower-order harmonics (low-frequency components) with smaller and smaller amplitudes at the higher frequencies are termed **red spectra**.[14] Such spectra are widespread in the physical sciences and represent a particular form of random noise. We define red spectra more rigorously later, as well as numerous variations on this plot, each of which has specific desirable characteristics.

Phase Spectrum

A similar plot can be made to display the phase angle (or displacement) for each of the harmonics. Such a plot is called the **phase spectrum**.[15] The phase spectrum for the example of Figure 8.10, is shown in Figure 8.12.

In general, phase spectra are more erratic than amplitude spectra, and as discussed later regarding stochastic processes, they are often ignored.

8.5 Take-Home Points

1. Series that are periodic over their length (T) are ones in which the next point that implicitly follows the last point in the series is equal to the first point in the series. Or, $f(t) = f(t + T) = f(t + nT)$. T is the period, but if the series varies in space, T is the wavelength.

[14] Red spectra have a very specific slope in the amplitude spectrum, so there are other "colors" often applied to spectra dominated by lower-frequency components, such as pink, etc., according to their actual slope. More about this later in Chapter 11, Spectral Analysis.

[15] Caution must be used here, since the term *phase spectrum* formally describes a plot of phase differences between two time series in *cross spectral analysis*, a topic that is covered later. As always, considerable care must be taken to avoid confusion in the use of the terms.

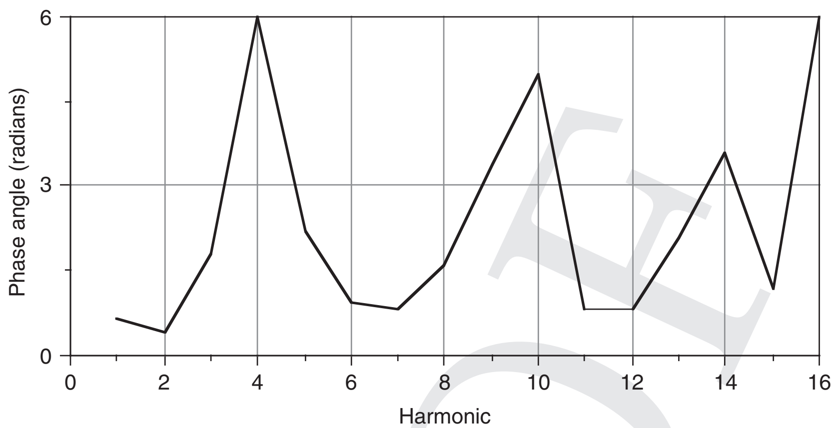

Figure 8.12 Phase angle for each harmonic of the example in Figure 8.10.

2. Sines and cosines are the most natural periodic functions, repeating every $2n\pi$, where n is any integer.
3. Any sinusoid with rotational frequency (f = 1/T), amplitude (A) and phase (φ), can be written as $A\cos(2\pi f + \varphi) = +a\cos(2\pi f) + b\sin(2\pi f)$, where $A = (a+b)^{1/2}$, $\varphi = tan^{-1}(b/a)$. Or, in terms of angular frequency, $\omega = 2\pi/T$.
4. Harmonics are frequencies that are integer multiples of a fundamental frequency. The fundamental frequency is $f_1 = 1/T$, i.e., it completes one full cycle over the length of the series. Second harmonic is $f_2 = 2f_1 = 2/T$, so it makes two full cycles over the length of the series. Generalized, the nth harmonic is $f_n = nf_1 = n/T$, making n full cycles over the length of the series.
5. A Fourier series is a sequence of sines and cosines that interpolates (exactly fits) a specific time series (y_i), written as $a_0 + \sum_{j=1}^{\leq n/2} [a_j \cos(j\omega t_i) + b_j \sin(j\omega t_i)] = y_i$. Fundamental to this series is that the harmonics are specific to the length (T) of the series being fit. Different series with different lengths have different "Fourier" harmonics.
6. Amplitude spectrum is the presentation of the a and b (Fourier) coefficients combined as an amplitude of the interpolated Fourier series as a function of harmonic, or A(ω), where $A = (a+b)^{1/2}$.

8.6 Questions

Pencil and Paper Questions

1. For a time series y_i of four evenly spaced data points (sampling interval Δt = 1s), write out each term of its Fourier series (actually write out each term of the sum). Show the exact solution for the a_i and b_i coefficients. What is the sampling rotational frequency?
2. Write out the Fourier series for a time series containing 100 data points, with spacing at Δt = 1s in standard matrix form (**Ax** = **b**), showing the first five and last three rows of the **A** matrix and **x** and **b** vectors, using angular frequency ω_i, including all details explicitly.